Advances in Directed Spanners

Grigory Yaroslavtsev
Penn State

Joint work with
Berman (PSU), Bhattacharyya (MIT), Grigorescu (GaTech), Makarychev (IBM), Raskhodnikova (PSU), Woodruff (IBM)

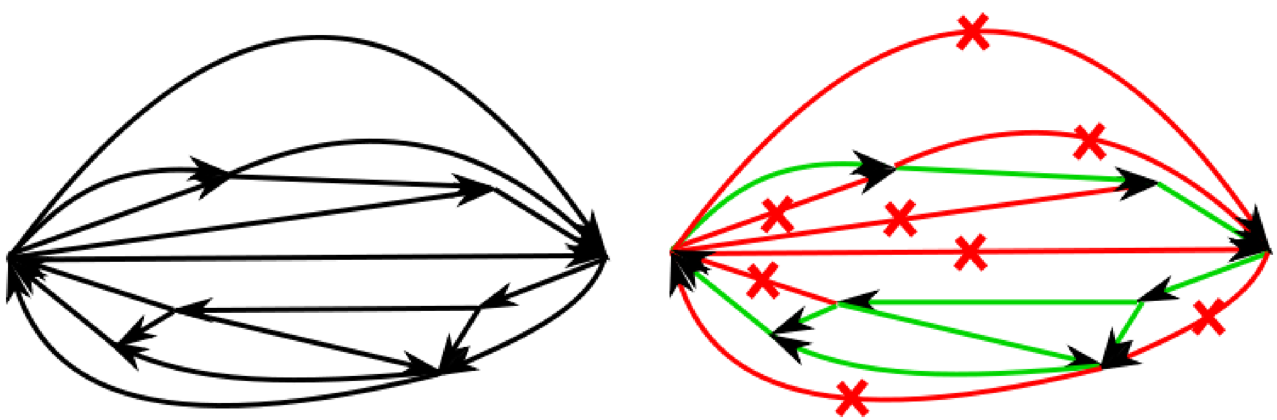
Directed Spanner Problem

- **k-Spanner** [Awerbuch '85, Peleg, Shäffer '89]

Subset of edges, preserving distances up to a factor k > 1 (stretch k).

- Graph G(V, E) with weights $w : E \to \mathbb{R}^{\geq 0}$

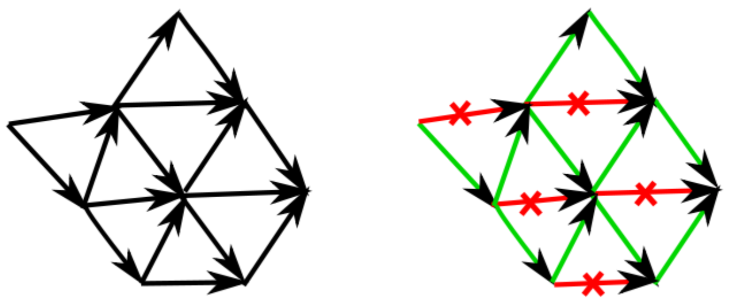
  H(V, $E_H \subseteq E$): $\forall (u, v) \in E \; dist_H(u, v) \leq k \cdot w(u, v)$



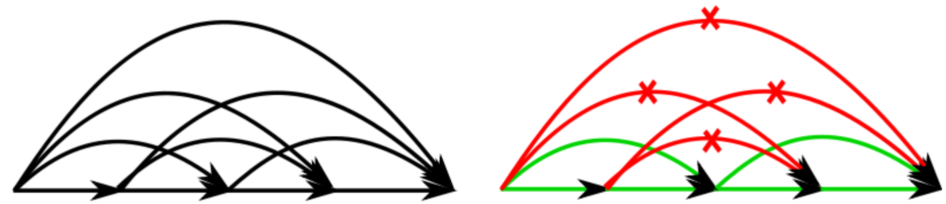
- **Problem:** Find the sparsest k-spanner of a directed graph.

