Network Coding Gaps for Completion Times of Multiple Unicasts

Nikolai Chukhin

Paper by: Haeupler, Wajc, Zuzic (2020)

15.05.2026

Content

Content

Multiple unicasts

Definition

A multiple-unicast instance is

$$M = (G, S), \qquad S = \{(s_i, t_i, d_i)\}_{i=1}^{k}.$$

The source s_i holds a packet of d_i sub-packets and wants to deliver it to t_i. Each edge e has capacity c_e per synchronous time step.

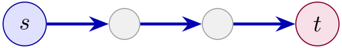
- A protocol has completion times $(T_1, \dots, T_k)$ if t_i decodes by time T_i.
- The makespan is $\max_i T_i$.

Routing vs. coding

Routing

Nodes only forward sub-packets they already know.

information behaves like a commodity.

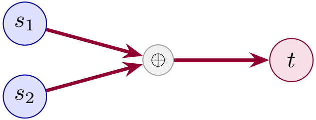


m_i moves as m_i

Network coding

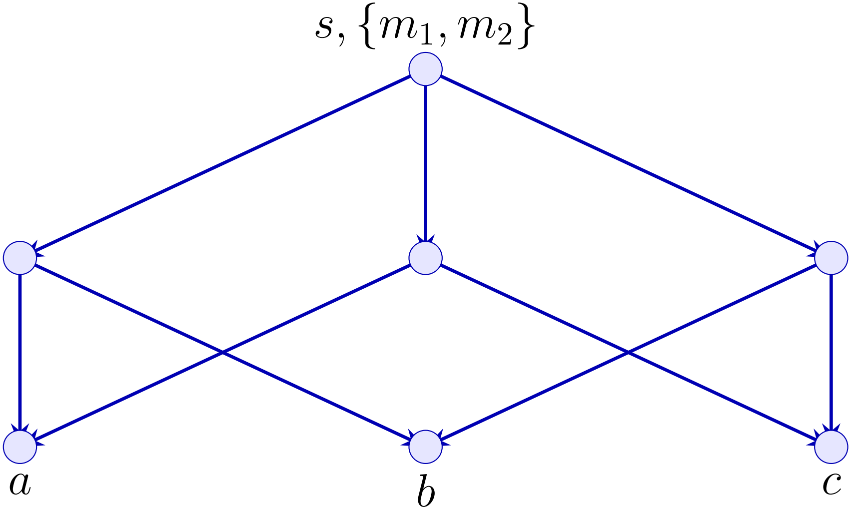
Nodes may send arbitrary functions of everything they have received.

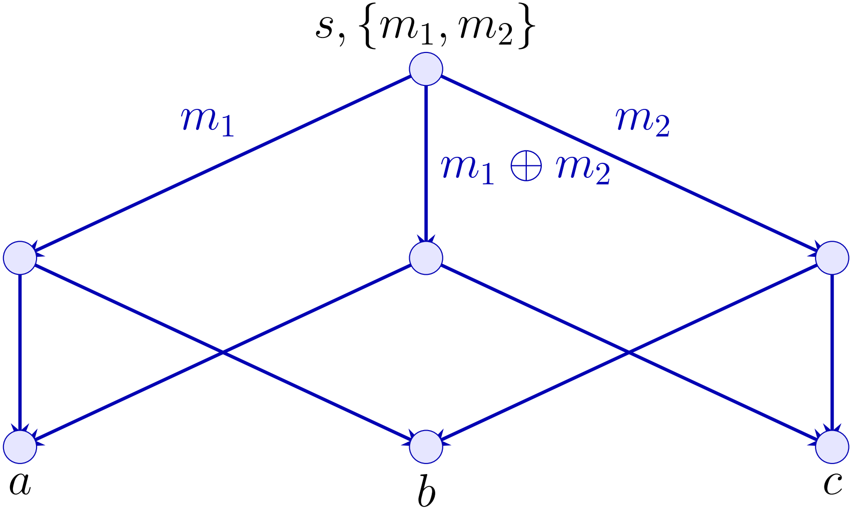
e.g. $m_1 \oplus m_2$.



$m_1 \oplus m_2$ can serve two decoders

[ACLY00]: coding can beat store-and-forward in networks.


$s, \{m_1, m_2\}$
a
b
c


$s, \{m_1, m_2\}$
m_1
$m_1 \oplus m_2$
m_2
a
b
c

Two basic lower bounds

Let T^*_{NC} be the optimal coding makespan and

$$\lambda_i = \min_{\substack{U \subseteq V \\ s_i \in U,\ t_i \notin U}} \sum_{e \in \delta(U)} c_e$$

be the s_i-t_i min-cut capacity. Then

$$T^*_{\mathrm{NC}} \geq \max_i \mathrm{dist}_G(s_i, t_i), \qquad T^*_{\mathrm{NC}} \geq \max_i \frac{d_i}{\lambda_i}.$$

Naive candidate

$$T^*_{\mathrm{NC}} = L_0 = \max \left\{ \max_i \mathrm{dist}_G(s_i, t_i), \max_i \frac{d_i}{\lambda_i} \right\}.$$

Counterexample

Completion-time coding gaps

Definition (Coding gap for a cost C)

For $C : \mathbb{R}^k_{\geq 0} \to \mathbb{R}_{\geq 0}$,

$$\text{gap}_C(M) = \frac{\min\{C(T) : T \text{ achievable by routing}\}}{\min\{C(T) : T \text{ achievable by coding}\}}.$$

The paper studies mainly *makespan*:

$$C(T_1, \ldots, T_k) = \|T\|_\infty = \max_i T_i.$$

What was known?

Completion-time setting

Before this work, no general makespan coding-gap bound was known.

What was known?

