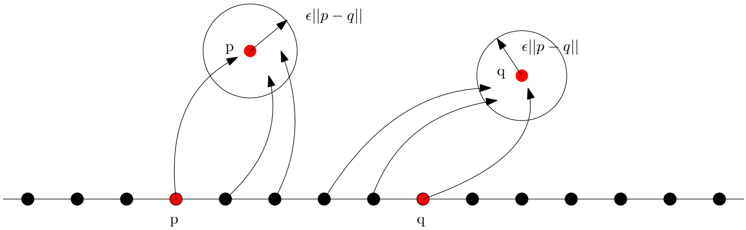
Locality-Sensitive Orderings

Authors: Timothy Chan, Sariel Har-Peled, Mitchell Jones (ITCS 2019)
Presenter: Anil Maheshwari
Carleton University
Ottawa, Canada

What we want to do?

Local Ordering Theorem (CHJ2019)

Consider a unit cube in d-dimensions. For $\epsilon > 0$, there is a family of $O(\frac{1}{\epsilon^d}\log(\frac{1}{\epsilon}))$ orderings of $[0,1)^d$ such that for any $p, q \in [0,1)^d$, there is an ordering in the family where all the points between p and q are within a distance of at most $\epsilon||p-q||_2$ from p or q.



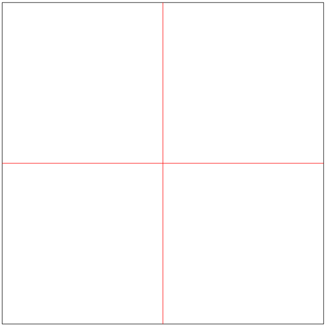
Tools & Techniques

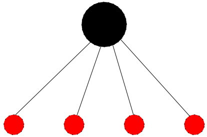
Old & New Concepts

1. Quadtree.
2. Linear orderings of points in a Quadtree.
3. Shifted Quadtrees and ANN.
4. Quadtree as union of ϵ-Quadtrees.
5. (Wonderful) Walecki Construction from 19th Century.
6. Locality-Sensitive Orderings.
7. Applications in ANN, Bi-chromatic ANN, Spanners, ...