Directed Spanners and Their Friends

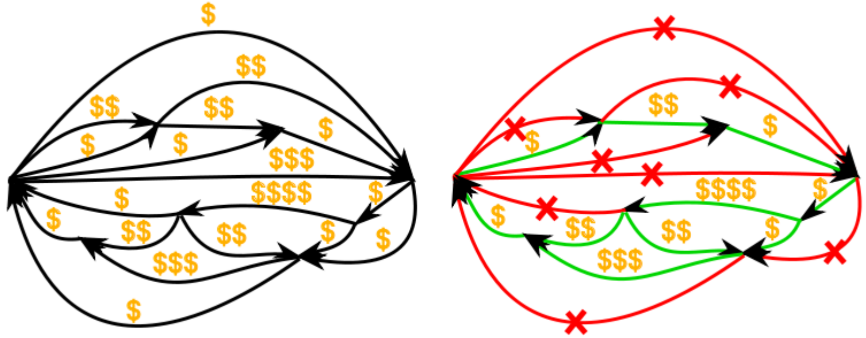
Unit lengths



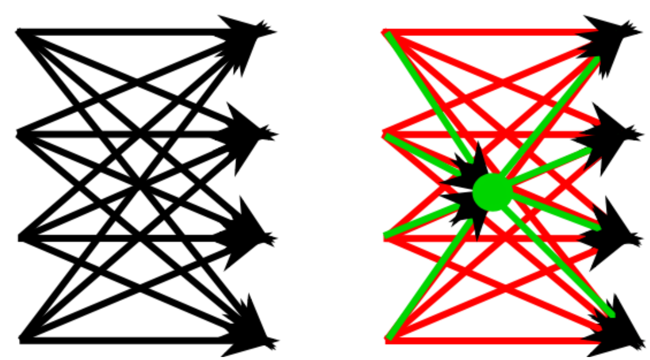
Transitive-closure spanner



Minimum cost spanner



Steiner spanner



Applications of spanners

- First application: simulating synchronized protocols in unsynchronized networks [Peleg, Ullman '89]
- Efficient routing [PU'89, Cowen '01, Thorup, Zwick '01, Roditty, Thorup, Zwick '02 , Cowen, Wagner '04]
- Parallel/Distributed/Streaming approximation algorithms for shortest paths [Cohen '98, Cohen '00, Elkin'01, Feigenbaum, Kannan, McGregor, Suri, Zhang '08]
- Algorithms for approximate distance oracles [Thorup, Zwick '01, Baswana, Sen '06]

Applications of directed spanners

- Access control hierarchies
  - Previous work: [Atallah, Frikken, Blanton, CCCS '05; De Santis, Ferrara, Masucci, MFCS'07]
  - Solution: TC-spanners [Bhattacharyya, Grigorescu, Jung, Raskhodnikova, Woodruff, SODA'09]
  - Steiner TC-spanners for access control: [Berman, Bhattacharyya, Grigorescu, Raskhodnikova, Woodruff, Y' ICALP'11]
- Property testing and property reconstruction [BGJRW'09; Raskhodnikova '10 (survey)]

Plan

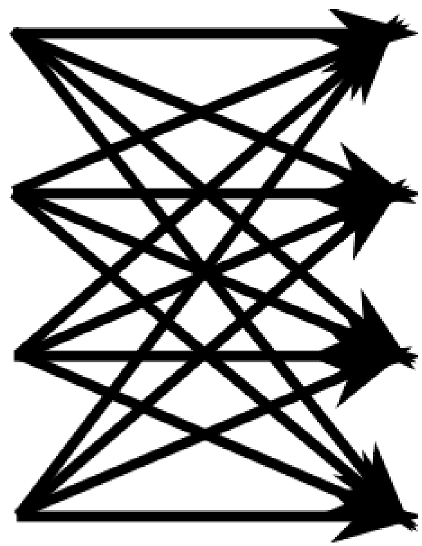
- Approximation algorithms
  - Undirected vs. Directed
  - Framework for directed case = Sampling + LP
  - Randomized rounding
    - Directed Spanner
    - Unit-length 3-spanner
    - Directed Steiner Forest
- Combinatorial bounds on TC-Spanners
  - Upper bounds for low-dimensional posets
  - Lower bounds via linear programming

Undirected vs. Directed

- Trivial lower bound: $\geq \boldsymbol{n-1}$ edges needed
- Every undirected graph has a (2t+1)-spanner with $\leq n^{1+1/t}$ edges. [Althofer, Das, Dobkin, Joseph, Soares '93]
- Kruskal-like greedy + girth argument => $n^{\frac{1}{t}}$ −approximation
- Time/space-efficient constructions of undirected approximate distance oracles [Thorup, Zwick, STOC '01]

Undirected vs Directed

- For some directed graphs $\Omega(n^2)$ edges needed for a k-spanner:



- No space-efficient directed distance oracles: some graphs require $\Omega(n^2)$ space. [TZ '01]

Unit-Length Directed k-Spanner

- O(n)-approximation: trivial (whole graph)

