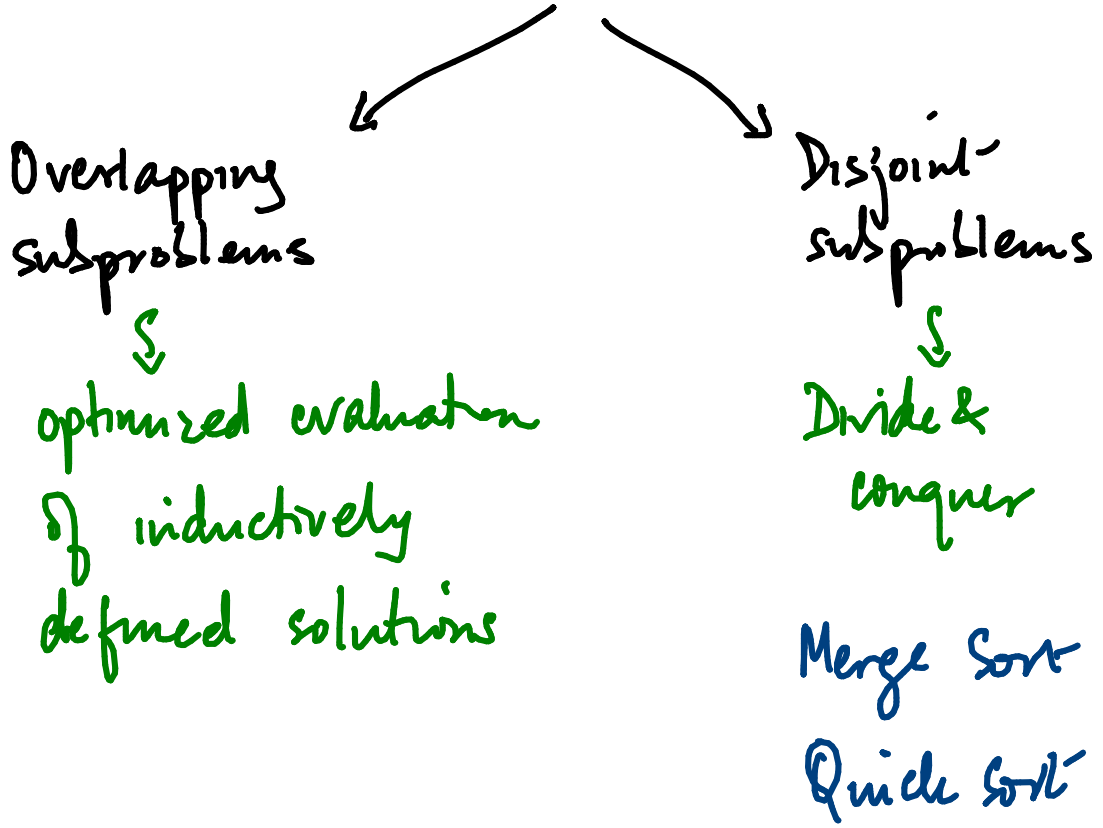
Greedy does not work ⇒ need to consider all subproblems



Recommendation systems / collaborative filtering

Customers like you also bought XYZ

Assume customers rate their purchases

Customers with similar rankings are "alike"

One proposal:

Items $i_1, i_2, \ldots, i_n$

Customer 1 ranks them $r_1 > r_2 \ldots > r_n$

2 " " $s_1 > s_2 \ldots > s_n$

Count how many disagreements on (i_j, i_k)

Simplification:

Take customer 1 as reference: $1, 2, \ldots, n$

customer 2 is a permutation: $\pi_1, \pi_2, \ldots, \pi_n$

Count inversions w.r.t. sorted order

Problem

Given a permutation of $1..n$, count inversions

(i.e. $i > j$ but i appears before j in permutation)

