Combination Network Coding

May 14 2010, at CUHK

Zongpeng Li
Department of Computer Science
University of Calgary

Supported by:
Natural Science and Engineering Research Council of Canada (NSERC)
Alberta Innovates — Technology Futures (AI-TF)

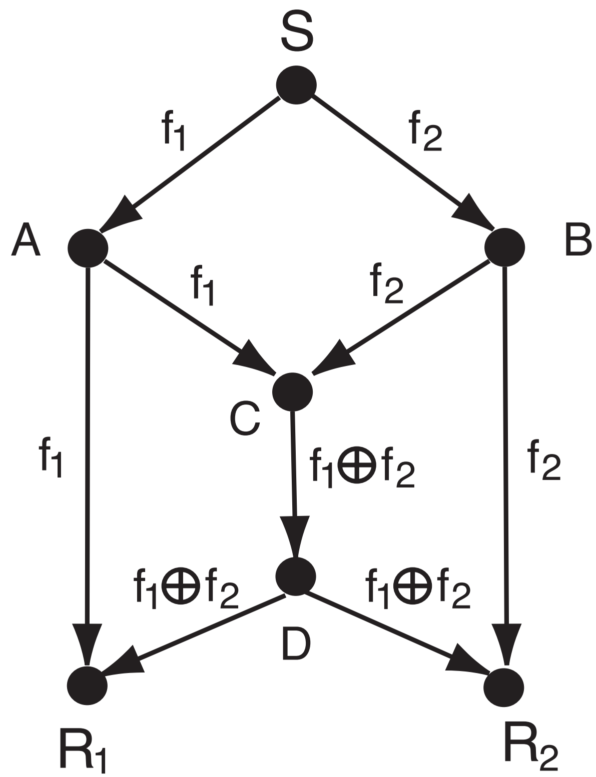
The Mathematics of Information Technology and Complex Systems (MITACS)