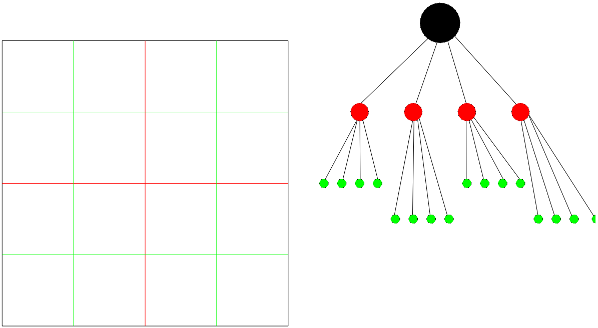
Quadtree of a point set

Quadtree of a point set

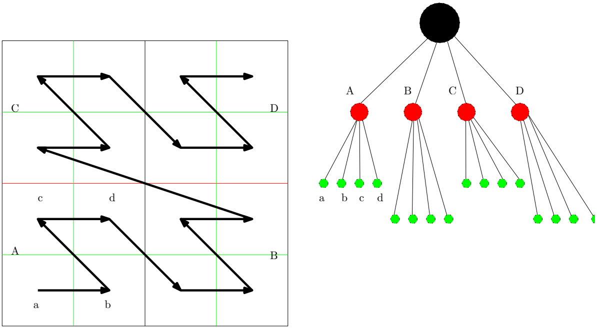




Quadtree of a point set



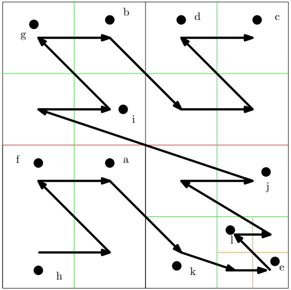
Quadtree of a point set

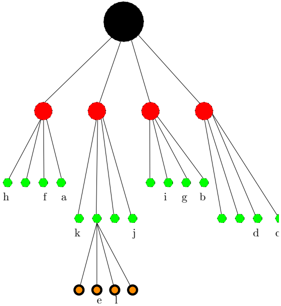


Linear order

DFS traversal of Quadtree

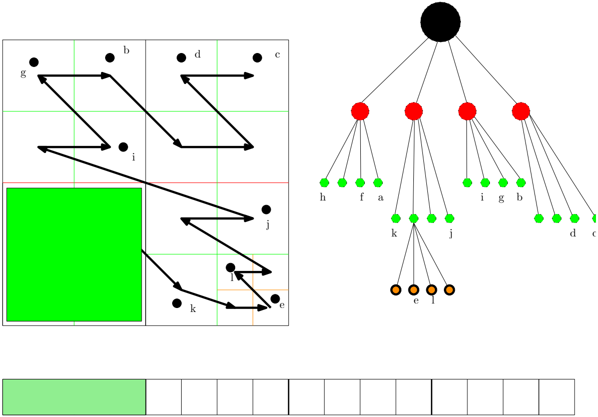
Obtain a linear order of points by performing the DFS traversal of the Quadtree.