Completion-time setting

Before this work, no general makespan coding-gap bound was known.

The well-studied version was the asymptotic large-packet regime: *throughput maximization*. Let $C(w)$ be the fastest makespan after multiplying all demands by w. The maximum throughput is

$$\sup_{w \to \infty} \frac{w}{C(w)}.$$

What was known?

Completion-time setting

Before this work, no general makespan coding-gap bound was known.

The well-studied version was the asymptotic large-packet regime: *throughput maximization*. Let $C(w)$ be the fastest makespan after multiplying all demands by w. The maximum throughput is

$$\sup_{w \to \infty} \frac{w}{C(w)}.$$

Multiple-unicast conjecture

In the throughput setting, no instance is known where coding beats routing; it is conjectured that the throughput coding gap is always 1 [HKL04, LL04].

Why it is hard

The throughput conjecture is tied to lower bounds in other models:

network coding $\Longrightarrow$ data structures, external memory, circuits [AHJKL06, AFKL19, FHI

In particular, a throughput coding-gap bound $o(\log k)$ for all k-unicast instances would imply explicit super-linear circuit lower bounds [AFKL19].
The best general throughput upper bound is $O(\log k)$.

Main results

Theorem (Upper bound, Haeupler–Wajc–Zuzic)

For every k-unicast instance,

$$\mathrm{gap}_{\ell_\infty}(M) \le O\left(\log k \cdot \log\left(\sum_i d_i / \min_i d_i\right)\right).$$

For similarly-sized packets this is $O(\log^2 k)$.

Main results

Theorem (Upper bound, Haeupler–Wajc–Zuzic)

For every k-unicast instance,

$$\mathrm{gap}_{\ell_\infty}(M) \leq O\left(\log k \cdot \log\left(\sum_i d_i / \min_i d_i\right)\right).$$

For similarly-sized packets this is $O(\log^2 k)$.

Theorem (Lower bound)

There are infinitely many k-unicast instances with unit demands and

$$\mathrm{gap}_{\ell_\infty}(M) \geq \Omega(\log^c k)$$

for an absolute constant $c > 0$.

Content

A first cut lemma for routing

Lemma

Let $M = (G, S)$ *be a simple instance: unit capacities and unit demands. Suppose* $F \subseteq E$ *and*

$$\mathrm{dist}_{G \setminus F}(s_i, t_i) \geq T \quad \textit{for every } i.$$

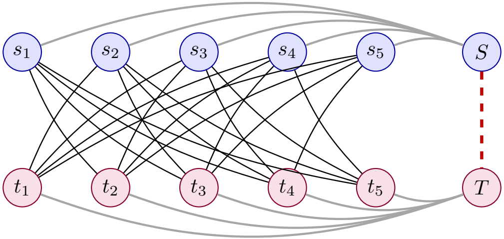
Then every ***routing*** *protocol has makespan at least*

$$\min\{T,\ k/|F|\}.$$

Proof.

Fix the routing paths. Either some session uses no edge of F, hence its path has length at least T. Or all k paths touch F, so one edge of F carries at least $k/|F|$ packets. □

Why the same statement fails for coding



- Delete ST: every own pair s_i, t_i becomes distance 5.
- Coding still finishes in 3 rounds.

$$S \text{ sends } \bigoplus_i m_i \text{ to } T, \qquad t_j \text{ already learns all } m_i \ (i \neq j),$$

so t_j recovers m_j by canceling all other messages.

A cut lemma that survives coding

Lemma

Let M be simple. Suppose $F \subseteq E$ and

$$\mathrm{dist}_{G \setminus F}(s_i, t_j) \geq T \quad \textit{for all } i, j.$$

*Then every **coding** protocol has makespan at least*

$$\min\{T,\ k/|F|\}.$$

Proof.

Give all sources to Alice and all sinks to Bob. This only makes the task easier. If a protocol finishes before time T, every bit learned by Bob must cross F. The cut carries at most $|F|T'$ bits in T' rounds, but Bob must learn k independent bits. Thus $|F|T' \geq k$. □

Moving cut

Definition (Moving cut)

A moving cut is an assignment of integer edge lengths

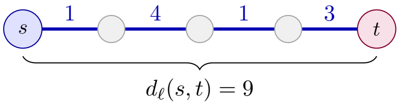
$$\ell : E \to \mathbb{Z}_{\geq 1}.$$

Its capacity is

$$C(\ell) = \sum_{e \in E} c_e(\ell_e - 1).$$

Its distance is at least T for a set of sessions if

$$d_\ell(s_i, t_j) \geq T \quad \text{for all sources } s_i \text{ and sinks } t_j.$$



Moving cut lower bound

Lemma

If $M = (G, S)$ admits a moving cut ℓ with

$$C(\ell) < \sum_i d_i \qquad \text{and} \qquad d_\ell(s_i, t_j) \geq T \quad \forall i, j,$$

then every coding protocol has makespan at least T.

Proof.

Assume a protocol finishes in $T - 1$ rounds. At time r, Alice spectates

$$A_r = \{v : \min_i d_\ell(s_i, v) \leq r\},$$

and Bob spectates $B_r = V \setminus A_r$.

Moving cut lower bound: the information count

Proof.

Bob sees a packet sent over $e = \{u, v\}$ from Alice's side to Bob's side only during rounds

$$r \in [\min_i d_\ell(s_i, u),\ \min_i d_\ell(s_i, v) - 1].$$

This interval has length at most $\ell_e - 1$.

Moving cut lower bound: the information count

Proof.

