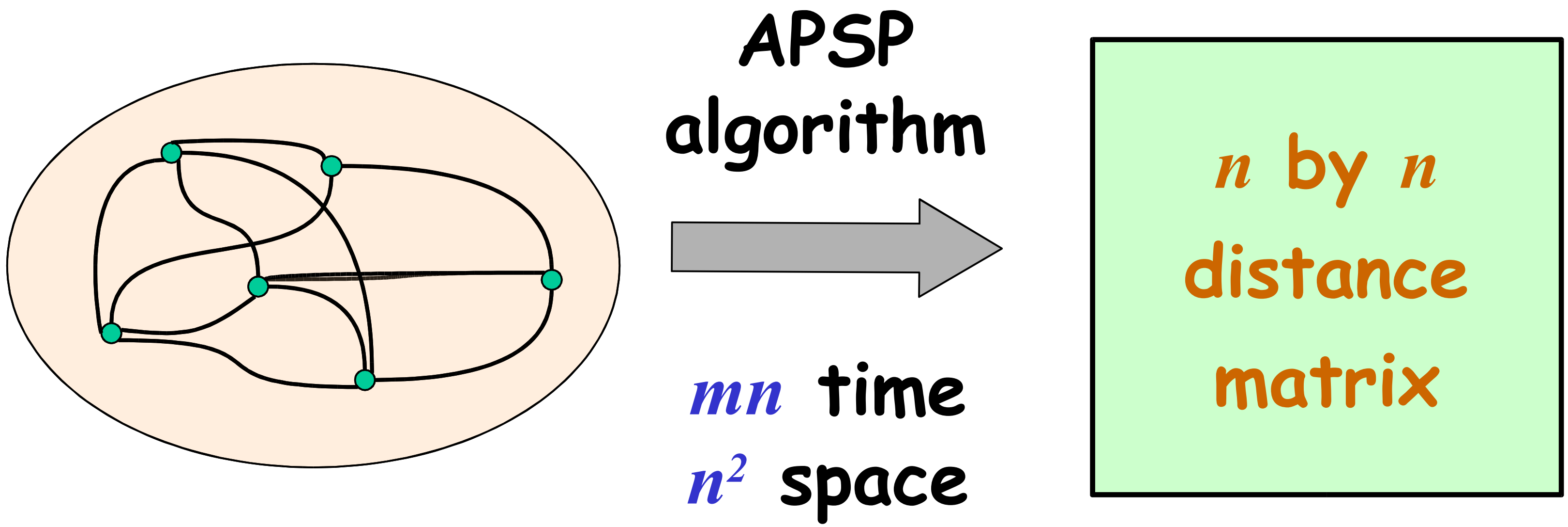
Approximating distances in graphs

Uri Zwick
Tel Aviv University

The 6th Max-Planck Advanced Course on the Foundations of Computer Science (ADFOCS)

All-Pairs Shortest Paths



Input: A weighted **undirected** graph $G=(V,E)$, where $|E|=m$ and $|V|=n$.

Output: An $n \times n$ distance matrix.

Approximate Shortest Paths

Let $\delta(u,v)$ be the distance from u to v.

Multiplicative error

An estimated distance $\delta'(u,v)$ is of **stretch** t iff

$$(\delta(u,v) \leq \delta'(u,v) \leq t \cdot \delta(u,v$$

Additive error

An estimated distance $\delta'(u,v)$ is of **surplus** t iff

$$\delta(u,v) \leq \delta'(u,v) \leq \delta(u,v) + t$$

Multiplicative and additive spanners

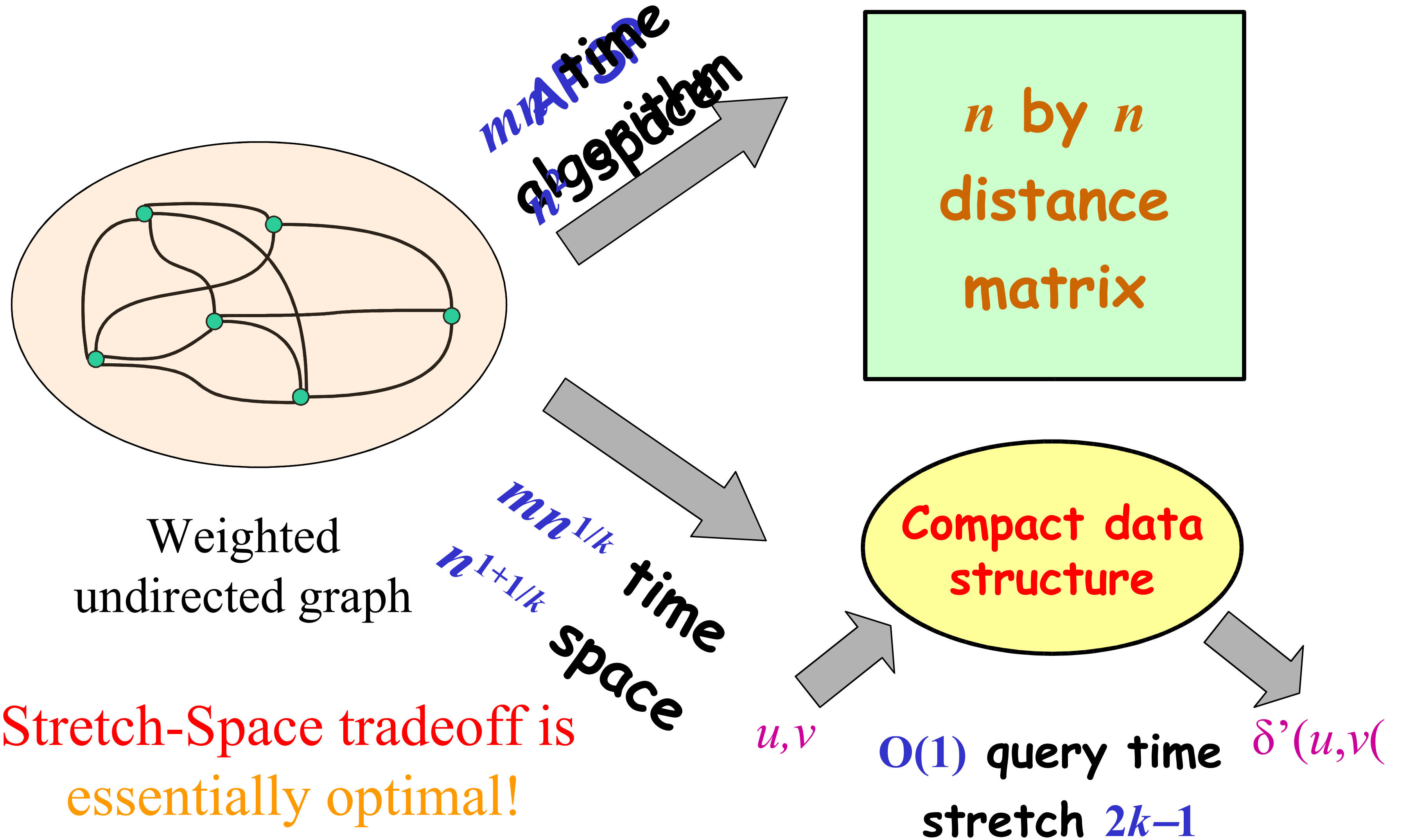
Let $G=(V, E)$ be a **weighted** undirected graph on n vertices. A subgraph $G'=(V, E')$ is a ***t*-spanner** of G iff for every $u,v\in V$ we have

$$\delta_{G'}(u,v) \le t\,\delta_G(u,v).$$

Let $G=(V, E)$ be a **unweighted** undirected graph on n vertices. A subgraph $G'=(V, E')$ is an **additive *t*-spanner** of G iff for every $u,v\in V$ we have

$$\delta_{G'}(u,v) \le \delta_G(u,v) + t.$$

Approximate Distance Oracles



Stretch-Space tradeoff is essentially optimal!

1. **All-pairs almost shortest paths** (unweighted)

   b. An $O(n^{5/2})$-time surplus-2 algorithm (ACIM'96)
   c. Additive 2-spanners with $O(n^{3/2})$ edges.
   d. An $O(n^{3/2}m^{1/2})$-time surplus-2 algorithm (DHZ'96)

2. **Multiplicative spanners** (weighted graphs)

   b. $(2k-1)$-spanners with $n^{1+1/k}$ edges (ADDJS'93)
   c. Linear time construction (BS'03)

3. **Approximate distance oracles** (weighted graphs)

   b. Stretch=$2k-1$ query time=$O(1)$ space=$O(kn^{1+1/k})$ (TZ'01)

5. **Spanners with sublinear distance errors** (unweighted)

   b. Additive error $O(d^{1/(k-1)})$ with $O(kn^{1+1/k})$ edges (TZ'05)

All-Pairs Almost Shortest Paths

unweighted, undirected graphs

Surplus	Time	Authors
0	mn	folklore
2	$n^{5/2}$	Aingworth-Chekuri-Indyk-Motwani '96
2	$n^{3/2}m^{1/2}$	Dor-Halperin-Zwick '96
2	$n^{7/3}$	"
$2(k-1)$	$n^{2-1/k}m^{1/k}$	"
$2(k-1)$	$n^{2+1/(3k-4)}$	"

$O(n^{5/2})$-time surplus-2 algorithm

unweighted, undirected graphs

1) Add each vertex v to A, independently, with probability $n^{-1/2}$. (Elements of A are "centers".) — $O(n)$

2) From every center $v \in A$, find a tree of shortest paths from v and add its edges to E'. — $O(m|A|) = O(n^{5/2})$

3) For every non-center $v \notin A$:

   a) If v has a neighbor $u \in A$, then add the single edge (u,v) to E'. — $O(m)$

   b) Otherwise, add all the edges incident to v to E'.

