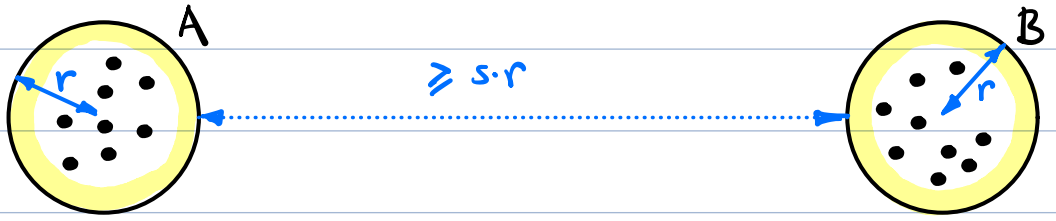
CMSC 754 - Computational Geometry
Lecture 15: Well-Separated Pair Decompositions (II)

Review:

- Given point set $P = \{p_1, \ldots, p_n\}$ in $\mathbb{R}^d$ ($d = O(1)$) and separation factor $s > 0$, two sets A & B are **s-well separated** if they can be enclosed in balls of some radius r that are separated by distance $\geq s \cdot r$



Def: Given $P \subseteq \mathbb{R}^d$ & $s > 0$, an **s-well separated pair decomposition** (s-WSPD) is collection of pairs:

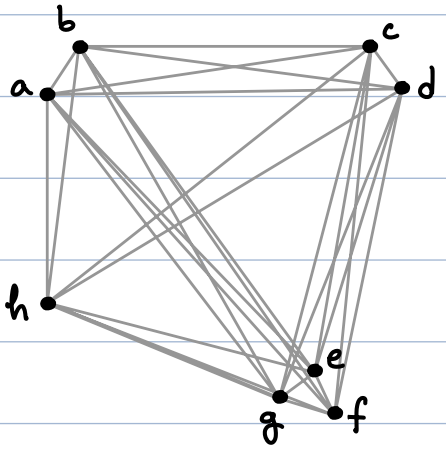
$$\{\{A_1, B_1\}, \ldots, \{A_m, B_m\}\} \quad \text{where } A_i, B_i \subseteq P$$

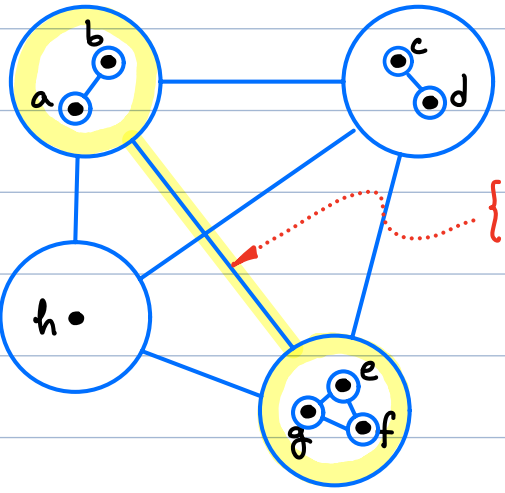
such that:

(i) A_i & B_i are s-well separated, $1 \leq i \leq m$
[$\Rightarrow A_i$ & B_i are disjoint]

(ii) $\bigcup_{i=1}^{m} A_i \otimes B_i = P \otimes P$
[$\Rightarrow$ all pairs are covered]

