On the Capacity of Information Networks

January 28, 2005

April Rasala Lehman

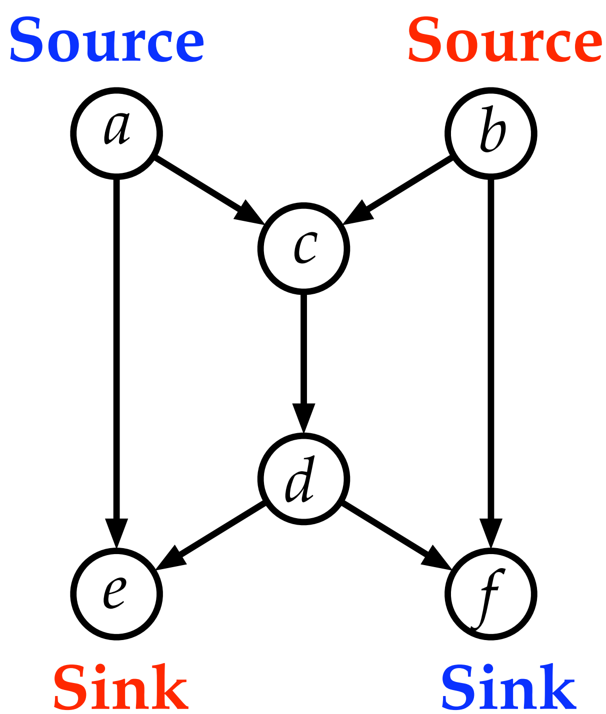
Joint work with Nicholas Harvey, Robert Kleinberg and Eric Lehman
MIT

"There is as yet no unified theory of network information flow. But there can be no doubt that a complete theory of communication networks would have wide implications for the theory of communication and computation."
- Cover & Thomas, *Elements of Information Theory*.

History of Network Coding

- Breakthrough [Ahlswede et al. '00].
  - Existence of multicast solution depends on min-cut condition.
- Algebraic framework [Koetter & Médard '03].
  - Led to a randomized, distributed, fault-tolerant algorithm for multicast [Ho et al. '03].
- Deterministic algorithms for multicast [Jaggi et al. '03, Harvey et al. '05].

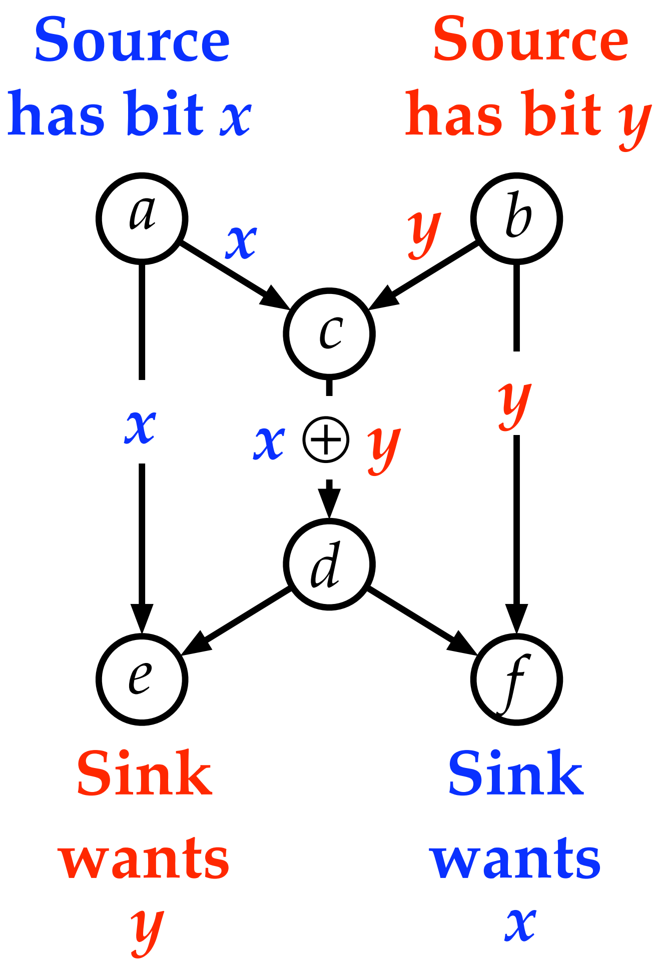
The Network Coding Problem



Given:

- Directed acyclic graph G.
- Integral capacity $c(u, v)$ for each edge (u, v).
- k-commodities:
  - Set of source nodes.
  - Set of sink nodes.

The Idea of Network Coding



- There is one message for each commodity.
  - Every source knows the message.
  - Every sink wants the message.
  - A message is a single symbol from an alphabet Σ.
- Each edge of capacity c can transmit c symbols from Σ.
- Question: Does there exist a solution?

This Talk: from Existence to Optimization

- Consider size of alphabet Σ.
  - Model of network coding that works for multicast doesn't work well in general.
  - Need a notion of "rate".
- What is the maximum achievable communication rate in a network?
  - Explore bounds based on cut conditions.
  - Develop entropy inequalities based on graph structure.
- What is the maximum rate in an *undirected* network?

Alphabet Size

Who Cares About Alphabet Size?

- Small alphabet means simple, efficiently-computable edge functions.
- Large alphabet implies large latency.
- Need $\Omega(\log |\Sigma|)$ bits of memory at each node to compute edge functions (naively).
- An upper bound on $|\Sigma|$ would imply that the network coding problem is decidable.

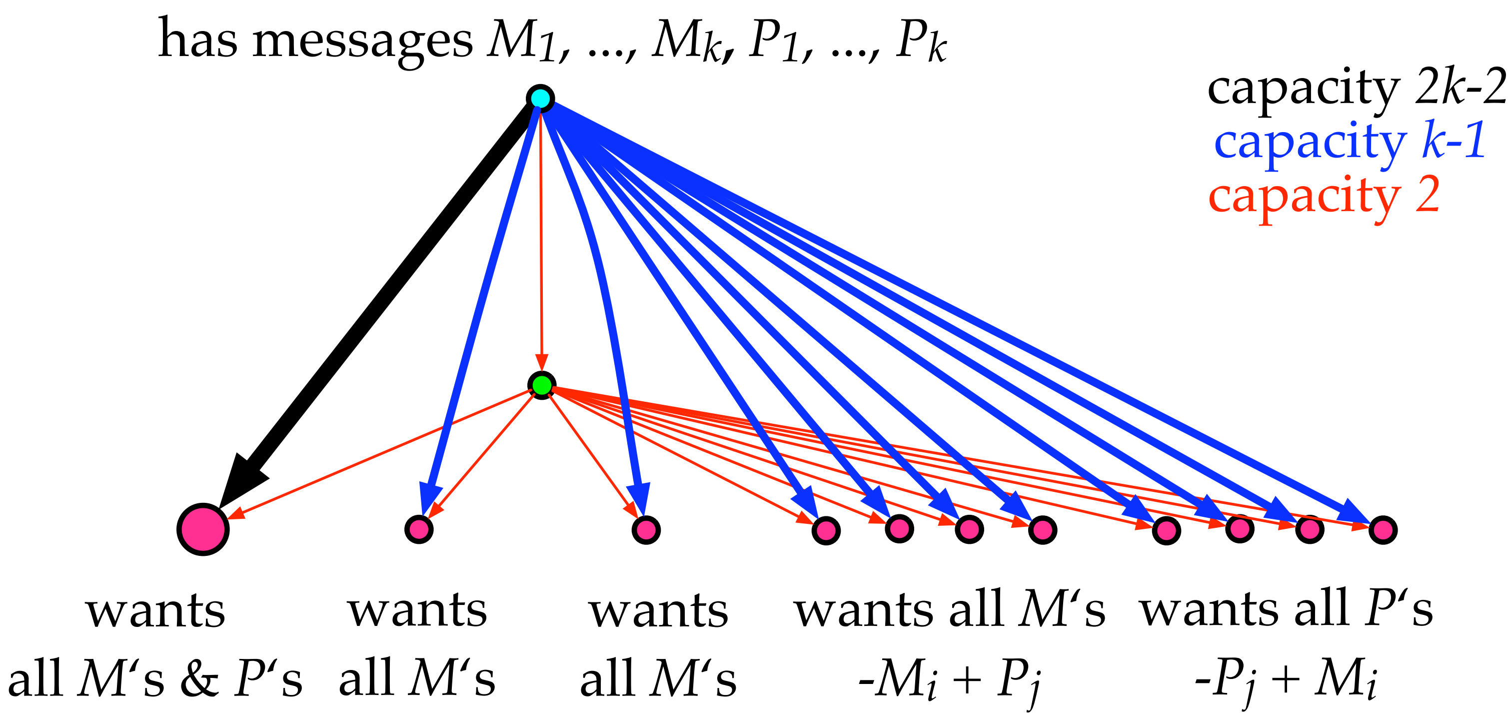
Our Results - The Bad News

- Sometimes an *enormous* alphabet is required!
  - An n-node network may require an alphabet of size:

$$|\Sigma| = 2^{e^{\Omega(n^{1/3})}}$$

  - Solution may exist but be hopelessly unwieldy!
- Nonmonotonicity:
  - Instance solvable with 4-symbol alphabet, but not with 1000-symbol alphabet!
  - Can't fix a single large alphabet size, e.g. 2^{64}.

Building Block: Network I_k



Lemma 1 *Solvable iff* $|\Sigma| = q^k$.

Doubly-Exponential Lower Bound

- Network I_k has $O(k^2)$ nodes and requires $|\Sigma|$ to be a perfect k-th power.
- Let J_n consist of disjoint networks

$$I_2 \quad I_3 \quad I_5 \quad I_7 \quad I_{11} \quad \dots \quad I_p$$

where p is largest prime less than $n^{1/3}$.

$\Rightarrow$ J_n has $O(n)$ nodes and there is a solution if and only if:

$$|\Sigma| = C^{2 \cdot 3 \cdot 5 \cdot 7 \cdot 11 \cdots p} = C^{e^{\Omega(n^{1/3})}} \geq 2^{e^{\Omega(n^{1/3})}}$$

Our Results - The Good News

If each edge can send *one* additional bit,
then the minimum alphabet size is $O(1)$.

- Our bad example is an artifact of using the network at 100.0% capacity.
- Are we wasting our time with this model?
- Tweak the model?
  - Messages are drawn from an alphabet Γ.
  - Each edge transmits one symbol from larger alphabet Σ.
  - Rate $= \frac{\log|\Gamma|}{\log|\Sigma|}$.

What is the Maximum Achievable Rate?

What is the Maximum Achievable Rate?

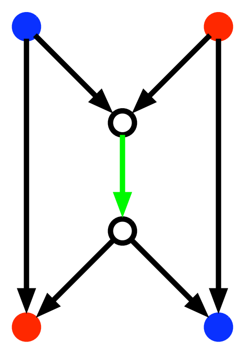
- Open problem except for multicast where max rate = min-cut between the source and any sink.
- Is there a cut-based upper bound on rate for the general problem?
- Do information theoretic tools give a better upper bound?

Sparsity

- Sparsity of a cut $A \subseteq E$ is:

$$\frac{\text{capacity of edges in cut } A}{\#\text{ commodities with no remaining source-sink path}}$$

- Sparsity of a graph is minimum sparsity over all cuts.
- There exist directed graphs in which the maximum rate $>$ sparsity.



