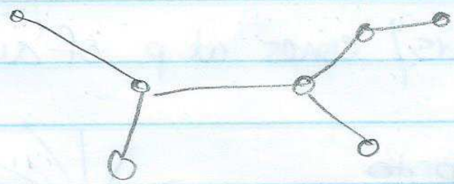
Appl'ns to Other Proximity Problems

Euclidean min spanning tree (EMST)



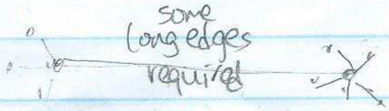
Exact alg'ms: # edges

$O(n^2)$ time since $m = O(n^2)$ $\left\{ \begin{array}{ll} \text{Kruskal} & O(m \log n) \\ \text{Prim} & O(n \log n + m) \\ \text{Karger-Klein-Tarjan} & O(m) \text{ rand} \\ \text{Chazelle '97} & O(m\, \alpha(n)) \text{ det.} \end{array} \right.$

$d = 2$ $O(n \log n)$

$d = 3$ $\tilde{O}(n^{4/3})$

$d > 3$ $\tilde{O}\left(n^{2 - 2/(\lceil d/2 \rceil + 1)}\right)$



idea - find sparse subgraph containing approx. EMST

Def Given n pts P in $\mathbb{R}^d$,

a c-spanner is a subgraph H of complete graph

s.t. $\forall p, q \in P$, $d(p,q) \le d_H(p,q) \le c\, d(p,q)$

Euclid. dist — shortest path dist in H — called stretch factor

EMST alg'm:

Compute c-spanner H with $O(n)$ edges

Compute MST T of H by Prim

$\Rightarrow O(n \log n)$ time

Analysis:

let T^* be the EMST.

Replace each edge $pq \in T^*$ w. shortest path from p to q in H

to get $\hat{T}^*$ (connected, spanning)

$\Rightarrow w(T) \le w(\hat{T}^*) \le c\, w(T^*) \Rightarrow$ factor $\le c$

Spanner Method 1: Yao graph ('82)

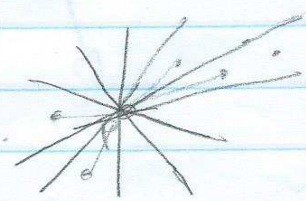
idea - rounding dir

for each $p \in P$

form $O(\frac{1}{\varepsilon^{d-1}})$ dirs/cones at p of angle ε

for each cone

add edge from p to nearest neighbor in cone.

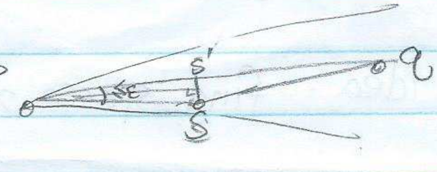


$\Rightarrow O(\frac{1}{\varepsilon^{d-1}} n)$ edges

Analysis:

Fix $p, q \in P$.

Let s be nearest neighbor of p in cone containing q



Know $d(p,s) \leq d(p,q)$

$d(s,q) < d(p,q)$ assuming $\varepsilon < \frac{\pi}{3}$

$$d_H(p,q) \leq d(p,s) + d_H(s,q)$$

$$\leq d(p,s) + c\, d(s,q)$$

by induction on edge length

$$\leq d(p,s) + c\, d(s,s') + c\, d(s',q)$$

$$\leq \left(1 + 2c \sin\frac{\varepsilon}{2}\right) d(p,s') + c\, d(s',q)$$

$$\leq c\, d(p,s') + c\, d(s',q)$$

$= c\, d(p,q)$ by setting $c = \frac{1}{1 - 2\sin \varepsilon/2}$

$= 1 + \Theta(\varepsilon)$.

runtime?

$O(\frac{1}{\varepsilon^{d-1}} n \log^{d-1} n)$ with data structures for range searching (and some modification...)

Spanner Method 2

idea - quadtree

Def (Callahan-Kosaraju '92)

Given n pts P in $\mathbb{R}^d$,

an ε-well-separated pair decomposition (WSPD)

is a collection of pairs $(S_1, T_1), \ldots, (S_m, T_m)$ s.t.