Bob sees a packet sent over $e = \{u, v\}$ from Alice's side to Bob's side only during rounds

$$r \in [\min_i d_\ell(s_i, u),\ \min_i d_\ell(s_i, v) - 1].$$

This interval has length at most $\ell_e - 1$.
Hence the total information Bob can receive from Alice is at most

$$\sum_e c_e(\ell_e - 1) = C(\ell) < \sum_i d_i.$$

But by time $T - 1$ Bob spectates every sink, since $d_\ell(s_i, t_j) \geq T$ for all i, j. Thus Bob recovers $\sum_i d_i$ independent bits from fewer bits, a contradiction. □

Content

Upper-bound proof: target

Suppose coding finishes the whole instance by time

$$T^* = \mathrm{OPT}_{\mathrm{NC}}(M).$$

Our goal is a routing protocol with makespan

$$O\left(T^* \log k \cdot \log(\sum_i d_i / \min_i d_i)\right)$$

Hop-bounded concurrent flow

For $T > 0$, let

$$P_i(T) = \{p : s_i \rightsquigarrow t_i,\ |p| \le T,\ p \text{ simple}\}.$$

Primal: $\mathrm{ConcurrentFlow}_M(T)$

$$\begin{aligned}
\max \quad & z \\
\text{s.t.} \quad & \sum_{p \in P_i(T)} f_i(p) \ge z d_i && \forall i \\
& \sum_{p \ni e} f_i(p) \le T c_e && \forall e \\
& f_i(p) \ge 0
\end{aligned}$$

Dual: $\mathrm{Cut}_M(T)$

$$\begin{aligned}
\min \quad & T \sum_e c_e \ell_e \\
\text{s.t.} \quad & \sum_{e \in p} \ell_e \ge h_i && \forall i,\ p \in P_i(T) \\
& \sum_i d_i h_i \ge 1 \\
& \ell_e, h_i \ge 0
\end{aligned}$$

If there is a routing protocol with T rounds, then $z = 1$.

High LP value gives routing

Proposition ([ST01, LMR94])

If $z, \{f_i(p)\}$ *is feasible for* $\text{ConcurrentFlow}_M(T)$*, then there is a routing protocol with makespan*

$$O(T/z).$$

High LP value gives routing

Proposition ([ST01, LMR94])

If $z, \{f_i(p)\}$ *is feasible for* $\text{ConcurrentFlow}_M(T)$*, then there is a routing protocol with makespan*

$$O(T/z).$$

Lemma

If the optimum of $\text{ConcurrentFlow}_M(T)$ *is at most* $1/10$*, then the coding makespan is at least*

$$\frac{T}{C \log k \cdot \log(\sum_i d_i / \min_i d_i)}$$

for a universal constant C*.*

Proof of the upper bound

Theorem

$$\mathrm{gap}_{\ell_\infty}(M) \leq O\left(\log k \cdot \log(\sum_i d_i / \min_i d_i)\right).$$

Proof of the upper bound

Theorem

$$\text{gap}_{\ell_\infty}(M) \leq O\left(\log k \cdot \log(\sum_i d_i / \min_i d_i)\right).$$

Proof.

Let $T^* = \text{OPT}_{\text{NC}}(M)$ and $T = (C+1)T^* \log k \cdot \log(\sum_i d_i / \min_i d_i)$. If $\text{ConcurrentFlow}_M(T)$ had value at most $1/10$, the low-LP theorem would imply a coding lower bound larger than T^*, impossible.

Proof of the upper bound

Theorem

$$\mathrm{gap}_{\ell_\infty}(M) \leq O\left(\log k \cdot \log(\sum_i d_i / \min_i d_i)\right).$$

Proof.

Let $T^* = \mathrm{OPT}_{\mathrm{NC}}(M)$ and $T = (C+1)T^* \log k \cdot \log(\sum_i d_i / \min_i d_i)$. If $\mathrm{ConcurrentFlow}_M(T)$ had value at most $1/10$, the low-LP theorem would imply a coding lower bound larger than T^*, impossible.

Thus the LP value is $\Omega(1)$. By the rounding proposition, there is a routing protocol with makespan

$$O(T) = O\left(T^* \log k \cdot \log(\sum_i d_i / \min_i d_i)\right).$$

□

Bucketing the weighted distances

Claim

Given $h_1, \ldots, h_k, d_1, \ldots, d_k \geq 0$ *with* $\sum_i d_i h_i \geq 1$, *then there is a nonempty* $I \subseteq [k]$ *with*

$$\min_{i \in I} h_i \geq \frac{1}{\alpha_{\text{gap}} \sum_{i \in I} d_i}, \qquad \alpha_{\text{gap}} = O\left(\log(\sum_i d_i / \min_i d_i)\right).$$

Bucketing the weighted distances

Claim

Given $h_1, \ldots, h_k, d_1, \ldots, d_k \geq 0$ *with* $\sum_i d_i h_i \geq 1$, *then there is a nonempty* $I \subseteq [k]$ *with*

$$\min_{i \in I} h_i \geq \frac{1}{\alpha_{\text{gap}} \sum_{i \in I} d_i}, \qquad \alpha_{\text{gap}} = O\left(\log(\sum_i d_i / \min_i d_i)\right).$$

Proof.