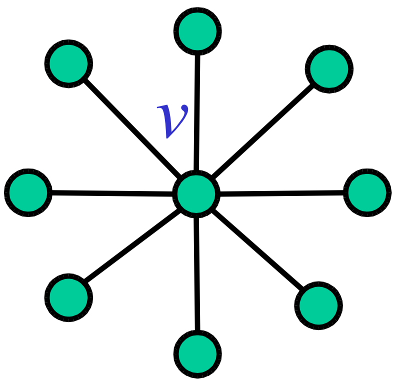
4) Solve the APSP problem on the subgraph $G'=(V, E')$. — $O(n|E'|) = O(n^{5/2})$

Number of edges in E'

- The expected # of edges added to E' in 2) is $O(n^{3/2})$.
- The expected # of edges added to E' in 3) is also $O(n^{3/2})$.

Consider a vertex v of degree d

If one of the neighbors of v is placed in A, then E' will contain only one edge incident on v.



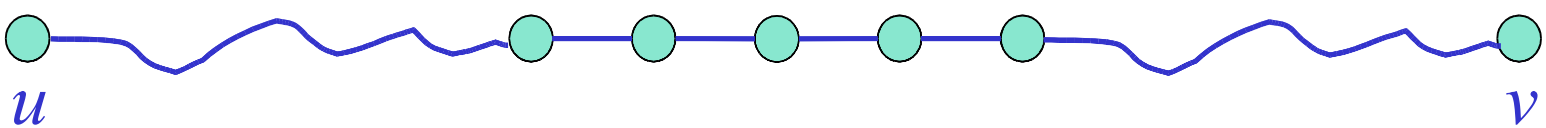
Hence, the expected number of edges incident to v added to E' is at most

$$d(1-n^{-1/2})^d + 1 \leq n^{1/2}$$

The surplus-2 algorithm

Correctness – Case 1

Case 1: No vertex on a shortest path from u to v has a neighboring center.



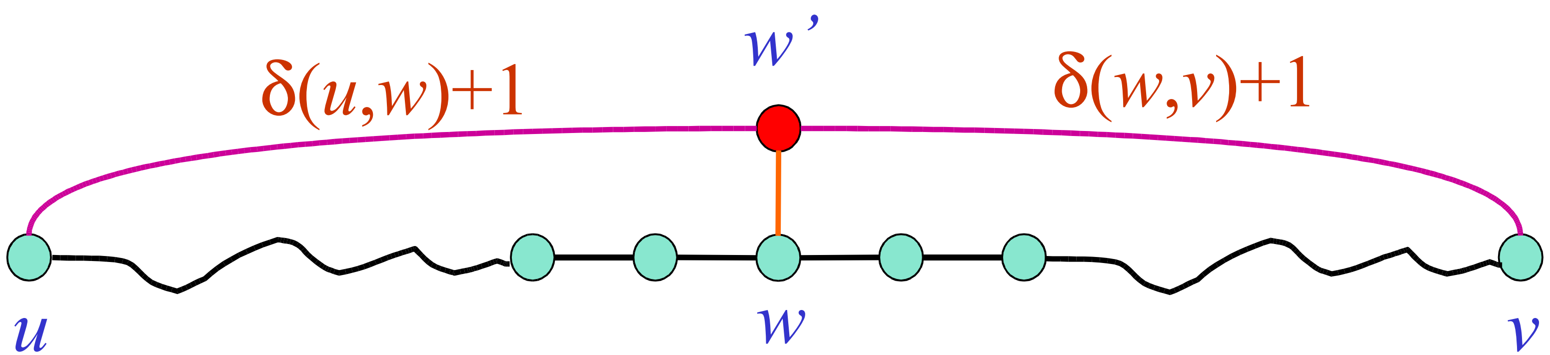
All the edges on the path are in E'.

We find a shortest path from u to v.

The surplus-2 algorithm

Correctness – Case 2

Case 2: At least one vertex on a shortest path from u to v has a neighboring center.



We find a path from u to v of surplus at most 2

Additive 2-spanners

Every unweighted undirected graph $G=(V, E)$ on n vertices has a subgraph $G'=(V, E')$ with $O(n^{3/2})$ edges such that for every $u,v \in V$ we have $\delta_{G'}(u,v) \leq \delta_G(u,v) + 2$.

$O(n^{3/2}m^{1/2})$-time surplus-2 algorithm

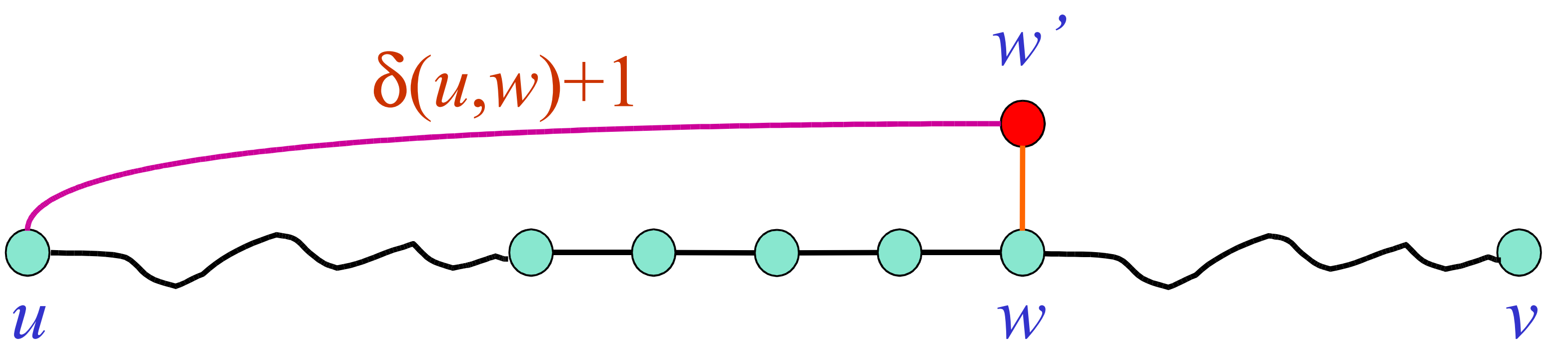
unweighted, undirected graphs

1) Add each vertex v to A, independently, with probability $(n/m)^{1/2}$. (Elements of A are "centers".)
2) From every center $v \in A$, find distances to all other vertices in the graph. (Do not add edges to E'.)
3) For every non-center $v \notin A$:
   a) If v has a neighbor $u \in A$, then add the single edge (u,v) to E'.
   b) Otherwise, add all the edges incident to v to E'.
4) For every non-center vertex $v \notin A$:
   a) Construct a set $F(v)=\{ (v,w) \mid w \in A \}$ of **weighted** edges. The weight of an edge (v,w) is $\delta(w,v)$.
   b) Find distances from v to all other vertices in the **weighted** graph $G'(v)=(V, E' \cup F(v))$.

$O(n^{3/2}m^{1/2})$-time surplus-2 algorithm

Correctness – Case 2

Case 2: At least one vertex on a shortest path from u to v has a neighboring center.



Consider the last vertex with a neighboring center.

We find a path from u to v of surplus at most 2

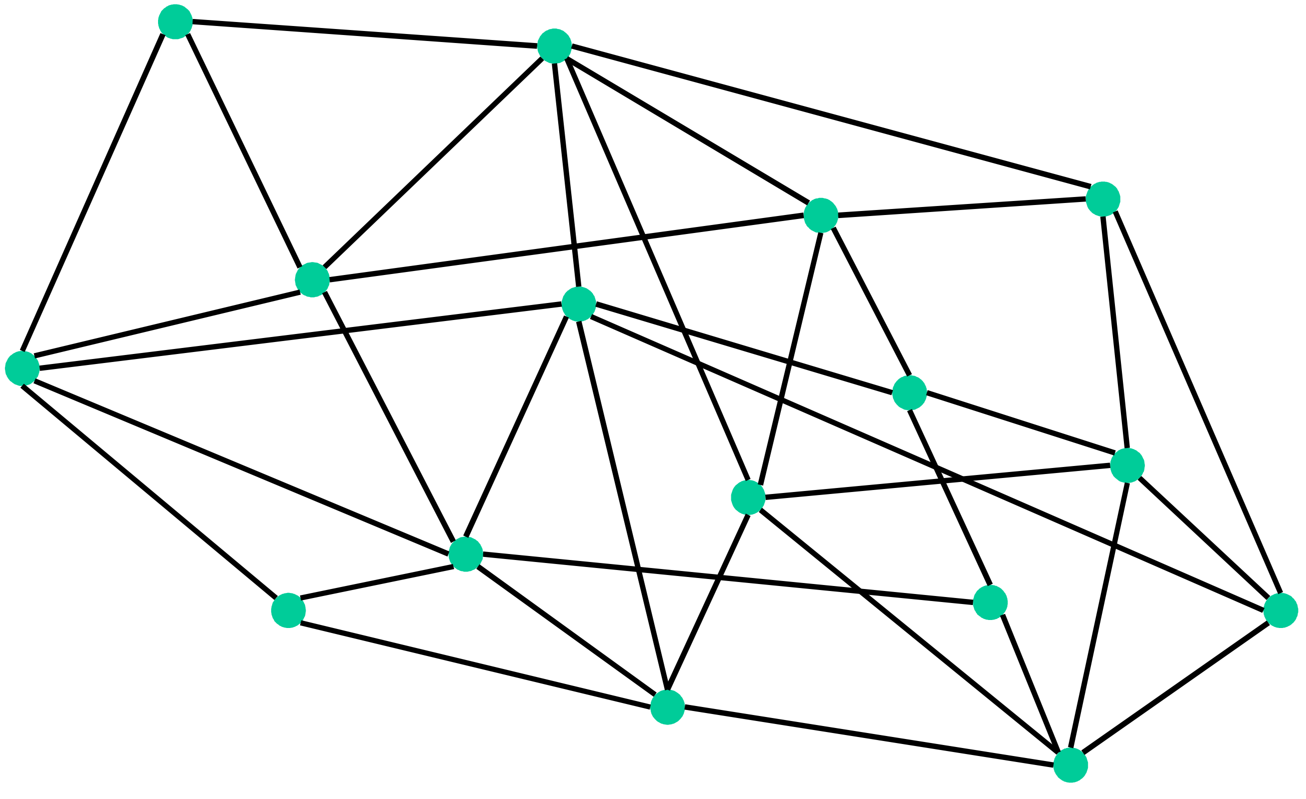
All-Pairs Almost Shortest Paths

weighted undirected graphs

Stretch	Time	Reference
1	mn	Dijkstra ’59
2	$n^{3/2}m^{1/2}$	Cohen-Zwick ’97
7/3	$n^{7/3}$	”
3	n^2	”

Some log factors ignores

Spanners



Given an **arbitrary** dense graph, can we always find a relatively **sparse subgraph** that approximates **all** distances fairly well?