HINT: DIVIDE & CONQUER

Split list into two equal parts

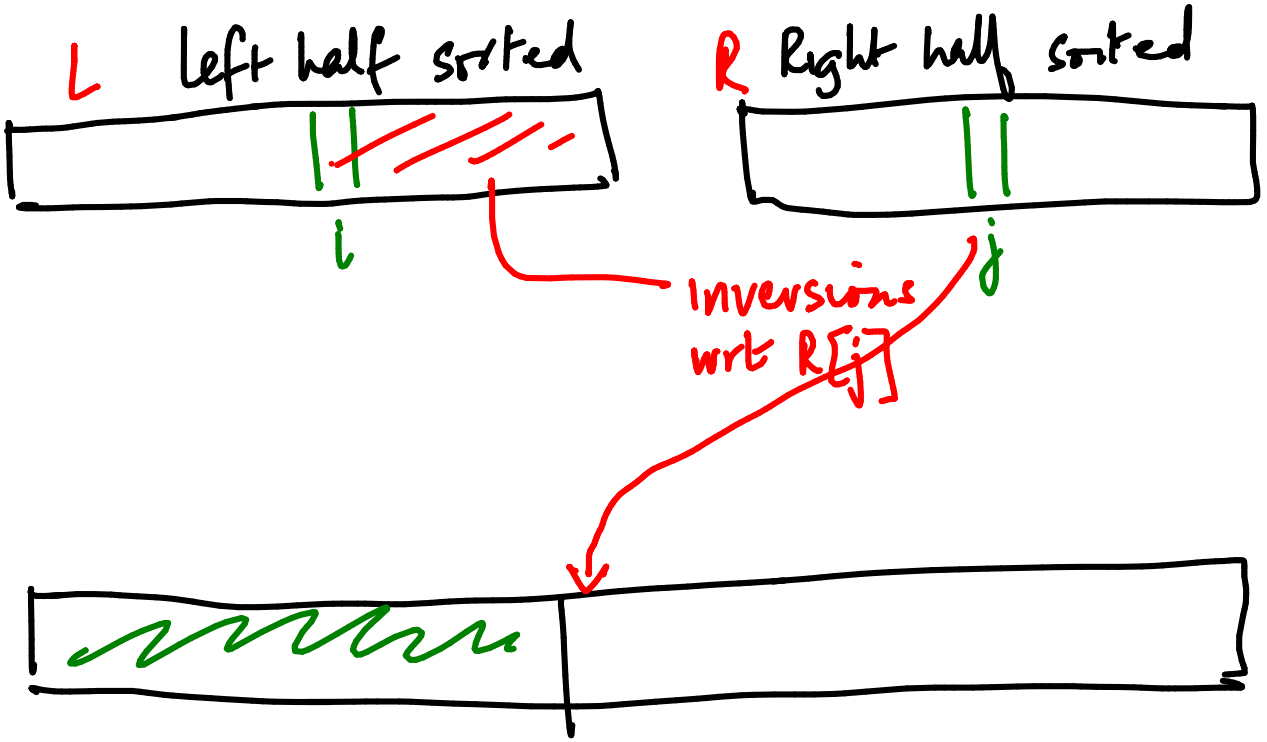
- Count inversions in each half (inductively)
- Count inversions across halves

  i in left, j in right s.t. i > j

Sort using merge sort

Combine step is merge

Each time we add to merged list from right list, we witness as many inversions as current size of left list



Inductively, know inversion counts in L, R

While merging, update count each time we pick an element from R — Overall $O(n \log n)$

Geometry

Given n 2D points $(x_1, y_1), (x_2, y_2) \ldots (x_n, y_n)$,

find the closest pair (usual Euclidean distance $\sqrt{(x_j - x_i)^2 + (y_j - y_i)^2}$)

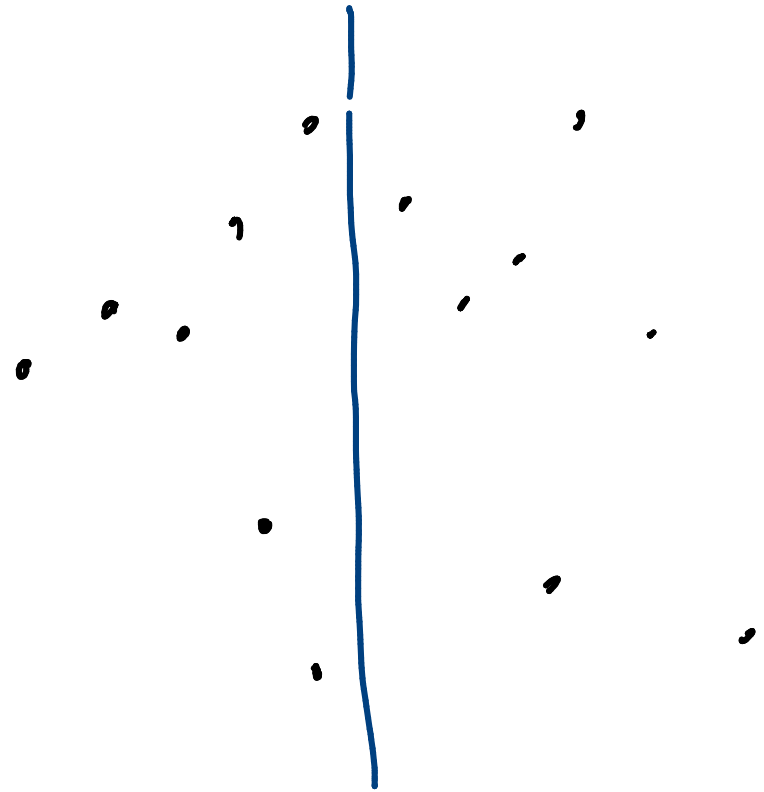
Trivial $O(n^2)$ soln

- Try every pair of points

Divide & Conquer?

