Improved Hardness of Approximation of Diameter in the CONGEST Model

Nikolai Chukhin
Paper by: Grossman, Khoury, Paz (DISC 2020)

08.05.2026

Content

1. The problem
2. Communication lower bounds
3. Proof I: diameter 3 vs 5
4. Robust lower bounds
5. Wrap-up

Content

1. The problem
2. Communication lower bounds
3. Proof I: diameter 3 vs 5
4. Robust lower bounds
5. Wrap-up

The CONGEST model

Definition (CONGEST [Pel00])

The input is an n-node graph $G = (V, E)$. Each node is a processor; each edge is a communication link. In every synchronous round, each node sends an $O(\log n)$-bit message to each neighbor.

- Local computation is free.
- The cost is the number of communication rounds.
- The network topology is also the input.

Diameter in a distributed network

Definition

$$\mathrm{diam}(G) = \max_{u,v \in V} \mathrm{dist}_G(u, v).$$

Brain Teaser: Why diameter is a natural problem in distributed computing?

Diameter in a distributed network

Definition

$$\mathrm{diam}(G) = \max_{u,v \in V} \mathrm{dist}_G(u, v).$$

Brain Teaser: Why diameter is a natural problem in distributed computing?

Theorem ([FHW12, HFQ+16])

Exact diameter in CONGEST *requires* $\Theta(n/\log n + D)$ *rounds.*

The easy baseline: a 1/2-approximation

Brain Teaser: How to compute $\frac{1}{2}$ approximation?

The easy baseline: a 1/2-approximation

Brain Teaser: How to compute $\frac{1}{2}$ approximation?

Pick an arbitrary node s, build a BFS tree, and return its depth:

$$R = \max_{v \in V} \operatorname{dist}(s, v).$$

Lemma

$$\frac{D}{2} \leq R \leq D.$$

Proof.

If a, b are endpoints of a diameter path, then

$$D = \operatorname{dist}(a, b) \leq \operatorname{dist}(a, s) + \operatorname{dist}(s, b) \leq 2R.$$

$\square$

What does "α-approximation" mean?

Definition (α-approximation)

For a graph G with $\mathrm{diam}(G) = D$, return $\widetilde{D}$ with $\alpha D \leq \widetilde{D} \leq D$.

Definition (α-approximation lower bound)

α-approximating diameter is $T(n)$-hard if no $o(T(n))$-round algorithm always returns α-approximation.

What does "α-approximation" mean?

Definition (α-approximation)

For a graph G with $\mathrm{diam}(G) = D$, return $\widetilde{D}$ with $\alpha D \leq \widetilde{D} \leq D$.

Definition (α-approximation lower bound)

α-approximating diameter is $T(n)$-hard if no $o(T(n))$-round algorithm always returns α-approximation.

Definition (Robust α-approximation)

α-approximating diameter is robustly $T(n)$-hard if for every constant $\beta > 0$, no $o(T(n))$-round algorithm always returns $\widetilde{D}$ such that:

$$\alpha D - \beta \leq \widetilde{D} \leq D.$$

The proof strategy

1. Reduce from two-party set disjointness.
2. Encode Alice's string inside one part of the graph.
3. Encode Bob's string inside the other part.
4. Make the diameter small iff the strings are disjoint.
5. Use the small cut to turn a fast CONGEST algorithm into a cheap protocol.

Content

1. The problem
2. Communication lower bounds
3. Proof I: diameter 3 vs 5
4. Robust lower bounds
5. Wrap-up

Set disjointness

Definition

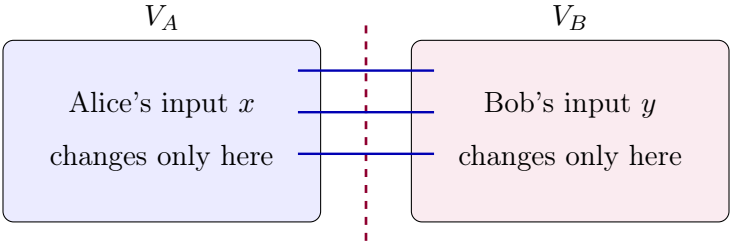
Alice gets $x \in \{0,1\}^K$, Bob gets $y \in \{0,1\}^K$. They must decide

$$\mathrm{DISJ}(x,y) = 1 \iff \forall h \in [K] \quad x_h y_h = 0.$$

Theorem ([Raz92, KN97])

Any randomized protocol for set disjointness with constant success probability requires $\Omega(K)$ bits.