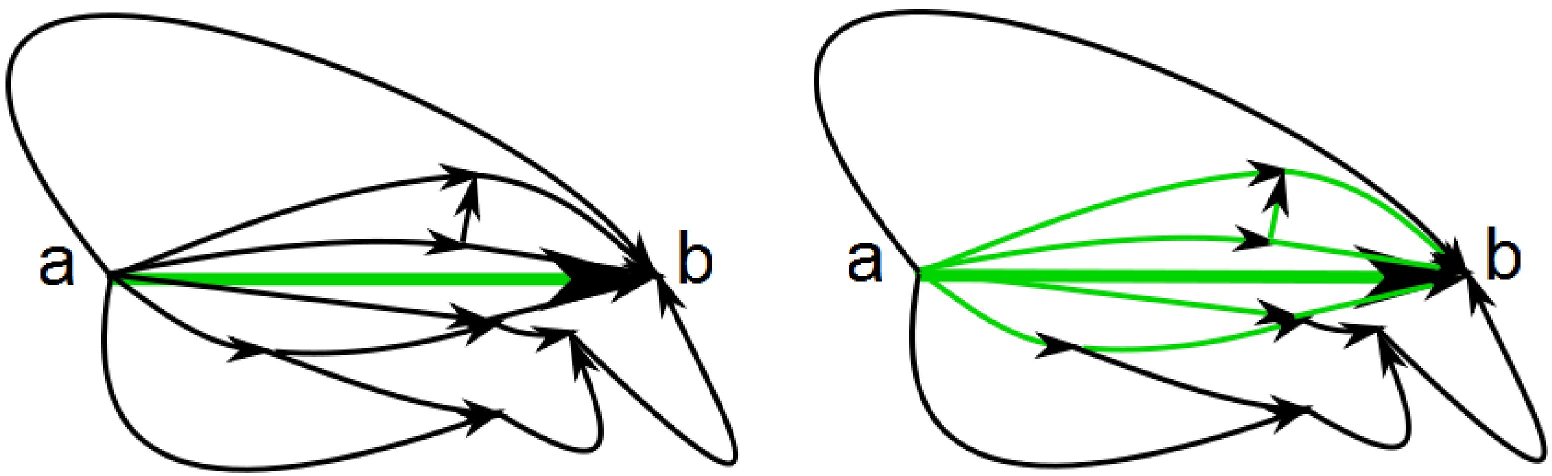
Stretch	$k = 2$	$k = 3$	$k \geq 4$
Previous work	$O(\log n)$ [KP94]	$\tilde{O}(n^{2/3})$ [EP00] $\tilde{O}(n^{2/3})$ [BGJRW09] $\tilde{O}(\sqrt{n})$ [BRR10] $\tilde{O}(\sqrt{n})$ [DK11]	$\tilde{O}(n^{1-\frac{1}{k}})$ [BGJRW09] $\tilde{O}(n^{1-\frac{1}{\lceil k/2 \rceil}})$ [BRR10] $\tilde{O}(n^{2/3})$[DK11]
Our work		$\tilde{O}(n^{1/3})$ + *undirected!*	$\tilde{O}(\sqrt{n})$
Integrality gap	$\Omega(\log n)$ [DK11]	$\Omega(\frac{1}{k} n^{1/3-\epsilon})$ [DK11]	
Hardness	$\Omega(\log n)$ NP-hard [K01]	$2^{\log^{1-\epsilon} n}$ quasi-NP-hard [EP00]	

Our $\tilde{O}(\sqrt{n})$-approximation

- Paths of stretch at most k for all edges =>
- Classify edges: **thick** and thin
- Take union of spanners for them
  - **Thick** edges: Sampling
  - Thin edges: LP + randomized rounding

Local Graph

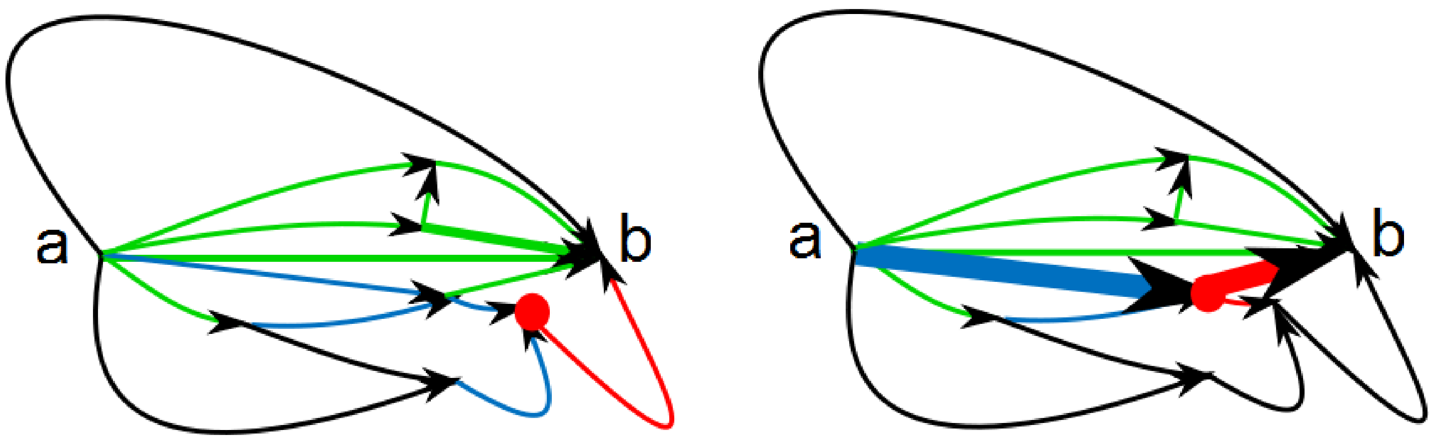
- Local graph for an edge (a,b): Induced by vertices on paths of stretch ≤ k from a to b