Sort $h_1 \geq h_2 \geq \cdots \geq h_k$ and write $d([j]) = \sum_{i \leq j} d_i$. If no prefix works, then $h_i < \frac{1}{\alpha\, d([i])}$. Multiplying by d_i and summing gives

$$1 \leq \sum_i d_i h_i < \frac{1}{\alpha} \sum_i \frac{d_i}{d([i])} \leq \frac{1}{\alpha}\left(1 + \ln \frac{\sum_i d_i}{\min_i d_i}\right),$$

a contradiction for this choice of α. □

Recall: the dual certificate

A feasible dual solution consists of edge lengths ℓ_e and weights h_i:

$$
\begin{aligned}
\min \quad & T \sum_e c_e \ell_e \\
\text{s.t.} \quad & \sum_{e \in p} \ell_e \geq h_i \qquad \forall i,\ p \in P_i(T) \\
& \sum_i d_i h_i \geq 1 \\
& \ell_e, h_i \geq 0.
\end{aligned}
$$

After bucketing, for some nonempty $I \subseteq [k]$, every $p \in P_i(T)$ with $i \in I$ satisfies

$$\sum_{e \in p} \ell_e \geq h_i \geq \frac{1}{\alpha_{\text{gap}} \sum_{q \in I} d_q}$$

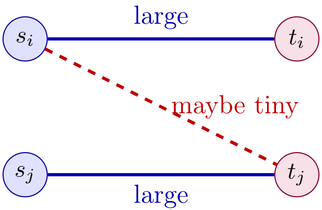
The pairwise-to-all-pairs obstacle

The dual LP only certifies large distance for the requested pairs:

$$d_\ell(s_i, t_i) \text{ is large for } i \in I.$$

Moving cuts need the stronger condition

$$d_\ell(s_i, t_j) \text{ is large for all } i, j.$$



Padded partitions

Let (X, d) be a metric space and let $\mathcal{P}$ be a partition of X.

Definition

- $\mathcal{P}$ is Δ-bounded if every part has diameter at most Δ.
- Denote by $\mathcal{P}(x)$ the part containing x.
- A distribution $\mathfrak{P}$ over Δ-bounded partitions is (β, Δ)-padded if, for a universal constant $\delta > 0$,

$$\Pr_{\mathcal{P}\sim\mathfrak{P}}[B(x,\gamma\Delta) \not\subseteq \mathcal{P}(x)] \leq \beta\gamma \qquad \forall x \in X,\ 0 \leq \gamma \leq \delta.$$

Theorem (Gupta–Krauthgamer–Lee [GKL03])

Every k-point metric admits a (β, Δ)-padded decomposition for every $\Delta > 0$, with $\beta = O(\log k)$.

Metric pruning lemma

Lemma

Let (X, d) be a metric space. Given n pairs (s_i, t_i) using at most k distinct points in $\cup_i \{s_i, t_i\}$ and satisfying

$$d(s_i, t_i) \geq T,$$

there is $I \subseteq [n]$, $|I| \geq n/9$, such that

$$d(s_i, t_j) \geq \frac{T}{O(\log k)} \quad \textit{for all } i, j \in I.$$

Metric pruning: proof

Proof.

Take a (β, Δ)-padded partition with

$$\Delta = T - 1, \qquad \beta = O(\log k).$$

Since $d(s_i, t_i) \geq T$, the two endpoints of each pair lie in different parts.

Metric pruning: proof

Proof.

Take a (β, Δ)-padded partition with

$$\Delta = T - 1, \qquad \beta = O(\log k).$$

Since $d(s_i, t_i) \geq T$, the two endpoints of each pair lie in different parts.
Let $\rho = \Delta/(2\beta)$. With probability at least $1/2$,

$$B(s_i, \rho) \text{ is inside one part.}$$

Now flip an independent fair coin for every part and put heads in U, tails in V. Keep pair i if $B(s_i, \rho) \subseteq U$ and $t_i \in V$.

Metric pruning: proof

Proof.

Take a (β, Δ)-padded partition with

$$\Delta = T - 1, \qquad \beta = O(\log k).$$

Since $d(s_i, t_i) \geq T$, the two endpoints of each pair lie in different parts.
Let $\rho = \Delta/(2\beta)$. With probability at least $1/2$,

$$B(s_i, \rho) \text{ is inside one part.}$$

Now flip an independent fair coin for every part and put heads in U, tails in V. Keep pair i if $B(s_i, \rho) \subseteq U$ and $t_i \in V$.
For every i, this happens with probability at least $1/8$, so the expected number kept is $n/8$. With constant probability at least $n/9$ are kept. For kept i, j, $B(s_i, \rho) \subseteq U$ and $t_j \in V$, hence $d(s_i, t_j) > \rho = T/O(\log k)$. □

From a dual solution to a moving cut

Lemma

If $\text{ConcurrentFlow}_M(T)$ *has optimum* $z^* \leq 1/10$*, then for some sub-instance* $\widetilde{I}$ *there is a moving cut* $\widetilde{\ell}$ *with*

$$C(\widetilde{\ell}) < \sum_{i \in \widetilde{I}} d_i, \qquad d_{\widetilde{\ell}}(s_i, t_j) \geq \frac{T}{O(\alpha_{\text{gap}} \log k)}.$$

Proof.

Let (ℓ_e, h_i) be a dual solution with $T \sum_e c_e \ell_e = z^* \leq 1/10$. Use the bucketing claim to get I with $h_i \geq \frac{1}{\alpha_{\text{gap}} \sum_{q \in I} d_q}$, for all $i \in I$. Define integer lengths

$$\widetilde{\ell}_e = 1 + \left\lfloor \ell_e T \sum_{q \in I} d_q \right\rfloor.$$

Proof.

$$\widetilde{\ell}_e = 1 + \left\lfloor \ell_e T \sum_{q\in I} d_q \right\rfloor .$$

The capacity is bounded by

$$C(\widetilde{\ell}) = \sum_e c_e(\widetilde{\ell}_e - 1) \le \sum_e c_e \ell_e \cdot T \sum_{q\in I} d_q = z^* \sum_{q\in I} d_q \le \frac{1}{10} \sum_{q\in I} d_q.$$

For $i \in I$, every $s_i t_i$ path p has

$$\widetilde{\ell}(p) > T/\alpha_{\text{gap}}.$$

Indeed, if $|p| > T$, then $\widetilde{\ell}_e \ge 1$ for every edge, so $\widetilde{\ell}(p) \ge |p| > T \ge T/\alpha_{\text{gap}}$. If $|p| \le T$, dual feasibility gives

$$\ell(p) \ge h_i \ge \frac{1}{\alpha_{\text{gap}} \sum_{q\in I} d_q}.$$

Hence

$$\widetilde{\ell}(p) \ge T \sum_{q\in I} d_q \cdot \ell(p) \ge T/\alpha_{\text{gap}}.$$

Finish the moving cut construction

Proof.