Lower-bound graph template



cut $C = E(V_A, V_B)$

Definition

A family $G(x, y)$ is a lower-bound family for predicate P if

$$P(G(x, y)) = \text{TRUE} \iff \text{DISJ}(x, y) = \text{TRUE},$$

and the dependence on x, y is separated by the partition $V_A \cup V_B$.

Simulation lemma

Theorem

If $G(x, y)$ is a lower-bound family for P, then every r-round CONGEST *algorithm deciding P gives a protocol of cost*

Brain Teaser: *????*

Simulation lemma

Theorem

If $G(x, y)$ is a lower-bound family for P, then every r-round CONGEST *algorithm deciding P gives a protocol of cost*

$$O(r \cdot |C| \cdot \log n).$$

Consequently,

$$r = \Omega\left(\frac{\mathrm{CC}_{\mathrm{DISJ}}(K)}{|C| \log n}\right) = \Omega\left(\frac{K}{|C| \log n}\right).$$

Proof idea

Alice simulates nodes in V_A, Bob simulates nodes in V_B. Only messages crossing the cut must be communicated.

Content

1. The problem
2. Communication lower bounds
3. Proof I: diameter 3 vs 5
4. Robust lower bounds
5. Wrap-up

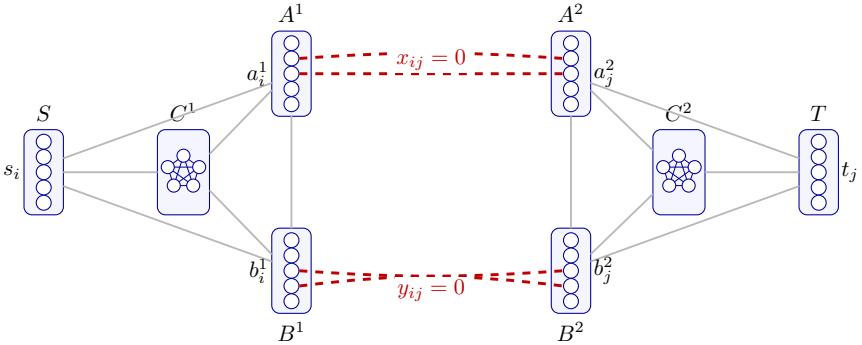
The first theorem

Theorem ([GKP20])

Any CONGEST *algorithm that distinguishes graphs of diameter* 3 *from graphs of diameter* 5 *requires* $\Omega(n/\log n)$ *rounds.*

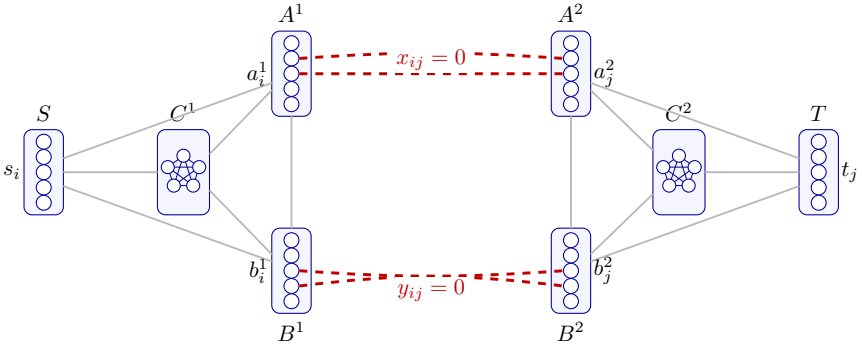
Perhaps surprisingly but we can output $\lfloor 2D/3 \rfloor < \widetilde{D} \leq D$ in $O(n^{1/2} + D)$ rounds [HPRW14]!