Spanners [PU’89,PS’89]

Let $G=(V,E)$ be a weighted undirected graph.

A subgraph $G'=(V,E')$ of G is said to be a t-spanner of G iff $\delta_{G'}(u,v) \leq t\,\delta_G(u,v)$ for every $u,v \in V$.

Theorem:

Every weighted undirected graph has a $(2k-1)$-spanner of size $O(n^{1+1/k})$. [ADDJS’93]

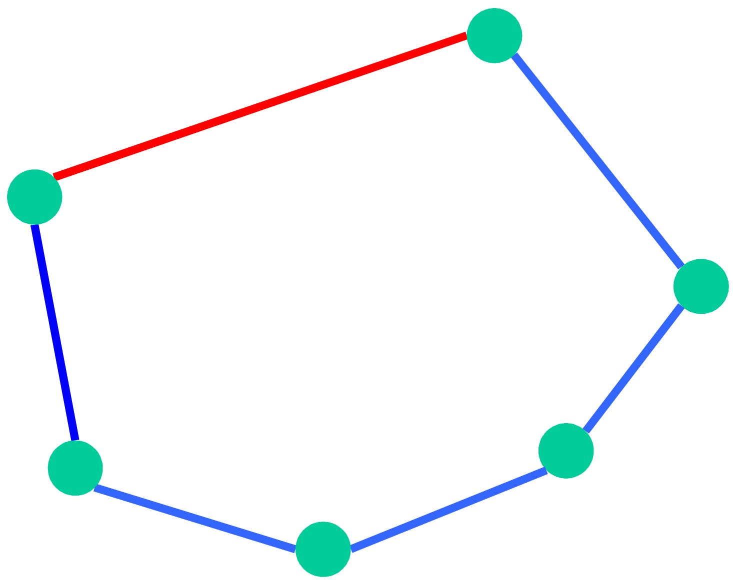
Furthermore, such spanners can be constructed deterministically in linear time. [BS’03] [RTZ’05]

The size-stretch trade-off is optimal if there are graphs with $\Omega(n^{1+1/k})$ edges and girth $2k+2$, as conjectured by Erdös and others.

A simple spanner construction algorithm

[Althöfer, Das, Dobkin, Joseph, Soares '93]

- Consider the edges of the graph in non-decreasing order of weight.
- Add each edge to the spanner if it does not close a cycle of size at most $2k$.
- The resulting graph is a $(2k-1)$-spanner.
- The resulting graph has girth $\geq 2k$. Hence the number of edges in it is at most $n^{1+1/k}$.



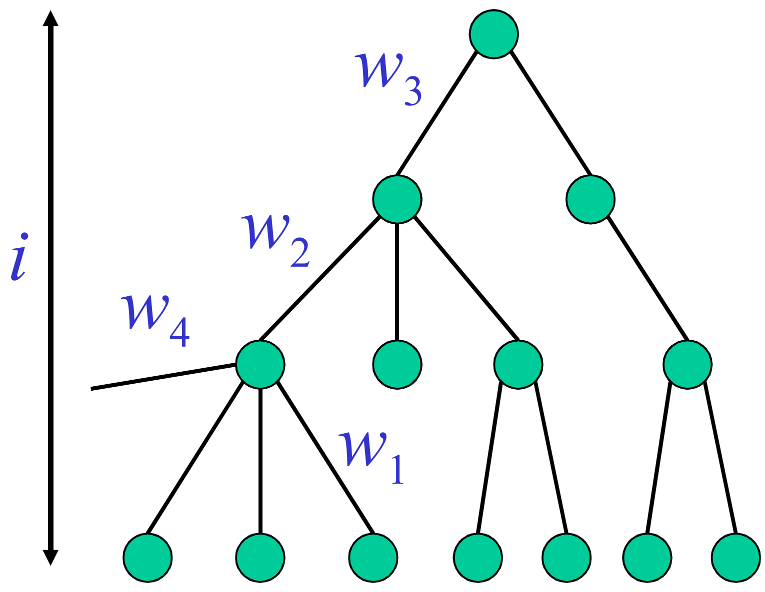
If $|\text{cycle}| \leq 2k$, then red edge can be removed.

Linear time spanner construction [BS'03]

- The algorithm is composed of k iterations.
- At each iteration some edges are added to the spanner and some edges and vertices are removed from the graph.
- At the end of the i-th iteration we have a collection of about $n^{1-i/k}$ trees of depth at most i that contain all the remaining vertices of the graph.

Tree properties

- The edges of the trees are spanner edges.
- The weights of the edges along every leaf-root path are non-increasing.
- For every surviving edge (u,v) we have $w(u,v) \geq w(u,p(u))$, where $p(u)$ is the parent of u.



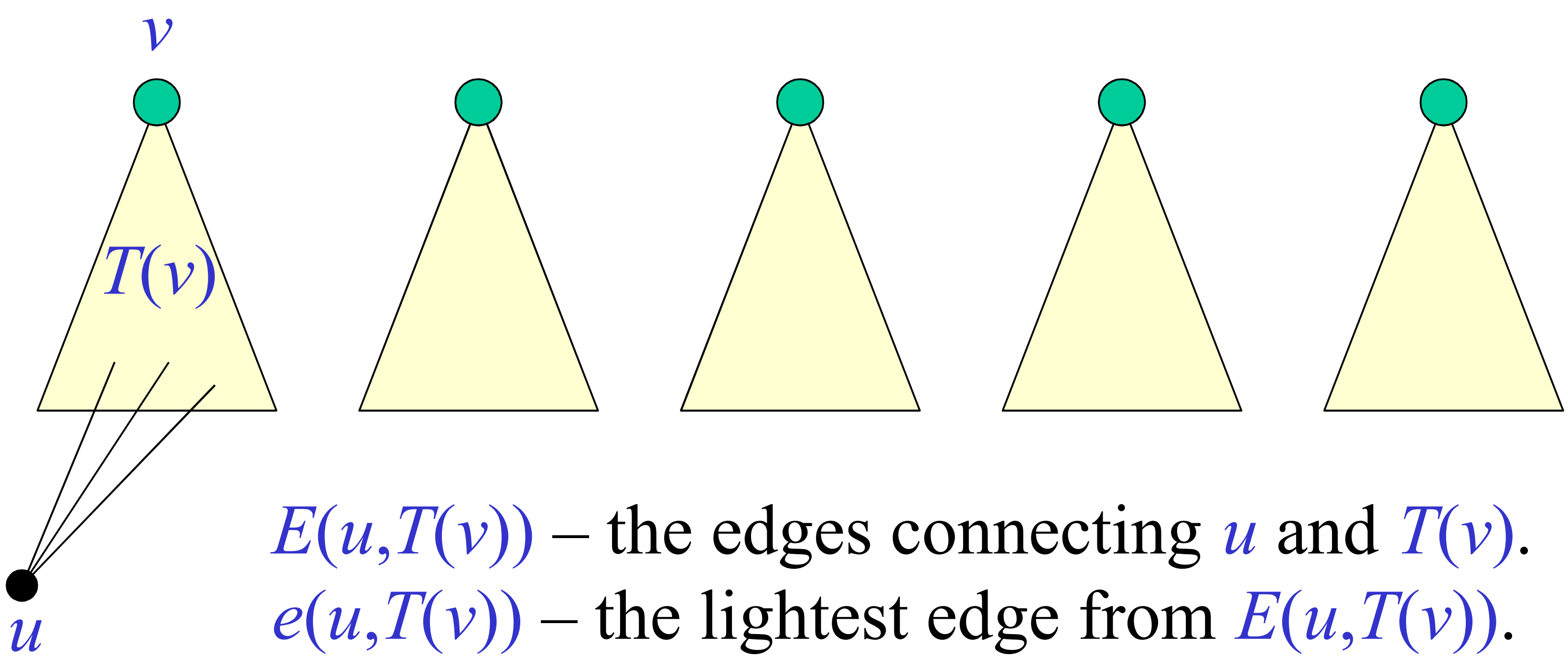
$$w_1 \geq w_2 \geq w_3$$

$$w_4 \geq w_2$$

Notation

A_i – roots of trees of the i-th iteration

$T(v)$ – the tree rooted at v



$E(u,T(v))$ – the edges connecting u and $T(v)$.