Apply the metric pruning lemma to the metric $d_{\widetilde{\ell}}$, repeating pair i exactly d_i times. It returns a submultiset whose total demand is at least

$$\frac{1}{9}\sum_{q\in I} d_q.$$

For its underlying index set $\widetilde{I}$,

$$d_{\widetilde{\ell}}(s_i,t_j) \geq \frac{T}{O(\alpha_{\text{gap}}\log k)} \qquad \forall i,j\in\widetilde{I}.$$

Since

$$C(\widetilde{\ell}) \leq \frac{1}{10}\sum_{q\in I} d_q < \sum_{i\in\widetilde{I}} d_i,$$

this is a valid moving cut for the sub-instance. □

Low LP value lower-bounds coding

Theorem

If the optimum of $\mathrm{ConcurrentFlow}_M(T)$ *is at most* $1/10$, *then*

$$\mathrm{OPT}_{\mathrm{NC}}(M) \geq \frac{T}{C \log k \cdot \log(\sum_i d_i / \min_i d_i)}.$$

Proof.

The previous lemma gives a moving cut for a sub-instance with distance $T/O(\alpha_{\mathrm{gap}} \log k)$. The moving cut lower bound applies to that sub-instance, and any protocol for M also solves the sub-instance. Finally,

$$\alpha_{\mathrm{gap}} = O\left(\log(\sum_i d_i / \min_i d_i)\right).$$

□

Content

Goal of this section

Now we will prove the lower-bound direction:

$$\exists c > 0 \text{ and infinitely many } k\text{-unicast instances } M \quad \text{gap}_{\ell_\infty}(M) \geq \Omega(\log^c k).$$

Goal of this section

Now we will prove the lower-bound direction:

$$\exists c > 0 \text{ and infinitely many } k\text{-unicast instances } M \quad \mathrm{gap}_{\ell_\infty}(M) \geq \Omega(\log^c k).$$

It is enough to build instances with

$$T_{\mathrm{NC}} \leq 3^{2^i}, \qquad T_{\textsc{Routing}} \geq 5^{2^i}, \qquad \log k \leq 2^{O(2^i)}.$$

Then

$$\frac{T_{\textsc{Routing}}}{T_{\mathrm{NC}}} \geq (5/3)^{2^i} \geq (\log k)^c.$$

Gap instances

Definition

A gap instance is $I = (G, S, F)$ with unit capacities, unit demands, disjoint terminals, and cut edges $F \subseteq E(G)$. It has parameters

$$(T_{\mathrm{NC}}, D_F, |F|, K, |E|, \rho, \sigma)$$

if:

- coding makespan is at most T_{NC};
- For all $i \in [K]$, we have $\mathrm{dist}_{G \setminus F}(s_i, t_i) \geq D_F$;
- $K = |S|$ and $|E(G)| \leq |E|$;
- $K/|F| \geq \rho$ and $|E|/|F| \leq \sigma$.

Observation

If I has parameters $(T_{\mathrm{NC}}, D_F, |F|, K, |E|, \rho, \sigma)$ and $D_F \leq \rho$, then the routing makespan is at least D_F.

Observation

If I has parameters $(T_{\mathrm{NC}}, D_F, |F|, K, |E|, \rho, \sigma)$ and $D_F \leq \rho$, then the routing makespan is at least D_F. Moreover, every routing protocol leaves at least

$$K\left(1 - \frac{D_F - 1}{\rho}\right)$$

sessions unfinished before time D_F.

why?

Observation

If I has parameters $(T_{\mathrm{NC}}, D_F, |F|, K, |E|, \rho, \sigma)$ and $D_F \leq \rho$, then the routing makespan is at least D_F. Moreover, every routing protocol leaves at least

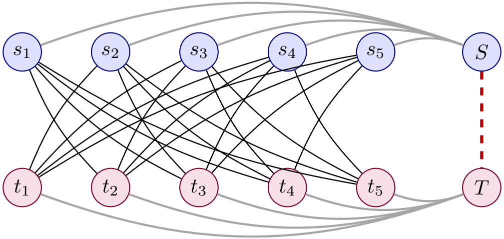
$$K\left(1 - \frac{D_F - 1}{\rho}\right)$$

sessions unfinished before time D_F.

Proof.

Before round D_F, any routed packet that reaches its sink must use an edge of F. The cut has $|F|$ unit-capacity edges, so fewer than $D_F|F|$ packets can cross it in the first $D_F - 1$ rounds. Since $K/|F| \geq \rho$, at most $(D_F - 1)|F| \leq K(D_F - 1)/\rho$ sessions finish before time D_F. □

The base gap: coding time 3, routing time 5



Observation

For every $K \geq 5$ this family has parameters

$$(T_{\mathrm{NC}}, D_F, |F|, K, |E|, \rho, \sigma) = (3, 5, 1, K, \Theta(K^2), K, \Theta(K^2)).$$

Colored bipartite graph

Definition

$B \in \mathcal{B}(n_1, n_2, m, k, g)$ is bipartite with parts $|V_1| = n_1, |V_2| = n_2$, left degree m, right degree k, and girth at least g. Each edge has:

$$\chi_1(e) \in [m] \quad \text{and} \quad \chi_2(e) \in [k].$$

Around each left vertex, edge colors cover $[m]$ and session colors are the same.
Around each right vertex, session colors cover $[k]$ and edge colors are the same.