Diameter 3 vs 5: construction picture



fixed edges
are matchings

Diameter 3 vs 5: construction picture



$\boxed{\mathrm{DISJ}(x, y) = 1 \Rightarrow \mathrm{diam}(G(x, y)) = 3}$ $\boxed{\mathrm{DISJ}(x, y) = 0 \Rightarrow \mathrm{diam}(G(x, y)) \geq 5}$

Finishing Theorem 4

The construction gives:

$$K = p^2, \qquad |C| = \Theta(p), \qquad n = \Theta(p).$$

Finishing Theorem 4

The construction gives:

$$K = p^2, \qquad |C| = \Theta(p), \qquad n = \Theta(p).$$

By the simulation lemma,

$$r = \Omega\left(\frac{K}{|C| \log n}\right) = \Omega\left(\frac{p^2}{p \log p}\right) = \Omega(n/\log n).$$

Content

1. The problem
2. Communication lower bounds
3. Proof I: diameter 3 vs 5
4. Robust lower bounds
5. Wrap-up

The unified theorem

Theorem ([GKP20])

Let $t \in \{3, 4, 6\}$. For every constant $\varepsilon > 0$ and small enough, any algorithm computing a

$$\left(\frac{t}{2t-1} + \varepsilon\right)$$

approximation to diameter in the CONGEST *model requires*

$$\Omega(n^{1/t}/\log n)$$

rounds.

Plugging in t

t	ratio $t/(2t-1)$	lower bound
3	$3/5$	$\Omega(n^{1/3}/\log n)$
4	$4/7$	$\Omega(n^{1/4}/\log n)$
6	$6/11$	$\Omega(n^{1/6}/\log n)$

The combinatorial ingredient

We need a balanced bipartite graph

$$H = (L, R, E_H), \qquad |L| = |R| = p,$$

with:

- diameter t,
- girth $2t$,
- many edges.

The combinatorial ingredient

We need a balanced bipartite graph

$$H = (L, R, E_H), \qquad |L| = |R| = p,$$

with:

- diameter t,
- girth $2t$,
- many edges.

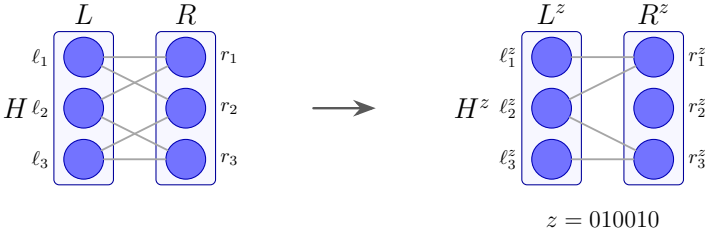
Theorem ([Sin66, Ben66])

For $t \in \{3, 4, 6\}$, there are such graphs with

$$K := |E_H| = \mathrm{Ex}(p, t) = \Omega\left(p^{1+\frac{1}{t-1}}\right).$$

Step 1: encode a string into H

Enumerate the edges $\pi : E_H \to [K]$. For $z \in \{0,1\}^K$, keep exactly the edges whose bits are zero.

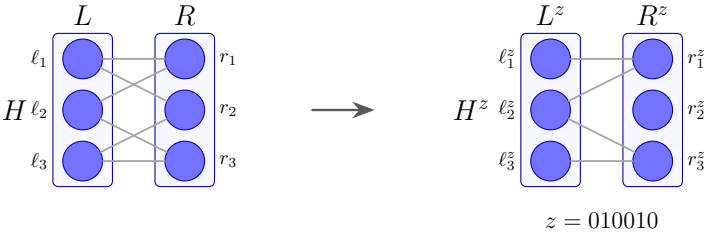


$$E(H^z) = \{(\ell_i^z, r_j^z) : (\ell_i, r_j) \in E_H,\ z_{\pi(\ell_i, r_j)} = 0\}.$$

Brain Teaser: What is a distance between pair of vertices when the edge between them is deleted?