Network Coding

- Coding in a Network
- Information flows 'mixed' at intermediate nodes

Multicast Network Coding: An Example

The classic butterfly network

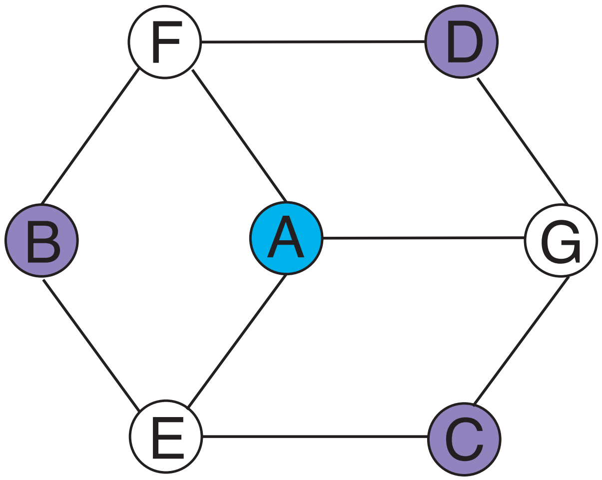


- Throughput with NC: 2
- Without NC (tree packing): < 2

Network Coding Benefits

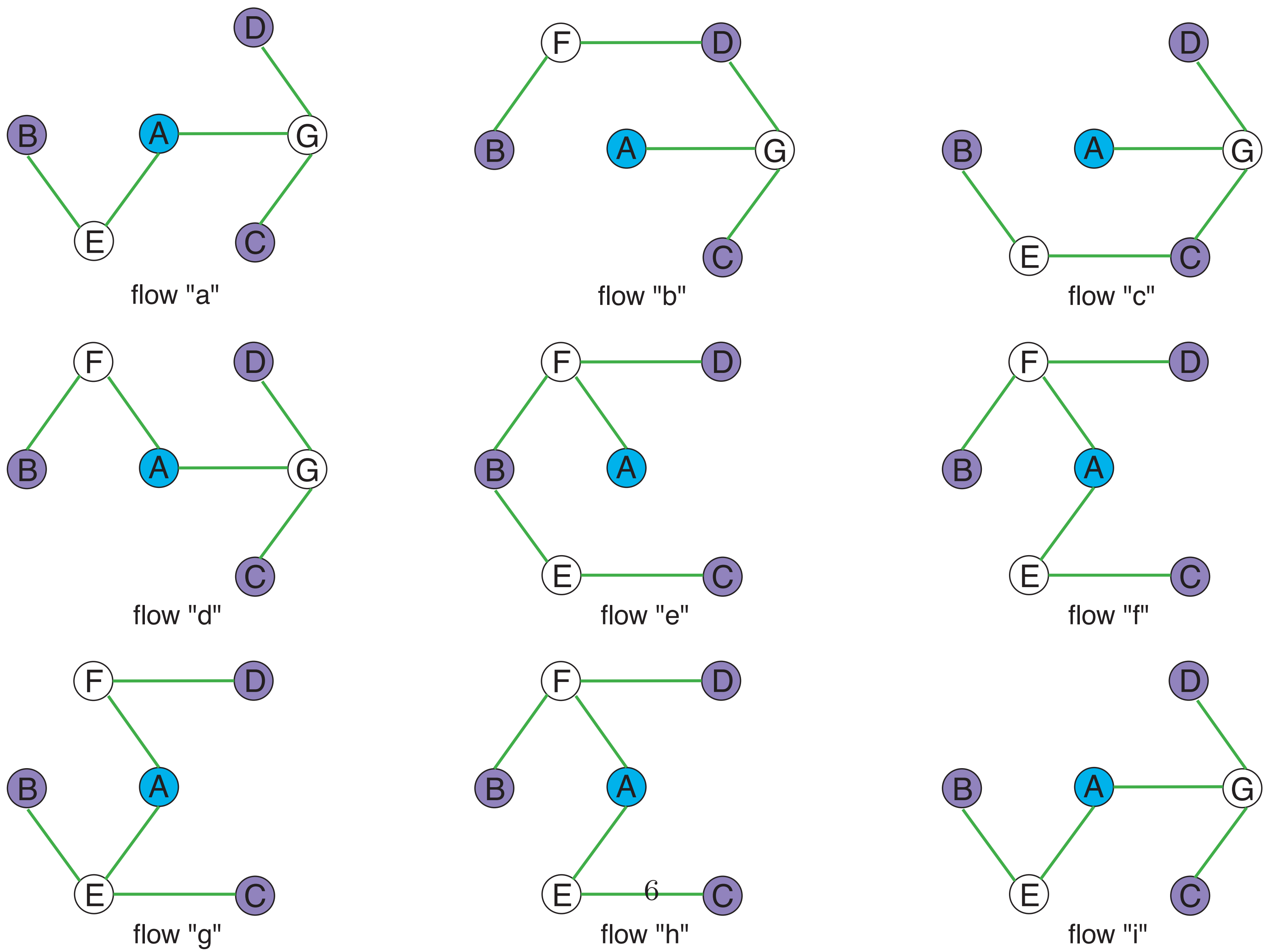
- Increasing throughput and network capacity
- Reducing routing cost
- Reducing energy consumption (wireless)
- Security
- Robustness, network error correction
- Data scheduling in P2P networks
- Reducing complexity of optimal routing problems

Network Coding: Increase Multicast Throughput



Network Coding: Increase Multicast Throughput

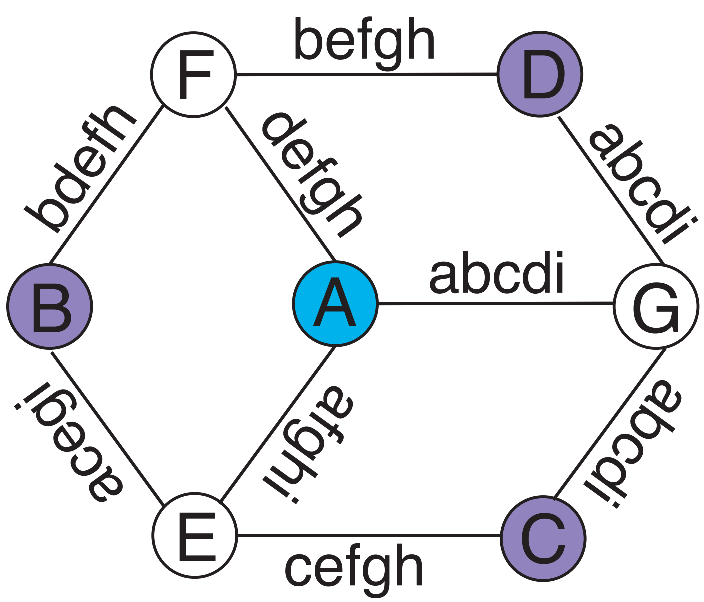
Multicast without NC: multicast tree packing