Sparsity = 1/2
Rate = 1

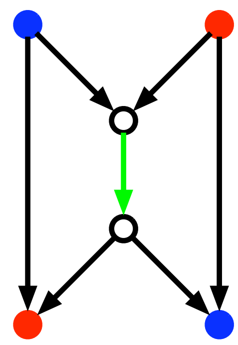
Meagerness

- A set of commodities P is *separated* by a cut if there is no remaining path from a source of *any* commodity in P to a sink of *any* commodity in P.

- The *meagerness* of a graph is the minimum over all sets of commodities P and cuts that separate P of

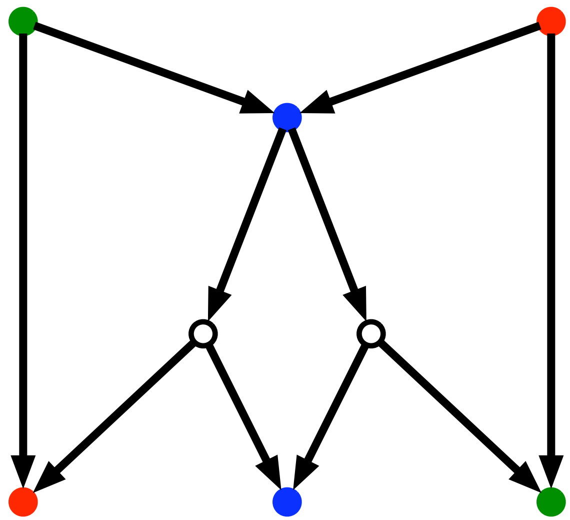
$$\frac{\text{capacity of edges in cut}}{|P|}$$

- The maximum rate $\leq$ meagerness in directed graphs.

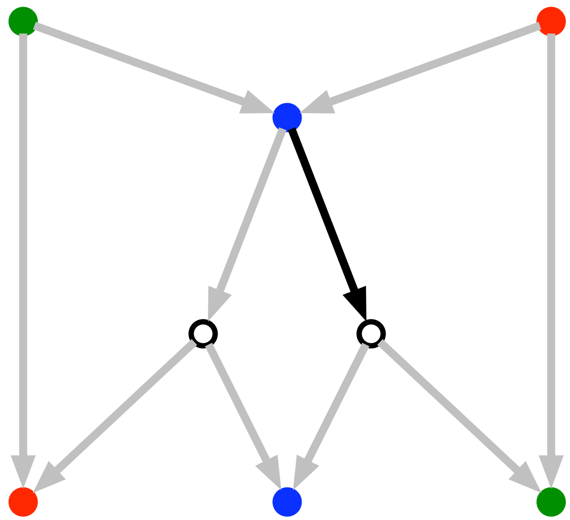


Meagerness = 1
Rate = 1

Sometimes Max Rate $<$ Meagerness

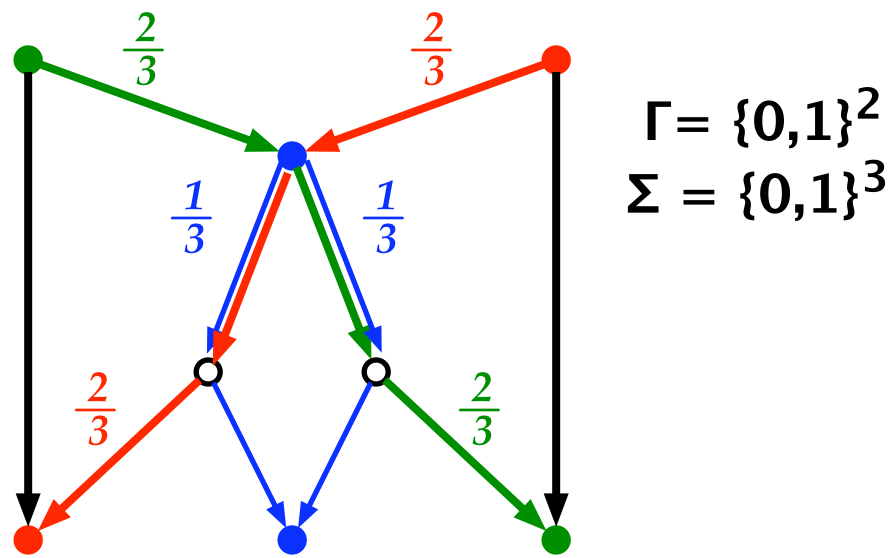


Sometimes Max Rate < Meagerness



- The meagerness is 1.

Sometimes Max Rate < Meagerness

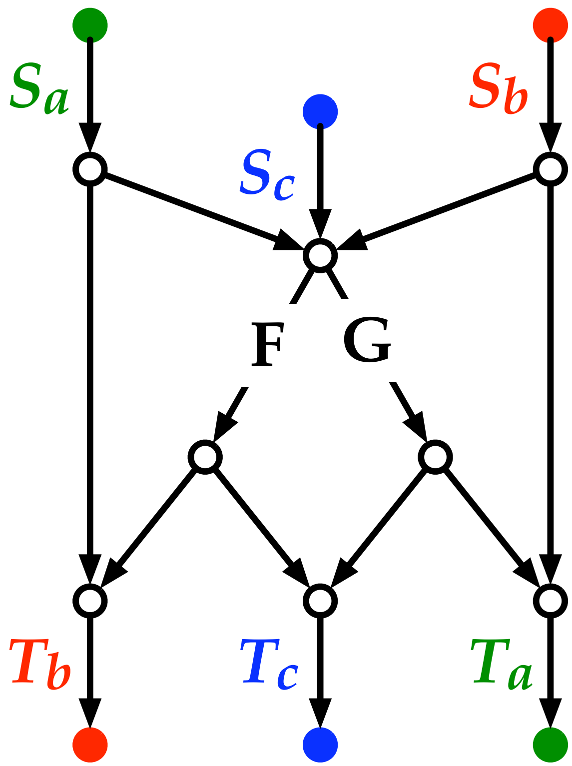


- The meagerness is 1.
- This flow solution has rate 2/3. Best possible?

Better Bounds Through Entropy

- Obtain strictly better bounds on rate through *entropy* arguments.
  - Show max rate 2/3 for previous example.
  - Implies meagerness is a loose upper bound on rate.
- Entropy of a random variable X is the information in X measured in bits.
  - The entropy of X is denoted $H(X)$.
  - The entropy of X and Y together is $H(X, Y)$.

Entropy View of Network Coding

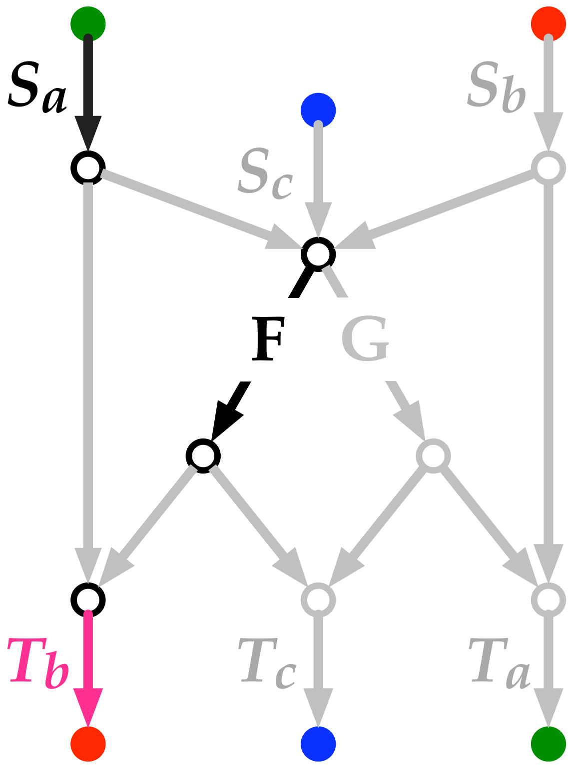


- Suppose messages are selected independently and uniformly from Γ.
- As a result, the symbol transmitted on each edge is a R.V.
- Structure of graph and properties of entropy imply constraints that a network code must satisfy.

Entropy and Network Coding

- Properties of entropy:
  - Nonnegative: $H(U) \geq 0$.
  - Nondecreasing: $H(U, x) \geq H(U)$.
  - Submodular: $H(U) + H(V) \geq H(U \cup V) + H(U \cap V)$.
- Constraints on a network coding solution:
  - Uniformity of sources: $H(S_A) = \log |\Gamma|$.
  - Independence of sources: $H(S_A, S_B) = H(S_A) + H(S_B)$.
  - sources = sinks: $H(S_A, U) = H(T_A, U)$ for all U.
  - Edge capacity: $H(e) \leq \log |\Sigma|$.

One More Condition: Downstreamness

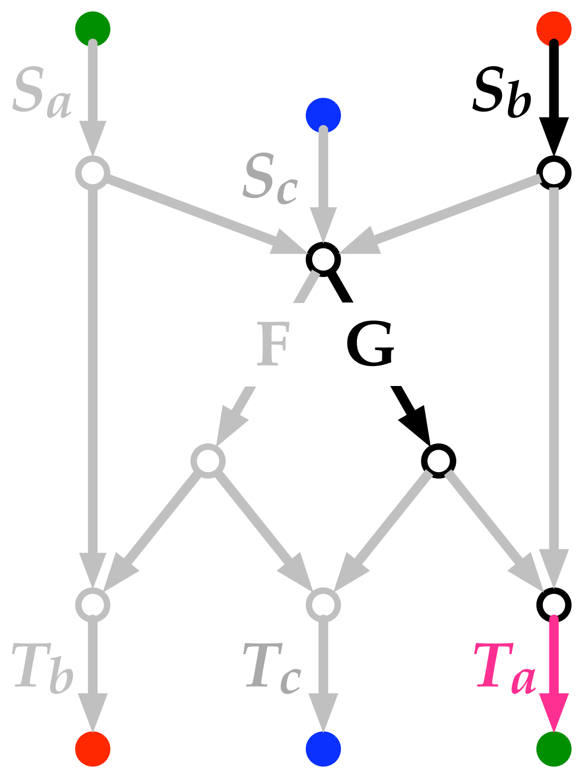


U is *downstream* of V if all paths from a source to an edge in U intersect V.

If U is downstream of V, $H(V) = H(U, V)$.

Ex 1: T_b is downstream of $\{S_a, F\}$.
$H(S_a, F) = H(S_a, T_b, F)$.

One More Condition: Downstreamness



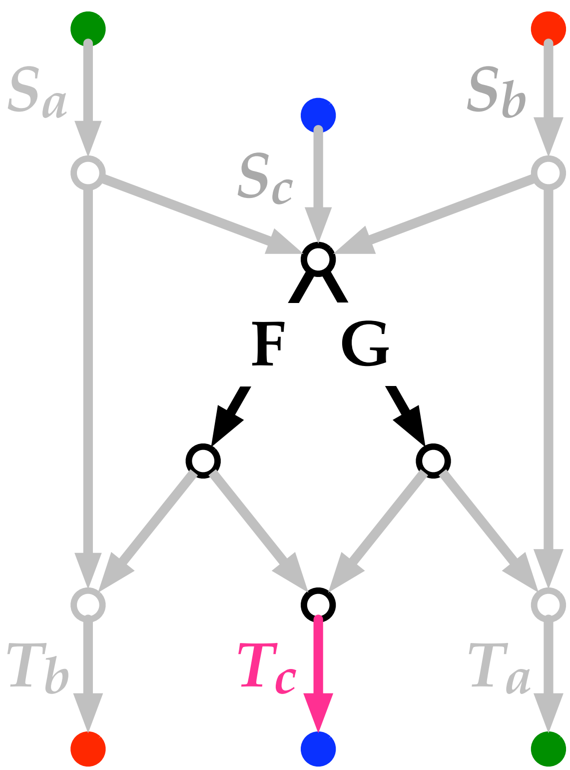
U is *downstream* of V if all paths from a source to an edge in U intersects V.