Step 1: encode a string into H

Enumerate the edges $\pi : E_H \to [K]$. For $z \in \{0,1\}^K$, keep exactly the edges whose bits are zero.

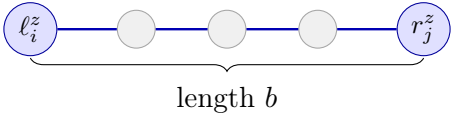


$$E(H^z) = \{(\ell_i^z, r_j^z) : (\ell_i, r_j) \in E_H,\ z_{\pi(\ell_i, r_j)} = 0\}.$$

If one deletes $(u, v) \in E(H^z)$, then the distance between them will be at least $2t - 1$.

Step 2: stretch every surviving edge

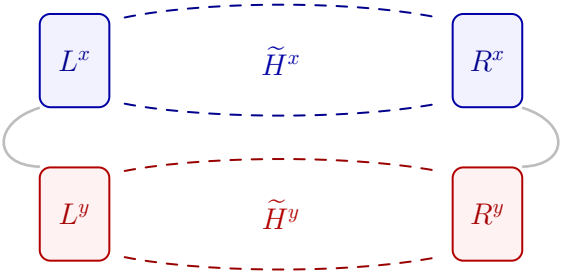
Choose an integer b. The graph $\widetilde{H}^z$ replaces every surviving edge by a path of length b:



Step 3: combine Alice and Bob copies

Given $x, y \in \{0,1\}^K$, build $G(x,y)$ from two $\widetilde{H}^x$ and $\widetilde{H}^y$.
Add matchings:

$$(\ell_i^x, \ell_i^y) \quad \text{and} \quad (r_i^x, r_i^y) \qquad \text{for all } i \in [p].$$

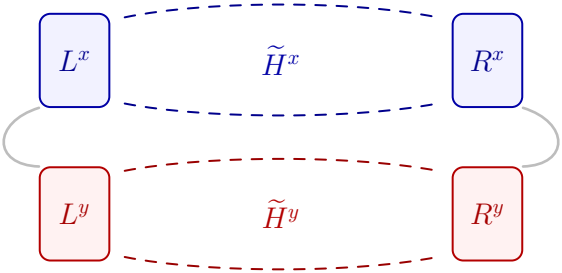


$$V_A = V_{\widetilde{H}^x}, \quad V_B = V_{\widetilde{H}^y}, \quad |C| = ???$$

Step 3: combine Alice and Bob copies

Given $x, y \in \{0,1\}^K$, build $G(x,y)$ from two $\widetilde{H}^x$ and $\widetilde{H}^y$.
Add matchings:

$$(\ell_i^x, \ell_i^y) \quad \text{and} \quad (r_i^x, r_i^y) \qquad \text{for all } i \in [p].$$



$$V_A = V_{\widetilde{H}^x}, \quad V_B = V_{\widetilde{H}^y}, \quad |C| = 2p.$$

The shadow graph H^z

Let $z = x \wedge y$.

- If x, y are disjoint, then $z = 0^K$ and $H^z \simeq H$.
- If x, y are not disjoint, then at least one edge of H is missing from H^z.

The shadow graph H^z

Let $z = x \wedge y$.

- If x, y are disjoint, then $z = 0^K$ and $H^z \simeq H$.
- If x, y are not disjoint, then at least one edge of H is missing from H^z.

Lemma

If x, y are disjoint, then $\mathrm{diam}(\tilde{H}^z) = tb$.

Brain Teaser: why?

The shadow graph H^z

Let $z = x \wedge y$.

- If x, y are disjoint, then $z = 0^K$ and $H^z \simeq H$.
- If x, y are not disjoint, then at least one edge of H is missing from H^z.