Network Coding: Increase Multicast Throughput

Multicast without NC: multicast tree packing

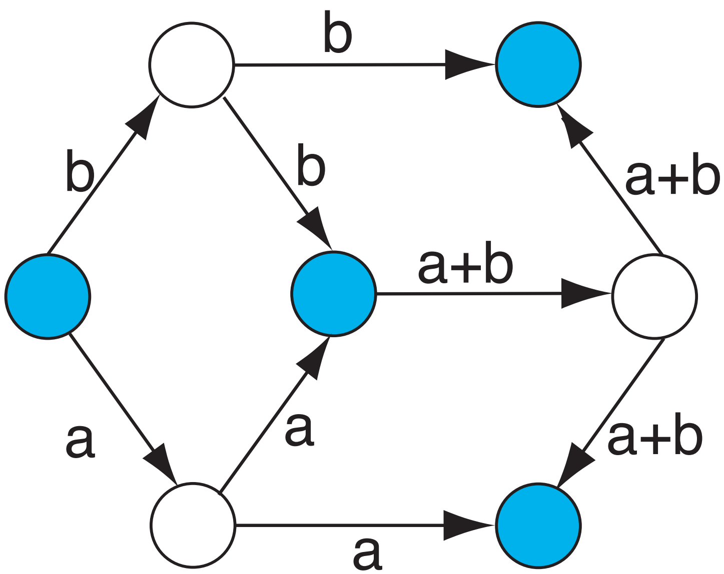
- Throughput: $0.2 \times 9 = 1.8$



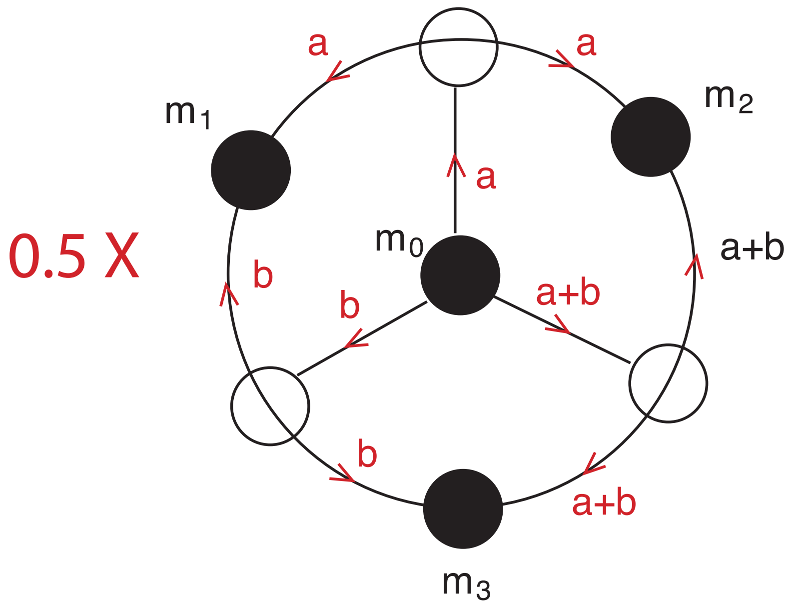
Network Coding: Increase Multicast Throughput

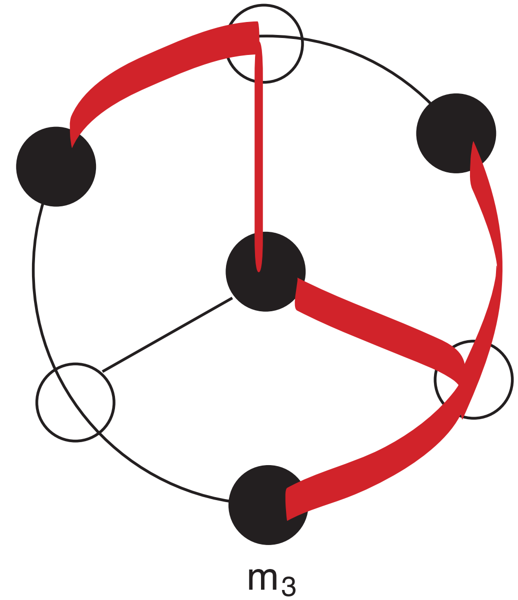
Multicast with NC: a union of network flows