If U is downstream of V, $H(V) = H(U, V)$.

Ex 1: T_b is downstream of $\{S_a, F\}$.
$H(S_a, F) = H(S_a, T_b, F)$.

Ex 2: T_a is downstream of $\{S_b, G\}$.
$H(S_b, G) = H(T_a, S_b, G)$.

One More Condition: Downstreamness



U is *downstream* of V if all paths from a source to an edge in U intersects V.

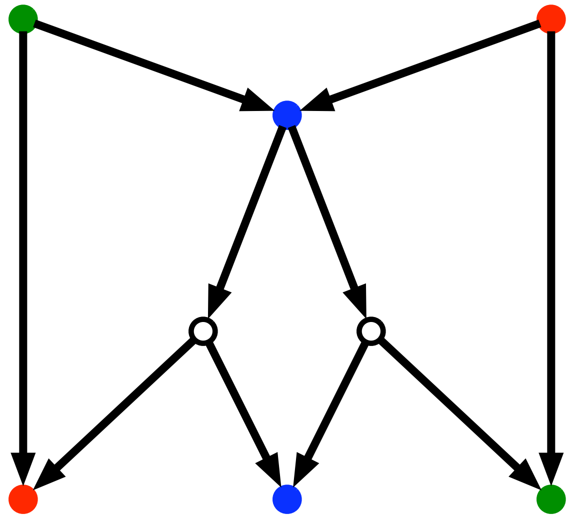
If U is downstream of V, $H(V) = H(U, V)$.

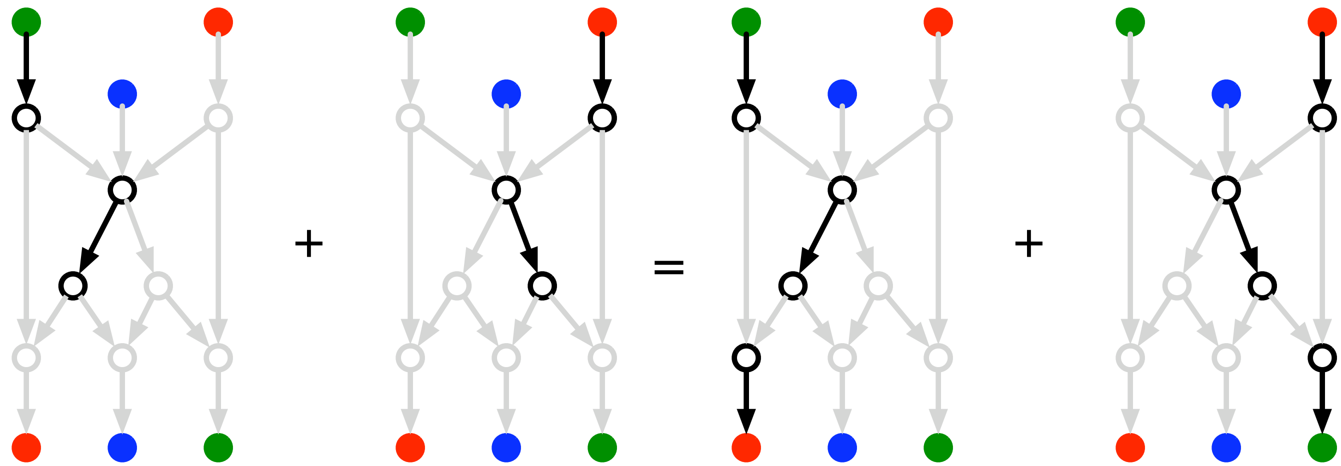
Ex 1: T_b is downstream of $\{S_a, F\}$.
$H(S_a, F) = H(S_a, T_b, F)$.

Ex 2: T_a is downstream of $\{S_b, G\}$.
$H(S_b, G) = H(T_a, S_b, G)$.

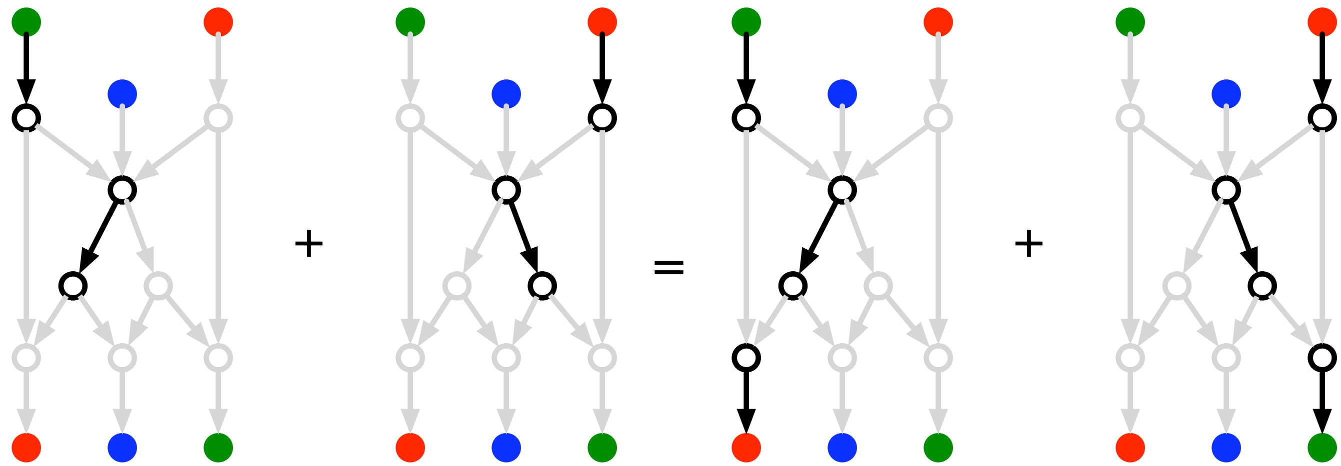
Ex 3: T_c is downstream of $\{F, G\}$.
$H(F, G) = H(T_c, F, G)$.