Assume all x_i, y_j are distinct

Sort by x-coord & split



Inductively, know closest pair on left & right, respectively

Compute distances across separator

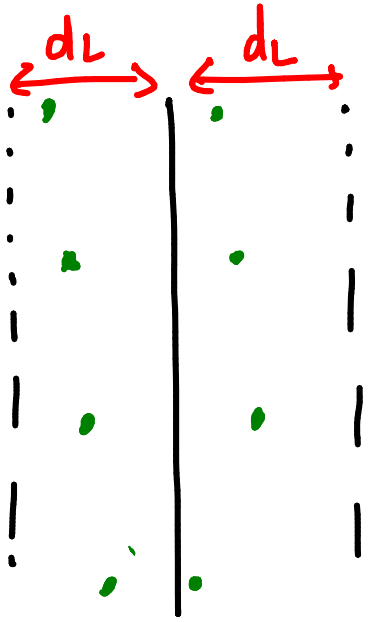
Sufficient to consider a "band"

Band:

Let d_L, d_R be min distances in left, right

Wlog suppose $d_L < d_R$

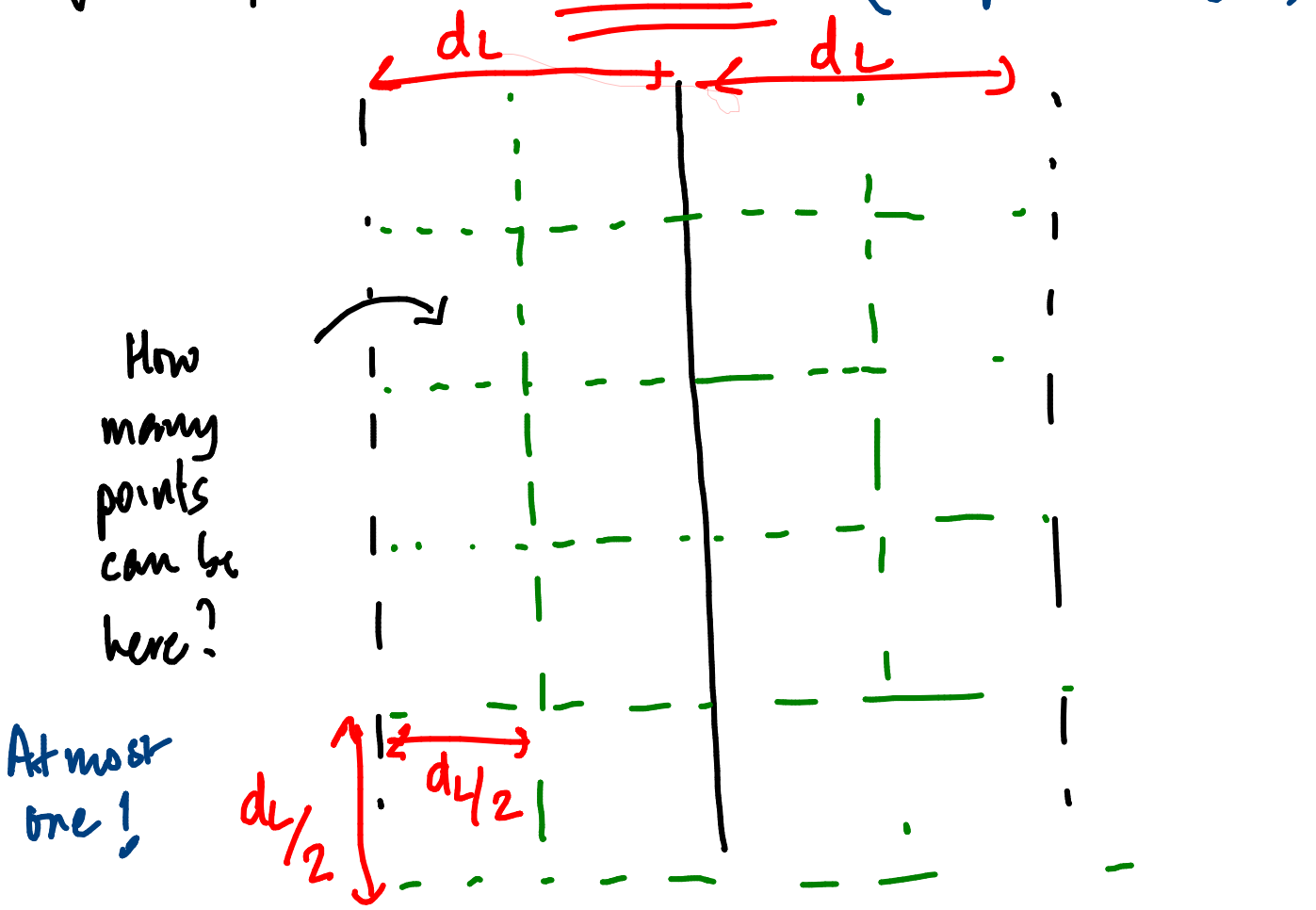
Need not consider points more than d_L away from separator

