



Fast k -Nearest Neighbour Search via Prioritized DCI

Ke Li

{ke.li, malik}@eecs.berkeley.edu

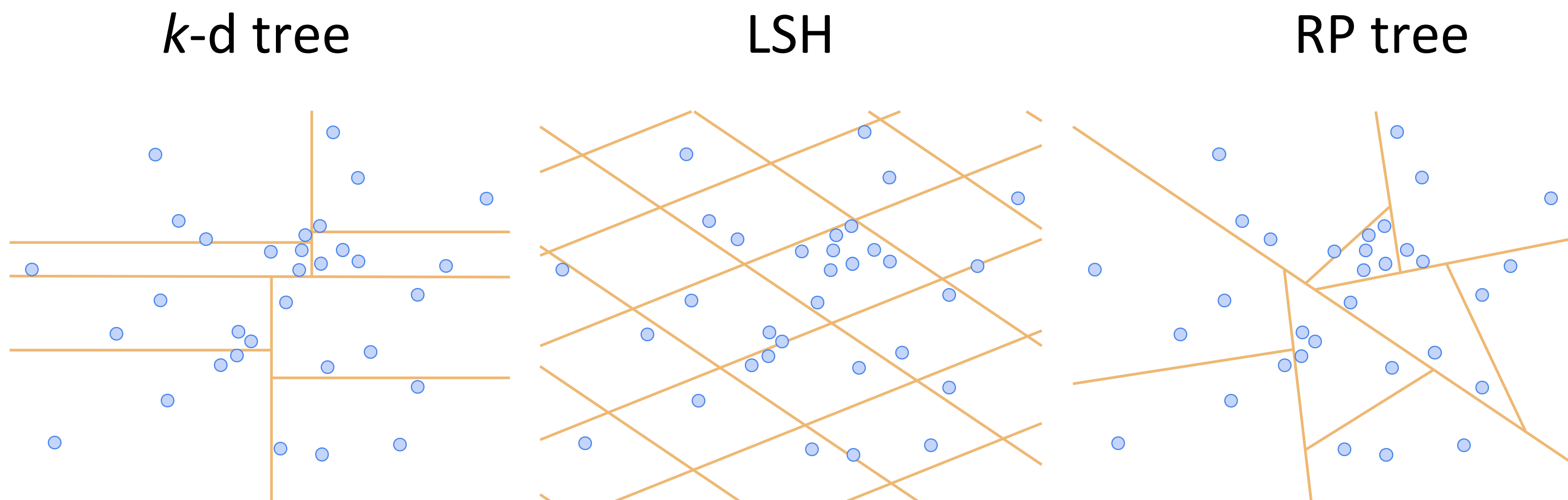
Jitendra Malik

Introduction

- Existing methods suffer from the curse of dimensionality.
- Most existing methods rely on a divide-and-conquer strategy known as space partitioning.
- We present a new algorithm that overcomes the curse of dimensionality, which has:
 - Time complexity: linear in ambient dimensionality, sub-linear in intrinsic dimensionality and size of the dataset.
 - Space complexity: independent of ambient dimensionality and linear in size of the dataset.