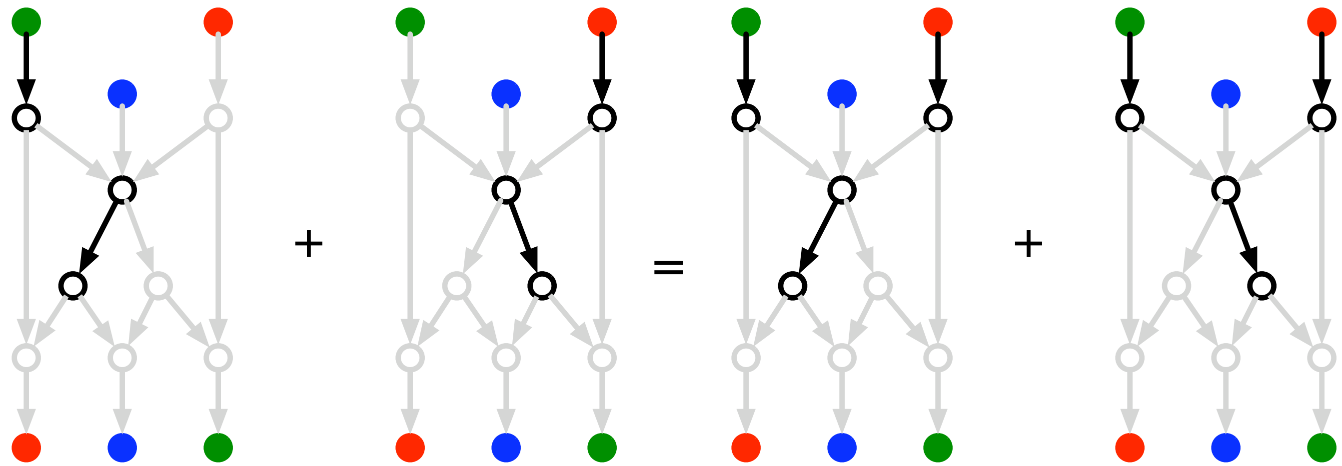
Proof: Max Rate = 2/3



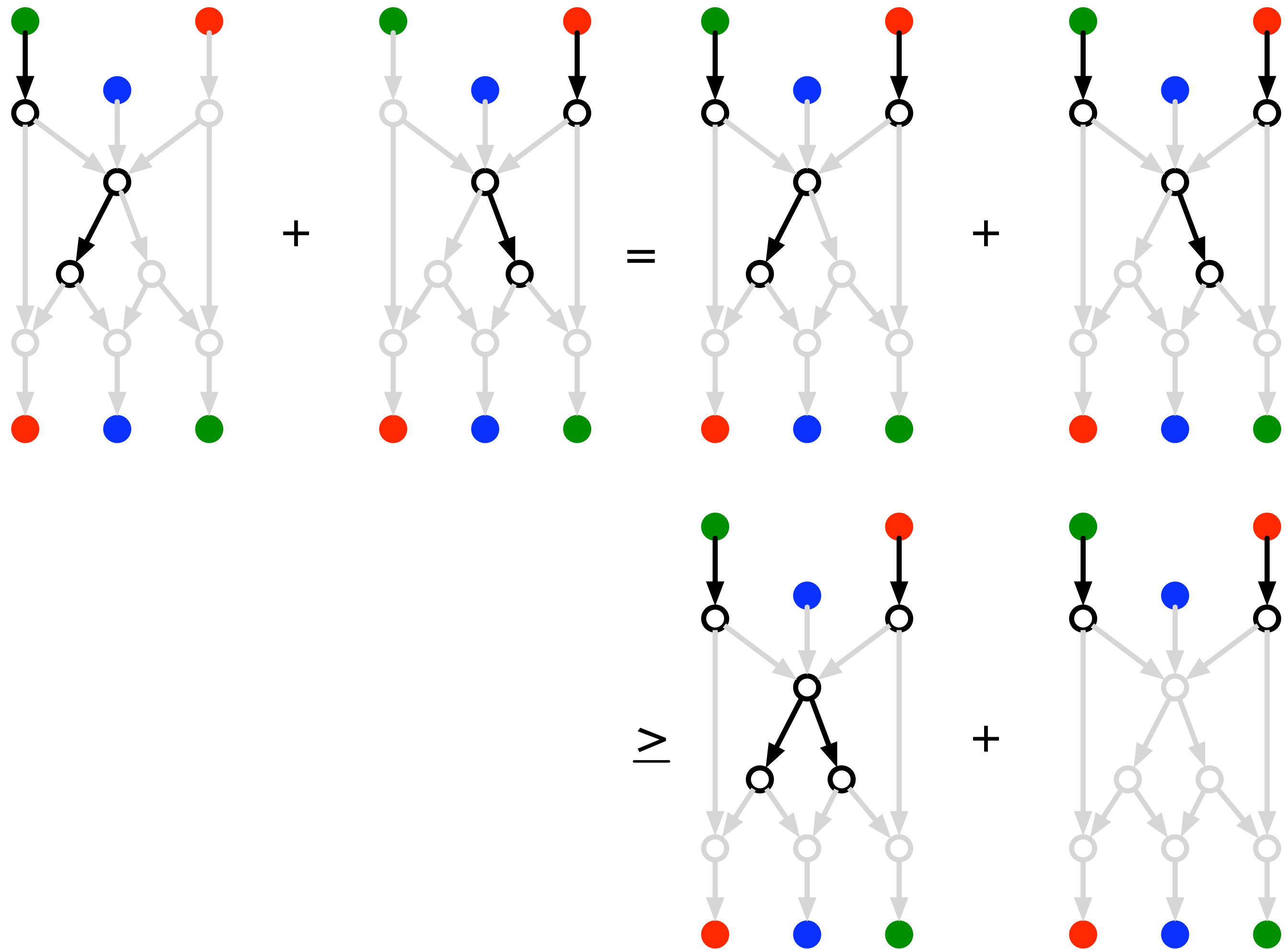


$$H(S_a, F) + H(S_b, G) = H(S_a, T_b, F) + H(T_a, S_b, G)$$

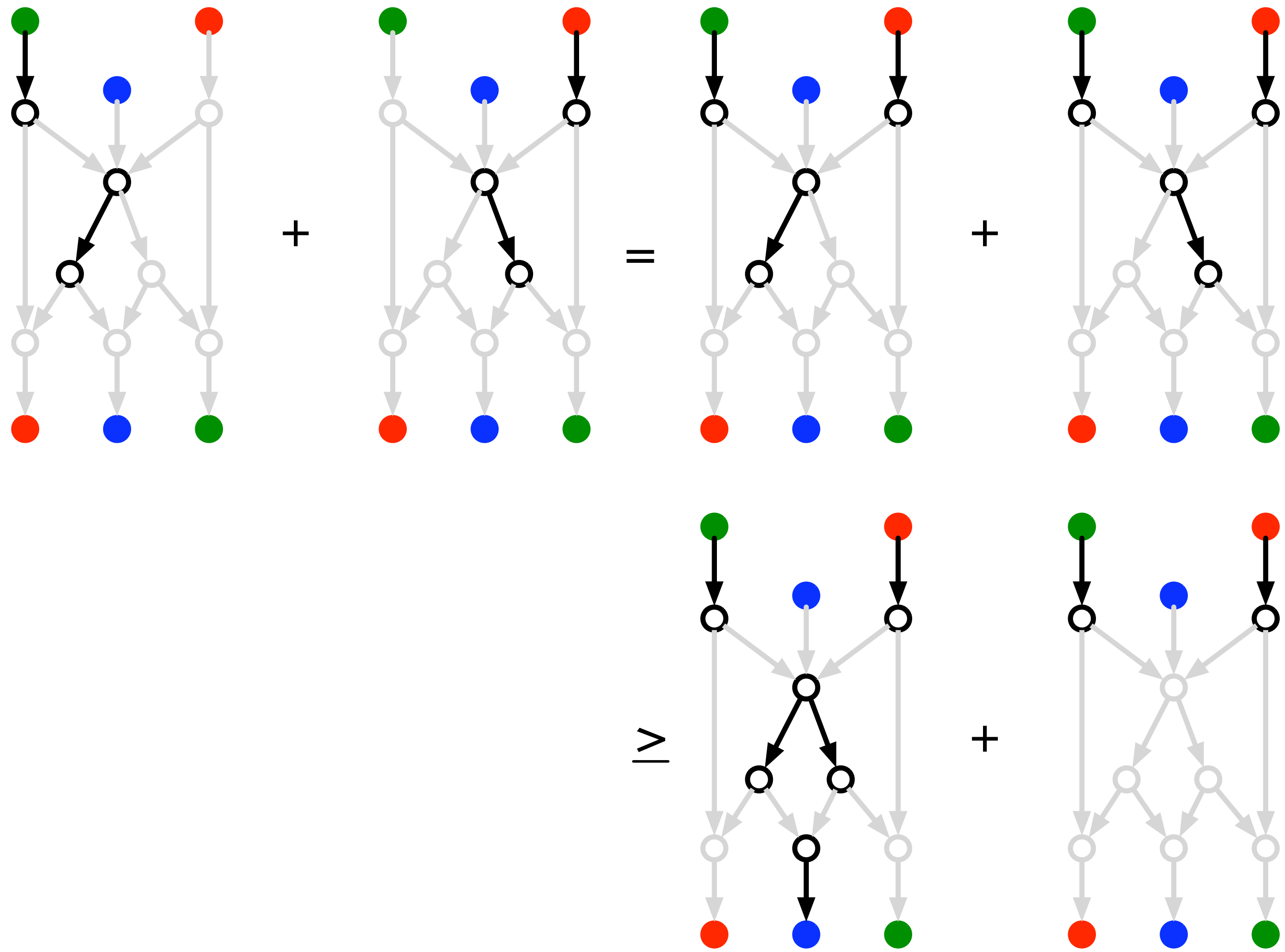




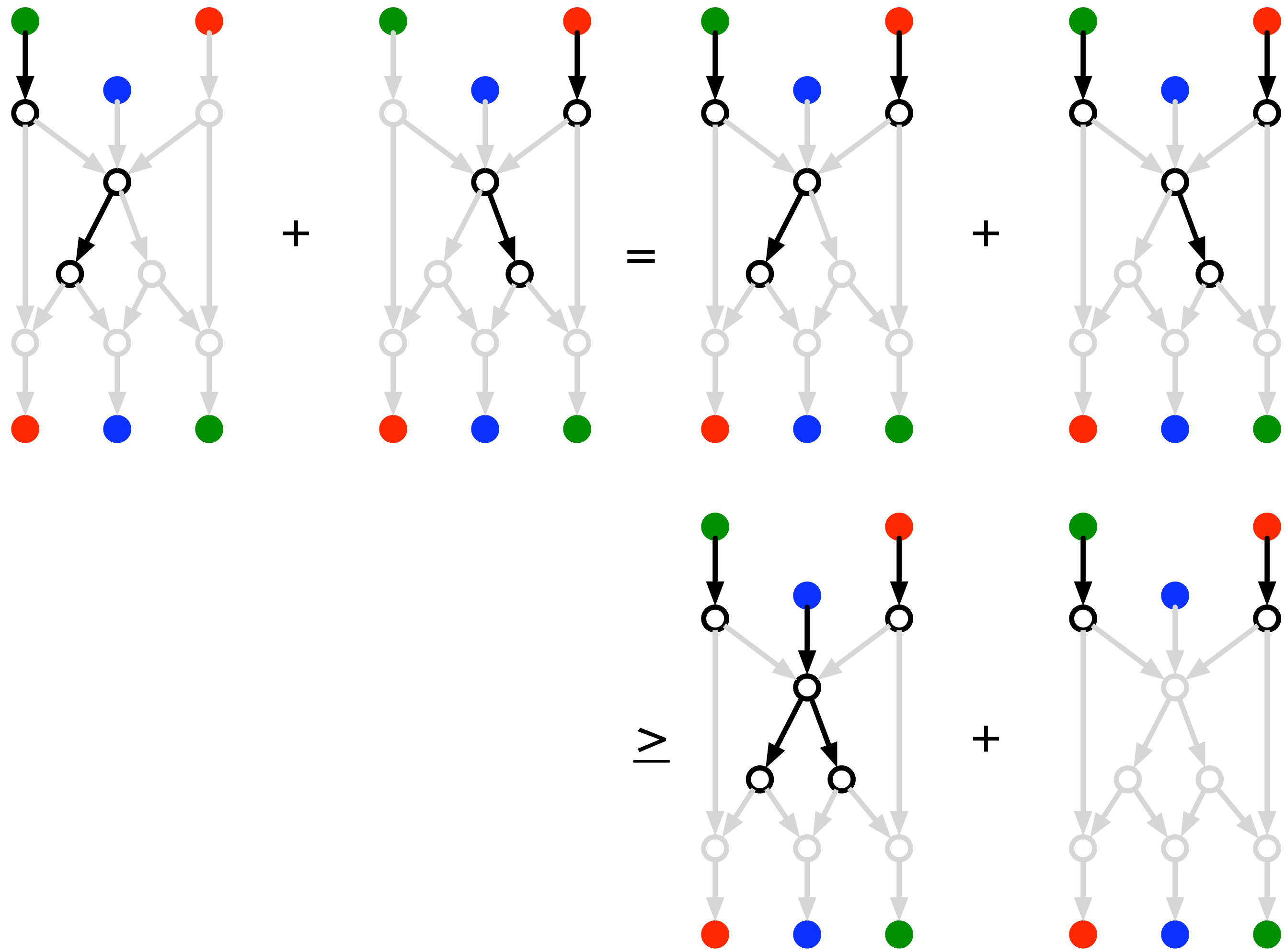
sources = sinks



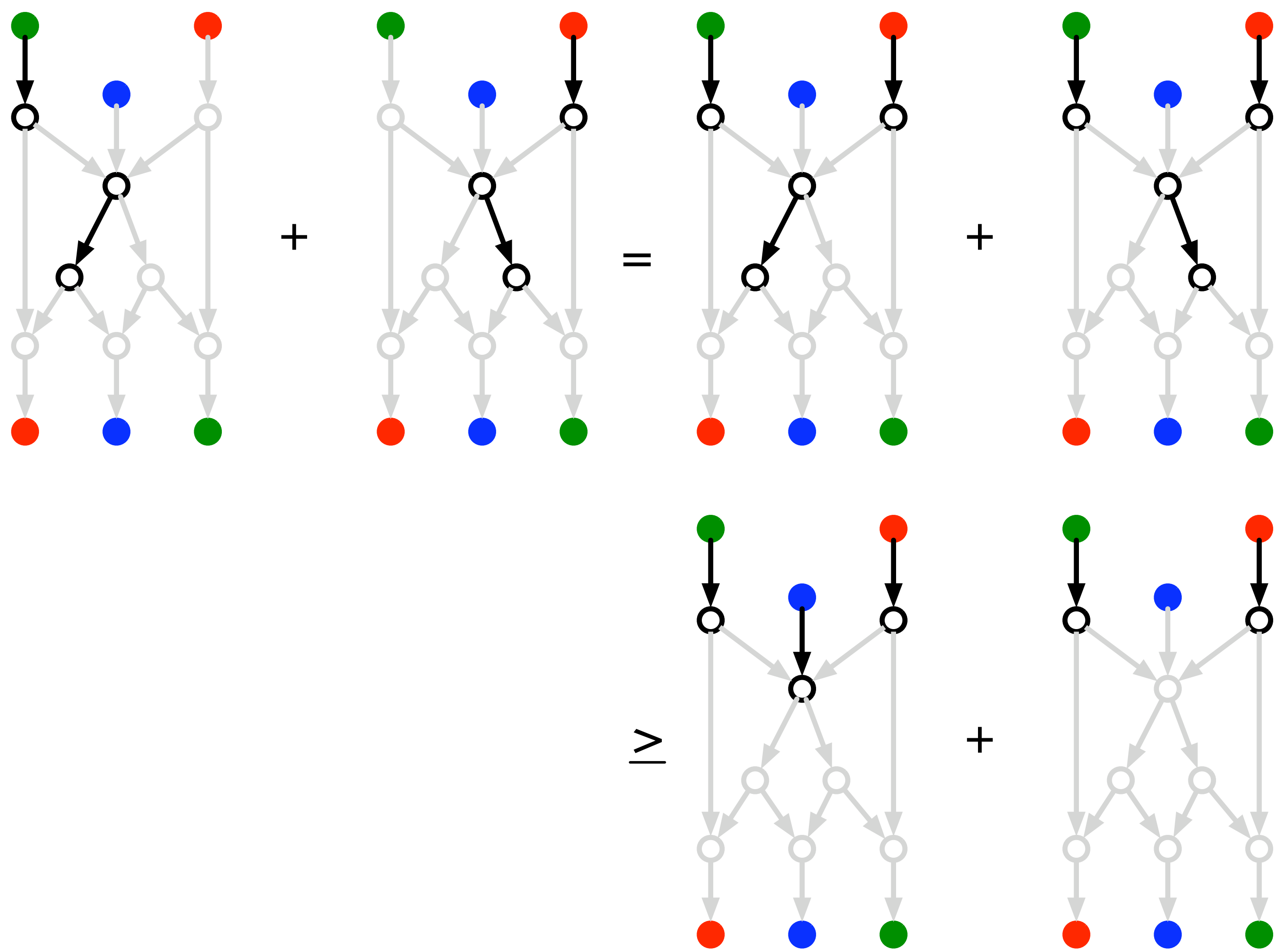
submodularity



downstreamness

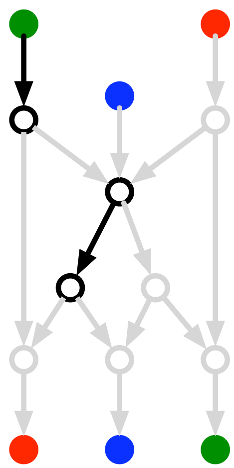


sources = sinks

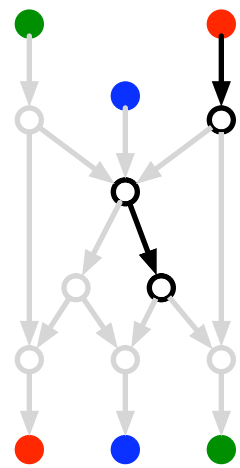


$$= 5 \log |\Gamma|$$

Max Rate = 2/3

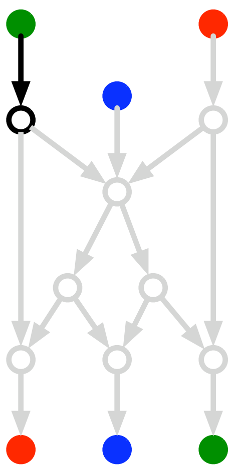


$+$

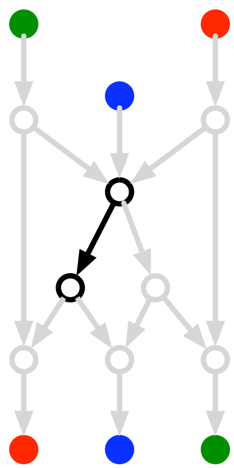


$\geq 5 \log |\Gamma|$

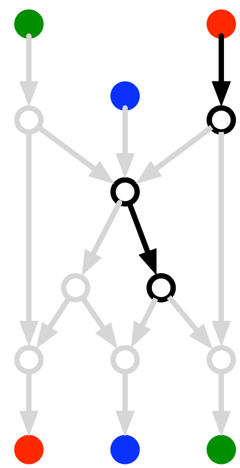
Max Rate = 2/3



+

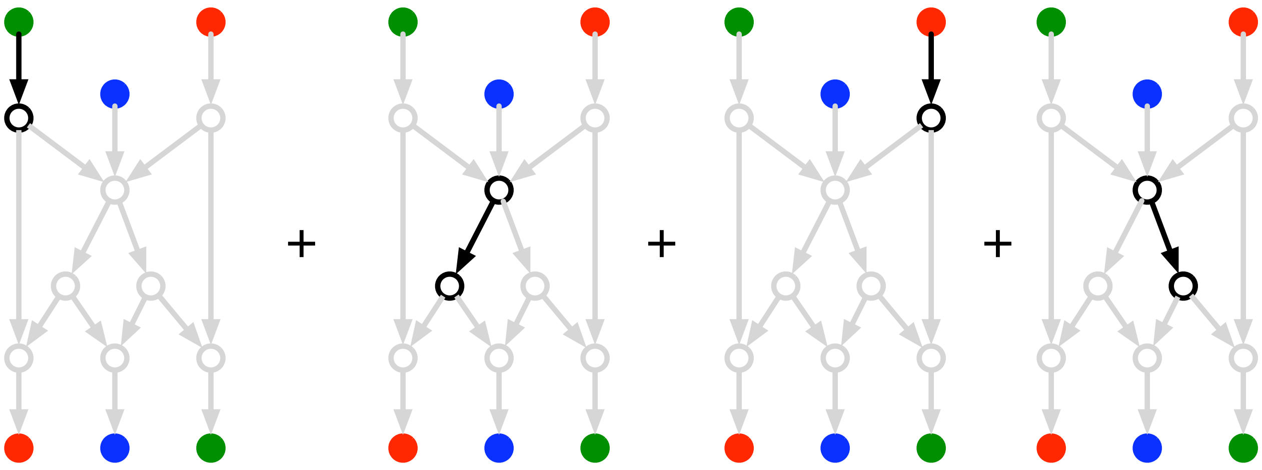


+



$\geq 5 \log |\Gamma|$

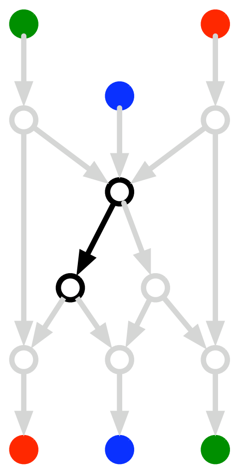
Max Rate = 2/3



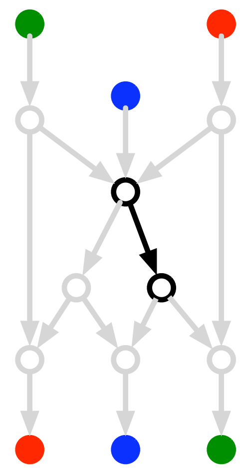
$\geq 5 \log |\Gamma|$

Max Rate = 2/3

$\log|\Gamma|$ +

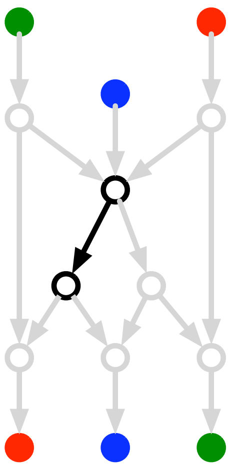


$+ \log|\Gamma| +$

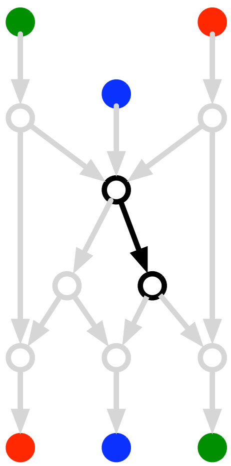


$\geq 5 \log|\Gamma|$

Max Rate = 2/3