- Throughput: 2
- Coding advantage: $2/1.8 = 10/9$



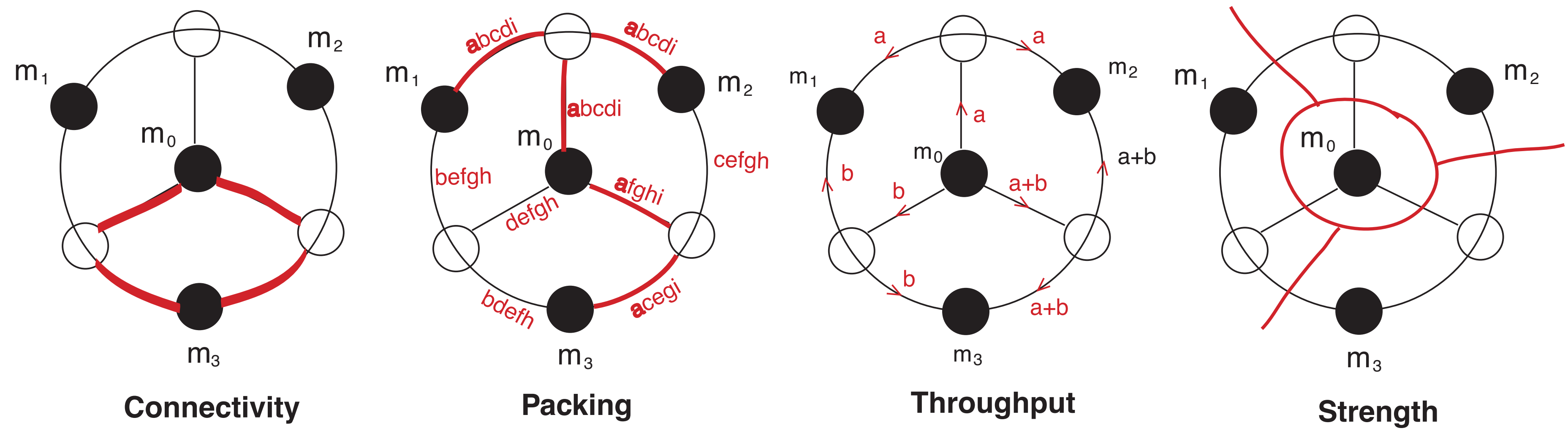
Network Coding: Reduce Multicast Cost





- Cost with NC: $0.5 \times 9 = 4.5$
- Without NC (min multicast tree): 5
- Cost advantage: $5/4.5 = 10/9$

A Bound of 2 for General Undirected Networks



Previous result:

- $\frac{1}{2}$connectivity $\leq$ packing $\leq$ throughput $\leq$ strength $\leq$ connectivity
- coding adv. $=$ throughput/packing ≤ 2

Real-world Observations

The coding advantage:

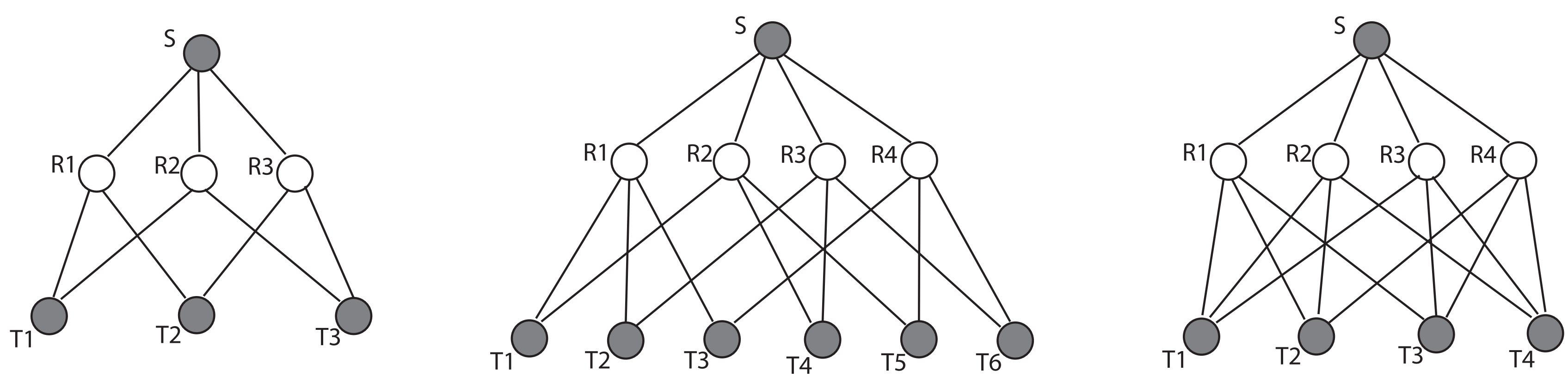
- Proved in theory: ≤ 2
- Largest value seen for networks of unbounded sizes: $8/7$
- Largest value seen for small contrived networks: $9/8$
- In hundreds of random networks: always 1

Prove a better upper-bound close to 1?

- General network coding: hard, open question
- One may first focus on special cases
  - "combination network coding"
  - "planar network coding"
  - . . .