28 pairs

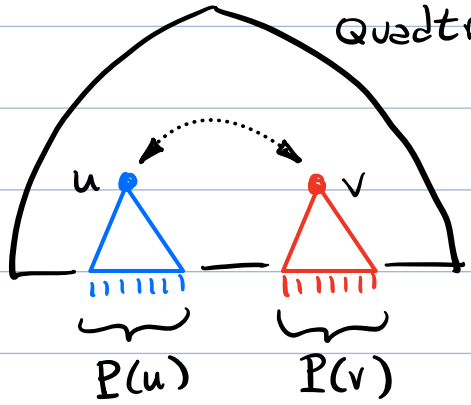


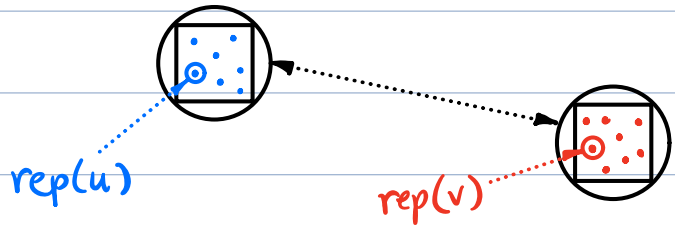


11 well-separated pairs

{{a,b},{e,f,g}}

- Given an n-element set P in $\mathbb{R}^d$ and $s \geq 1$, an s-WSPD of size $O(s^d n)$ can be constructed in time $O(n \log n + s^d n)$
- Construction uses a quadtree. Given node u:
  - $P(u)$ = points in u's subtree
  - $rep(u) \in P(u)$ arbitrary representative
- Each WSP given as node pair $\{u,v\}$. WSP = $P(u) \otimes P(v)$

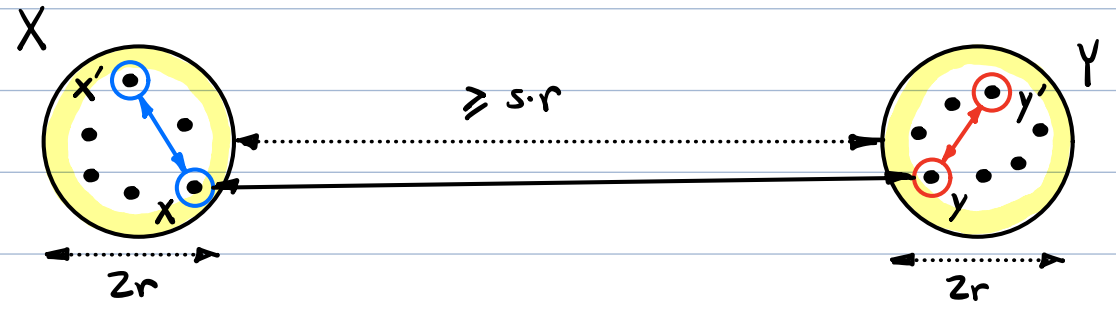




Utility Lemma: Given an s-WSP $\{X, Y\}$ and $x, x' \in X$ & $y, y' \in Y$:

(i) $\|x - x'\| \leq (2/s) \cdot \|x - y\|$

(ii) $\|x' - y'\| \leq (1 + 4/s) \cdot \|x - y\|$



Intuition: (i) Points in same dumbbell head much closer than in opposite heads.

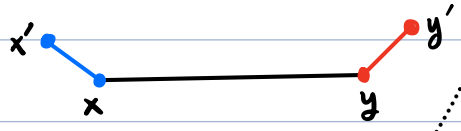
(ii) Distances on opposite sides similar

Proof: (i) $\|x-x'\| \le 2 \cdot r$

$= 2 \cdot r \frac{s \cdot r}{s \cdot r} \le \frac{2 \cdot r}{s \cdot r} \|x-y\|$ (since $\|x-y\| \ge s \cdot r$)

$= (2/s)\|x-y\|$ ✓

(ii) Since $\|x-y\| \ge s \cdot r$, $4r \le (4/s)\|x-y\|$



By triangle inequality:

$\|x'-y'\| \le \|x'-x\| + \|x-y\| + \|y-y'\|$

$\le 2r + \|x-y\| + 2r$

$= \|x-y\| + 4r$

$\le \|x-y\| + (4/s) \cdot \|x-y\|$

$= (1+4/s)\|x-y\|$ ✓ □

Applications:

① $(1+\varepsilon)$ approx. to diameter

② Exact closest pair

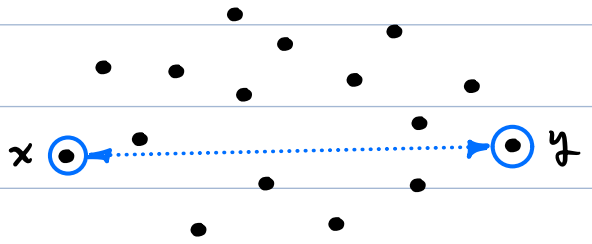
③ Euclidean t-spanner (for $t > 1$)