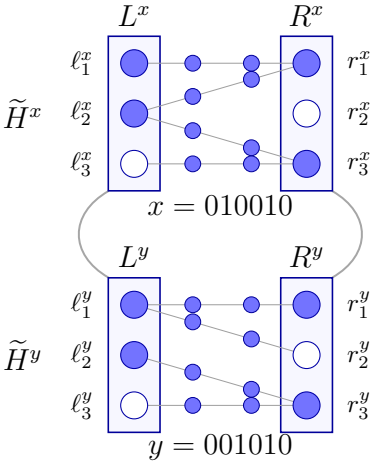
Lemma

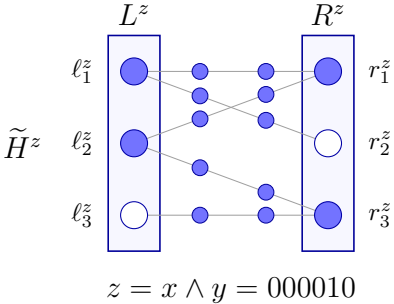
If x, y are disjoint, then $\mathrm{diam}(\tilde{H}^z) = tb$.

Proof.

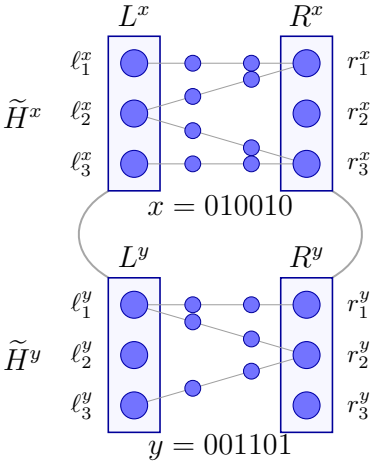
Trivially. $\square$

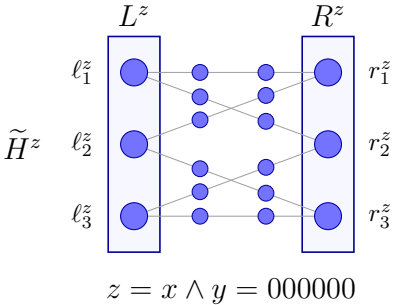
Example: non-disjoint inputs





Example: disjoint inputs





Non-disjoint case

Claim

If x, y are not disjoint, then $\mathrm{diam}(\widetilde{H}^z) \geq (2t-1)b$.

Brain Teaser: Why?

Non-disjoint case

Claim

If x, y are not disjoint, then $\operatorname{diam}(\widetilde{H}^z) \geq (2t-1)b$.

Proof.

Pick an edge $(\ell_i, r_j) \in E_H$ with

$$x_{\pi(\ell_i, r_j)} = y_{\pi(\ell_i, r_j)} = 1.$$

Then (ℓ_i^z, r_j^z) is missing from H^z.
Since H has girth $2t$, every path from ℓ_i^z to r_j^z in H^z has length at least $2t-1$.
And $\tilde{H}^z$ is just H^z but stretched. $\square$

Transfer the distance to $G(x, y)$

Claim

For every $\ell_i^x \in L^x$ *and* $r_j^x \in R^x$,

$$\mathrm{dist}_{G(x,y)}(\ell_i^x, r_j^x) \geq \mathrm{dist}_{\widetilde{H}^z}(\ell_i^z, r_j^z).$$

Brain Teaser: why?

Transfer the distance to $G(x, y)$

Claim

For every $\ell_i^x \in L^x$ *and* $r_j^x \in R^x$,

$$\mathrm{dist}_{G(x,y)}(\ell_i^x, r_j^x) \geq \mathrm{dist}_{\widetilde{H}^z}(\ell_i^z, r_j^z).$$

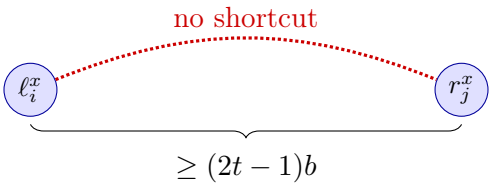
Proof.