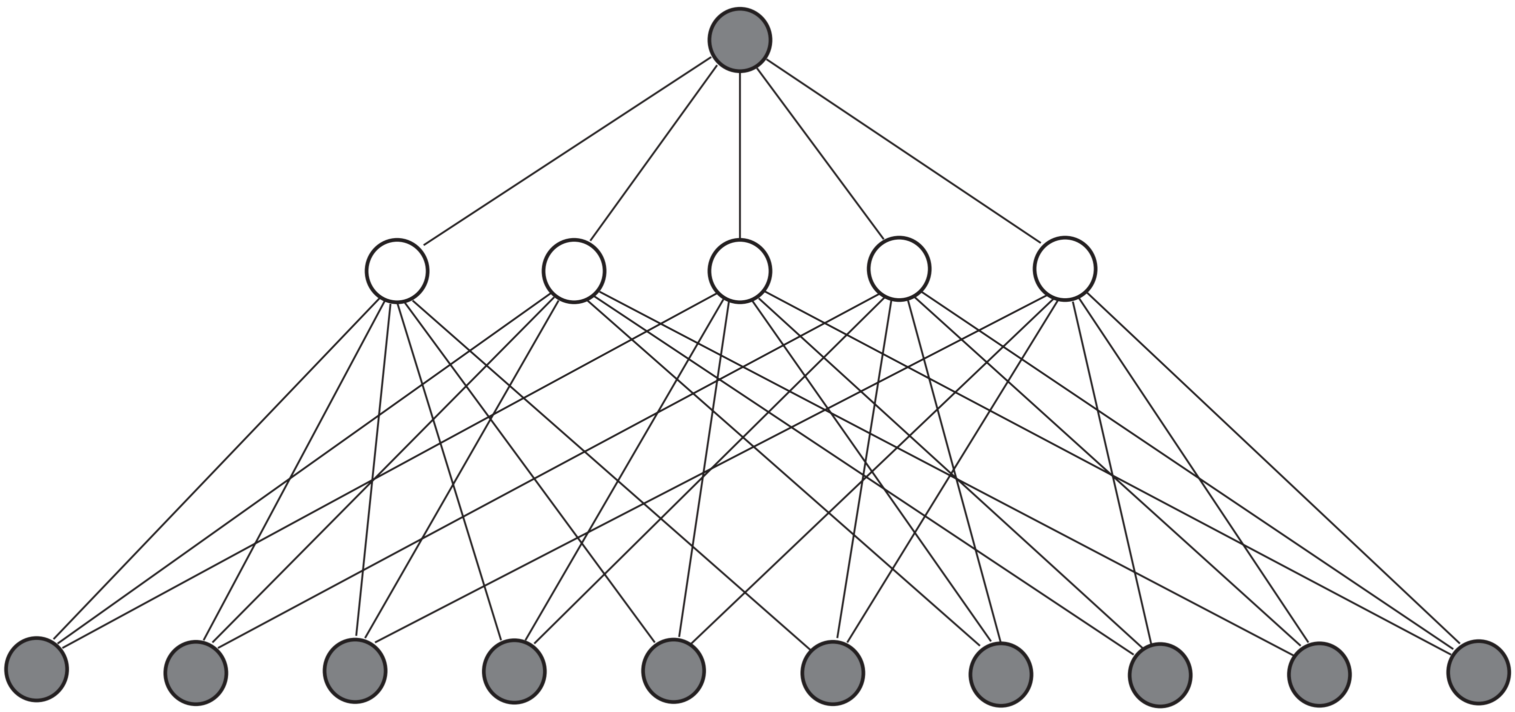
The combination network $C(n, k)$

- $C(n, k)$: 3-layer topology
- Layer 1: 1 sender
- Layer 2: n relays, each connect to sender
- Layer 3: $\binom{n}{k}$ receivers, each connect to a diff. set of k relays



The combination network $C(n, k)$

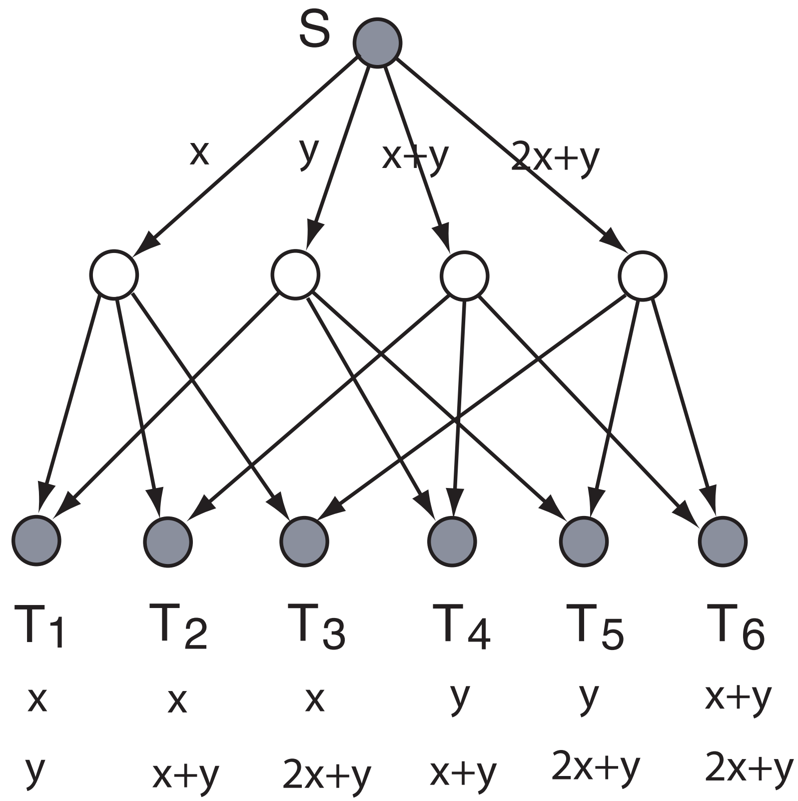
C(5,3)



