Divide and Conquer

Closest Pair of Points

Input unsorted set of points $\{P_1, P_2, P_3, \dots, P_n\} = \{P_i\}_{i=0}^{n}$

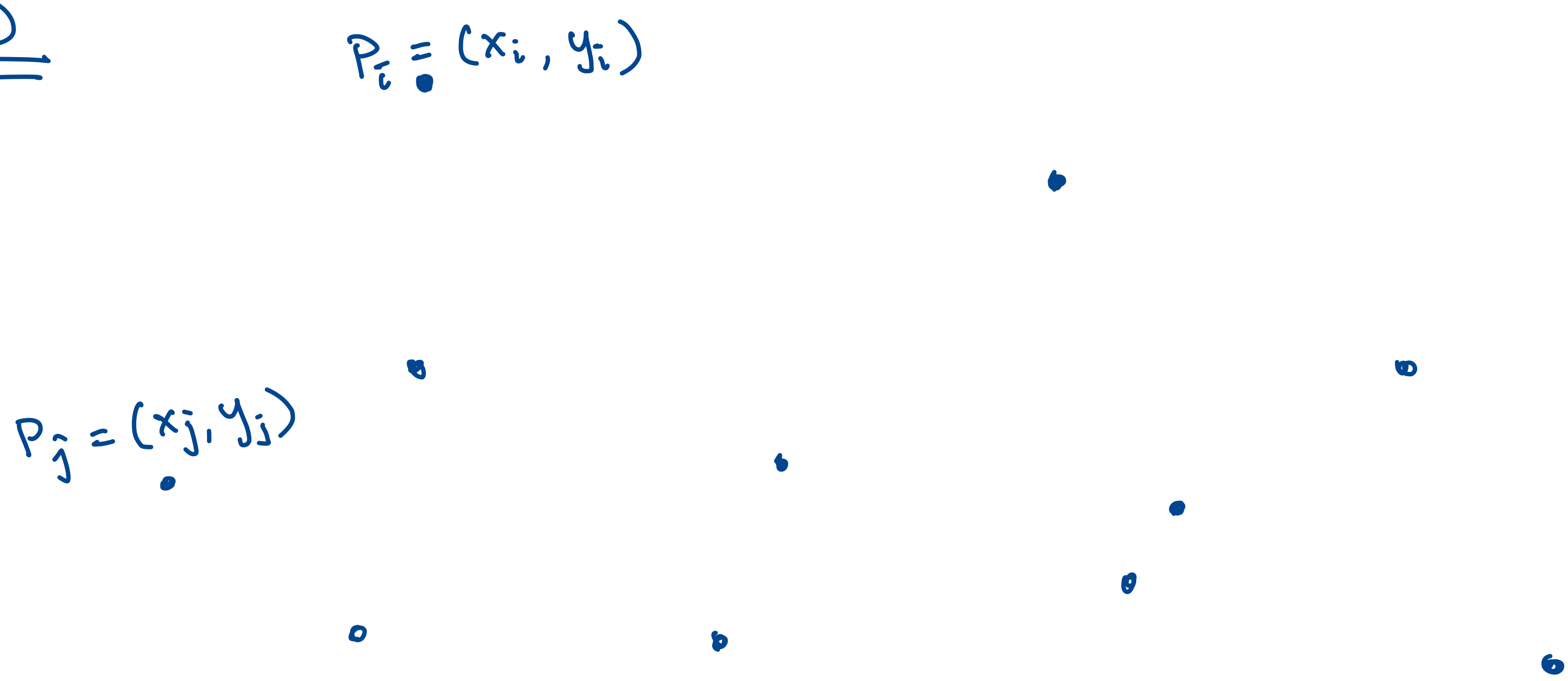
Output the closest pair of points.

1D:

Closest Pair : Sort the points, compare P_i' to P_{i+1}' in sorted list.

Runtime : $O(n \log n)$

2D



$$d(P_i, P_j) = \sqrt{(x_j - x_i)^2 + (y_j - y_i)^2}$$

Return the closest pair.

Naive : Compair $\forall i,j \; d(P_i, P_j)$, Find the smallest.