Lemma (BGS17)

For all $r, m, g \geq 3$, such $B \in \mathcal{B}(n_1, n_2, m, k, 2g)$ exist with

$$n_1, n_2 \leq (9mk)^{g+3}.$$

Product of two gap instances

- For each $i \in \{1, 2\}$, let $I_i = (G_i, S_i, F_i)$ be a gap instance $(T_i^{\mathrm{NC}}, D_i, |F_i|, K_i, |E_i|, \rho_i, \sigma_i)$.

Product of two gap instances

- For each $i \in \{1, 2\}$, let $I_i = (G_i, S_i, F_i)$ be a gap instance $(T_i^{\mathrm{NC}}, D_i, |F_i|, K_i, |E_i|, \rho_i, \sigma_i)$.
- Let $B \in \mathcal{B}(n_1, n_2, 2(|E_1| - |F_1|), K_2, g)$ with girth $g > 2D_1D_2$.

Product of two gap instances

- For each $i \in \{1,2\}$, let $I_i = (G_i, S_i, F_i)$ be a gap instance $(T_i^{\mathrm{NC}}, D_i, |F_i|, K_i, |E_i|, \rho_i, \sigma_i)$.
- Let $B \in \mathcal{B}(n_1, n_2, 2(|E_1| - |F_1|), K_2, g)$ with girth $g > 2D_1D_2$.
- We refer to I_1 as the outer instance and to I_2 as the inner instance.

Product of two gap instances

- For each $i \in \{1, 2\}$, let $I_i = (G_i, S_i, F_i)$ be a gap instance $(T_i^{\mathrm{NC}}, D_i, |F_i|, K_i, |E_i|, \rho_i, \sigma_i)$.
- Let $B \in \mathcal{B}(n_1, n_2, 2(|E_1| - |F_1|), K_2, g)$ with girth $g > 2D_1D_2$.
- We refer to I_1 as the outer instance and to I_2 as the inner instance.
- The product $I_+ = T(I_1, I_2, B)$ does this:

Picture!

Product of two gap instances

- For each $i \in \{1,2\}$, let $I_i = (G_i, S_i, F_i)$ be a gap instance $(T_i^{\text{NC}}, D_i, |F_i|, K_i, |E_i|, \rho_i, \sigma_i)$.
- Let $B \in \mathcal{B}(n_1, n_2, 2(|E_1| - |F_1|), K_2, g)$ with girth $g > 2D_1D_2$.
- We refer to I_1 as the outer instance and to I_2 as the inner instance.
- The product $I_+ = T(I_1, I_2, B)$ does this:

$$\begin{aligned} T_+^{\text{NC}} &= ?, \\ D_+ &= \text{later}, \\ |F_+| &= ?, \\ K_+ &= ?, \\ |E_+| &= ?, \\ \rho_i &= ?, \\ \sigma_i &= ?. \end{aligned}$$

Product of two gap instances

- For each $i \in \{1,2\}$, let $I_i = (G_i, S_i, F_i)$ be a gap instance $(T_i^{\mathrm{NC}}, D_i, |F_i|, K_i, |E_i|, \rho_i, \sigma_i)$.
- Let $B \in \mathcal{B}(n_1, n_2, 2(|E_1| - |F_1|), K_2, g)$ with girth $g > 2D_1D_2$.
- We refer to I_1 as the outer instance and to I_2 as the inner instance.
- The product $I_+ = T(I_1, I_2, B)$ does this:

$$
\begin{aligned}
T_+^{\mathrm{NC}} &= T_1^{\mathrm{NC}} \cdot T_2^{\mathrm{NC}}, \\
D_+ &= \text{later}, \\
|F_+| &= n_1 \cdot |F_1| + n_2 \cdot |F_2|, \\
K_+ &= n_1 K_1, \\
|E_+| &= T_2^{\mathrm{NC}} n_1 |F_1| + n_2 |F_2|, \\
\rho_i &= \rho_1 \frac{1}{1 + 2\sigma_1/\rho_2}, \\
\sigma_i &= \sigma_2 \frac{1 + T_2^{\mathrm{NC}}/2}{1 + \rho_2/(2\sigma_1)}.
\end{aligned}
$$

Product: coding composes

Lemma

The product instance admits a coding protocol with makespan

$$T_+^{\mathrm{NC}} = T_1^{\mathrm{NC}} T_2^{\mathrm{NC}}.$$

Proof.

Take a coding protocol for I_1 of length T_1^{NC}. One round of that protocol sends one unit across every arc corresponding to an edge of G_1.

Each non-cut edge of I_1 has been replaced by a session in a copy of I_2. So simulate one outer round by running the T_2^{NC}-round coding protocol for all relevant copies of I_2.

After T_1^{NC} outer rounds, each consisting of T_2^{NC} inner rounds, every outer session is delivered.

$$T_{\mathrm{NC}}(I_+) \leq T_1^{\mathrm{NC}} T_2^{\mathrm{NC}}.$$

□

Product: routing distance multiplies

Lemma

If B has girth $g = 2D_1D_2$, then in $G_+ \setminus F_+$, $\mathrm{dist}(s_i, t_i) \geq D_1D_2$ for every session of the product.

Proof.

Let p be a shortest source-sink path in $G_+ \setminus F_+$, minimized over all product sessions. Project p to the closed walk q in the bipartite graph B.

Product: routing distance multiplies

Lemma

If B has girth $g = 2D_1D_2$, then in $G_+ \setminus F_+$, $\text{dist}(s_i, t_i) \geq D_1D_2$ *for every session of the product.*

Proof.