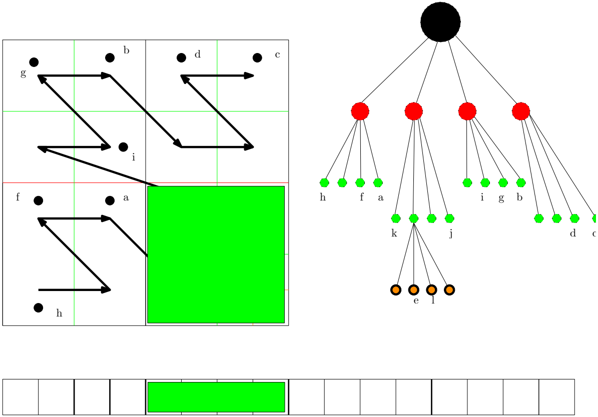


h	f	a	k	e	l	j	i	g	b	d	c

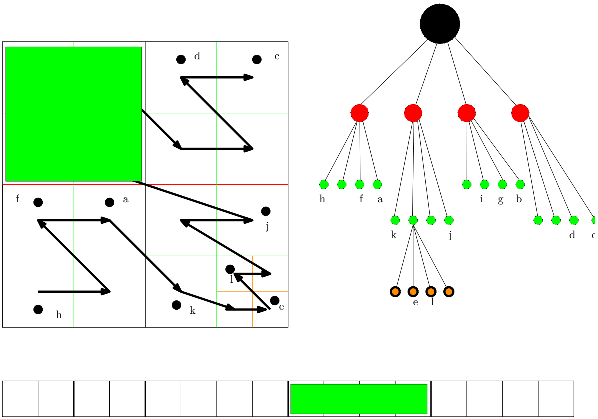
Quadtree Cells & DFS order



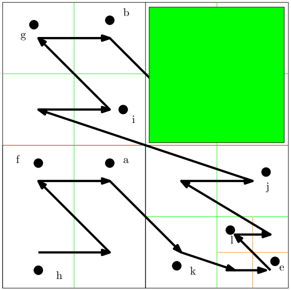
Quadtree Cells & DFS order

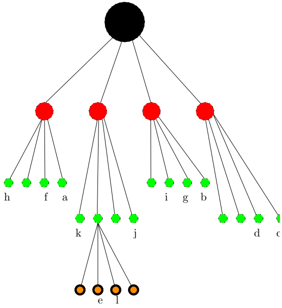


Quadtree Cells & DFS order

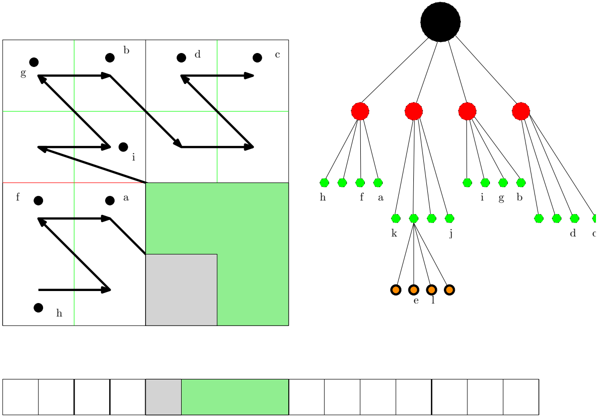


Quadtree Cells & DFS order

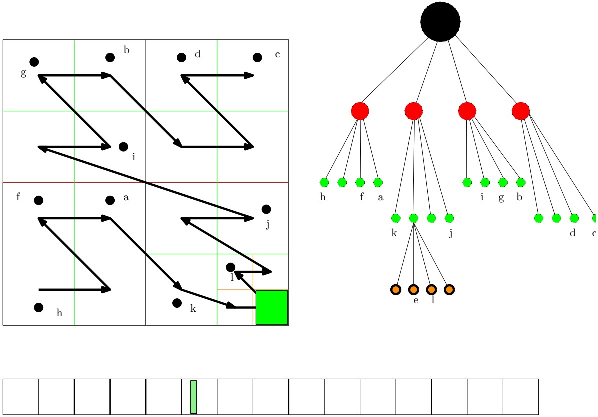




Quadtree Cells & DFS order



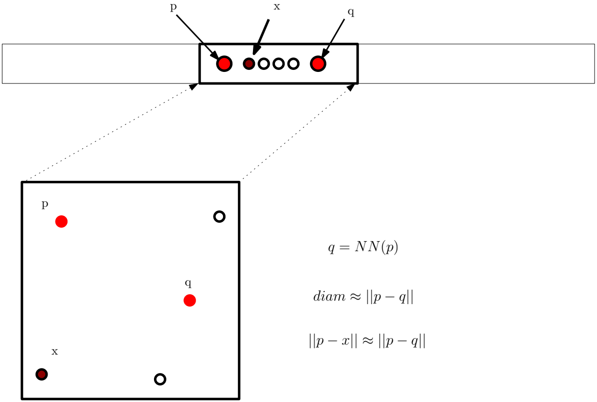
Quadtree Cells & DFS order



Approximate NN from Linear Order

Approximate NN

Let q be nearest-neighbor of p. Assume that there is a cell containing p and q in Quadtree with diameter $\approx ||p - q||$.



$q = NN(p)$

$diam \approx ||p - q||$

$||p - x|| \approx ||p - q||$

Quadtrees of Shifted Point Sets

Assume all points in $P \in [0,1)^d$.
Construct $D = 2\lceil \frac{d}{2} \rceil + 1$ copies of P.

Shifted Point Sets

For $i = 0, \ldots, D$, define shifted point sets
$P_i = \{p_j + (\frac{i}{D+1}, \frac{i}{D+1}, \ldots, \frac{i}{D+1}) | \forall p_j \in P\}$

Let Quadtrees of $P_0, P_1, \ldots, P_D$ be $T_0, T_1, \ldots, T_D$.

Chan (DCG98)

For any pair of points $p, q \in P$, there exists a Quadtree $T \in \{T_0, T_1, \ldots, T_D\}$ such that the cell containing p, q in T has diameter $c||p - q||$ (for some constant $c \geq 1$).

Dynamic ANN

Chan's ANN Algorithm:

1. Construct linear (dfs) order for each of the Quadtrees $T_0, T_1, \ldots, T_D$.
2. For each point p, find its neighbor in each of the linear orders that minimizes the distance.
3. Let q be the neighbor of p with the minimum distance.
4. Report q as the ANN of p.

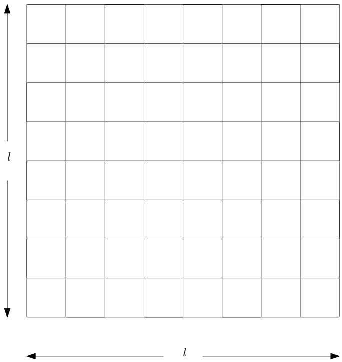
Chan (1998, 2006)

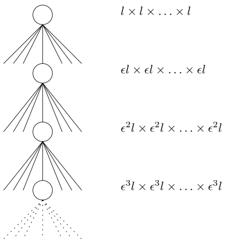
For fixed dimension d, in $O(n \log n)$ preprocessing time and $O(n)$ space, we can find a c-approximate nearest neighbor of any point in P in $O(\log n)$ time ($c = f(d)$).

ϵ-Quadtree

ϵ-Quadtree

For a constant $\epsilon > 0$, recursively partition a cube $[0,1)^d$ evenly into $\frac{1}{\epsilon^d}$ sub-cubes ($\epsilon = 1/2 \implies$ Standard Quadtree).