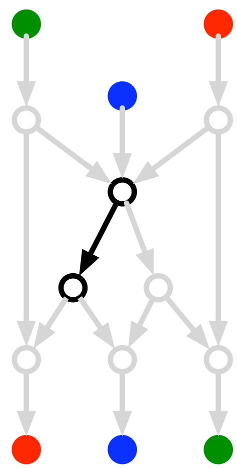
+



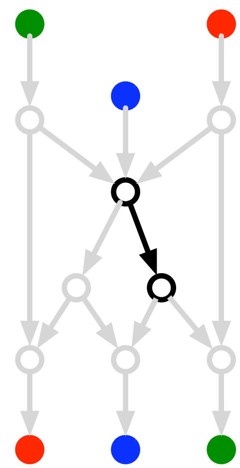
$\geq$ **3 log |Γ|**

Max Rate = 2/3

$2 \log |\Sigma| \geq$



$+$



$\geq 3 \log |\Gamma|$

What is the Maximum Rate?

- Simple cut-based characterizations of max rate unsatisfactory.
  - Sparsity is wrong for directed graphs.
  - Meagerness is a loose upper bound.
- Do the entropy conditions give a tight upper bound on rate?
  - Unknown in general.
  - Many inequalities and many ways to combine; get giant LP.

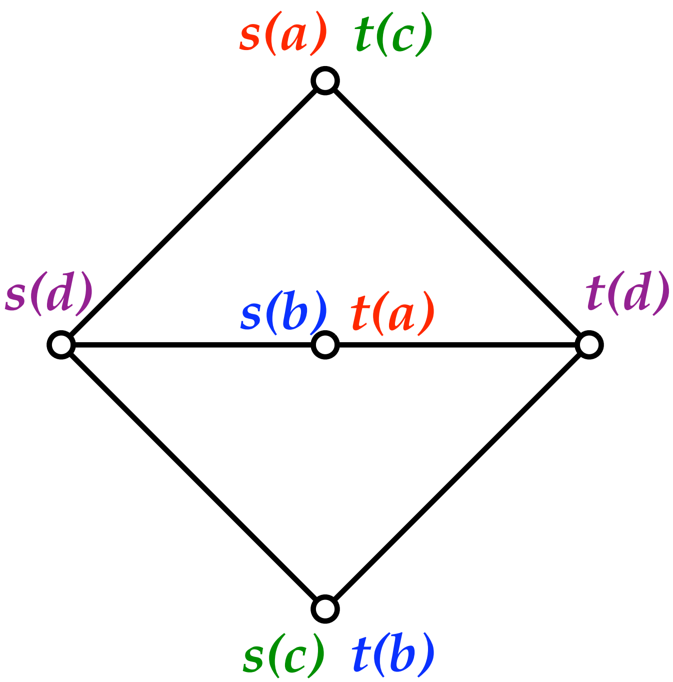
Further Results: Coding in Undirected Graphs

- How do we even model this?
  - Rule out cyclic dependencies between edge functions.
  - Edge capacity bounds information flow in two directions.
- Entropy conditions carry over, e.g. downstreamness.
- Sparsity is a loose upper bound on rate.

Conjecture: In an undirected graph, the maximum multicommodity flow = the maximum network coding rate.

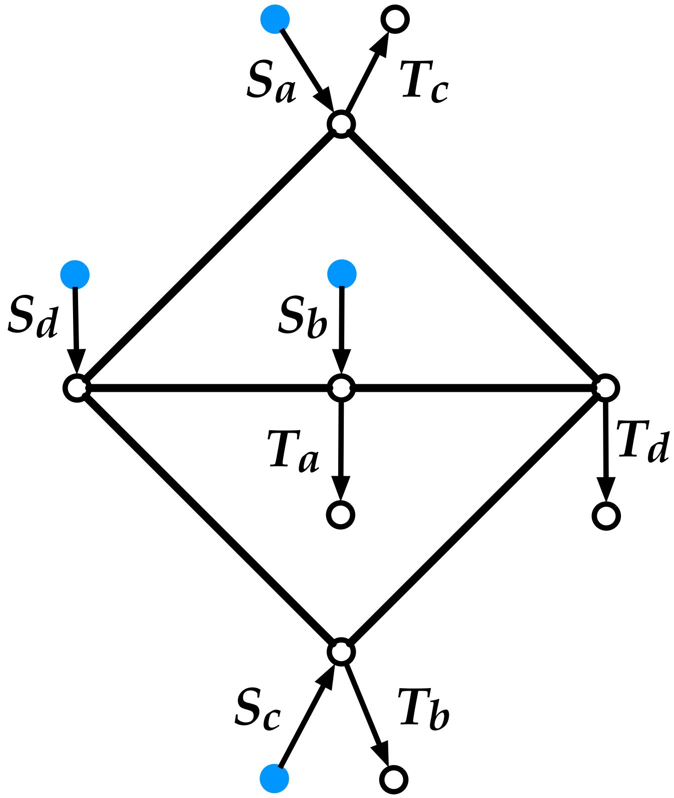
- We prove for an infinite class of "interesting" graphs.