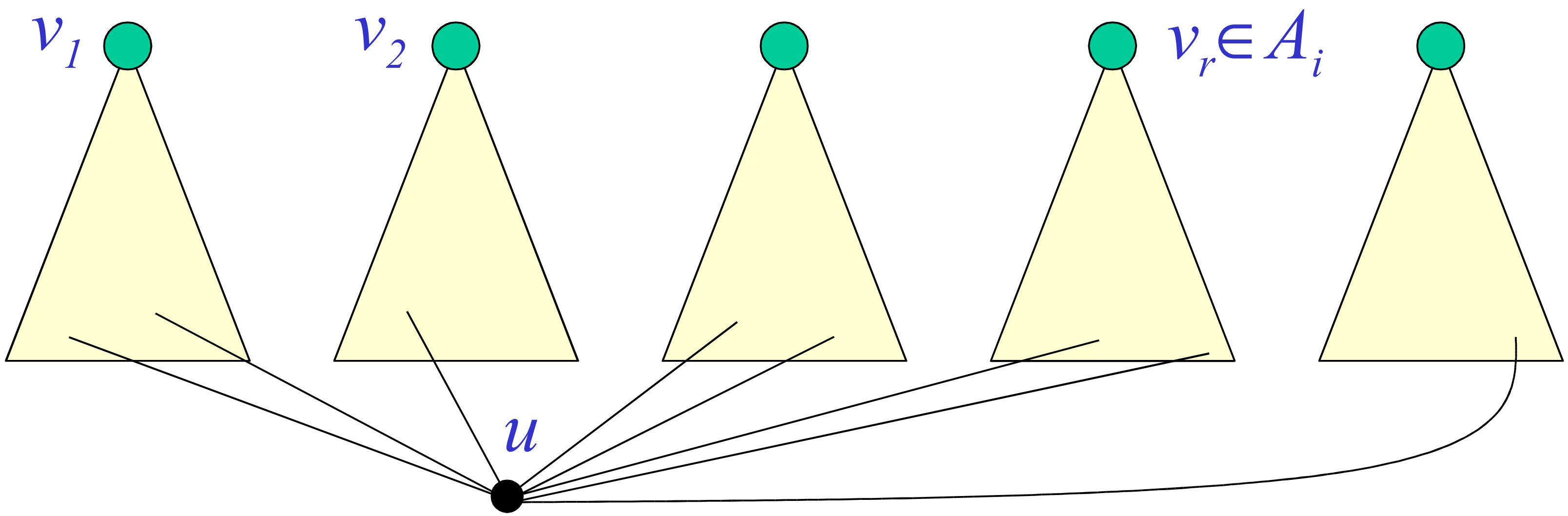
$e(u,T(v))$ – the lightest edge from $E(u,T(v))$.

$w(u,T(v))$ – the weight of $e(u,T(v))$ (or ∞).

The i-th iteration

Each vertex $v \in A_{i\text{-}1}$ is added to A_i with probability $n^{-1/k}$.

In the last iteration $A_k \leftarrow \emptyset$.



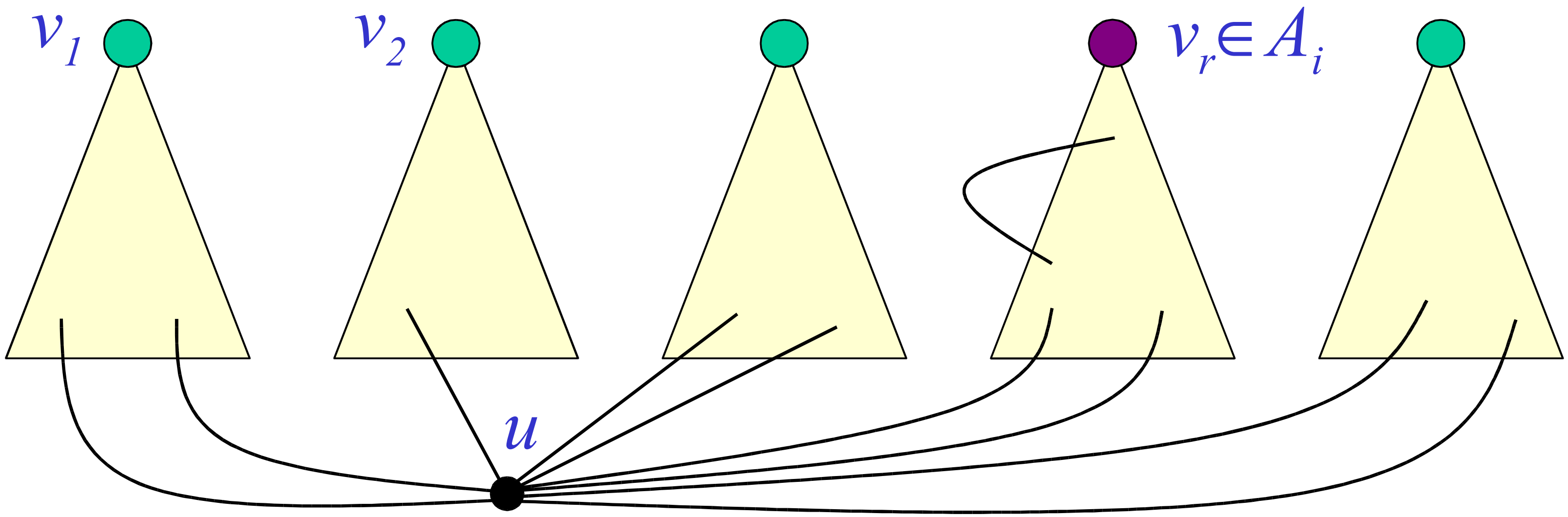
Let $v_1, v_2, \ldots$ be the vertices of $A_{i\text{-}1}$ such that

$w(u,T(v_1)) \leq w(u,T(v_2)) \leq \ldots$

Let $r=r(u)$ be the minimal index for which $v_r \in A_i$.

If there is no such index, let $r(u) = |A_{i-1}|$.

The i-th iteration (cont.)



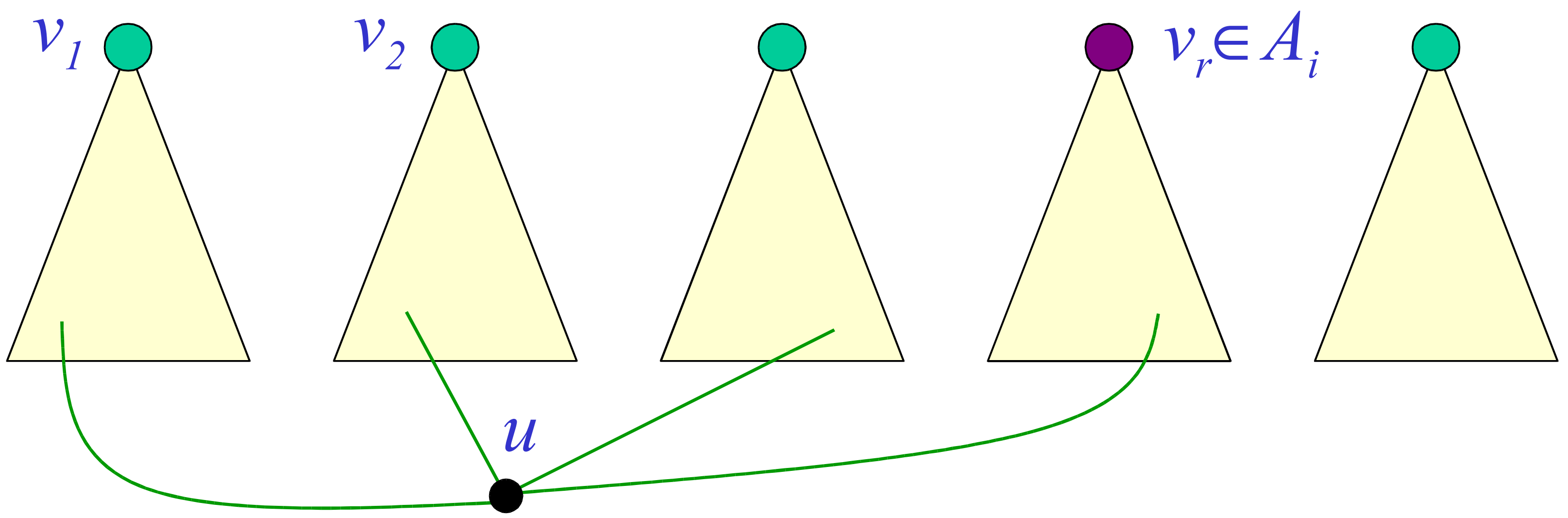
For every vertex u that belongs to a tree whose root is in $A_{i-1} - A_i$:

For every $1 \leq j \leq r$:
Add $e(u,T(v_i))$ to the spanner.
Remove $E(u,T(v_i))$ from the graph

Remove edges that connect vertices in the same tree.

Remove vertices that have no remaining edges.

How many edges are added to the spanner?



$$\mathrm{E}[r(u)] \le n^{1/k}$$

Hence, the expected number of edges added to the spanner in each iteration is at most $n^{1+1/k}$.

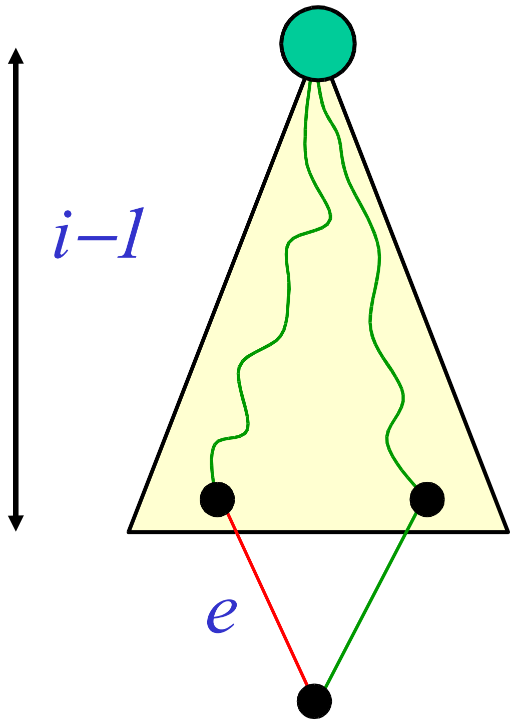
What is the stretch?