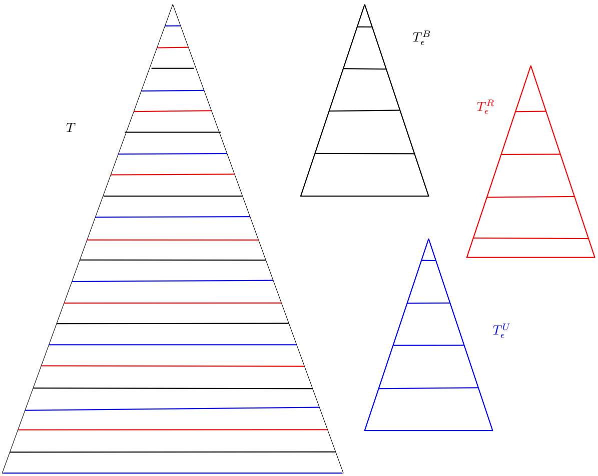




Quadtree as union of ϵ-Quadtrees

Partitioning a Quadtree T into $\log \frac{1}{\epsilon}$ ϵ-Quadtrees

Let $\epsilon = 2^{-3}$. $T = T_\epsilon^B \cup T_\epsilon^R \cup T_\epsilon^U$.



Ordering cells of a node of an ϵ-Quadtree

Let $\epsilon = 2^{-3}$. Any two cells are neighbors in at least one of the 8 orders.

A	B	C	D
E	F	G	H
I	J	K	L
M	N	O	P

ABPCODNEMFLGKHJI

BCADPEOFNGMHLIKJ

CDBEAFPGOHNIMJLK

DECFBGAHPIOJNKML

EFDGCHBIAJPKOLNM

FGEHDICJBKALPMON

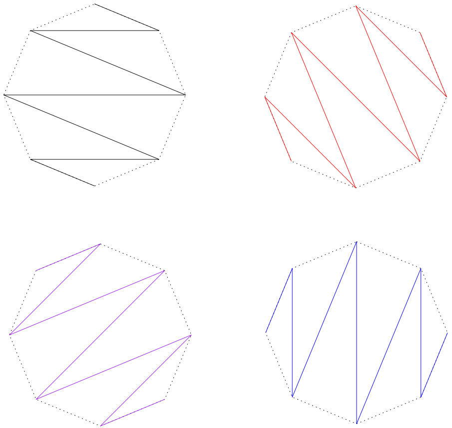
GHFIEJDKCLBMANPO

HIGJFKELDMCNBOAP

(Wonderful) Walecki Result

Walecki Theorem

A complete graph on n vertices can be partitioned into $\lceil \frac{n}{2} \rceil$ Hamiltonian paths.



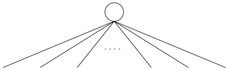
Linear orders of points of $P \in [0,1)^d$

DFS Traversal of an ϵ-Quadtree T_ϵ

1. #children of any node of $T_\epsilon = O(1/\epsilon^d)$.
2. Construct $O(1/\epsilon^d)$ linear orders of cells using Walecki's construction.
3. Generate $O(1/\epsilon^d)$ permutations of points in P by performing DFS traversal of T_ϵ with respect to each linear order.

Structure of Cells

A	B	C	D
E	F	G	H
I	J	K	L
M	N	O	P



A	B	P	C	O	D	N	E	M	F	L	G	K	H	J	I

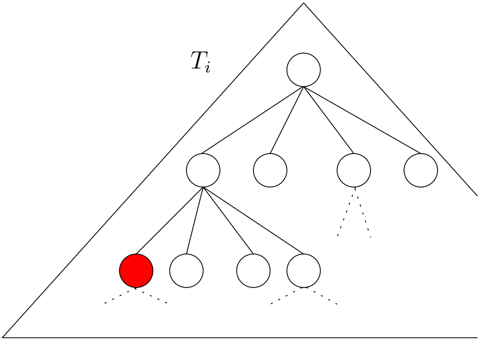
What have we learnt so far?

1. Point set $P \in [0,1)^d$.
2. Shifted points sets $P_0, P_1, \ldots, P_D$ and their Quadtrees $T_0, T_1, \ldots, T_D$.
3. Each Quadtree T_i partitioned into $\log \frac{1}{\epsilon}$ ϵ-Quadtrees.
4. Linear orders of cells of a node in an ϵ-Quadtree.
5. Permutations of points of P obtained from DFS (for each linear order) of ϵ-Quadtrees.
6. Total #Permutations $= O(D \times \log \frac{1}{\epsilon} \times \frac{1}{\epsilon^d}) = O(\frac{1}{\epsilon^d} \log \frac{1}{\epsilon})$.
7. These permutations satisfy “locality” condition.