④ $(1+\varepsilon)$ approx. to Euclidean MST

① Approximating the diameter

- Given a point set $P = \{p_1, \ldots, p_n\}$ in $\mathbb{R}^d$, define

$$\text{diam}(P) = \max_{x,y \in P} \|x - y\|$$



- $\mathbb{R}^2$ – can solve in $O(n \log n)$ (Scan convex hull)
- $\mathbb{R}^d$ – Brute force: $O(n^2)$ – Too slow!
- Given $\varepsilon > 0$, an ε-approx. diameter is $x', y' \in P$ s.t.

$$\|x' - y'\| \geq \text{diam}(P)/(1+\varepsilon)$$

Approx. Diameter (Given P + $\varepsilon > 0$)

- Set $s = 4/\varepsilon$
- Compute s-WSPD for P ⋯⋯▸ $O(n \log n + s^d n)$
- for each WSP $\{u, v\}$ ⋯⋯▸ $O(s^d n)$

  $\text{dist}_{u,v} = \|\text{rep}(u) - \text{rep}(v)\|$
- Return max. $\text{dist}_{u,v}$

  (+ rep(u), rep(v) as points)

Time: $O(n \log n + s^d n)$

(But $s^d = (4/\varepsilon)^d = O(1/\varepsilon^d)$, since $d = O(1)$)

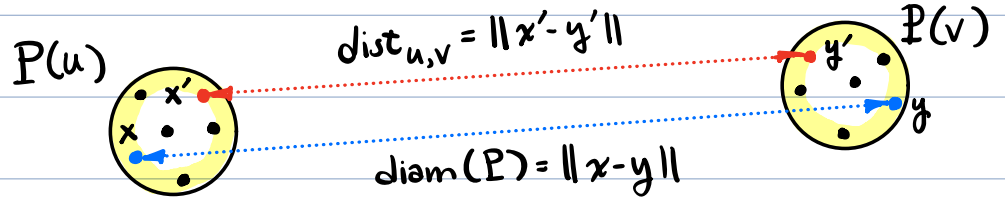
$= O(n \log n + n/\varepsilon^d)$

Correctness:

Claim: $\exists$ WSP u, v s.t. $dist_{u,v} \geq diam(P)/(1+\varepsilon)$

Proof:

- Let x, y be true diameter pair: $diam(P) = \|x-y\|$
- $\exists$ wsp $\{u, v\}$ s.t. $x \in P(u)$ + $y \in P(v)$



- Let $x' = rep(u)$ $y' = rep(v)$

  $\Rightarrow dist_{u,v} = \|x'-y'\|$
- By utility lemma:

  $\|x-y\| \leq (1+4/s)\,\|x'-y'\|$
- Since $s = \varepsilon/4$, $4/s = \varepsilon$
- Therefore:

$$diam(P) = \|x-y\| \leq (1+4/s)\,\|x'-y'\|$$
$$= (1+\varepsilon)\,\|x'-y'\| = (1+\varepsilon)\,dist_{u,v}$$
$$\Rightarrow dist_{u,v} \geq diam(P)/(1+\varepsilon)$$

□

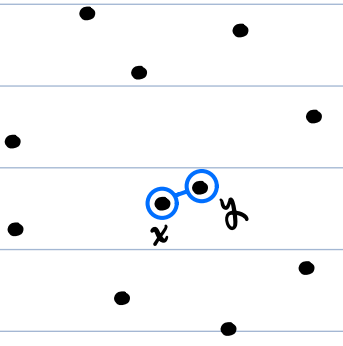
Since alg. outputs max $dist_{u,v}$

$\Rightarrow$ approx. diam $\geq diam(P)/(1+\varepsilon)$ ✓

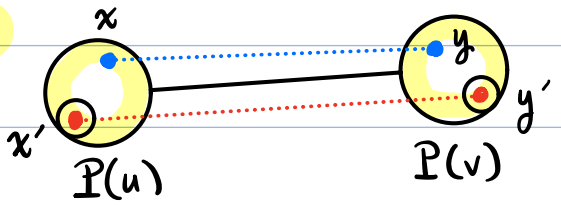
② Closest Pair: (natural complement to diameter)

- Given $P = \{p_1, \ldots, p_n\}$ in $\mathbb{R}^d$, define:

$$\text{closest}(P) = \min_{x,y \in P} \|x-y\|$$



- What if we try same approach?
  - Build s-WSPD
  - Return closest pair of reps
    - Let x, y be true closest
    - Let u, v be WSP
    - $x' = \text{rep}(u)$ $y' = \text{rep}(v)$
    - $\|x'-y'\|$ is a good approx to closest pair



- Something's wrong - $x+y$ cannot be closest because $\|x-x'\| < \|x-y\|$!