Trivially. □

Non-disjoint conclusion

Combining the two claims:

$$\mathrm{diam}(G(x,y)) \geq (2t-1)b.$$



Disjoint case: the shadow graph is intact

If x and y are disjoint, then

$$z = x \wedge y = 0^K.$$

Disjoint case: the shadow graph is intact

If x and y are disjoint, then

$$z = x \wedge y = 0^K.$$

Therefore:

$$H^z \simeq H, \qquad \operatorname{diam}(H^z) = t, \qquad \operatorname{diam}(\widetilde{H}^z) = tb.$$

Disjoint case: the shadow graph is intact

If x and y are disjoint, then

$$z = x \wedge y = 0^K.$$

Therefore:

$$H^z \simeq H, \qquad \operatorname{diam}(H^z) = t, \qquad \operatorname{diam}(\widetilde{H}^z) = tb.$$

What remains

Show that $G(x, y)$ is not much larger:

$$\operatorname{diam}(G(x, y)) \leq tb + t + 1.$$

Brain Teaser: why?

Disjoint case: the shadow graph is intact

If x and y are disjoint, then

$$z = x \wedge y = 0^K.$$

Therefore:

$$H^z \simeq H, \qquad \operatorname{diam}(H^z) = t, \qquad \operatorname{diam}(\widetilde{H}^z) = tb.$$

What remains

Show that $G(x, y)$ is not much larger:

$$\operatorname{diam}(G(x, y)) \leq tb + t + 1.$$

Follow path in H^z, when needed use matching.

Brain Teaser: right?)

Routing between boundary vertices

In the disjoint case, every edge of H appears in at least one copy:

$$x_h = 0 \quad \text{or} \quad y_h = 0.$$

So we can follow a shortest path in H. Whenever the next stretched edge is in the other copy, use one matching edge to switch copies.

Cost

At most t stretched paths plus at most $t+1$ matching switches:

$$tb + t + 1.$$

Auxiliary lemmas

Lemma

If t is odd, then the distance between any two nodes $u, v \in L$ in H is at most $t - 1$. Similarly, the distance between any two nodes $u, v \in R$ in H is at most $t - 1$.

Brain Teaser: Why?

Auxiliary lemmas

Lemma

If t is odd, then the distance between any two nodes $u, v \in L$ in H is at most $t - 1$. Similarly, the distance between any two nodes $u, v \in R$ in H is at most $t - 1$.

Proof.

The distance between every two nodes in H is at most its diameter t, but the distance between every two nodes in the same side of the bipartition is even. Hence, it is at most $t - 1$. □

Auxiliary lemmas

Lemma

If t is odd, then the distance between any two nodes $u, v \in L$ in H is at most $t - 1$. Similarly, the distance between any two nodes $u, v \in R$ in H is at most $t - 1$.

Proof.

The distance between every two nodes in H is at most its diameter t, but the distance between every two nodes in the same side of the bipartition is even. Hence, it is at most $t - 1$. □

Lemma

If t is even, then the distance between any pair of nodes $u \in L$ and $v \in R$ in H is at most $t - 1$.

Proof.

Similarly. □

Which boundary pairs are close?

- If t is odd, then any two boundary nodes on the same side

$$L^x \cup L^y \quad \text{or} \quad R^x \cup R^y$$

are at distance at most $(t-1)b + t$ in $G(x, y)$.

Which boundary pairs are close?

- If t is odd, then any two boundary nodes on the same side

$$L^x \cup L^y \quad \text{or} \quad R^x \cup R^y$$

  are at distance at most $(t-1)b + t$ in $G(x, y)$.
- If t is even, then any two boundary nodes on different sides

$$L^x \cup L^y \quad \text{and} \quad R^x \cup R^y$$

  are at distance at most $(t-1)b + t$.

From arbitrary nodes to boundary nodes

For two arbitrary nodes u, v:

$$\text{either they can both move to the } L \text{ side with total cost } \leq b + 1,$$

or they can both move to the R side with total cost $\leq b + 1$.