- Paths of stretch ≤ k only use edges in local graphs
- **Thick** edges: ≥ $\sqrt{n}$ vertices in their local graph. Otherwise thin.

Sampling [BGJRW'09, DK11]

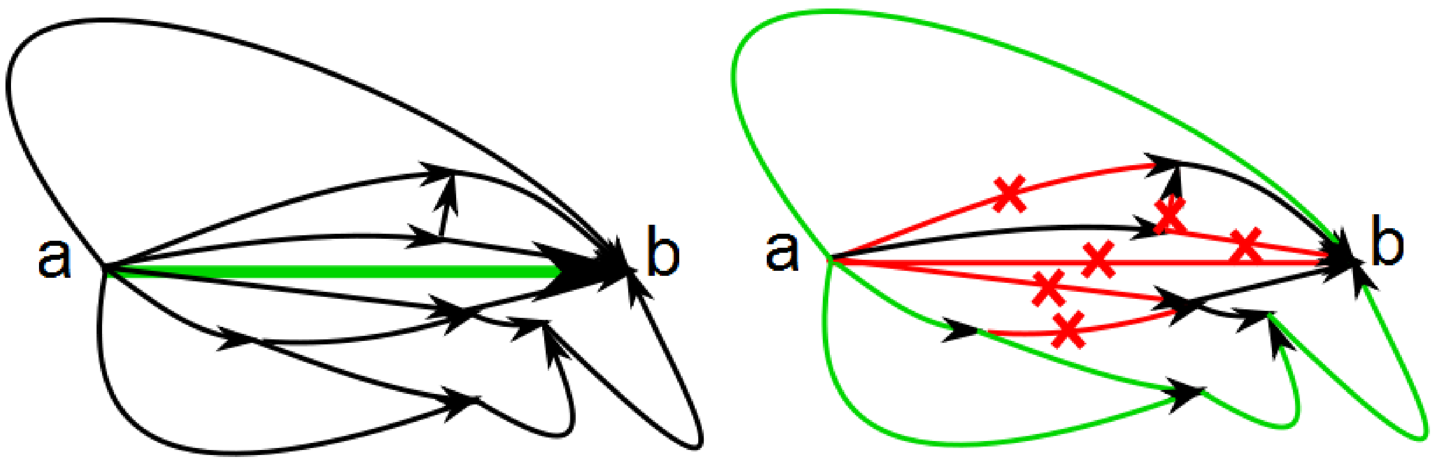
- Pick $O(\sqrt{\boldsymbol{n}}\,\mathbf{ln}\,\boldsymbol{n})$ seed vertices at random
- Take in- and out- shortest path trees for each



- Handles all **thick** edges ($\geq \sqrt{n}$ vertices in their local graph) w.h.p.
- # of edges $\leq 2(n-1)\boldsymbol{O}(\sqrt{\boldsymbol{n}}\,\mathbf{ln}\,\boldsymbol{n}) \leq OPT \cdot \tilde{O}(\sqrt{n})$.

Key Idea: Antispanners

- Antispanner – subset of edges, whose removal destroys all paths from **a** to **b** of stretch at most k



- Graph is spanner <=> hits all antispanners
- Enough to hit all minimal antispanners for all thin edges
- If E_H is not a spanner for an edge **(a,b)** => $E \setminus E_H$ is an antispanner, can be minimized greedily

Linear Program (~dual to [DK'11])

Hitting-set LP: $\sum_{e \in E} x_e \to min$

$$\sum_{e \in \mathbf{A}} x_e \geq 1$$

for all minimal antispanners **A** for all thin edges.