Okamura-Seymour Example



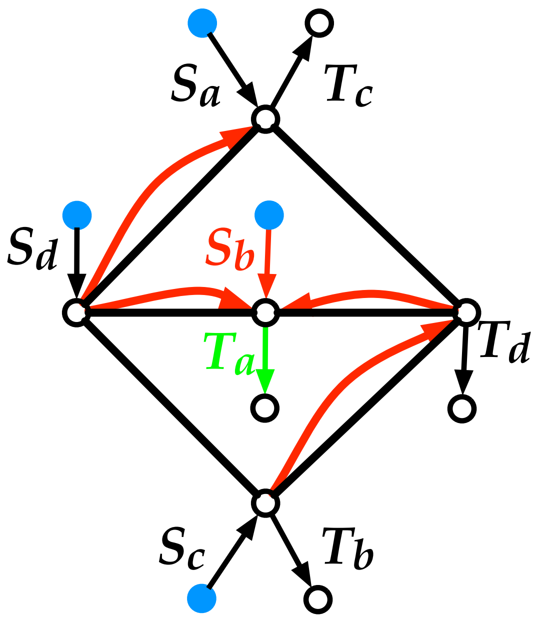
- 4 commodities.
- Each edge has capacity 1.
- Sparsity 1.
- Maximum multicommodity flow 3/4.
- Maximum rate with network coding is also 3/4!

Okamura-Seymour Example



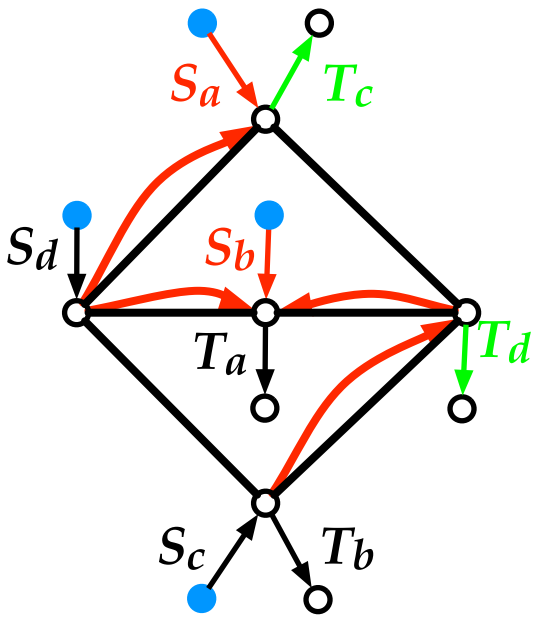
- Add new sources and sinks and the corresponding edges.
- Each source transmits one symbol from Γ.
- Each edge transmits one symbol from Σ.
- Want to show $\frac{\log|\Gamma|}{\log|\Sigma|} \leq 3/4$.
- Use three different edge-cuts.

Okamura-Seymour Example - Cut #1



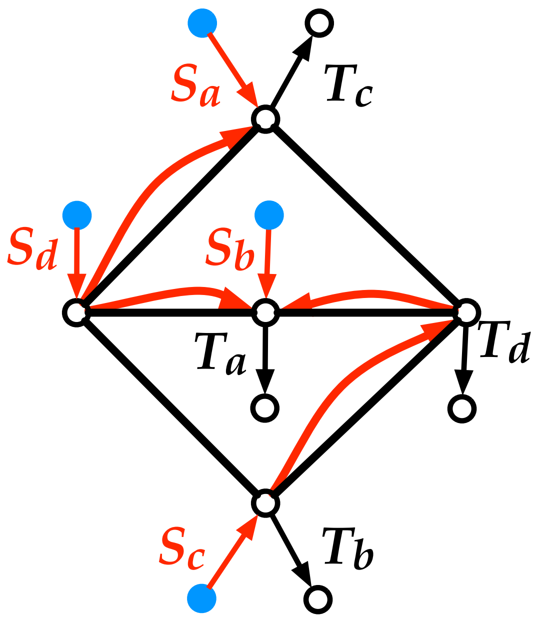
$$H(S_b, U) \;=\; H(T_a, S_b, U)$$

Okamura-Seymour Example - Cut #1



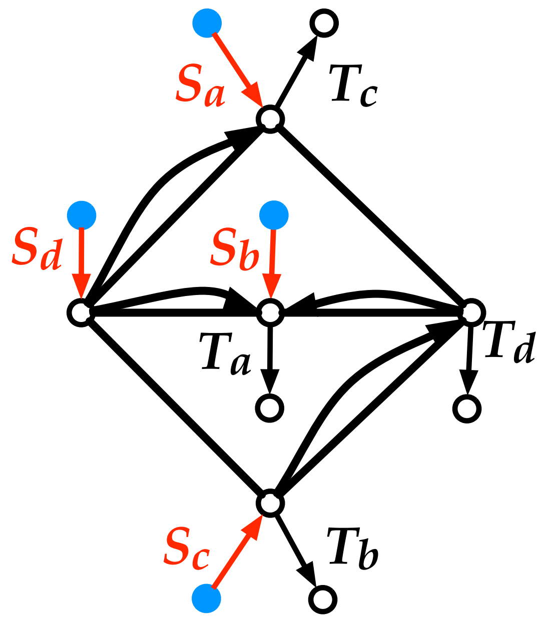
$$\begin{aligned} H(S_b, U) &= H(T_a, S_b, U) \\ &= H(S_a, S_b, U) \\ &= H(S_a, S_b, T_c, T_d, U) \end{aligned}$$

Okamura-Seymour Example - Cut #1



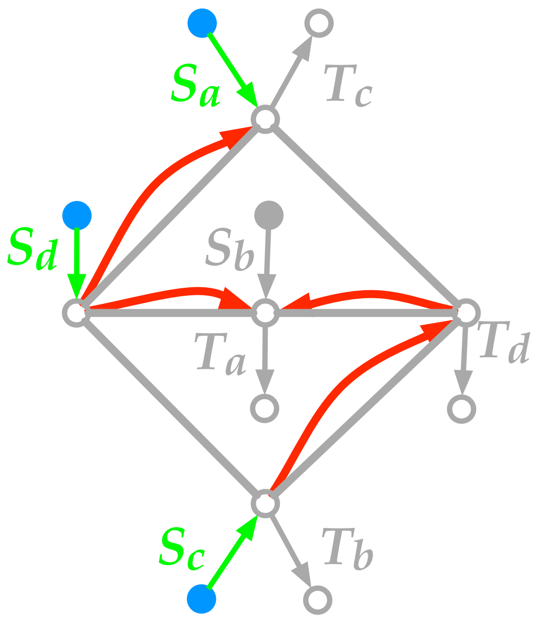
$$
\begin{aligned}
H(S_b, U) &= H(T_a, S_b, U) \\
&= H(S_a, S_b, U) \\
&= H(S_a, S_b, T_c, T_d, U) \\
&= H(S_a, S_b, S_c, S_d)
\end{aligned}
$$

Okamura-Seymour Example - Cut #1



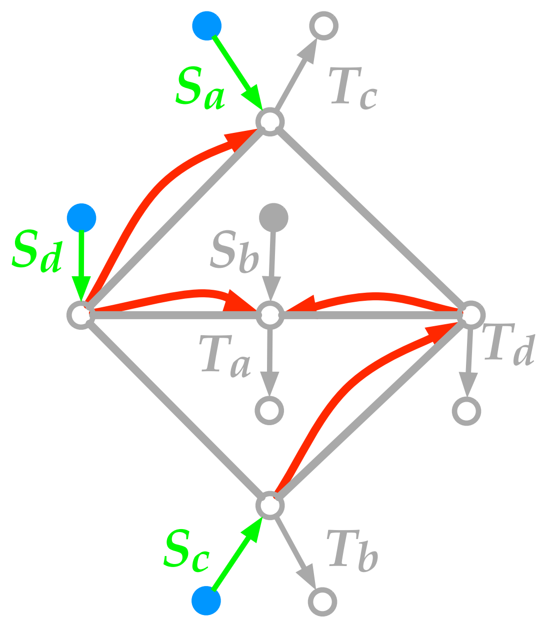
$$
\begin{aligned}
H(S_b, U) &= H(T_a, S_b, U) \\
&= H(S_a, S_b, U) \\
&= H(S_a, S_b, T_c, T_d, U) \\
&= H(S_a, S_b, S_c, S_d)
\end{aligned}
$$

Okamura-Seymour Example - Cut #1

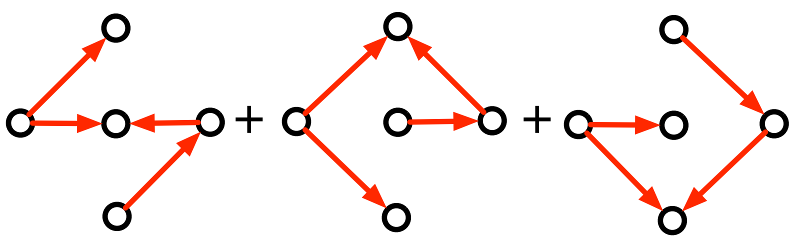


$$
\begin{aligned}
H(S_b, U) &= H(T_a, S_b, U) \\
&= H(S_a, S_b, U) \\
&= H(S_a, S_b, T_c, T_d, U) \\
&= H(S_a, S_b, S_c, S_d) \\
H(S_b) + H(U) &\geq 4 \log |\Gamma| \\
H(U) &\geq 3 \log |\Gamma|
\end{aligned}
$$

Okamura-Seymour Example - Cut #1

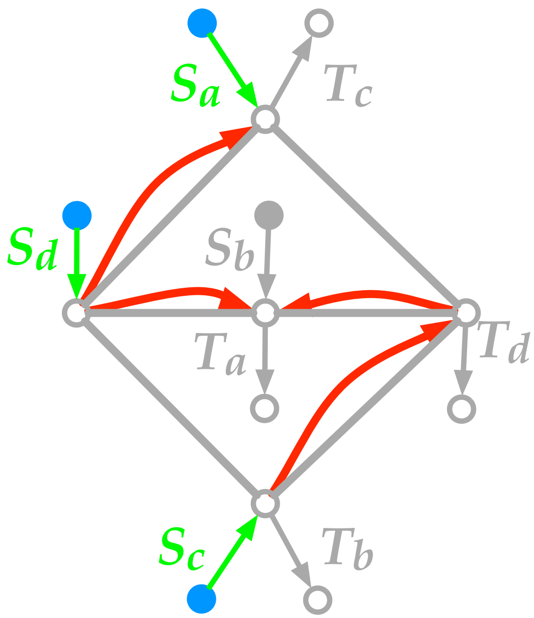


$$
\begin{aligned}
H(S_b, U) &= H(T_a, S_b, U) \\
&= H(S_a, S_b, U) \\
&= H(S_a, S_b, T_c, T_d, U) \\
&= H(S_a, S_b, S_c, S_d) \\
H(S_b) + H(U) &\geq 4 \log |\Gamma| \\
H(U) &\geq 3 \log |\Gamma|
\end{aligned}
$$