Let e be an edge deleted in the i-th iteration.

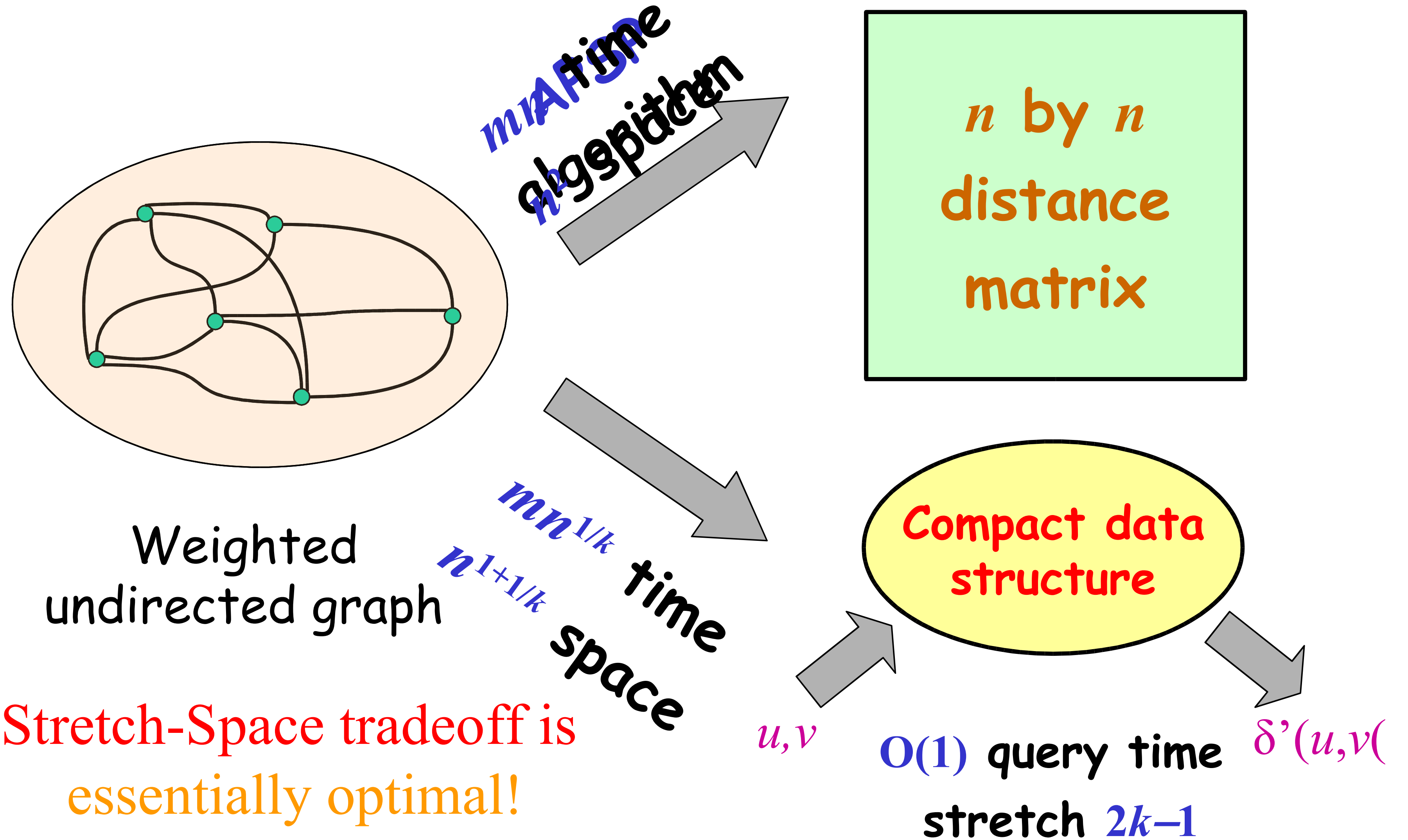
The spanner contains a path of at most $2(i-1)+1$ edges between the endpoints of e.

The edges of the path are not heavier than e

Hence, stretch $\leq 2k-1$

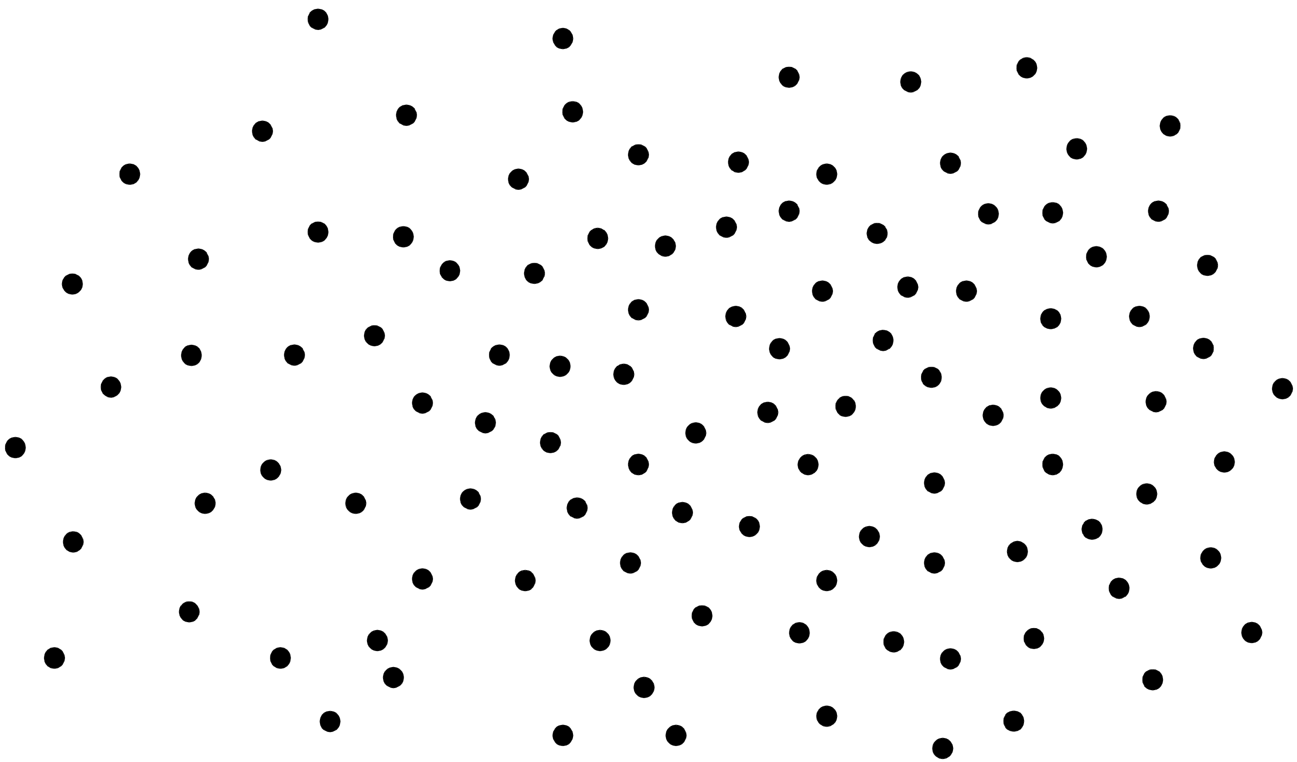


Approximate Distance Oracles (TZ'01)



Approximate Distance Oracles [TZ'01]