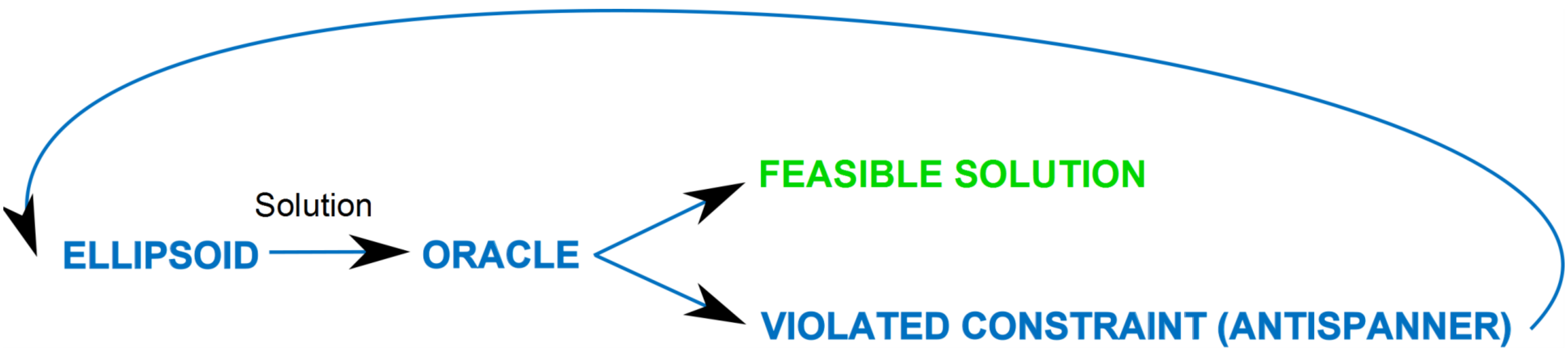
- # of minimal antispanners may be exponential in $\sqrt{n}$ => Ellipsoid + Separation oracle
- We will show: $\leq \sqrt{n}^{\sqrt{n}} = e^{\frac{1}{2}\sqrt{n}\ln n}$ minimal antispanners for a fixed thin edge
- Assume that we guessed OPT = the size of the sparsest k-spanner (at most n^2 values)

Oracle

Hitting-set LP: $\sum_{e \in E} x_e \leq OPT$

$$\sum_{e \in \mathbf{A}} x_e \geq 1$$

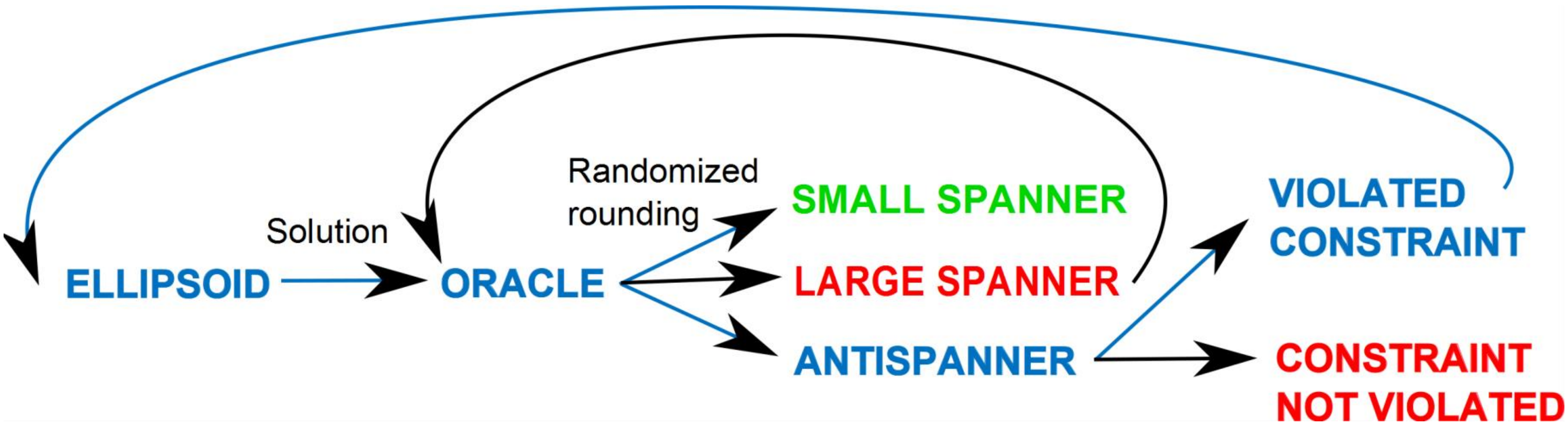
for all minimal antispanners $\mathbf{A}$ for all thin edges.