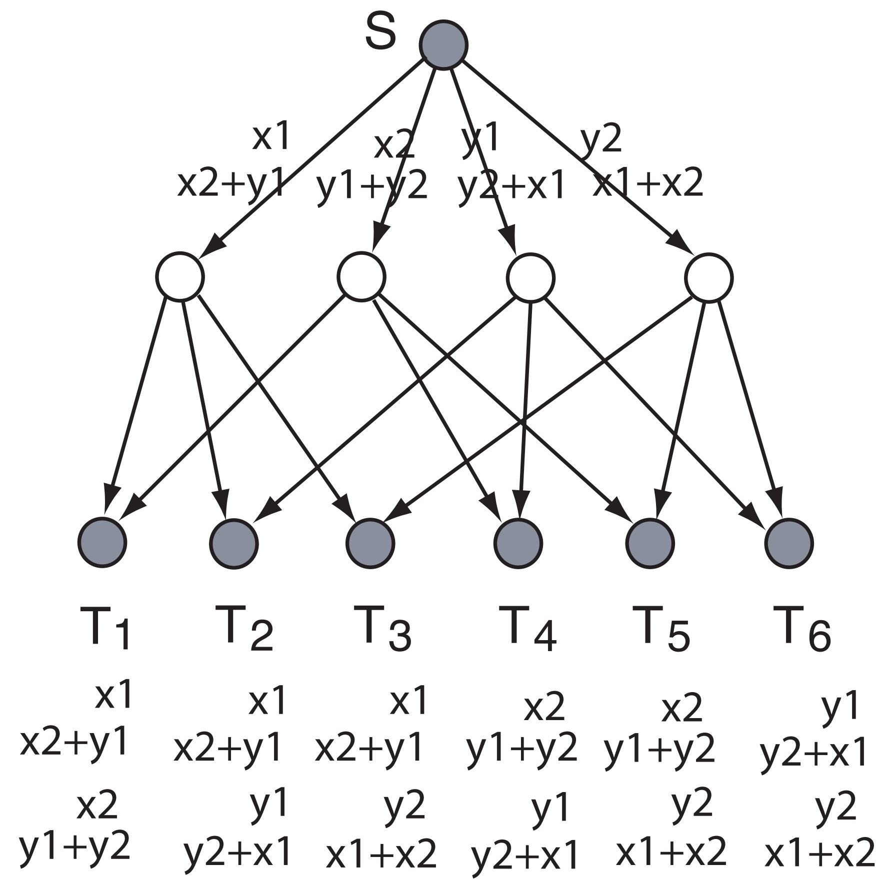
Combination Network Coding (CNC)

- CNC: network coding where information flows propagate along a $C(n, k)$ topology
- CNC can be applied to combination networks and general networks
- Among the first network coding schemes studied
- An important class of network coding scheme, relatively well-understood

Combination Network Coding (CNC)



algebraic coding



block coding

Coding Advantage in Uniform $C(n, k)$

Network	$\|V\|$	$\|M\|$	$\|E\|$	$\chi(N)$	$\pi(N)$	$\frac{\chi(N)}{\pi(N)}$	# of trees
butterfly	7	3	9	2	1.875	1.067	17
$C(3, 2)$	7	4	9	2	1.8	1.111	26
$C(4, 3)$	9	5	16	3	2.667	1.125	1,113
$C(4, 2)$	11	7	16	2	1.778	1.125	1,128
$C(5, 4)$	11	6	25	4	3.571	1.12	75,524
$C(5, 2)$	16	11	25	2	1.786	1.12	119,104
$C(5, 3)$	16	11	35	3	–	–	49,956,624

Why Undirected Networks?

- A simple and classic network model
- Already known: in directed networks, coding adv is unbounded
- A sense of "fair play"
- Undirected — future of the Internet?

Questions

- By how much can CNC increase throughput?
- By how much can CNC reduce multicast cost?
- Coding advantage *vs.* cost advantage, general network coding
- What's special about CNC?
- How does the coding advantage of CNC depend on n and k?
- Is the highest cost advantage of CNC realized in a uniform-cost network or a hetero-cost network?