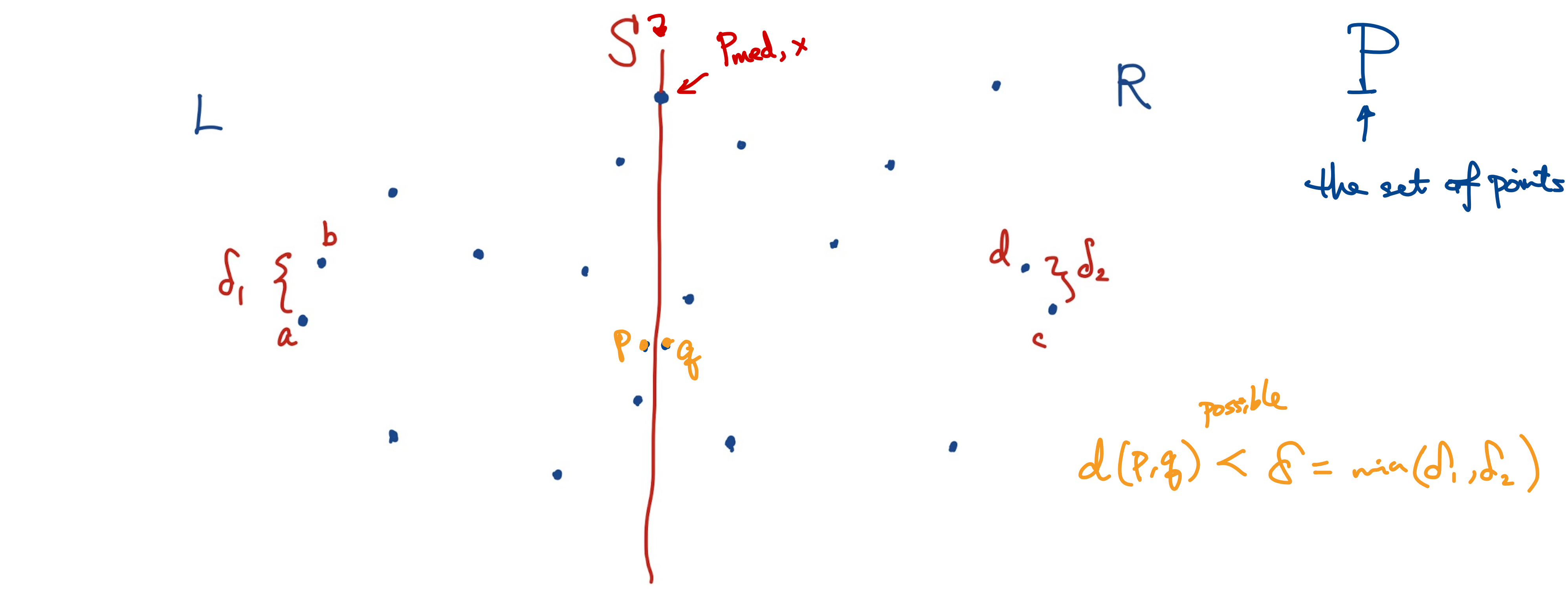
There are $\binom{n}{2}$ pairs, which is roughly $O(n^2)$.

Question : Can we beat $O(n^2)$?

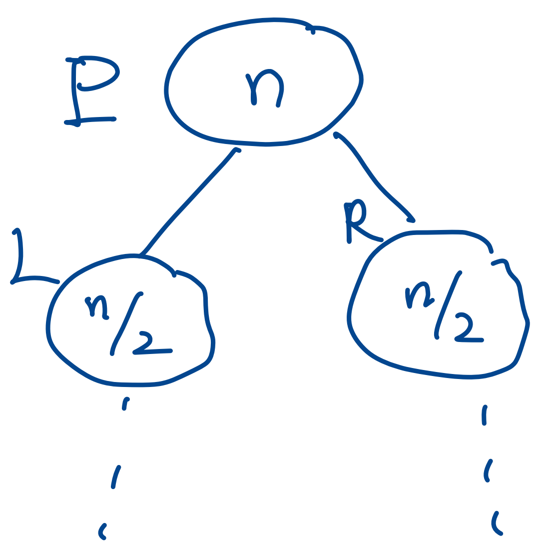
Answer : Yes. $O(n \log n)$ using D & C.

$$T(n) = 2T(n/2) + O(n)$$

Two subproblem with size $n/2$.



* P sorted by x, P_x, and P sorted by y, P_y. $- O(n \log n)$
* Find $P_{med,x}$. Partition P into L and R. $- O(n)$.
* Find the closest pair in L and R recursively.