- We use a randomized oracle => in both cases oracle fails with exponentially small probability.

Randomized Oracle = Rounding

- Rounding: Take **e** w.p. $p_e = \min(\sqrt{\boldsymbol{n}}\,\mathbf{ln}\,\boldsymbol{n} \cdot x_e, 1)$



- **SMALL SPANNER**: We have a set of edges of size $\leq \sum_e x_e \cdot \tilde{O}(\sqrt{n}) \leq OPT \cdot \tilde{O}(\sqrt{n})$ w.h.p.
- Pr[**LARGE SPANNER**] $\leq e^{-\Omega(\sqrt{n})}$ by Chernoff.
- Pr[**CONSTRAINT NOT VIOLATED**] $\leq e^{-\Omega(\sqrt{n})}$ (next slide)

Pr[CONSTRAINT NOT VIOLATED]

- Set S: $\forall e \in E$ we have $\Pr[e \in S] = \min(\sqrt{n} \ln n \, x_e, 1)$
- For a fixed minimal antispanner $\mathbf{A}$, such that $\sum_{e \in A} x_e \geq 1$:

$$\Pr[S \cap \mathbf{A} = \emptyset] \leq \prod_{e \in \mathbf{A}} (1 - \sqrt{n} \ln n \, x_e) \leq e^{-\sqrt{n} \ln n \sum_{e \in A} x_e} \leq \boldsymbol{e^{-\sqrt{n} \ln n}}$$

- #minimal antispanners for a fixed edge $(\boldsymbol{s}, \boldsymbol{t}) \leq$

#different shortest path trees with root **s** in a local graph

$$\leq \sqrt{n}^{\sqrt{n}} = e^{\frac{1}{2}\sqrt{n} \ln n} \text{ (for a thin edge)}$$

- #minimal antispanners $\leq |E| \boldsymbol{e^{\frac{1}{2}\sqrt{n} \ln n}}$ => union bound:

Pr[CONSTRAINT NOT VIOLATED] $\leq |E| \boldsymbol{e^{-\frac{1}{2}\sqrt{n} \ln n}}$

Unit-length 3-spanner

- $\tilde{O}(n^{1/3})$-approximation algorithm
  - Sampling $\tilde{O}(n^{1/3})$ times
  - Dual LP + Different randomized rounding (simplified version of [DK'11])
- Rounding scheme (vertex-based):
  - For each vertex $u \in V$: sample $r_u \in [0,1]$
  - Take all edges (u, v) if

$$\min(r_u, r_v) \leq \tilde{O}(n^{1/3}) x_{(u,v)}$$

  - Feasible solution => 3-spanner w.h.p. (see paper)

Approximation wrap-up

- Sampling + LP with randomized rounding
- Improvement for **Directed Steiner Forest**:
  - Cheapest set of edges, connecting pairs (s_i, t_i)
  - Previous: Sampling + similar LP [Feldman, Kortsarz, Nutov, SODA '09] . Deterministic rounding gives $\tilde{O}(n^{4/5+\epsilon})$-approximation
  - We give $\tilde{O}(n^{2/3+\epsilon})$-approximation via **randomized rounding**

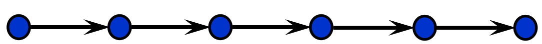
Approximation wrap-up

- **$\tilde{O}(\sqrt{n})$**-approximation for Directed Spanner
- Small local graphs => better approximation
- Can we do better for general graphs?
  - Hardness: only excludes **polylog(n)**-approximation
  - Integrality gap: $\Omega(n^{1/3-\epsilon})$ **[DK'11]**
- Can we do better for specific graphs
  - Planar graphs (still NP-hard)?

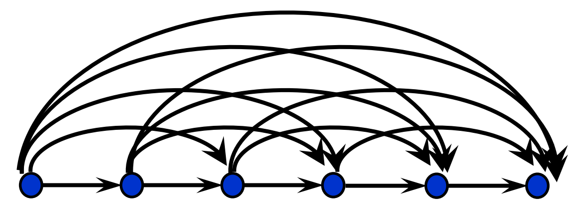
Transitive-Closure Spanners

Transitive closure TC(G) has an edge from u to v iff

G has a path from u to v



G

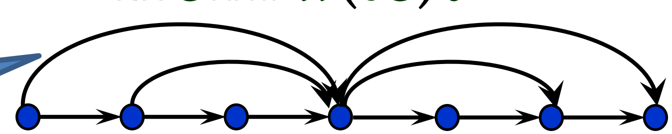


TC(G)

H is a ***k*-TC-spanner** of G if H is a subgraph of TC(G) for which $distance_H(u,v) \leq k$ iff G has a path from u to v