Locality Property

Locality-Sensitive Orderings

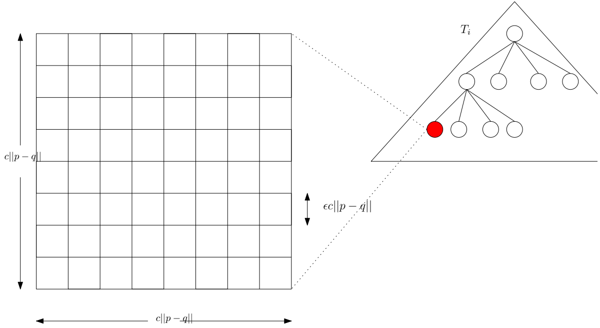
Let the Quadtree $T_i \in \{T_0, T_1, \ldots, T_D\}$ has a cell containing p and q with diameter $\approx ||p - q||$.



Locality Property

Locality-Sensitive Orderings

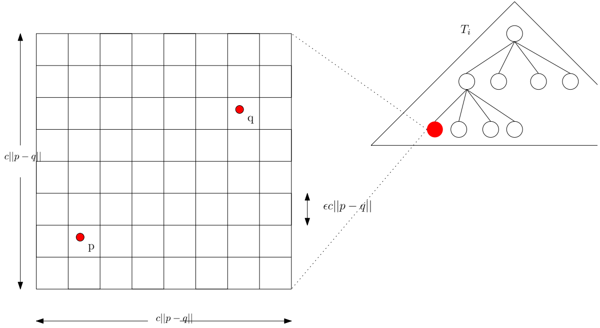
Let the Quadtree $T_i \in \{T_0, T_1, \ldots, T_D\}$ has a cell containing p and q with diameter $\approx ||p - q||$.



Locality Property

Locality-Sensitive Orderings

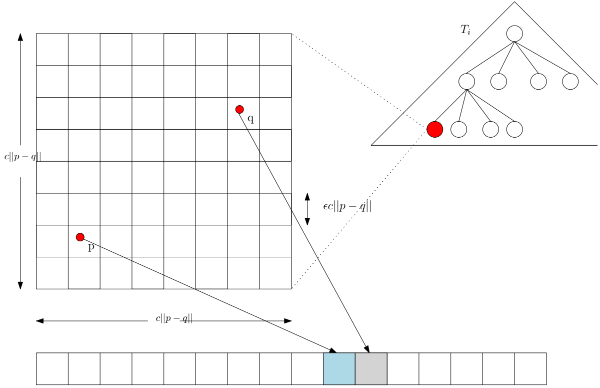
Let the Quadtree $T_i \in \{T_0, T_1, \ldots, T_D\}$ has a cell containing p and q with diameter $\approx ||p - q||$.



Locality Property

Locality-Sensitive Orderings

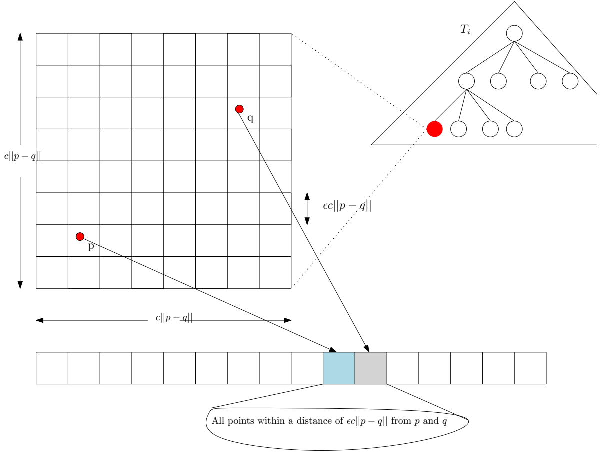
Let the Quadtree $T_i \in \{T_0, T_1, \ldots, T_D\}$ has a cell containing p and q with diameter $\approx ||p - q||$.



Locality Property

Locality-Sensitive Orderings

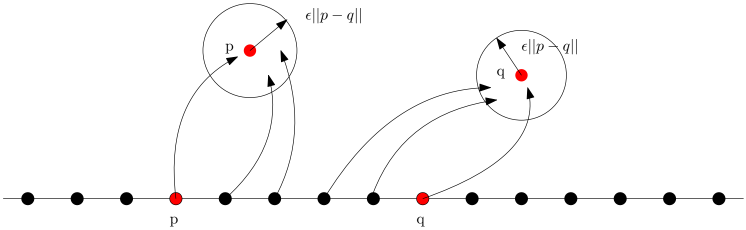
Let the Quadtree $T_i \in \{T_0, T_1, \ldots, T_D\}$ has a cell containing p and q with diameter $\approx ||p - q||$.