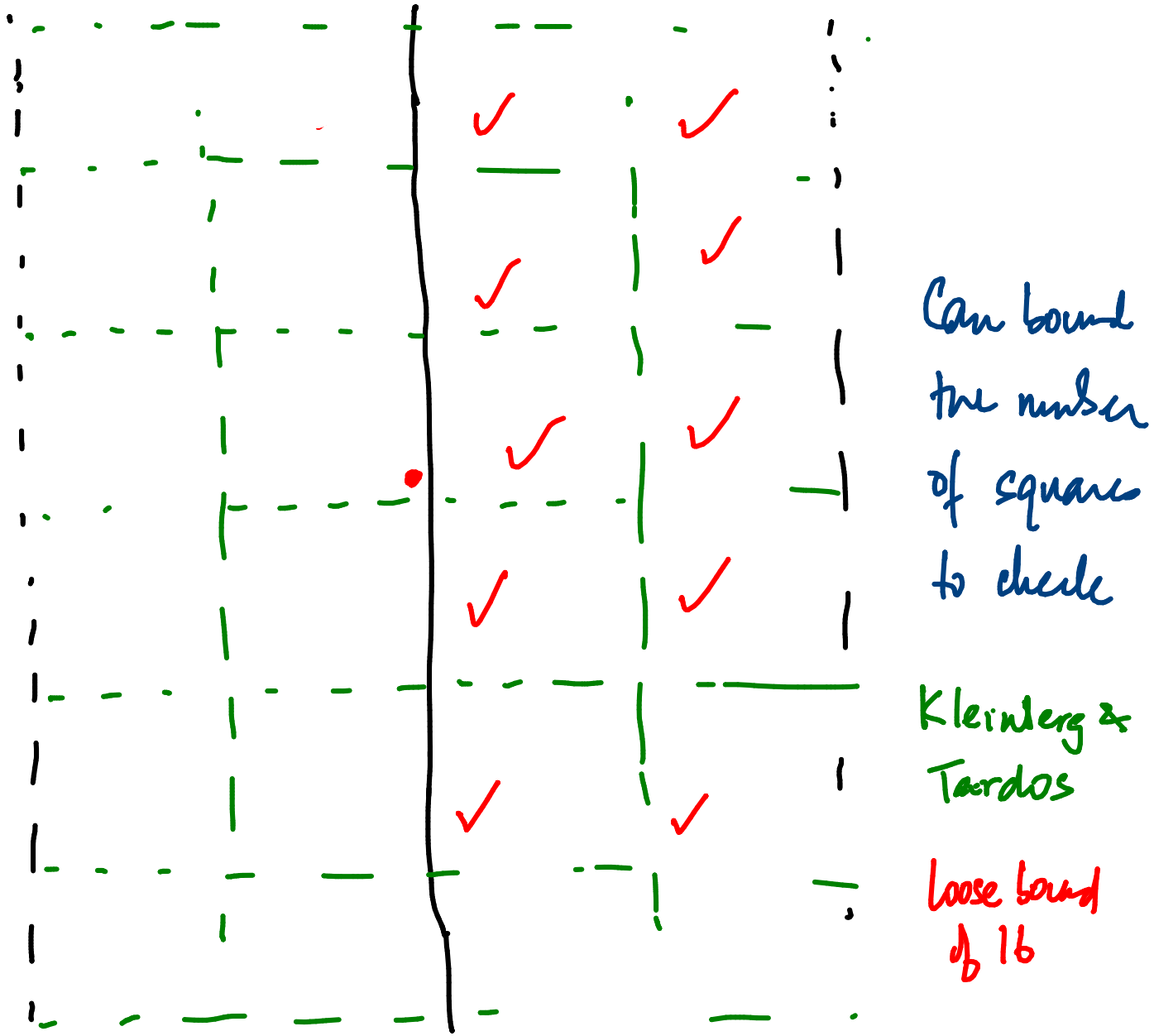


But...

How do we ensure

Within band, limit comparisons made for a given point to a constant (independent of n)




Can bound
the number
of squares
to check
Kleinberg &
Tardos
loose bound
of 16

Merge step - process points in y-order

Originally sorted P $\langle$ P_x, P_y