H is a ***k*-TC-spanner** of G if H is a k-spanner of TC(G)

Shortcut edge consistent with ordering



2-TC-spanner of G

[Bhattacharyya, Grigorescu, Jung, Raskhodnikova, Woodruff, SODA'09], generalizing [Yao 82; Chazelle 87; Alon, Schieber 87, …]

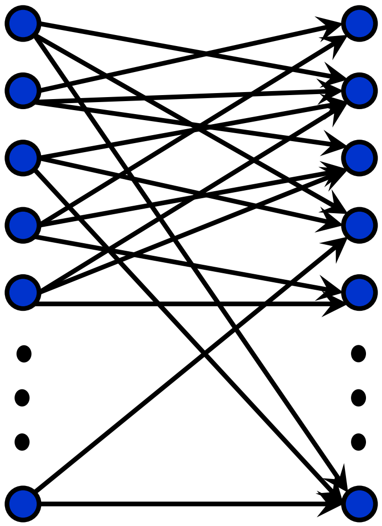
Applications of TC-Spanners

- Data structures for storing partial products [Yao, ’82; Chazelle ’87, Alon, Schieber, 88]
- Constructions of unbounded fan-in circuits [Chandra, Fortune, Lipton ICALP, STOC‘83]
- Property testers for monotonicity and Lipschitzness [Dodis et al. ‘99,BGJRW’09; Jha, Raskhodnikova, FOCS ‘11]
- Lower bounds for reconstructors for monotonicity and Lipshitzness [BGJJRW’10, JR’11]
- Efficient key management in access hierarchies

Follow references in [Raskhodnikova ’10 (survey)]

Bounds for Steiner TC-Spanners

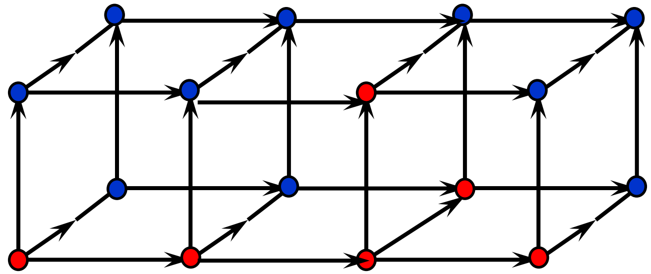
- No non-trivial upper bound for arbitrary graphs



FACT: For a random directed bipartite graph of density ½, any Steiner 2-TC-spanner requires $\tilde{\Omega}(n^2)$ edges.

Low-Dimensional Posets

- [ABFF 09] access hierarchies are **low-dimensional posets**
- Poset $\equiv$ DAG
- Poset *G* has dimension *d* if *G* can be embedded into a hypergrid of dimension *d* and *d* is minimum.



$e: G \to G'$ is a **poset embedding** if for all $x, y \in G$, $x \preccurlyeq_G y$ iff $e(x) \preccurlyeq_{G'} e(y)$.

Hypergrid $[m]^d$ has ordering $(x_1, \ldots, x_d) \preccurlyeq (y_1, \ldots, y_d)$ iff $x_i \leq y_i$ for all *i*

Main Results

Stretch *k*	Upper Bound	Lower Bound
2	$O\left(n \log^d n\right)$	$\Omega\left(n\left(\frac{\alpha \log n}{d}\right)^d\right)$ where α is a constant
≥ 3	$O(n \log^{d-1} n \log\log n)$ for constant *d* [DFM 07]	$\Omega(n \log^{\lceil (d-1)/k \rceil} n)$ for constant *d*

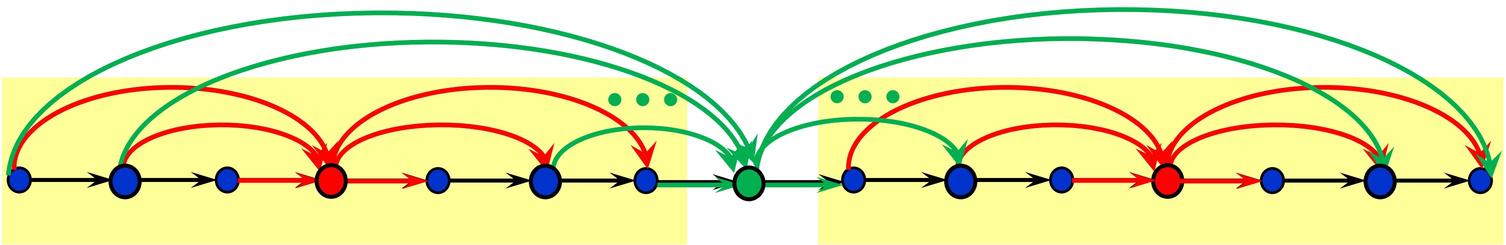
d is the poset dimension

$k = 3$. Nothing better known for larger *k*.