Let p be a shortest source-sink path in $G_+ \setminus F_+$, minimized over all product sessions.
Project p to the closed walk q in the bipartite graph B.
If q contains a non-degenerate cycle, then $|q| \geq g$. Every time q enters a right node of B, the path p spends at least one real edge inside an inner copy. Thus
$|p| \geq |q|/2 \geq g/2 = D_1D_2$.

Projection to a tree

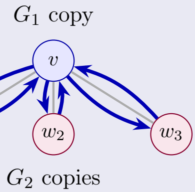
Proof.

Otherwise, the edges used by q span a tree $\mathcal{T}$ in B. Root $\mathcal{T}$ at the left node $v \in V_1(B)$ containing the terminals of p. Suppose $\mathcal{T}$ has depth 1.

Projection to a tree

Proof.

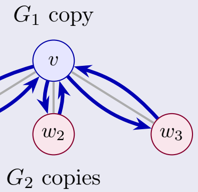
Otherwise, the edges used by q span a tree $\mathcal{T}$ in B. Root $\mathcal{T}$ at the left node $v \in V_1(B)$ containing the terminals of p. Suppose $\mathcal{T}$ has depth 1.



Projection to a tree

Proof.

Otherwise, the edges used by q span a tree $\mathcal{T}$ in B. Root $\mathcal{T}$ at the left node $v \in V_1(B)$ containing the terminals of p. Suppose $\mathcal{T}$ has depth 1.

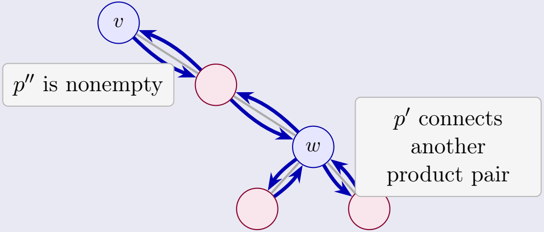


Then every path through a leaf node is a source-sink walk of length at least D_2. Contract every such inner walk back to the corresponding non-cut outer edge. The result is a surviving s_i–t_i walk in the outer instance, so it uses at least D_1 non-cut outer edges. Therefore $|p| \geq D_1 D_2$.

Tree of larger depth

Proof.

Now suppose $\mathcal{T}$ has depth more than 1. Then the Euler tour enters a subtree rooted at some left node $w \neq v$ and later leaves it.



The corresponding subpath p' connects a different product session, while the rest contains a nonempty piece p'' passing through a right node. By minimality, $|p'| \geq |p|$; but p contains both p' and p'', so $|p| \geq |p'| + |p''| > |p'|$, a contradiction. $\square$

Product parameters

Lemma

Let I_i have parameters $(T_i^{\mathrm{NC}}, D_i, |F_i|, K_i, |E_i|, \rho_i, \sigma_i)$. Assume $|E_i|/|F_i| \geq 2$ and $T_i^{\mathrm{NC}} \geq 2$. Then $I_+ = T(I_1, I_2, B)$ has parameters

$$
\begin{aligned}
T_+^{\mathrm{NC}} &= T_1^{\mathrm{NC}} T_2^{\mathrm{NC}}, & D_+ &= D_1 D_2, \\
|F_+| &= n_1|F_1| + n_2|F_2|, & K_+ &= n_1 K_1, \\
|E_+| &= T_2^{\mathrm{NC}} n_1 |F_1| + n_2|E_2|,
\end{aligned}
$$

and

$$
\rho_+ = \rho_1 \frac{1}{1 + 2\sigma_1/\rho_2}, \qquad \sigma_+ = \sigma_2 \frac{1 + T_2^{\mathrm{NC}}/2}{1 + \rho_2/(2\sigma_1)}.
$$

Recursive family

Definition

For $R \geq 5$, define $I(0, R)$ to be the base instance with

$$(T^{\mathrm{NC}}_{0,R}, D_{0,R}, F_{0,R}, K_{0,R}, E_{0,R}, \rho_{0,R}, \sigma_{0,R}) = (3, 5, 1, R, \Theta(R^2), R, \Theta(R^2)).$$

For $i + 1 > 0$ define

$$I(i + 1, R) = T(I_1, I_2),$$

where

$$I_1 = I(i, 3R), \qquad I_2 = I(i, \sigma_{i,3R}).$$

Observation

For all $i \geq 0$,

$$T^{\mathrm{NC}}_{i,R} = 3^{2^i}, \qquad D_{i,R} = 5^{2^i}.$$

Controlling the size: first recurrences

Lemma

For every $i \geq 0$ *and* $R \geq 5$*:*

$$\rho_{i,R} \geq R,$$

$$\sigma_{i+1,R} \leq 3^{2^i}\, \sigma_{i,\sigma_{i,3R}},$$

and

$$\log E_{i+1,R} \leq O(5^{2^{i+1}}) \log\big(E_{i,3R} E_{i,\sigma_{i,3R}}\big).$$

Proof.

The first two inequalities follow directly using $\rho_2 = \sigma_1$. The colored graph has

$$\max(n_1, n_2) \leq (2(|E_1| - |F_1|)K_2)^{O(D_1 D_2)} \leq (|E_1||E_2|)^{O(5^{2^{i+1}})} \implies$$

$$E_{i+1,R} = T_2^{\mathrm{NC}} n_1 |F_1| + n_2 |E_2| \leq T_2^{\mathrm{NC}} \max(n_1, n_2) |E_1||E_2|.$$

□

Bounding $\sigma_{i,R}$

Lemma

For all $i \geq 0$ and $R \geq 5$,

$$\log \sigma_{i,R} \leq 2^{O(2^i)} \log R.$$

Proof.

$$\sigma_{i+1,R} \leq 3^{2^i} \sigma_{i,\sigma_{i,3R}} \implies \log \sigma_{i+1,R} \leq O(2^i) + \log \sigma_{i,\sigma_{i,3R}}.$$

Then

$$\begin{aligned} \log \sigma_{i+1,R} &\leq O(2^i) + \frac{1}{c}(c2)^{2^i} \log \sigma_{i,3R} \\ &\leq O(2^i) + \frac{1}{c}(c2)^{2^i} \cdot c2^{2^i} \log(3R) \\ &\leq \frac{1}{c}(c2)^{2^{i+1}} \log R \end{aligned}$$

for a sufficiently large constant c. □

Bounding the number of edges

Lemma

For all $i \geq 0$ *and* $R \geq 5$,

$$\log E_{i,R} \leq 2^{O(2^i)} \log R.$$

Proof.