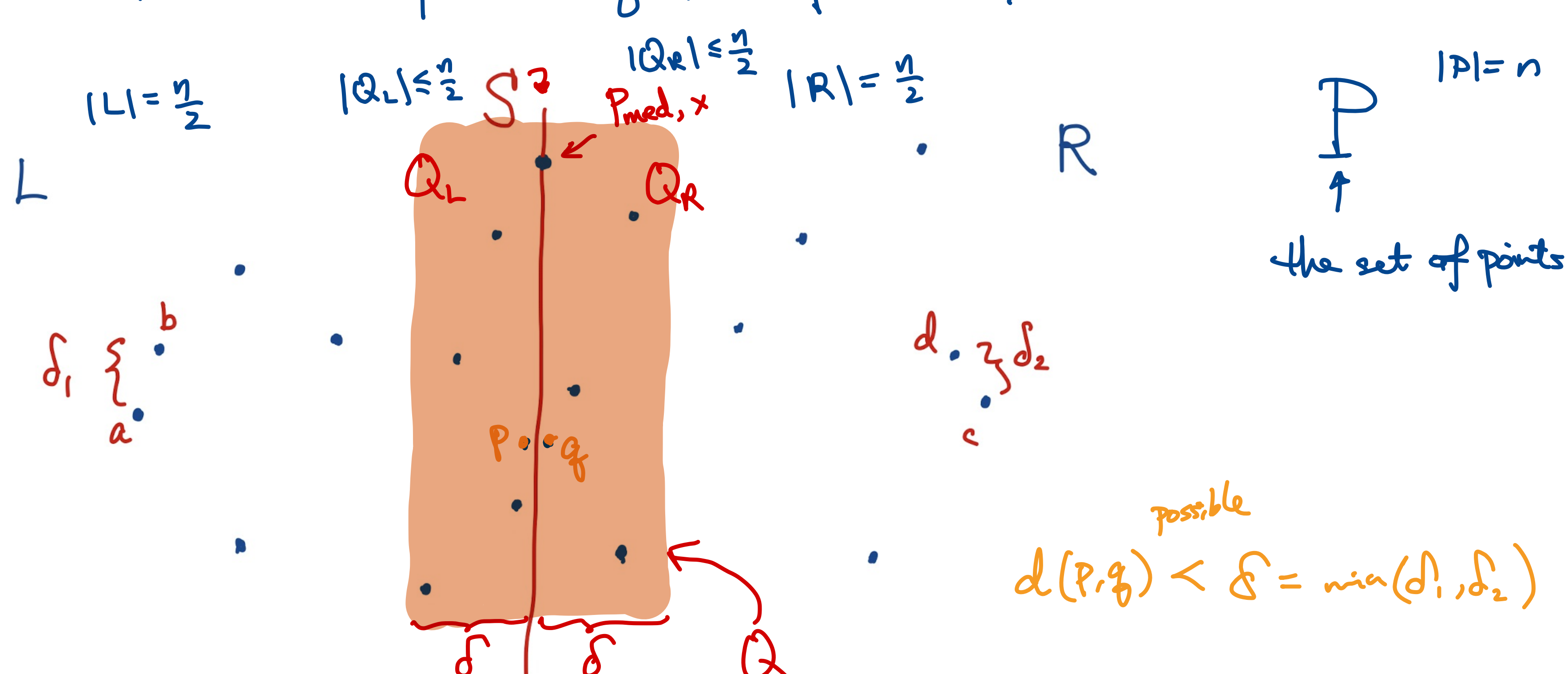


* In L, the pair (a,b) is the closest with distance δ_1.
* In R, the pair (c,d) is the closest with distance δ_2.
* $\delta = \min(\delta_1, \delta_2)$.

We're trying to see if a pair (p,q) exists with $d(p,q) < \delta$.

If it does, $p \in L$ and $q \in R$

- How far can p and q be from S? Answer: At most δ.



$p \in Q_L$ and $q \in Q_r$. Otherwise, $d(p,q) > \delta$.

* Can we check all pairs (p, q) s.t. $p \in Q_L$ and $q \in Q_R$?

No, because we may still have $\sim \underbrace{\left(\frac{n}{2}\right)^2}_{O(n^2)}$ pairs to check.

$\Rightarrow$ We will show that it's enough to compare the distance for any $p \in Q$ to at most 15 other points in Q.