2-TC-Spanner for $[m]^d$

- d = 1 (so, n = m)

2-TC-spanner with $\leq m \log m = n \log n$ edges



We show this is tight upto α^d for a constant α.

- d > 1 (so, $n = m^d$)

2-TC-spanner with $\leq (m \log m)^d = n \left(\frac{\log n}{d}\right)^d$ edges by taking d-wise Cartesian product of 2-TC-spanners for a line.

Lower Bound Strategy

- Write in IP for a minimal 2-TC-spanner.
- OPT $\geq LP = LP_{dual}$
- It is crucial that the integrality gap of the primal is small.
- **Idea:** Construct some feasible solution for the dual => lower bound on OPT.
- OPT $\geq LP = LP_{dual} \geq LB$

IP Formulation

- {0,1}-program for Minimal *2*-TC-spanner:

minimize

$$\sum_{u,v:u \preccurlyeq v} x_{uv}$$

subject to:

$$x_{uw} \geq p_{uwv} \qquad \forall u \preccurlyeq w \preccurlyeq v$$

$$x_{wv} \geq p_{uwv} \qquad \forall u \preccurlyeq w \preccurlyeq v$$

$$\sum_{w:u \preccurlyeq w \preccurlyeq v} p_{uwv} \geq 1 \qquad \forall u \preccurlyeq v$$

Dual LP

- Take fractional relaxation of IP and look at its dual:

$$\text{maximize} \quad \sum_{u,v:u \preccurlyeq v} y_{uv}$$

$$\text{subject to:} \quad \sum_{w:v \preccurlyeq w} q_{uvw} + \sum_{w:w \preccurlyeq u} r_{wuv} \leq 1 \quad \forall u \preccurlyeq v$$

$$y_{uv} \leq q_{uwv} + r_{uwv} \qquad \forall u \preccurlyeq w \preccurlyeq v$$

$$0 \leq y_{uv}, q_{uwv}, r_{uwv} \leq 1 \qquad \forall u \preccurlyeq w \preccurlyeq v$$

Constructing solution to dual LP

- Now we use fact that poset is a hypergrid! For $u \preccurlyeq v$, set $y_{uv} = \frac{1}{V(v-u)}$, where $V(v - u)$ is volume of box with corners *u* and *v*.

maximize $$\sum_{u,v:u \preccurlyeq v} y_{uv}$$

$> (m \ln m)^d$

subject to:

$$\sum_{w:v \preccurlyeq w} q_{uvw} + \sum_{w:w \preccurlyeq u} r_{wuv} \le 1 \quad \forall u \preccurlyeq v$$

$\le (4\pi)^d$

$$y_{uv} = q_{uwv} + r_{uwv} \quad \forall u \preccurlyeq w \preccurlyeq v$$

Set $q_{uwv} = y_{uv} \frac{V(w-u)}{V(w-u)+V(v-w)}$, $r_{uwv} = y_{uv} \frac{V(v-w)}{V(w-u)+V(v-w)}$

Constructing solution to dual LP

- Now we use fact that poset is a hypergrid! For $u \preccurlyeq v$, set $y_{uv} = \frac{1}{(4\pi)^d V(v-u)}$, where $V(v-u)$ is volume of box with corners u and v.

maximize

$$\sum_{u,v:u \preccurlyeq v} y_{uv}$$

$> (m \ln m)^d / (4\pi)^d$

subject to:

$$\sum_{w:v \preccurlyeq w} q_{uvw} + \sum_{w:w \preccurlyeq u} r_{wuv} \leq 1 \quad \forall u \preccurlyeq v$$

≤ 1

$$y_{uv} = q_{uwv} + r_{uwv} \qquad \forall u \preccurlyeq w \preccurlyeq v$$

Set $q_{uwv} = y_{uv} \frac{V(w-u)}{V(w-u)+V(v-w)}$, $r_{uwv} = y_{uv} \frac{V(v-w)}{V(w-u)+V(v-w)}$

Wrap-up

- Upper bound for Steiner 2-TC-spanner
- Lower bound for 2-TC-spanner for a hypergrid.
  - Technique: find a feasible solution for the dual LP
- Lower bound of $\Omega(n \log^{\lceil (d-1)/k \rceil} n)$ for $k \geq 3$.
  - Combinatorial
  - Holds for randomly generated posets, not explicit.
- OPEN PROBLEM:
  - Can the LP technique give a better lower bound for $k \geq 3$?