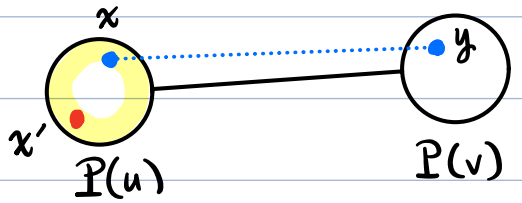
- Idea - set s to isolate closest pair as WSP.

Lemma: If $s > 2$, any s-WSPD has a pair $\{\{x\}, \{y\}\}$ where $x + y$ are the closest pair.

That is, $x + y$ are singletons in WSP

Proof: (by contradiction)

- Let $\{u, v\}$ be WSP such that $x \in P(u)$ + $y \in P(v)$.
  Suppose either set has another point, say $x' \in P(u)$



- By Utility Lemma,

$$\|x - x'\| \leq (2/s) \cdot \|x - y\|$$

$$< \|x - y\| \quad (\text{since } s > 2)$$

$\Rightarrow$ x, y are not closest pair $\rightarrow$ contradiction!

□

Closest Pair (exact) (Given P)

- Set $s > 2$ (eg. $s = 2.01$)
- Build s-WSPD for P ⋯▸ $O(n \log n + s^d n) = O(n \log n)$
- for each WSP $\{u, v\}$ ⋯▸ $O(s^d n) = O(n)$

  $dist_{u,v} \leftarrow \|rep(u) - rep(v)\|$
- return $\min_{u,v} dist_{u,v}$

  (+ rep(u), rep(v) is closest pair)

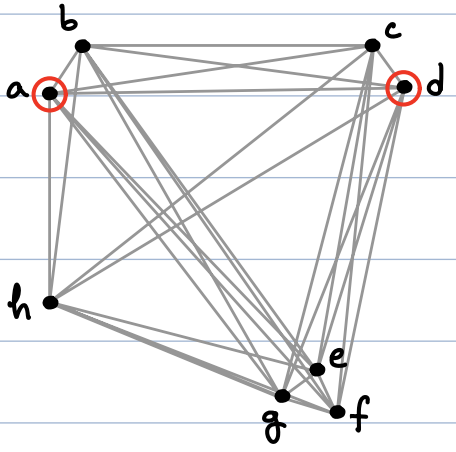
③ Euclidean t-Spanners

Given a point set P in $\mathbb{R}^d$ + $t \geq 1$, a Euclidean t-spanner is a graph G on P with edge weights $w(x,y) = \|x-y\|$, s.t. for all $x, y \in P$

$$\|x-y\| \leq d_G(x,y) \leq t \cdot \|x-y\|$$

where $d_G(x,y)$ is length of shortest x-y path in G

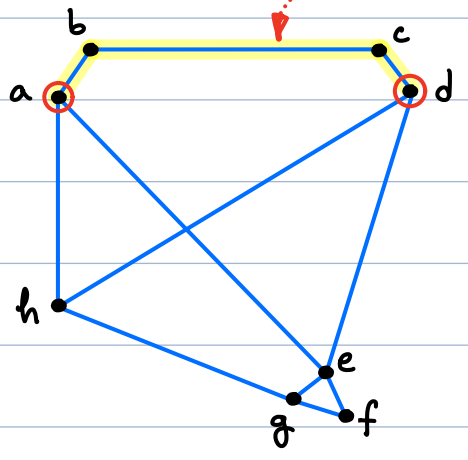
Example:



$t = 1.2$

G:

$d(a,d) \leq t \cdot \|a-d\|$



- Ideally, G has $O(n)$, not $O(n^2)$ edges.

- In $\mathbb{R}^2$, Delaunay Triangulation is a 1.998-spanner

- We'll show how to construct a $(1+\varepsilon)$-spanner [$t = 1+\varepsilon$, for $\varepsilon > 0$] for P with $O(n/\varepsilon^d)$ edges.