The Case Against Space Partitioning

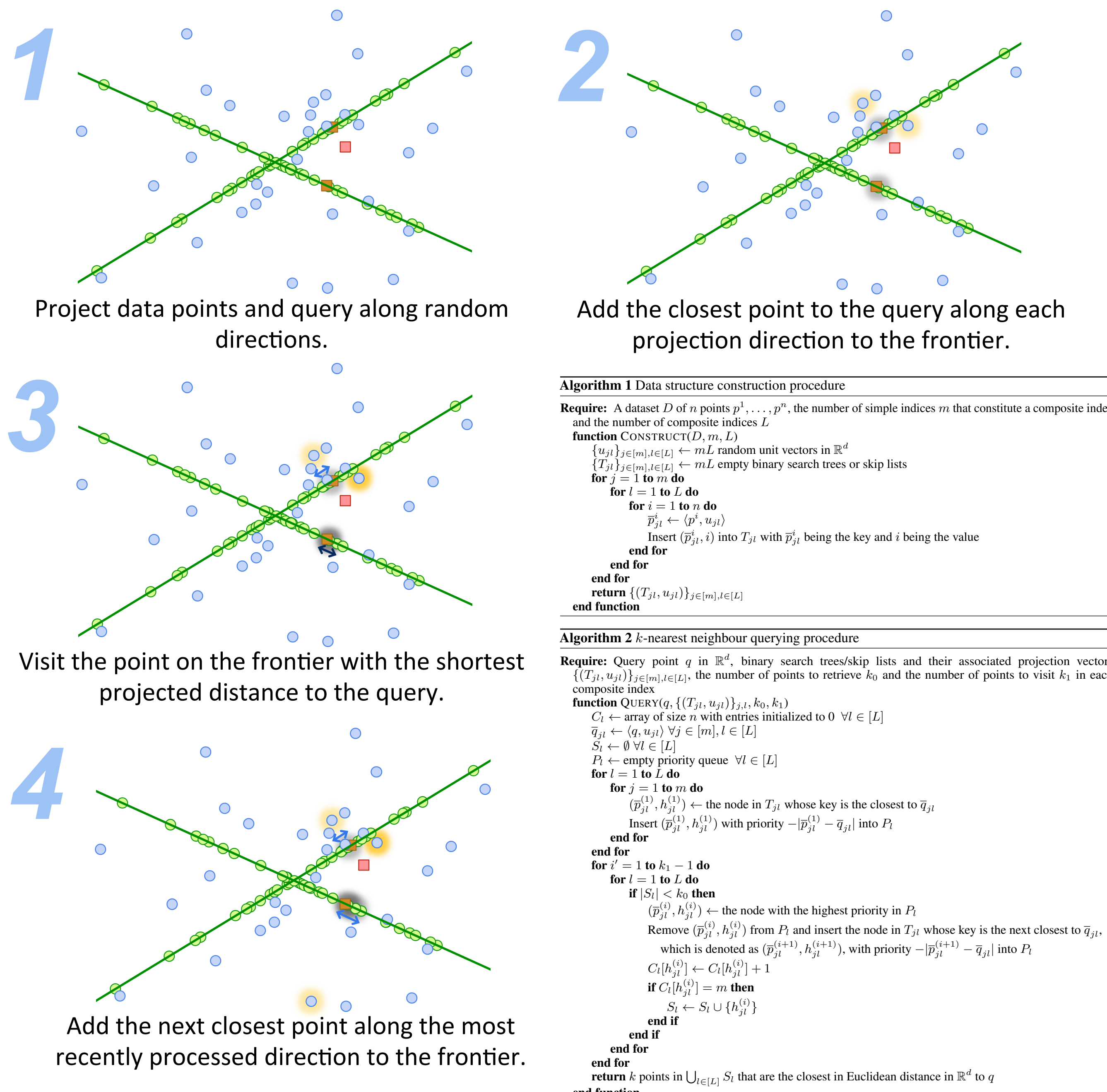
- Most existing methods rely on space partitioning:



- Problems:

- As dimensionality increases, volume of space grows exponentially \Rightarrow either the number or the size of cells must grow exponentially.
- “Field of view” is limited to the cell containing the query; algorithm unaware of points in adjacent cells.
- As dimensionality increases, surface area grows faster than volume \Rightarrow points likely to be near cell boundaries.
- Choosing good partitioning is non-trivial. Once chosen, cannot adapt to changes in data density.