From arbitrary nodes to boundary nodes

For two arbitrary nodes u, v:

$$\text{either they can both move to the } L \text{ side with total cost } \leq b+1,$$

or they can both move to the R side with total cost $\leq b+1$.
Similarly, one can move them to opposite sides with total cost $\leq b+1$.

From arbitrary nodes to boundary nodes

For two arbitrary nodes u, v:

$$\text{either they can both move to the } L \text{ side with total cost } \leq b+1,$$

or they can both move to the R side with total cost $\leq b+1$.
Similarly, one can move them to opposite sides with total cost $\leq b+1$.

Disjoint diameter bound

$$\mathrm{diam}(G(x,y)) \leq (t-1)b + t + (b+1) = tb + t + 1.$$

The robust diameter gap

We have built graphs with:

$$x, y \text{ disjoint} \Rightarrow D \leq tb + t + 1,$$

$$x, y \text{ not disjoint} \Rightarrow D \geq (2t - 1)b.$$

The robust diameter gap

We have built graphs with:

$$x, y \text{ disjoint} \Rightarrow D \leq tb + t + 1,$$

$$x, y \text{ not disjoint} \Rightarrow D \geq (2t - 1)b.$$

An α-approximation distinguishes the two cases if

$$\alpha \cdot (2t - 1)b > tb + t + 1.$$

The robust diameter gap

We have built graphs with:

$$x, y \text{ disjoint} \Rightarrow D \leq tb + t + 1,$$

$$x, y \text{ not disjoint} \Rightarrow D \geq (2t - 1)b.$$

An α-approximation distinguishes the two cases if

$$\alpha \cdot (2t - 1)b > tb + t + 1.$$

Hence

$$\alpha = \frac{t}{2t - 1} + \varepsilon,$$

since $b \gg t$.

Lower-bound calculation

Recall:

$$K = |E_H| = \Omega\left(p^{1+\frac{1}{t-1}}\right), \qquad |C| = \Theta(p).$$

Lower-bound calculation

Recall:

$$K = |E_H| = \Omega\left(p^{1+\frac{1}{t-1}}\right), \qquad |C| = \Theta(p).$$

For constant b,

$$n = \Theta(Kb) = \Theta(K), \qquad p = O\left(n^{\frac{t-1}{t}}\right).$$

Lower-bound calculation

Recall:

$$K = |E_H| = \Omega\left(p^{1+\frac{1}{t-1}}\right), \qquad |C| = \Theta(p).$$

For constant b,

$$n = \Theta(Kb) = \Theta(K), \qquad p = O\left(n^{\frac{t-1}{t}}\right).$$

Therefore:

$$r = \Omega\left(\frac{K}{|C|\log n}\right) = \Omega\left(\frac{n}{n^{(t-1)/t}\log n}\right) = \Omega(n^{1/t}/\log n).$$

One last issue

Brain Teaser: Any problems?

One last issue

Deleting many edges could disconnect H^x!

Brain Teaser: How to fix?

One last issue

Deleting many edges could disconnect H^x!

Fix

Choose a tree S of H and make its edges always present.

Brain Teaser: Why it works?

One last issue

Deleting many edges could disconnect H^x!

Fix

Choose a spanning tree S of H and make its edges always present.

- $|S| = O(p)$.
- $K = \Omega(p^{1+\frac{1}{t-1}})$, so removing those input positions does not change K asymptotically.

One last issue

Deleting many edges could disconnect H^x!

Fix

Choose a shortest-path tree S of H and make its edges always present.

- $|S| = O(p)$.
- $K = \Omega(p^{1+\frac{1}{t-1}})$, so removing those input positions does not change K asymptotically.
- The graph becomes connected, in fact of constant diameter for constant t, b.