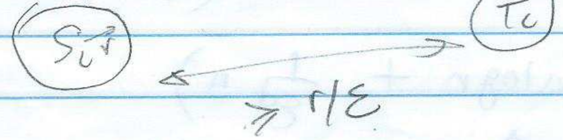
a) $\forall p, q \in P$ $(p \neq q)$,

$\exists$ unique i s.t. $p \in S_i$, $q \in T_i$ or vice versa

b) $\forall i$, S_i and T_i are well-separated

i.e. $\exists$ 2 balls enclos. S_i and T_i of radius r

that are of dist $\geq \frac{r}{\varepsilon}$



(many appl's - bichrom. closest pair, approximating all dists, N-body problems, ...)

note: $\sum^m (|S_i| + |T_i|)$ could be large

Thm $\exists O(\varepsilon)$-WSPD with $m = O(\frac{1}{\varepsilon^d} n)$ pairs.

Pf: idea - quadtree (compressed, but unbalanced)

WSPD(A, B): // given quadtree cells A, B of same side length

  if $|P \cap A| \le 1$ & $|P \cap B| \le 1$, done

  shrink A, B into quadtree cells $\hat{A}$, $\hat{B}$ of same side length r

    (with $P \cap \hat{A} = P \cap A$, $P \cap \hat{B} = P \cap B$)

  if $\hat{A}$, $\hat{B}$ have dist $\ge r/\varepsilon$,

    output pair $(P \cap \hat{A}, P \cap \hat{B})$ & return

  divide $\hat{A}$ into 2^d subcells $\{A_i\}$

    $\hat{B}$ " " " $\{B_j\}$

  for $i = 1, \ldots, 2^d$

    for $j = 1, \ldots, 2^d$

      WSPD(A_i, B_j)

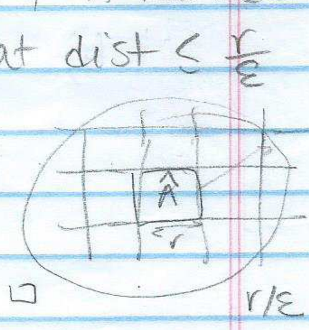
Analysis:

  note that $\hat{A}$ <u>or</u> $\hat{B}$ corresponds to a node in compressed quadtree — $O(n)$ nodes

  fix a node $\hat{A}$ of compressed quadtree, say $\hat{A}$ has side length r

  # quadtree cells $\hat{B}$ of side length r at dist $\le \frac{r}{\varepsilon}$

  $$= O\left(\frac{1}{\varepsilon^d}\right)$$



  $\Rightarrow$ # recursive calls $= O\left(\frac{1}{\varepsilon^d} n\right)$ $\square$

  runtime $O\left(\underbrace{n \log n}_{\text{time to build compressed quadtree}} + \underbrace{\frac{1}{\varepsilon^d} n}_{\text{time for WSPD}}\right)$

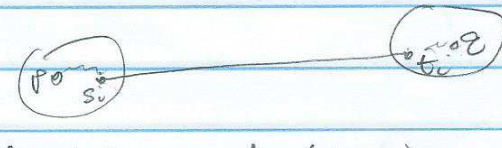
  $\left(T(n) = \sum T(n_i) + O(n - \max n_i)\right)$

RMK: other trade-offs (max deg, hop diameter, total weight, ...)

RMK: for EMST, Arya, Mount '16 $\tilde{O}\left(n \log n + \frac{1}{\varepsilon^2} n\right)$

Obs Let $(S_1, T_1), \ldots, (S_m, T_m)$ be an ε-WSPP.
Pick an arbitrary $s_i \in S_i$, $t_i \in T_i$.
Then $\{s_1 t_1, \ldots, s_m t_m\}$ is a $(1+O(\varepsilon))$-spanner.

Pf. Fix $p, q \in P$.
Say $p \in S_i$, $q \in T_i$.



$$
\begin{aligned}
d_H(p,q) &\le d_H(p, s_i) + d(s_i, t_i) + d_H(t_i, q) \\
&\underset{\text{by induction}}{\le} c\, d(p, s_i) + d(s_i, t_i) + c\, d(t_i, q) \\
&\le (c+1)\, d(p, s_i) + d(p,q) + (c+1)\, d(t_i, q) \\
&\le (1 + 2(c+1)\varepsilon)\, d(p,q) \\
&\le c\, d(p,q) \quad \text{by setting } c = \frac{1+2\varepsilon}{1-2\varepsilon} \\
&\qquad\qquad\qquad\qquad\qquad\quad = 1 + \Theta(\varepsilon)
\end{aligned}
$$

$\square$