This step takes at most $15n = O(n)$.

$\Rightarrow$ In $O(n)$ time, we have found (p^*, q^*) s.t. $p^* \in L$ and $q^* \in R$ and they are closest left-righ pair.

Let $\delta_3 = d(p^*, q^*)$

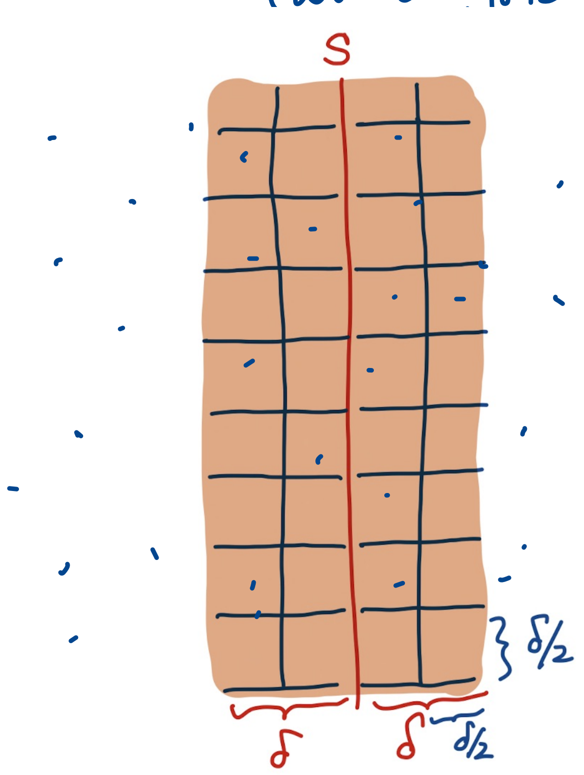
$\Rightarrow$ Output : $\min(\delta_1, \delta_2, \delta_3)$

Recurrsion : $T(n) = T(n/2) + T(n/2) + O(n)$

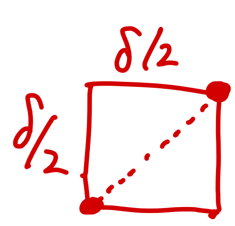
$\Rightarrow$ solves to $O(n \log n)$

NEED TO SHOW HOW TO FIND (p^*, q^*).

* Put a grid of size $\delta/2$ on the Q.



Question: How many points can a box have?

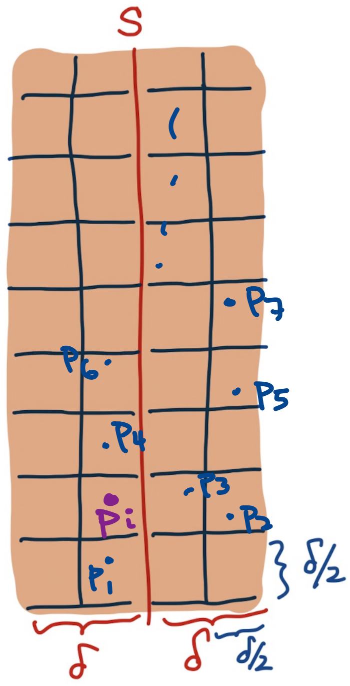


If a box have more than 1 point, their distance is at most $\sqrt{(\delta/2)^2 + (\delta/2)^2} = \sqrt{\frac{\delta^2}{4} + \frac{\delta^2}{4}} = \sqrt{\frac{\delta^2}{2}} = \frac{\delta}{\sqrt{2}} \approx \frac{\delta}{1.414} < \delta$

Any box is either to the left of S or to the right. So, if it has 2 points in it, the distance would be $< \delta$. Then we would have found this in the recursive call.

Obs: Any box has at most one point.

Let P' be the set of points in the strip storted by y coord.



$P' = \{P_1, P_2, \ldots \}$

Claim: We only need to find the distance to the next 15 points in this list.

$$P' = \{ \ldots, P_i, \underbrace{\ldots\ldots}_{15 \text{ points.}}, \underset{16^{th}}{P_j} \}$$

(15 can be reduced)

Reason: Let say q is 16 positions aways from p in P'.

Obs: At most 1 point per box.

$\Rightarrow$ there are at least 3 row betwee p and q.

$$\Rightarrow d(p,q) \geq 3 \cdot \delta/2 = \delta,$$

So (p,q) cannot be the closest pairs