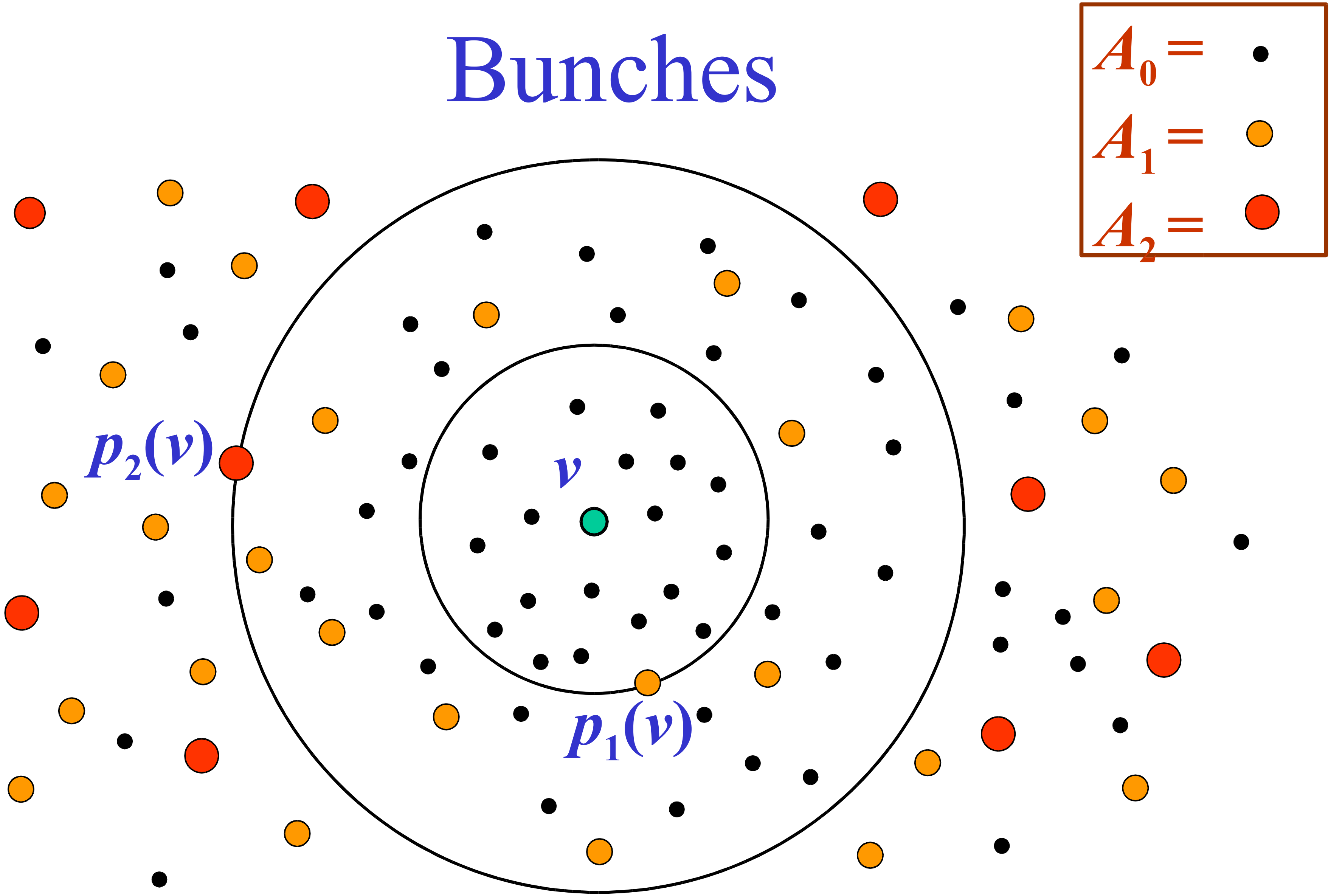
A hierarchy of centers



$$A_0 \leftarrow V \,;\, A_k \leftarrow \varnothing$$

$$A_i \leftarrow \text{sample}(A_{i-1}, n^{-1/k})$$

Bunches



$$B(v) \leftarrow \bigcup_i \{ w \in A_i - A_{i+1} \mid \delta(w,v) < \delta(A_{i+1},v) \}$$

Lemma: $\mathrm{E}[|B(v)|] \leq kn^{1/k}$

Proof: $|B(v) \cap A_i|$ is stochastically dominated by a geometric random variable with parameter $p=n^{-1/k}$.

The data structure

Keep for every vertex $v \in V$:

- The centers $p_1(v), p_2(v), \ldots, p_{k-1}(v)$
- A **hash table** holding $B(v)$

For every $w \in V$, we can check, in **constant time**, whether $w \in B(v)$, and if so, what is $\delta(v,w)$.

Query answering algorithm

Algorithm dist$_k(u,v)$

$w \leftarrow u$, $i \leftarrow 0$

while $w \notin B(v)$

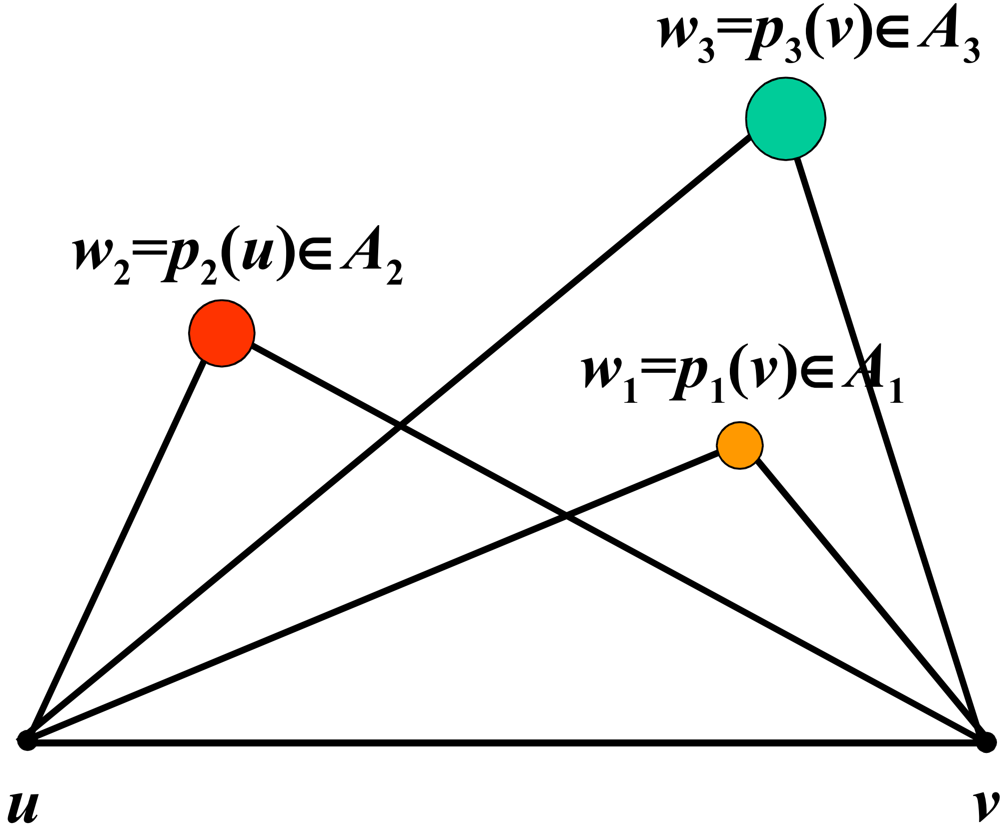
{ $i \leftarrow i+1$

$(u,v) \leftarrow (v,u)$

$w \leftarrow p_i(u)$ }

return $\delta(u,w) + \delta(w,v)$

Query answering algorithm



Analysis

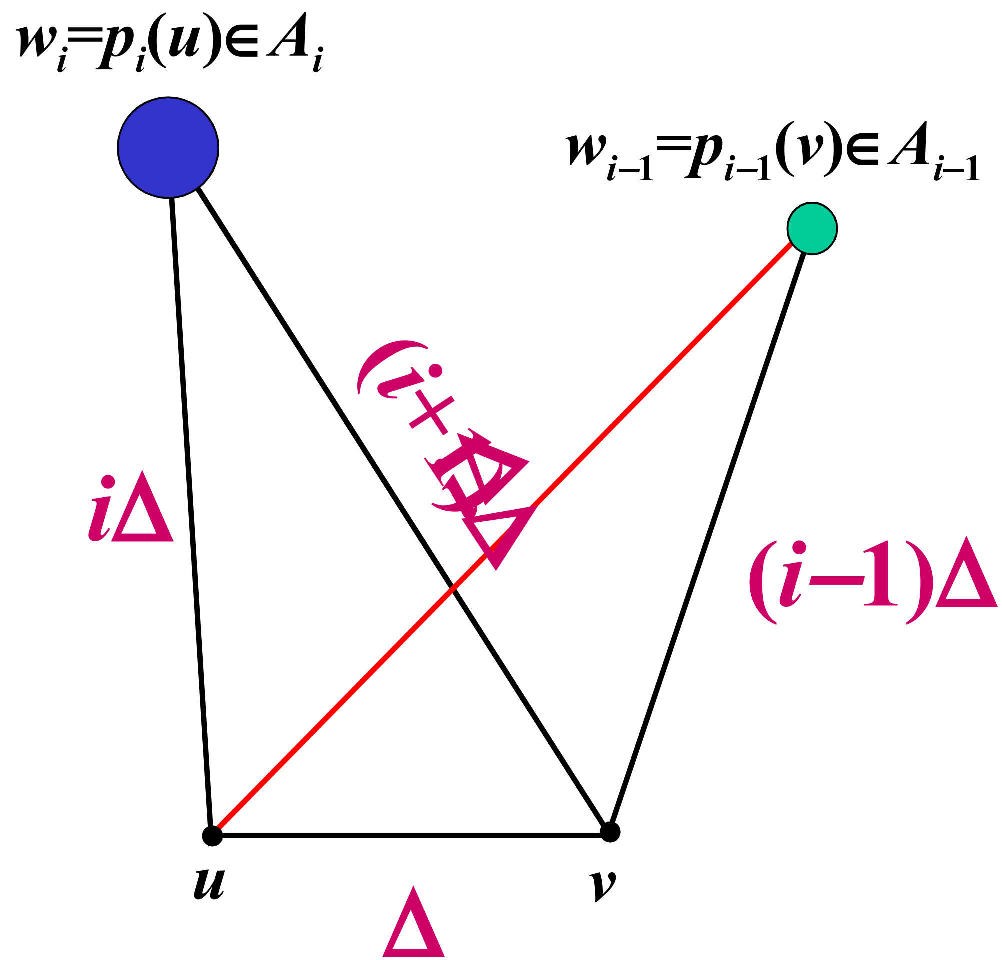
Claim 1:

$\delta(u,w_i) \leq i\Delta$, *i* even

$\delta(v,w_i) \leq i\Delta$, *i* odd

Claim 2:

$$\delta(u,w_i)+\delta(w_i,v) \leq (2i+1)\Delta \leq (2k-1)\Delta$$

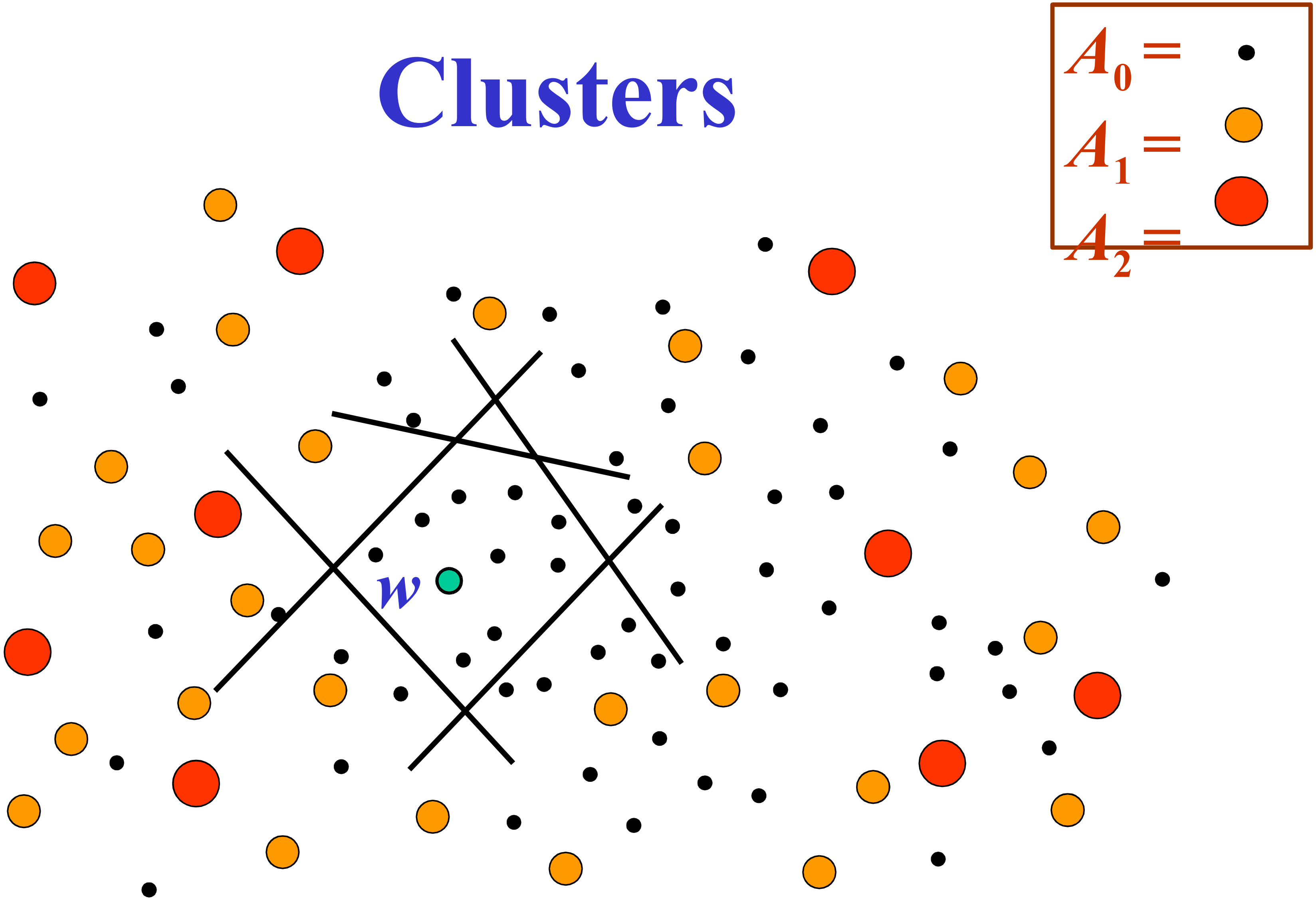


Where are the spanners?

Define clusters, the “duals” of bunches.

For every $u \in V$, put in the spanner a tree of shortest paths from u to all the vertices in the cluster of u.

Clusters



$$C(w) \leftarrow \{v \in V \mid \delta(w,v) < \delta(A_{i+1},v)\} \quad , \quad w \in A_i - A_{i+1}$$

Bunches and clusters

$$w \in B(v) \quad \Leftrightarrow \quad v \in C(w)$$

$$C(w) \leftarrow \{v \in V \mid \delta(w,v) < \delta(A_{i+1},v)\} \quad ,$$

$$\textit{if } w \in A_i - A_{i+1}$$

$$B(v) \leftarrow \bigcup_i \{w \in A_i - A_{i+1} \mid \delta(w,v) < \delta(A_{i+1},v)\}$$

Additive Spanners

Let $G=(V,E)$ be a unweighted undirected graph.

A subgraph $G'=(V, E')$ of G is said to be an additive t-spanner of G iff $\delta_{G'}(u,v) \le \delta_G(u,v) + t$ for every $u,v \in V$.

Theorem: Every unweighted undirected graph has an additive 2-spanner of size $O(n^{3/2})$. [ACIM '96] [DHZ '96]

Theorem: Every unweighted undirected graph has an additive 6-spanner of size $O(n^{4/3})$. [BKMP '04]

Major open problem

Do all graphs have additive spanners with only $O(n^{1+\varepsilon})$ edges, for every $\varepsilon>0$?

Spanners with sublinear surplus

Theorem:

For every $k>1$, every undirected graph $G=(V,E)$ on n vertices has a subgraph $G'=(V,E')$ with $O(n^{1+1/k})$ edges such that for every $u,v\in V$, if $\delta_G(u,v)=d$, then $\delta_{G'}(u,v)=d+O(d^{1-1/(k-1)})$.

$$d \;\Rightarrow\; d + O(d^{1-1/(k-1)})$$

Extends and simplifies a result of Elkin and Peleg (2001)

All sorts of spanners

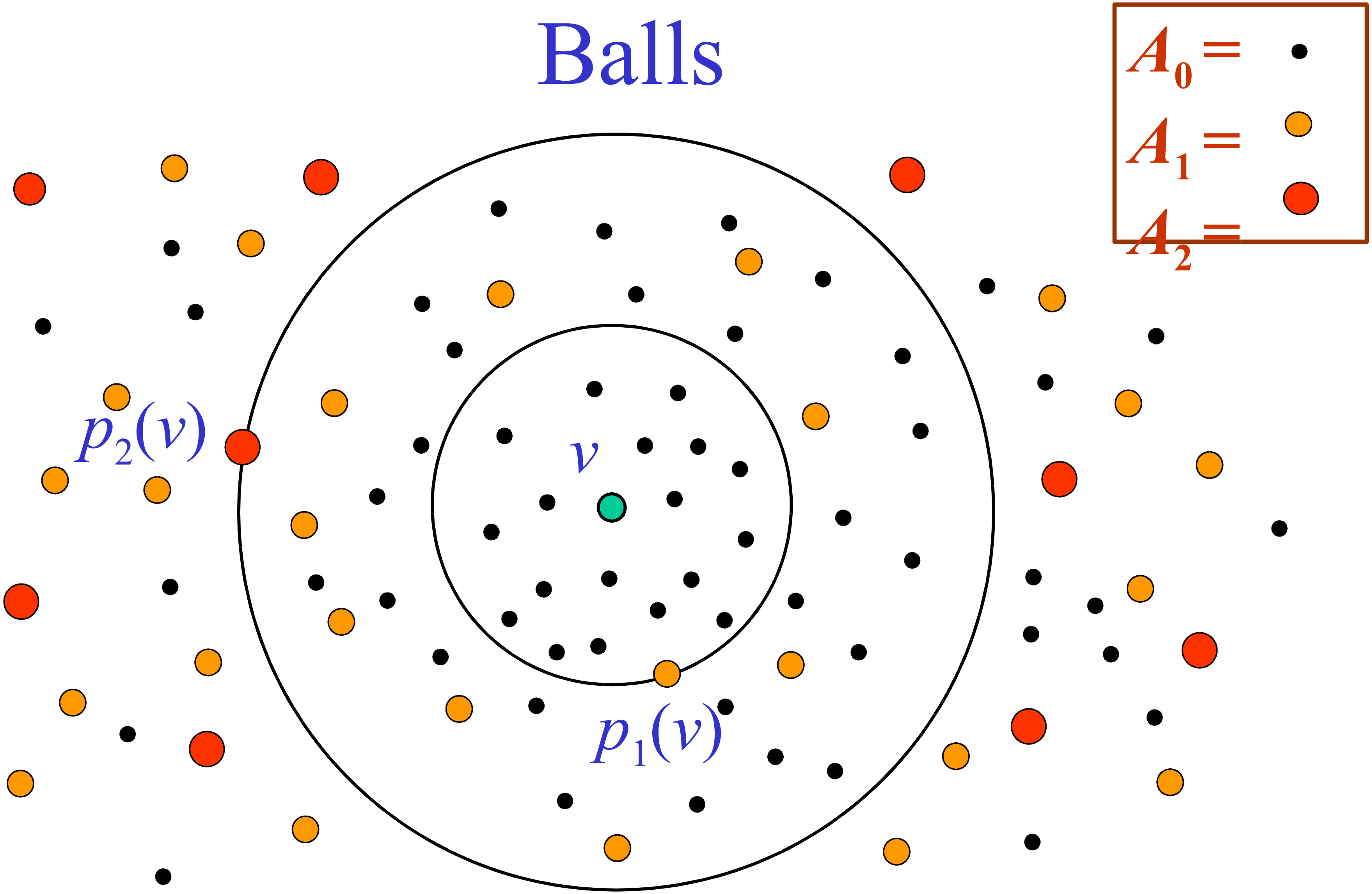
A subgraph $G'=(V,E')$ of G is said to be a functional f-spanner if G iff $\delta_{G'}(u,v) \leq f(\delta_G(u,v))$ for every $u,v \in V$.

size	$f(d)$	reference
$n^{1+1/k}$	$(2k-1)d$	[ADDJS '93]
$n^{3/2}$	$d + 2$	[ACIM '96] [DHZ '96]
$n^{4/3}$	$d + 6$	[BKMP '04]
$\beta n^{1+\delta}$	$(1+\varepsilon)d + \beta(\varepsilon,\delta)$	[EP '01]
$n^{1+1/k}$	$d + O(d^{1-1/(k-1)})$	[TZ '05]

The construction of the approximate distance oracles, when applied to unweighted graphs, produces spanners with sublinear surplus!

We present a slightly modified construction with a slightly simpler analysis.

Balls



$$Ball(u) = \{v \in V \mid \delta(u,v) < \delta(u, A_{i+1})\} \;\;,\;\; u \in A_i - A_{i+1}$$

$$Ball[u] = Ball(u) \cup \{p_{i+1}(u)\} \;\;,\;\; u \in A_i - A_{i+1}$$

Spanners with sublinear surplus

Select a hierarchy of centers $A_0 \supset A_1 \supset \ldots \supset A_{k-1}$.