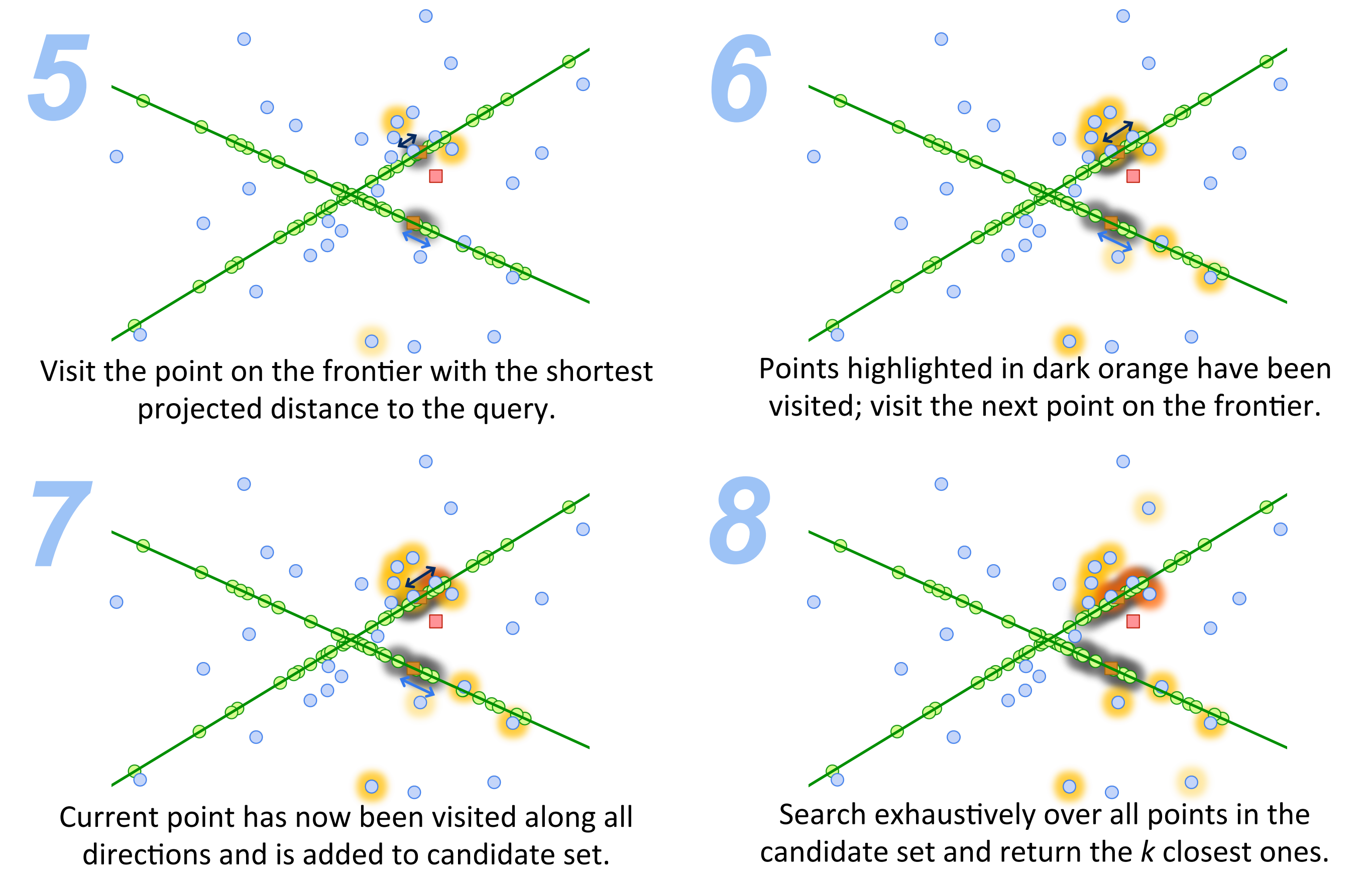
Algorithm



Questions?