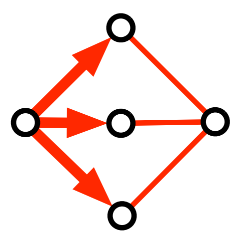


$\geq 9 \log |\Gamma|$

Okamura-Seymour Example - Cut #1

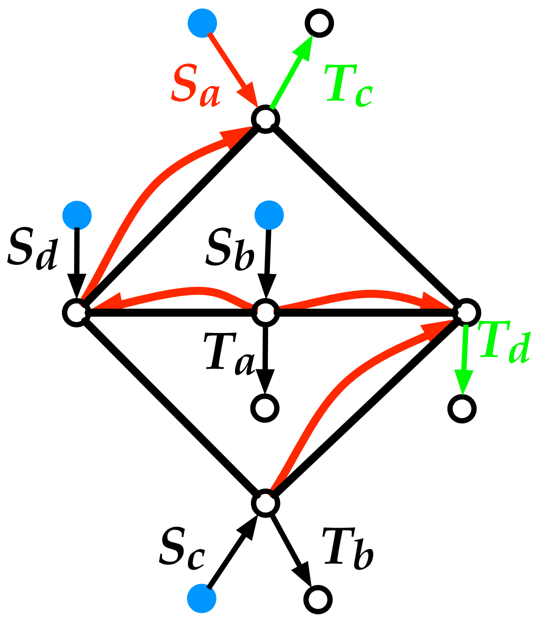


$$
\begin{aligned}
H(S_b, U) &= H(T_a, S_b, U) \\
&= H(S_a, S_b, U) \\
&= H(S_a, S_b, T_c, T_d, U) \\
&= H(S_a, S_b, S_c, S_d) \\
H(S_b) + H(U) &\geq 4 \log |\Gamma| \\
H(U) &\geq 3 \log |\Gamma|
\end{aligned}
$$



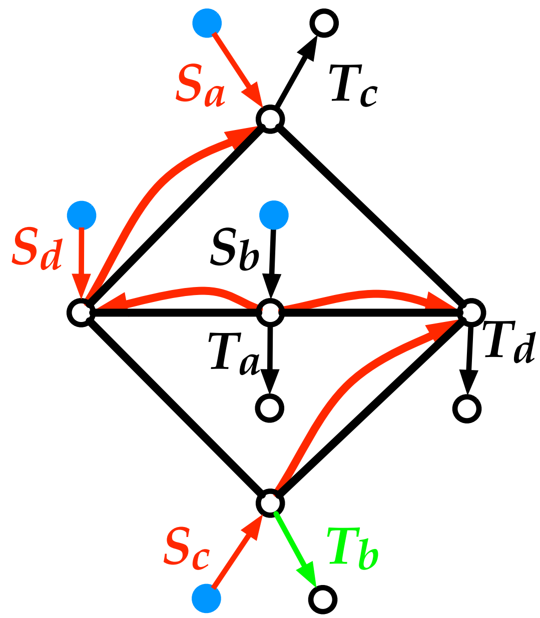
$\geq 9 \log |\Gamma|$

Okamura-Seymour Example - Cut #2



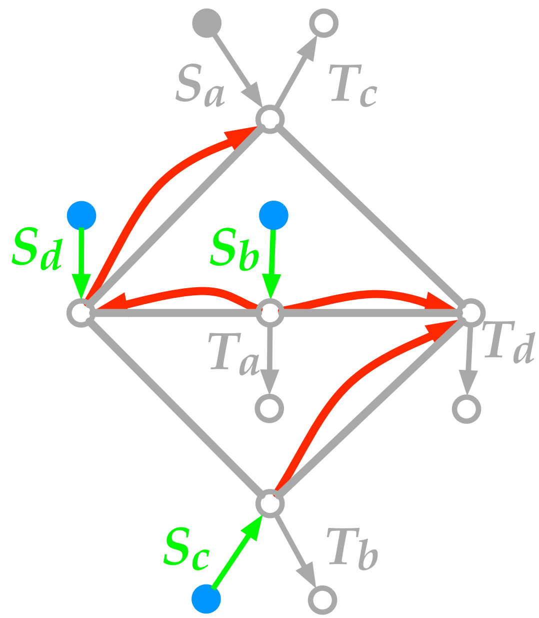
$$H(S_a, V) \;=\; H(S_a, T_c, T_d, V)$$

Okamura-Seymour Example - Cut #2



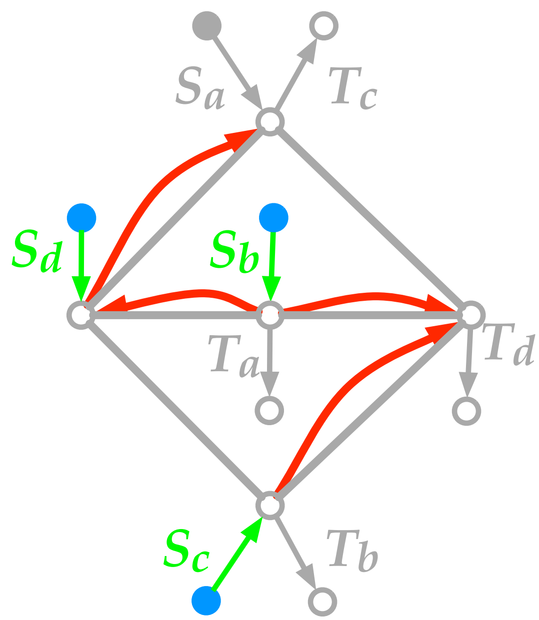
$$
\begin{aligned}
H(S_a, V) &= H(S_a, T_c, T_d, V) \\
&= H(S_a, T_b, S_c, S_d, V)
\end{aligned}
$$

Okamura-Seymour Example - Cut #2

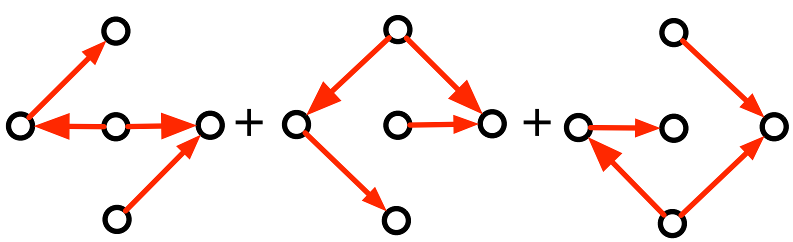


$$
\begin{aligned}
H(S_a, V) &= H(S_a, T_c, T_d, V) \\
&= H(S_a, T_b, S_c, S_d, V) \\
&= H(S_a, S_b, S_c, S_d) \\
H(V) &\geq 3 \log |\Gamma|
\end{aligned}
$$

Okamura-Seymour Example - Cut #2

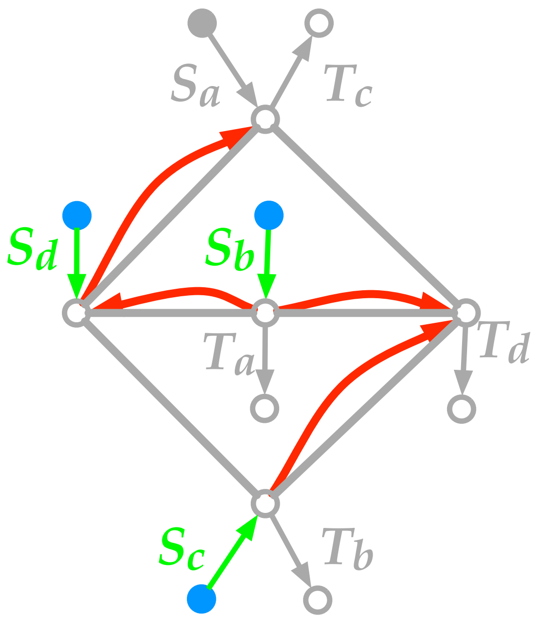


$$
\begin{aligned}
H(S_a, V) &= H(S_a, T_c, T_d, V) \\
&= H(S_a, T_b, S_c, S_d, V) \\
&= H(S_a, S_b, S_c, S_d) \\
H(V) &\geq 3 \log |\Gamma|
\end{aligned}
$$

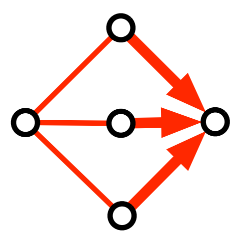


$\geq 9 \log |\Gamma|$

Okamura-Seymour Example - Cut #2

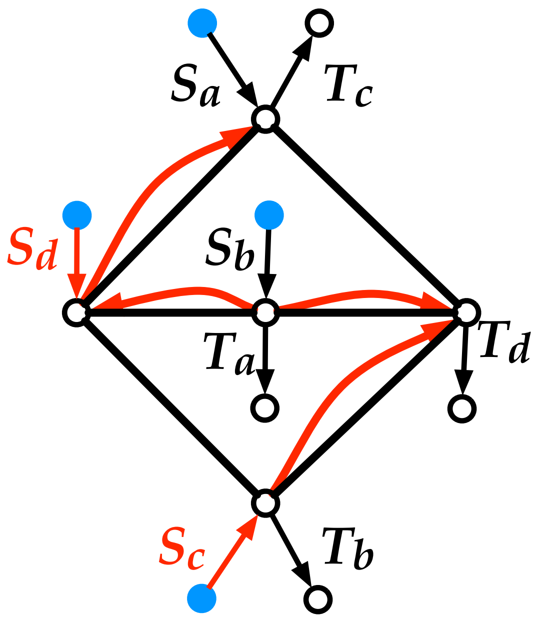


$$
\begin{aligned}
H(S_a, V) &= H(S_a, T_c, T_d, V) \\
&= H(S_a, T_b, S_c, S_d, V) \\
&= H(S_a, S_b, S_c, S_d) \\
H(V) &\geq 3 \log |\Gamma|
\end{aligned}
$$