For every $u \in V$, add to the spanner a shortest paths tree of $\mathrm{Ball}[u]$.

The path-finding strategy

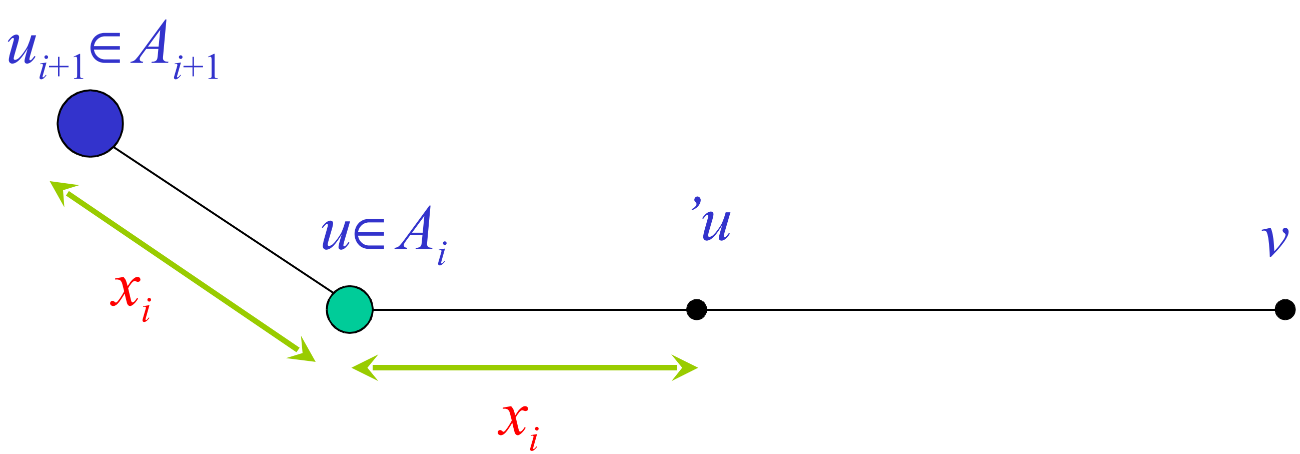
Suppose we are at $u \in A_i$ and want to go to v.

Let Δ be an integer parameter.

If the first $x_i = \Delta^i - \Delta^{i-1}$ edges of a shortest path from u to v are in the spanner, then use them.

Otherwise, head for the $(i+1)$-center u_{i+1} nearest to u.

► The distance to u_{i+1} is at most x_i. (As $u' \notin \mathrm{Ball}(u)$.)

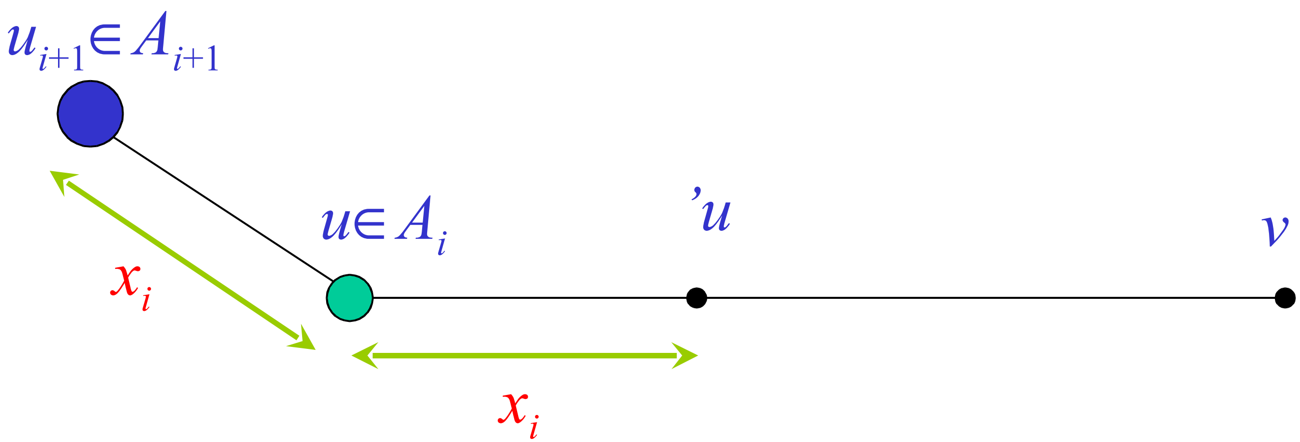


The path-finding strategy

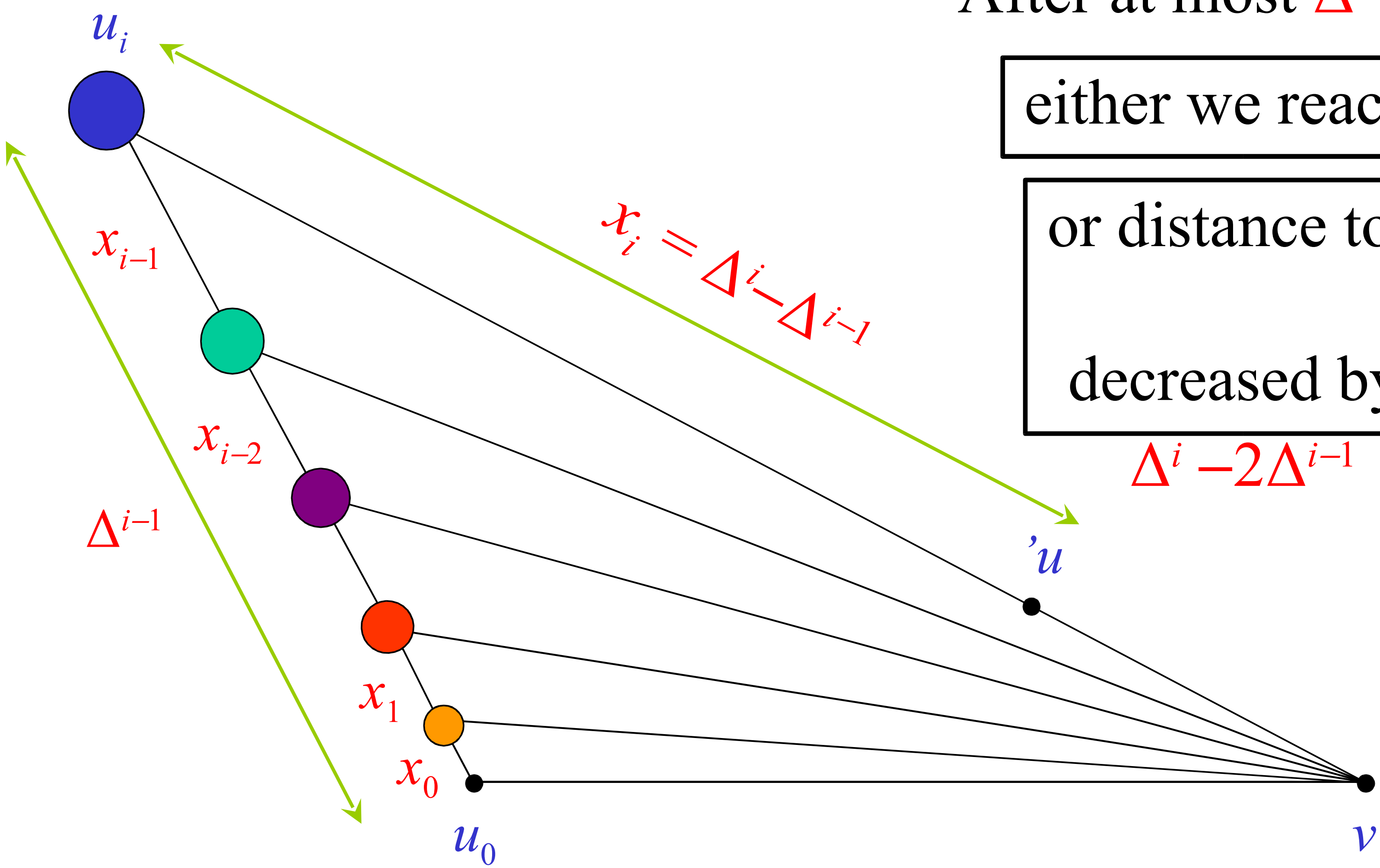
We either reach v, or at least make $x_i = \Delta^i – \Delta^{i–1}$ steps in the right direction.

Or, make at most $x_i = \Delta^i – \Delta^{i–1}$ steps, possibly in a wrong direction, but reach a center of level $i+1$.

If $i=k–1$, we will be able to reach v.



The path-finding strategy



After at most Δ^i steps:

either we reach v

or distance to v decreased by $\Delta^i - 2\Delta^{i-1}$

The path-finding strategy

After at most Δ^i steps:

either we reach v → Surplus $2\Delta^{i-1}$

or distance to v decreased by $\Delta^i - 2\Delta^{i-1}$ → Stretch

$$\frac{\Delta^i}{\Delta^i - 2\Delta^{i-1}} = 1 + \frac{2}{\Delta - 2}$$

The surplus is incurred only once!

$$\delta'(u,v) \le \left(1 + \frac{2}{\Delta - 2}\right) \cdot \delta(u,v) + 2\Delta^{k-2}$$

Sublinear surplus

$$\delta'(u,v) \leq \left(1+\frac{2}{\Delta-2}\right)\cdot\delta(u,v) + 2\Delta^{k-2}$$

$$\delta(u,v)=d \quad , \quad \Delta=\left\lceil d^{1/(k-1)}+2\right\rceil$$

$$\Downarrow$$

$$\delta'(u,v) \leq d+O(d^{1-\frac{1}{k-1}})$$