Main Result

The potential for CNC to increase multicast throughput or reduce multicast cost is tightly bounded by a factor of 9/8.

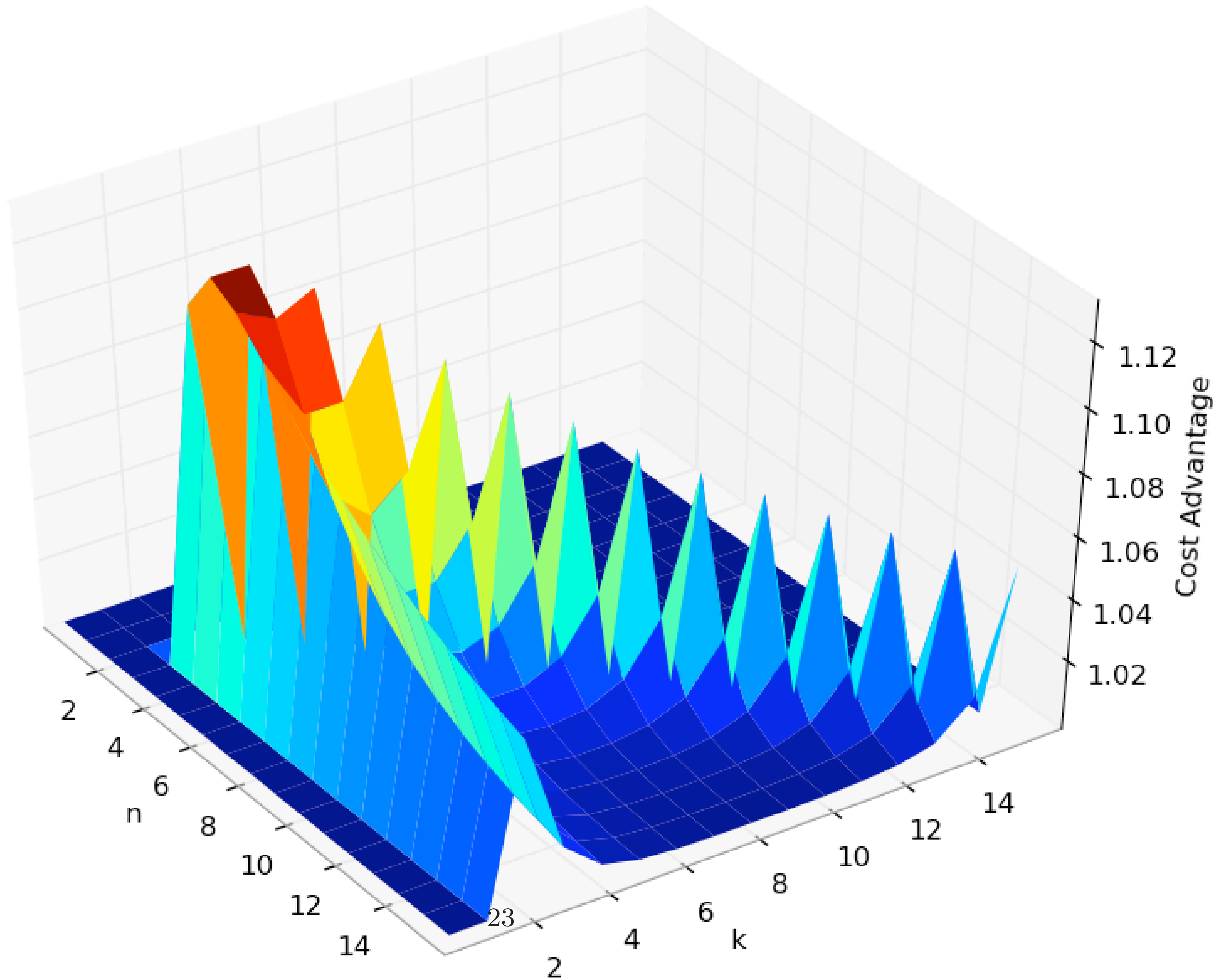
Proof Outline

- STEP 1. The cost advantage of CNC is at most 9/8 under uniform link cost
- STEP 2. The cost advantage of CNC under heterogenous link cost cannot be higher
- STEP 3. In general, coding adv (under hetero-capacity) $\leq$ cost adv (under hetero-cost)

STEP 1. Uniform Cost: Cost Adv $\leq 9/8$

- Minimum tree cost: $\binom{n}{k} + n - k + 1$
- Minimum CNC cost: $\binom{n}{k} + \frac{n}{k}$
- The cost advantage, closed-form representation: $\frac{\binom{n}{k}+n-k+1}{\binom{n}{k}+\frac{n}{k}}$
- Mathematically prove: $\frac{\binom{n}{k}+n-k+1}{\binom{n}{k}+\frac{n}{k}} \leq 9/8$
- Maximum value of $9/8$ attained at $C(4,2)$ and $C(4,3)$.