Build-Spanner (Given P + $t > 0$)

- set: $s \leftarrow \frac{4(t+1)}{t-1}$
- Initialize G to graph of P with no edges
- Build s-WSPD for P
- for each WSP $\{u, v\}$:
  - add edge $(rep(u), rep(v))$ to G
- return G

Time: Let $t = 1+\varepsilon \Rightarrow s = \frac{4(2+\varepsilon)}{\varepsilon} \leq \frac{12}{\varepsilon} = O(1/\varepsilon)$ $[0 < \varepsilon < 1]$

$\Rightarrow O(n \log n + n/\varepsilon^d)$

Num. edges $= O(s^d n) = O(n/\varepsilon^d)$

Correctness:

- We'll show that for all $x, y \in P$,

  there exists a path of length $\leq t\,\|x-y\|$

  $\Rightarrow d_G(x,y) \leq t \cdot \|x-y\|$.
- Proof by induction on num. of edges in the path.

- **Basis:** Edge (x,y) is in $G \Rightarrow d_G(x,y) = \|x-y\|$ ✓

- **Step:** Let u, v be WSP s.t. $x \in P(u)$, $y \in P(v)$

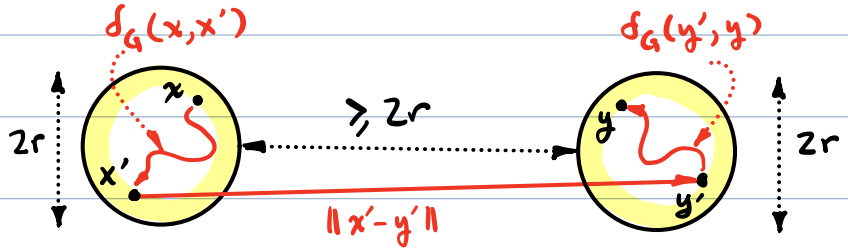
  Let $x' = rep(u)$ $y' = rep(v)$

Note: It is possible that $x = x'$ and/or $y = y'$

There exists a path in G:

$$x \rightsquigarrow x' \longrightarrow y' \rightsquigarrow y$$

$\delta_G(x,x')$ edge (x',y') $\delta_G(y',y)$



By the induction hypothesis:

$$\delta_G(x,x') \leq t\,\|x-x'\|$$
$$\delta_G(y',y) \leq t\,\|y'-y\|$$

$\Rightarrow$ Path length is at most:

$$\delta_G(x,x') + \|x'-y'\| + \delta_G(y',y)$$
$$\leq t\,\|x-x'\| + \|x'-y'\| + t\cdot\|y'-y\|$$
$$= t\left(\|x-x'\| + \|y'-y\|\right) + \|x'-y'\|$$

By WSPD utility lemma:

$$\left.\begin{matrix}\|x-x'\| \\ \|y'-y\|\end{matrix}\right\} \leq \frac{2}{s}\|x,y\|$$

$$\|x'-y'\| \leq \left(1+\tfrac{4}{s}\right)\|x-y\|$$

$\Rightarrow$ Path length is at most:

$$t\left(\left(\tfrac{2}{s}\right)\|x-y\| + \left(\tfrac{2}{s}\right)\|x-y\|\right) + \left(1+\tfrac{4}{s}\right)\|x-y\|$$

$$= \left(\frac{4t}{s} + 1 + \frac{4}{s}\right)\|x-y\|$$

$$= \left(1 + \frac{4(t+1)}{s}\right)\|x-y\|$$

Since $s = 4(t+1)/(t-1)$ this is:

$$= (1 + (t-1))\,\|x-y\|$$

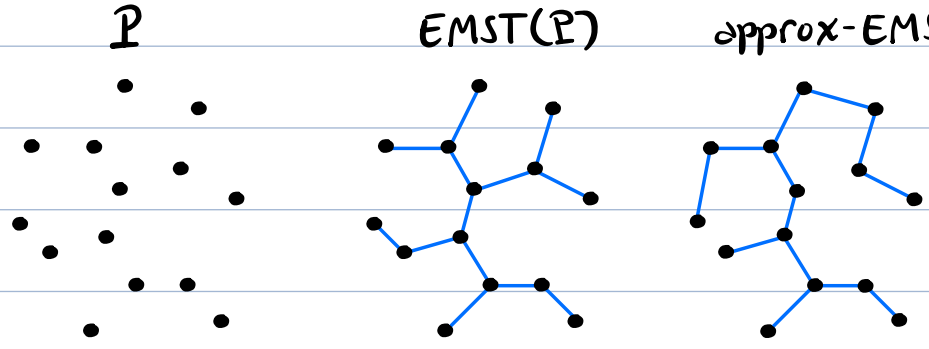
$$= t \cdot \|x-y\| \quad \checkmark \qquad \square$$

Note: This was not a lucky guess. We chose s to make it work.

To obtain a $(1+\varepsilon)$ spanner, set $t = 1+\varepsilon$.

④ Approximate Euclidean Min. Spanning Tree (EMST)

- EMST is the MST of the Euclidean Graph
- In $\mathbb{R}^2$, can compute in $O(n \log n)$ time as MST of Delaunay triangulation.
- In $\mathbb{R}^d$, best is: $O\left(n^{2-2/(\lceil d/2 \rceil + 1)}\right) \approx O(n^2)$ [if d is large]



A spanning tree T on P is an ε-approx. EMST if

$$wt(T) \leq (1+\varepsilon)\, wt(EMST(P))$$

where:

$$wt(T) = \sum_{(x,y)\in T} \|x-y\|$$

Our construction follows from:

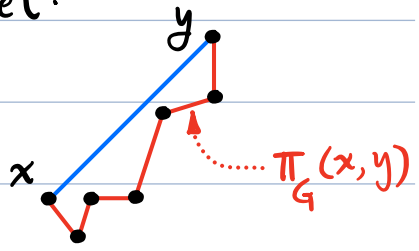
Lemma: If G is a $(1+\varepsilon)$-Euclidean spanner for P, then $MST(G)$ is an ε-EMST for P.

Proof: For each $x, y \in P$, let:

$\pi_G(x,y)$ = edges of shortest path $x \rightsquigarrow y$ in G

Define graph H on P with edge set:

$$\bigcup_{(x,y)\in EMST(P)} \pi_G(x,y)$$



That is, for each edge of (true) EMST, add all edges of its spanner path.

Since G is a $(1+\varepsilon)$-spanner:

$$wt(H) \leq \sum_{x,y\in EMST} (1+\varepsilon)\|x-y\| = (1+\varepsilon) \sum_{x,y\in EMST} \|x-y\|$$

$$= (1+\varepsilon)\, wt(EMST(P))$$

Also observe that H is connected, because it connects all pairs $(x,y) \in EMST$.

Since H is a connected subgraph of G,

$$wt(MST(G)) \leq wt(MST(H))$$

G has more edges to pick from

Combining:

$$\begin{aligned} wt(MST(G)) &\leq wt(MST(H)) \\ &\leq wt(H) \\ &\leq (1+\varepsilon)\, wt(EMST(P)) \end{aligned}$$

$\Rightarrow$ The MST of G is an ε-approx. EMST. □

Thus, we have the following algorithm:

Approx-EMST (Given P and ε)
- $G \leftarrow (1+\varepsilon)$-spanner of P
- return $MST(G)$

Time: $O(n \log n + n/\varepsilon^d)$ to compute spanner

Spanner has $m = n/\varepsilon^d$ edges

MST of a graph with n vert. + m edges computable in $O(n \log n + m)$ time

$\Rightarrow$ Total time: $O(n \log n + n/\varepsilon^d)$

Summary:

- WSPD Utility Lemma (useful for apps.)
- Applications: (all in $\mathbb{R}^d$, $d = O(1)$)
  - $(1+\varepsilon)$ approx to diameter (farthest pair)
    - $O(n \log n + n/\varepsilon^d)$ time
  - (Exact) Closest pair
    - $O(n \log n)$ time
  - $(1+\varepsilon)$ spanner
    - $O(n \log n + n/\varepsilon^d)$ time; $O(n/\varepsilon^d)$ edges
  - $(1+\varepsilon)$ approx EMST
    - $O(n \log n + n/\varepsilon^d)$ time