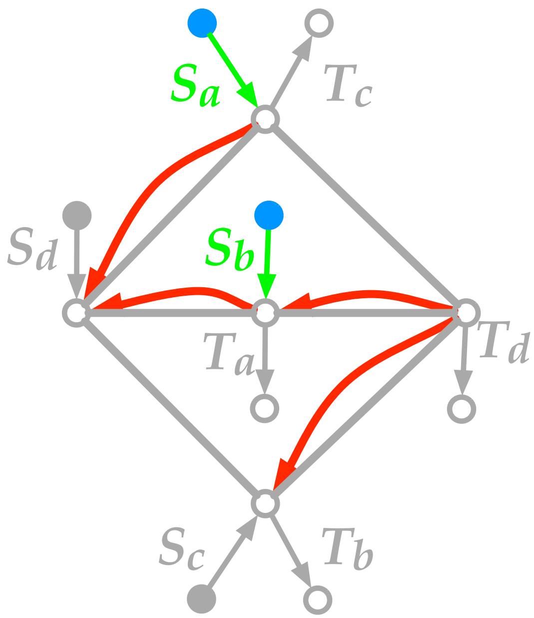


$\geq 9 \log |\Gamma|$

Okamura-Seymour Example - Cut #3

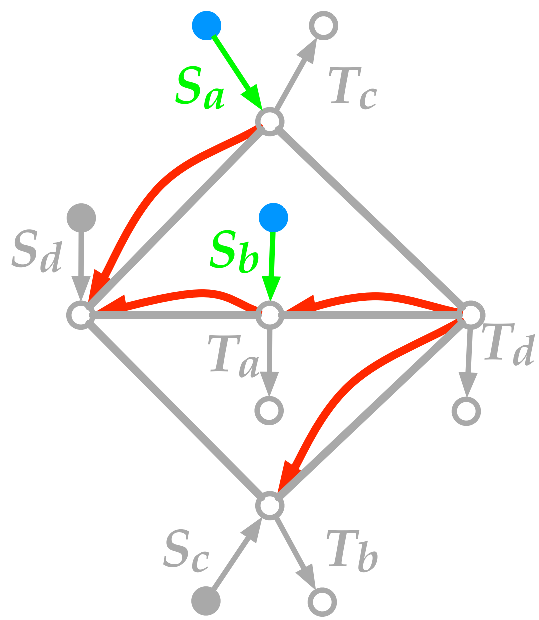


Okamura-Seymour Example - Cut #3

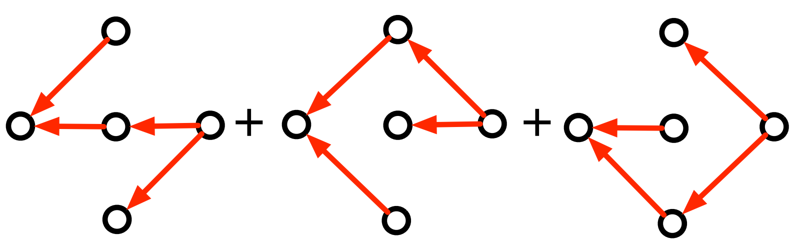


$$
\begin{aligned}
H(S_c, S_d, W) &= H(T_b, S_c, S_d, W) \\
&= H(T_a, S_b, S_c, S_d, W) \\
&= H(S_a, S_b, S_c, S_d) \\
H(W) &\geq 2 \log |\Gamma|
\end{aligned}
$$

Okamura-Seymour Example - Cut #3

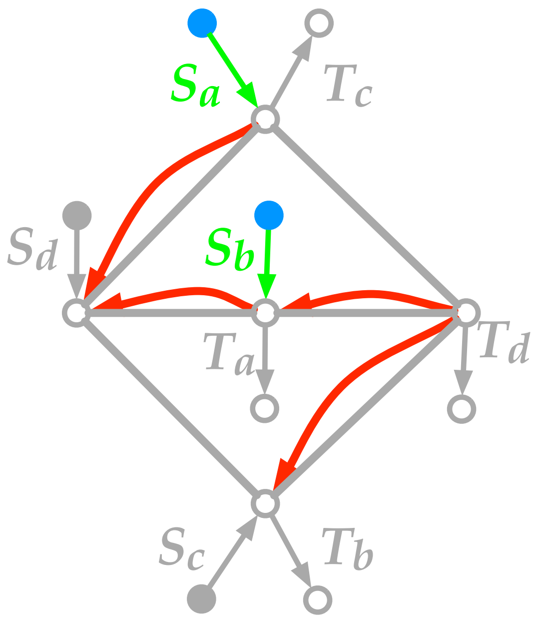


$$H(S_c, S_d, W) = H(T_b, S_c, S_d, W)$$
$$= H(T_a, S_b, S_c, S_d, W)$$
$$= H(S_a, S_b, S_c, S_d)$$
$$H(W) \geq 2 \log |\Gamma|$$

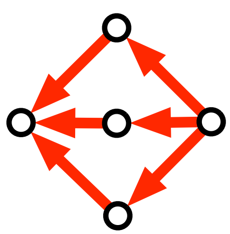


$\geq 6 \log |\Gamma|$

Okamura-Seymour Example - Cut #3

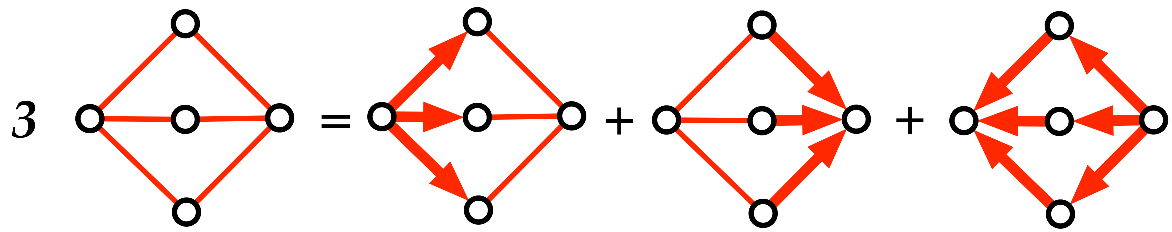


$$
\begin{aligned}
H(S_c, S_d, W) &= H(T_b, S_c, S_d, W) \\
&= H(T_a, S_b, S_c, S_d, W) \\
&= H(S_a, S_b, S_c, S_d) \\
H(W) &\geq 2 \log |\Gamma|
\end{aligned}
$$



$\geq 6 \log |\Gamma|$

Putting It Together



$$
\begin{aligned}
3(6\log|\Sigma|) &\geq 9\log|\Gamma| + 9\log|\Gamma| + 6\log|\Gamma| \\
18\log|\Sigma| &\geq 24\log|\Gamma| \\
\frac{3}{4} &\geq \frac{\log|\Gamma|}{\log|\Sigma|}
\end{aligned}
$$

Network Coding vs. Multicommodity Flow

- Only comparable when each commodity has a single source and single sink.
- For this example, shown:

$$\text{max flow rate} = \text{max network coding rate}$$

- Open: Is this true for all undirected graphs?

Additional Results

- Can prove the conjecture for all instances defined on bipartite graphs such that
  - Length 1 for all edges is dual optimal.
  - Distance between each source and sink is 2.
- Operational downstreamness: A set of edges U is operationally downstream of a set V if for all network coding solutions there exists a function mapping the symbols transmitted on edges in V to edges in U.
  - In undirected graphs, we have a graph theoretic condition that characertizes operational downstreamness.
  - In directed graphs, the graph theoretic condition implies operational downstreamness.

Summary

- Capacity of information networks is poorly understood.
- Model for multicast is not appropriate for more general problems.
- Introduce a notion of rate.
- What is the maximum rate?
  - Directed graphs: meagerness is a loose upper bound.
  - Undirected graphs: sparsity is a loose upper bound.
- Introduced entropy relationships based on graph structure.
  - Do these exactly characterize the rate?

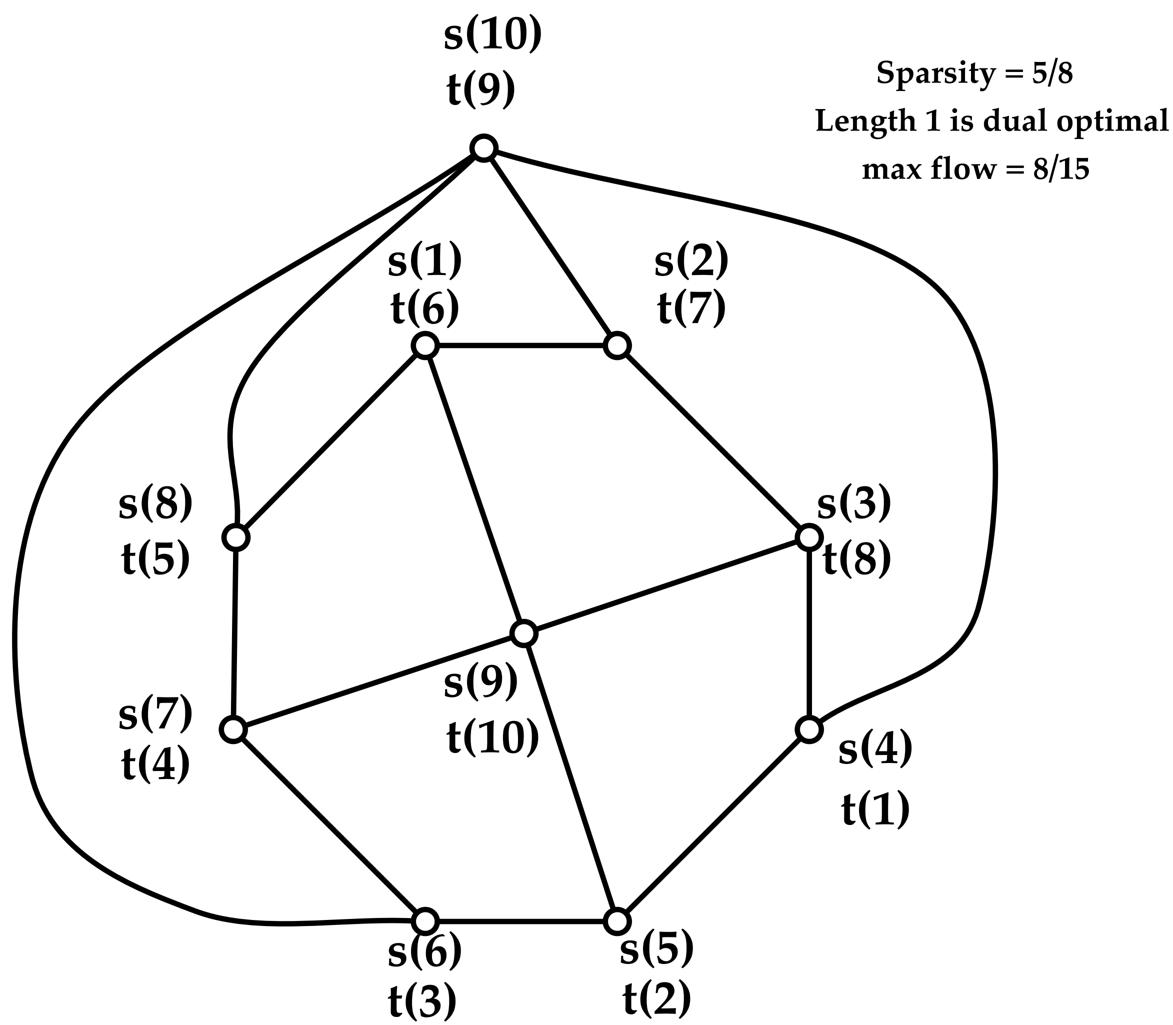
Related Work

- By Monday, details will be available at:

  `http:\\theory.csail.mit.edu/~arasala/thesis.pdf`

- Song, Yeung and Cai '03
  - For directed acyclic graphs, used similar entropy constraints to characterize an outer-bound on the feasible rate region.
- Jain et al. '05
  - Developed similar entropy constraints for the general problem.
  - Independently derived same results for undirected graphs.

Can you solve this?


s(10)
t(9)
Sparsity = 5/8
Length 1 is dual optimal
max flow = 8/15
s(1)
t(6)
s(2)
t(7)
s(8)
t(5)
s(3)
t(8)
s(9)
t(10)
s(7)
t(4)
s(4)
t(1)
s(6)
t(3)
s(5)
t(2)