Use the recurrence

$$\log E_{i+1,R} \leq O(5^{2^{i+1}}) \left(\log E_{i,3R} + \log E_{i,\sigma_{i,3R}}\right).$$

Inductively,

$$\begin{aligned} \log E_{i+1,R} &\leq O(5^{2^{i+1}}) \left(2^{O(2^i)} \log(3R) + 2^{O(2^i)} \log \sigma_{i,3R}\right) \\ &\leq 2^{O(2^{i+1})} \log R, \end{aligned}$$

where the last step uses the bound on $\sigma_{i,3R}$. □

Proof of the polylog lower bound

Theorem

There is an absolute $c > 0$ and infinitely many k-unicast instances with

$$\text{gap}_{\ell_\infty}(M) \geq \Omega(\log^c k).$$

Proof.

Take $I(i, R)$ with $R = 5^{2^i}$. It has coding makespan at most

$$T_{i,R}^{\text{NC}} = 3^{2^i},$$

and because $D_{i,R} = 5^{2^i} \leq \rho_{i,R}$, its routing makespan is at least

$$D_{i,R} = 5^{2^i}.$$

Thus the gap is at least $(5/3)^{2^i}$.

Polylog lower bound: in terms of k

Proof.

The number of sessions k is at most the number of edges $E_{i,R}$, and

$$\log E_{i,R} \leq 2^{O(2^i)} \log R \leq 2^{O(2^i)}.$$

Therefore

$$(5/3)^{2^i} \geq \log^c E_{i,R} \geq \log^c k$$

for some constant $c > 0$. □

Open directions

- Close the gap between $O(\log^2 k)$ and $\Omega(\log^c k)$?
- Extend completion-time coding gaps beyond multiple unicasts (Anton's idea)?
- Derive lower bounds in other models?

Bibliography

R. Ahlswede, N. Cai, S.-Y. R. Li, and R. W. Yeung.
Network information flow.
IEEE Transactions on Information Theory, 2000.

M. Adler, N. J. Harvey, K. Jain, R. Kleinberg, and A. R. Lehman.
On the capacity of information networks.
SODA, 2006.

P. Afshani, C. B. Freksen, L. Kamma, and K. G. Larsen.
Lower bounds for multiplication via network coding.
ICALP, 2019.

A. Al-Bashabsheh and A. Yongacoglu.
On the k-pairs problem.
ISIT, 2008.

M. Braverman, S. Garg, and A. Schvartzman.
Coding in undirected graphs is either very helpful or not helpful at all.
ITCS, 2017.

Bibliography

A. Farhadi, M. Hajiaghayi, K. G. Larsen, and E. Shi.
Lower bounds for external memory integer sorting via network coding.
STOC, 2019.

A. Gupta, R. Krauthgamer, and J. R. Lee.
Bounded geometries, fractals, and low-distortion embeddings.
FOCS, 2003.

Bibliography

Y. Aumann and Y. Rabani.
An $O(\log k)$ approximate min-cut max-flow theorem and approximation algorithm.
SIAM Journal on Computing, 1998.

N. J. Harvey, R. D. Kleinberg, and A. R. Lehman.
Comparing network coding with multicommodity flow for the k-pairs communication problem.
Technical report, 2004.

N. J. Harvey, R. Kleinberg, and A. R. Lehman.
On the capacity of information networks.
IEEE/ACM Transactions on Networking, 2006.

K. Jain, V. V. Vazirani, R. Yeung, and G. Yuval.
On the capacity of multiple unicast sessions in undirected graphs.
IEEE/ACM Transactions on Networking, 2006.

Bibliography

G. Kramer and S. A. Savari.
Edge-cut bounds on network coding rates.
Journal of Network and Systems Management, 2006.

M. Langberg and M. Medard.
On the multiple unicast network coding conjecture.
Allerton, 2009.

Z. Li and B. Li.
Network coding: The case of multiple unicast sessions.
Allerton, 2004.

N. Linial, E. London, and Y. Rabinovich.
The geometry of graphs and some of its algorithmic applications.
Combinatorica, 1995.

Bibliography

F. T. Leighton, B. M. Maggs, and S. B. Rao.
Packet routing and job-shop scheduling in $O(\text{congestion} + \text{dilation})$ steps.
Combinatorica, 1994.

B. Haeupler, D. Wajc, and G. Zuzic.
Network coding gaps for completion times of multiple unicasts.
arXiv:1905.02805, 2020.

A. Srinivasan and C.-P. Teo.
A constant-factor approximation algorithm for packet routing and balancing local vs. global criteria.
SIAM Journal on Computing, 2001.

T. Xiahou, Z. Li, C. Wu, and J. Huang.
A geometric perspective to multiple-unicast network coding.
IEEE Transactions on Information Theory, 2014.

X. Yin, Z. Li, Y. Liu, and X. Wang.
A reduction approach to the multiple-unicast conjecture in network coding.
IEEE Transactions on Information Theory, 2018.