Main Theorem

(CHJ 2019)

Consider a unit cube $[0,1)^d$. For $\epsilon > 0$, there is a family of $O(\frac{1}{\epsilon^d}\log(\frac{1}{\epsilon}))$ orderings of $[0,1)^d$ such that for any $p, q \in [0,1)^d$, there is an ordering in the family where all the points between p and q are within a distance of at most $\epsilon||p-q||_2$ from p or q.



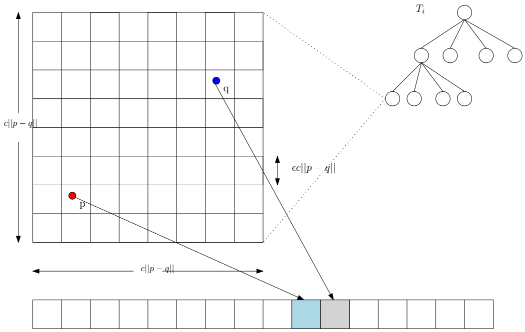
Applications

1. Approximate Bichromatic NN
2. Geometric Spanners
3. (Points) Fault-Tolerant Spanners
4. Approximate EMST
5. Approximate NN
6. Dynamization of all of the above
7. ...

Bichromatic NN

Approximate Bichromatic NN

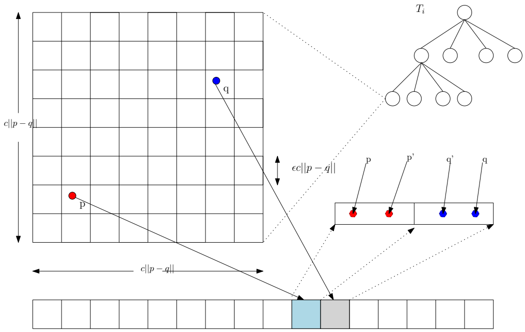
Let p and q constitute a red-blue Nearest Neighbor of the point set.



Bichromatic NN

Approximate Bichromatic NN

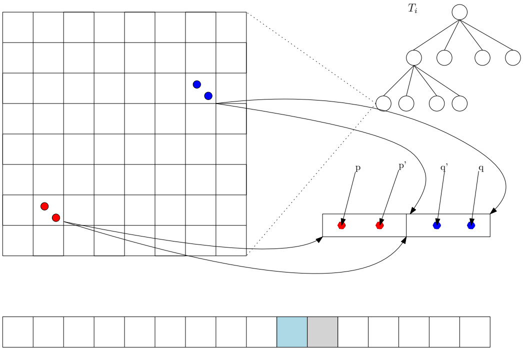
Let p and q constitute a red-blue Nearest Neighbor of the point set.



Bichromatic NN

Approximate Bichromatic NN

Let p and q constitute a red-blue Nearest Neighbor of the point set.



$$||p' - q'|| \leq ||p' - p|| + ||p - q|| + ||q - q'|| \leq (1 + 2\epsilon)||p - q||$$

Bichromatic ANN Algorithm

Input: Bichromatic point set $R \cup B \in [0, 1)^d$.
Output: Bichromatic ANN pair (r, b), $r \in R, b \in B$.

For each of $D = O(d)$ quadtrees of shifted point sets &
For each of the $\log \frac{1}{\epsilon}$ ϵ-quadtrees

1. Construct $O(\frac{1}{\epsilon^d})$ Walecki's orderings.
2. For each ordering, perform DFS traversal of the ϵ-quadtrees, resulting in a permutation of points in P.
3. Among all pairs of consecutive red-blue points in all the permutations, find the pair (r, b) that minimizes $||r - b||$.
4. Report (r, b) as Bichromatic ANN.

Bichromatic ANN

Bichromatic ANN Theorem (CHJ19)

Let R and B be two sets of points in $[0,1)^d$ and let $\epsilon \in (0,1)$ be a parameter. Then one can maintain a $(1+\epsilon)$-approximation to the bichromatic closest pair in $R \times B$ under updates (i.e., insertions and deletions) in $O(\log n \log^2 \frac{1}{\epsilon} / \epsilon^d)$ time per operation, where n is the total number of points in the two sets. The data structure uses $O(n \log \frac{1}{\epsilon} / \epsilon^d)$ space, and at all times maintains a pair of points $r \in R$, $b \in B$, such that $||r-b|| \leq (1+\epsilon) d(R,B)$, where $d(R,B) = \min_{r \in R, b \in B} ||r-b||$.

Variants of linear orders/permutations are used to construct dynamic structures for ANN, Geometric Spanners, Approximate EMST, etc.

Thank-you