Content

1. The problem
2. Communication lower bounds
3. Proof I: diameter 3 vs 5
4. Robust lower bounds
5. Wrap-up

Takeaways

Result	Round lower bound	Robust?
diameter 3 vs 5	$\Omega(n/\log n)$	no
$(3/5+\varepsilon)$-approximation	$\Omega(n^{1/3}/\log n)$	yes
$(4/7+\varepsilon)$-approximation	$\Omega(n^{1/4}/\log n)$	yes
$(6/11+\varepsilon)$-approximation	$\Omega(n^{1/6}/\log n)$	yes

State of the art

Approx.	Bound	Ref. and Comments
Exact	$\widetilde{\Theta}(n)$	[FHW12, HFQ+16]
2 vs. 3	$\widetilde{\Omega}(n)$	[HW12]
$\lfloor 2D/3 \rfloor < \widetilde{D} \leq D$	$O(n^{1/2} + D)$	[HPRW14]
$2/3 + \varepsilon$	$\widetilde{\Omega}(n)$	[ACHK16]
3 vs. 5	$\widetilde{\Omega}(n)$	[GKP20]
$3/5 + \varepsilon$	$\widetilde{\Omega}(n^{1/3})$	[GKP20]
$4/7$	$O(n^{1/3} + D)$	[ACHD+21]
$4/7 + \varepsilon$	$\widetilde{\Omega}(n^{1/4})$	[GKP20]
$6/11 + \varepsilon$	$\widetilde{\Omega}(n^{1/6})$	[GKP20]
$1/2$	$O(D)$	Folklore

Bibliography I

Bertie Ancona, Keren Censor-Hillel, Mina Dalirrooyfard, Yuval Efron, and Virginia Vassilevska Williams.
Distributed distance approximation.
In *24th International Conference on Principles of Distributed Systems (OPODIS 2020)*, volume 184 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 30:1–30:17. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2021.

Amir Abboud, Keren Censor-Hillel, and Seri Khoury.
Near-linear lower bounds for distributed distance computations, even in sparse networks.
In *Distributed Computing - 30th International Symposium (DISC)*, volume 9888 of *Lecture Notes in Computer Science*, pages 29–42. Springer, 2016.

Bibliography II

Clark T. Benson.
Minimal regular graphs of girths eight and twelve.
Canadian Journal of Mathematics, 18:1091–1094, 1966.

Silvio Frischknecht, Stephan Holzer, and Roger Wattenhofer.
Networks cannot compute their diameter in sublinear time.
In *Proceedings of the Twenty-Third Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 1150–1162. SIAM, 2012.

Ofer Grossman, Seri Khoury, and Ami Paz.
Improved hardness of approximation of diameter in the CONGEST model.
In *34th International Symposium on Distributed Computing (DISC 2020)*, volume 179 of *LIPIcs*, pages 19:1–19:16. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2020.

Bibliography III

Qiang-Sheng Hua, Haoqiang Fan, Lixiang Qian, Ming Ai, Yangyang Li, Xuanhua Shi, and Hai Jin.
Brief announcement: A tight distributed algorithm for all pairs shortest paths and applications.
In *Proceedings of the 28th ACM Symposium on Parallelism in Algorithms and Architectures (SPAA)*, pages 439–441. ACM, 2016.

Stephan Holzer, David Peleg, Liam Roditty, and Roger Wattenhofer.
Brief announcement: Distributed 3/2-approximation of the diameter.
In *Distributed Computing - 28th International Symposium (DISC)*, volume 8784 of *Lecture Notes in Computer Science*, pages 562–564. Springer, 2014.

Bibliography IV

Stephan Holzer and Roger Wattenhofer.
Optimal distributed all pairs shortest paths and applications.
In *Proceedings of the ACM Symposium on Principles of Distributed Computing (PODC)*, pages 355–364. ACM, 2012.

Eyal Kushilevitz and Noam Nisan.
Communication Complexity.
Cambridge University Press, 1997.

David Peleg.
Distributed Computing: A Locality-Sensitive Approach.
Society for Industrial and Applied Mathematics, 2000.

Bibliography V

Alexander A. Razborov.
On the distributional complexity of disjointness.
Theoretical Computer Science, 106(2):385–390, 1992.

R. Singleton.
On minimal graphs of maximum even girth.
Journal of Combinatorial Theory, 1:306–332, 1966.