STEP 1. Uniform Cost: Cost Adv $\leq 9/8$



STEP 2. Hetero-Cost *vs* Uniform-Cost

- Want to prove: $\frac{W^h_{tree}}{W^h_{NC}} \leq \frac{W^u_{tree}}{W^u_{NC}}$
- Equivalent to: $\frac{W^h_{tree}}{W^u_{tree}} \leq \frac{W^h_{NC}}{W^u_{NC}}$
- (Assume $w(e) \geq 1, \forall e$:) cost inflation of opt tree is no worse than cost inflation of network coding
- Intuition: many candidate trees, can pick best tree to avoid more costly links.
- The hard part: which is the "best tree", or a "good tree"? How to bound its cost inflation?

STEP 2. Hetero-Cost *vs* Uniform-Cost

- Key idea: build a set of trees, claim that one of them must be "good"!

  – even though we don't know which.

- Technique: throughput-critical packing of multicast trees

STEP 2. Hetero-Cost *vs* Uniform-Cost

Definition: A critical packing is a tree packing scheme that exactly saturates every link of a multicast network.

Theorem: Every $C(n, k)$ network under uniform capacity has a critical packing of minimum multicast trees.

One of the trees in the critical packing must be a 'good' tree.

- Because average cost inflation of all trees in the critical packing = cost inflation of CNC

STEP 3. Coding Adv $\leq$ Cost Adv

Max-throughput multicast with network coding, LP:

Maximize $\quad d^{NC}$

Subject to:

$$\begin{cases} d^{NC} \leq f_i(\overrightarrow{T_i S}) & \forall T_i \in \mathcal{T} \\ f_i(\overrightarrow{uv}) \leq c(\overrightarrow{uv}) & \forall T_i \in \mathcal{T}, \forall\, \overrightarrow{uv} \neq \overrightarrow{T_i S} \\ \sum_{v \in N(u)} (f_i(\overrightarrow{uv}) - f_i(\overrightarrow{vu})) = 0 & \forall T_i \in \mathcal{T}, \forall u \\ c(\overrightarrow{uv}) + c(\overrightarrow{vu}) \leq c(uv) & \forall uv \neq T_i S \end{cases}$$

$$c(\overrightarrow{uv}), f_i(\overrightarrow{uv}), d^{NC} \geq 0 \qquad \forall T_i, \forall\, \overrightarrow{uv}$$

STEP 3. Coding Adv $\leq$ Cost Adv

By analyzing a Lagrange dual:

Theorem: A multicast rate d is feasible in an undirected multicast network (G, c) with network coding, if and only if for every link cost vector $w \in \mathcal{Q}_+^{E_G}$,

$$\frac{|G|_w}{\min_{d^{NC}(f)=1} |f|_w} \geq d$$

.

STEP 3. Coding Adv $\leq$ Cost Adv

Max-throughput multicast with tree packing, LP:

Maximize $\quad \sum_{t \in \mathcal{T}} f(t)$

Subject to:

$$\sum_{t \in \mathcal{T}: e \in t} f(t) \leq c(e) \qquad \forall e \in E_G \qquad \longleftrightarrow w(e)$$

$$f(t) \geq 0 \qquad \forall t \in \mathcal{T}$$

STEP 3. Coding Adv $\leq$ Cost Adv

By analyzing a Lagrange dual:

Theorem: A multicast rate d is feasible in an undirected multicast network (G, c) with tree packing, if and only if for every link cost vector $w \in \mathcal{Q}_+^{E_G}$,

$$\frac{|G|_w}{\min_{t \in \mathcal{T}} |t|_w} \geq d.$$

Furthermore, for the max-throughput d^*_{tree}, there exists a corresponding cost vector w^*_{tree}, such that equality holds.

STEP 3. Coding Adv $\leq$ Cost Adv

Combining the two previous theorems:

Theorem: In any undirected multicast network topology G, for a given link capacity vector $c \in \mathcal{Q}_+^{E_G}$, there always exists a link cost vector $w \in \mathcal{Q}_+^{E_G}$, such that the cost adv of NC in (G, w) is at least as high as the coding adv of NC in (G, c).

Same topology: Coding Adv $\leq$ Cost Adv

3-Step Proof Done

- We now finished the 3 steps for proving that 9/8 is an upper-bound for the coding adv and cost adv of CNC
- The bound is tight since it's achieved in known networks

Conclusion

CNC can increase multicast throughput by at most 1/8, can reduce multicast cost by at most 1/9, in undirected networks.