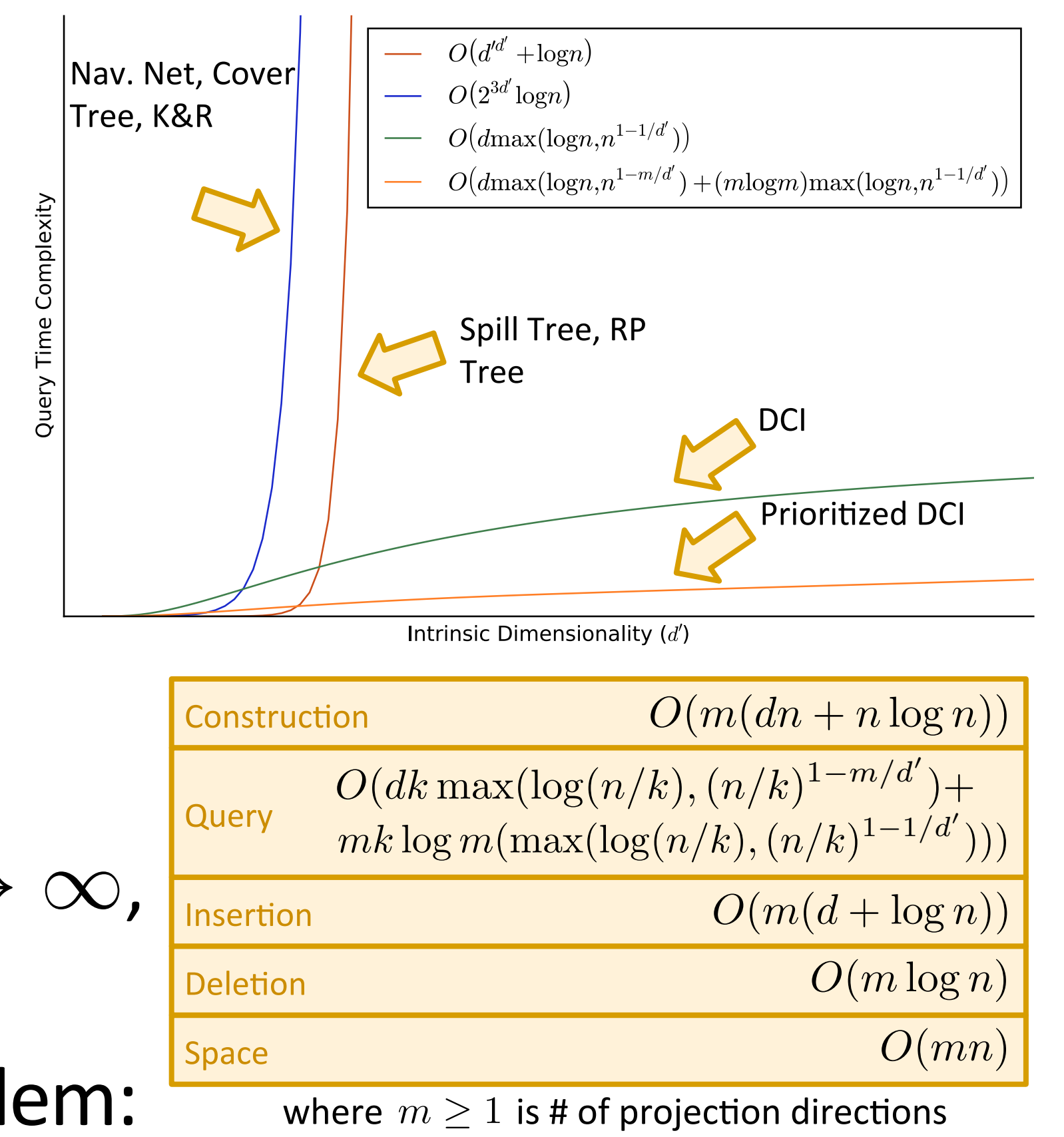
Sorry if we (the authors) aren't around – we have another poster on “Implicit Maximum Likelihood Estimation” at this workshop and are probably at that poster. Feel free to come by our other poster (which is probably next to this poster) and ask us about this poster.

Algorithm (cont'd)



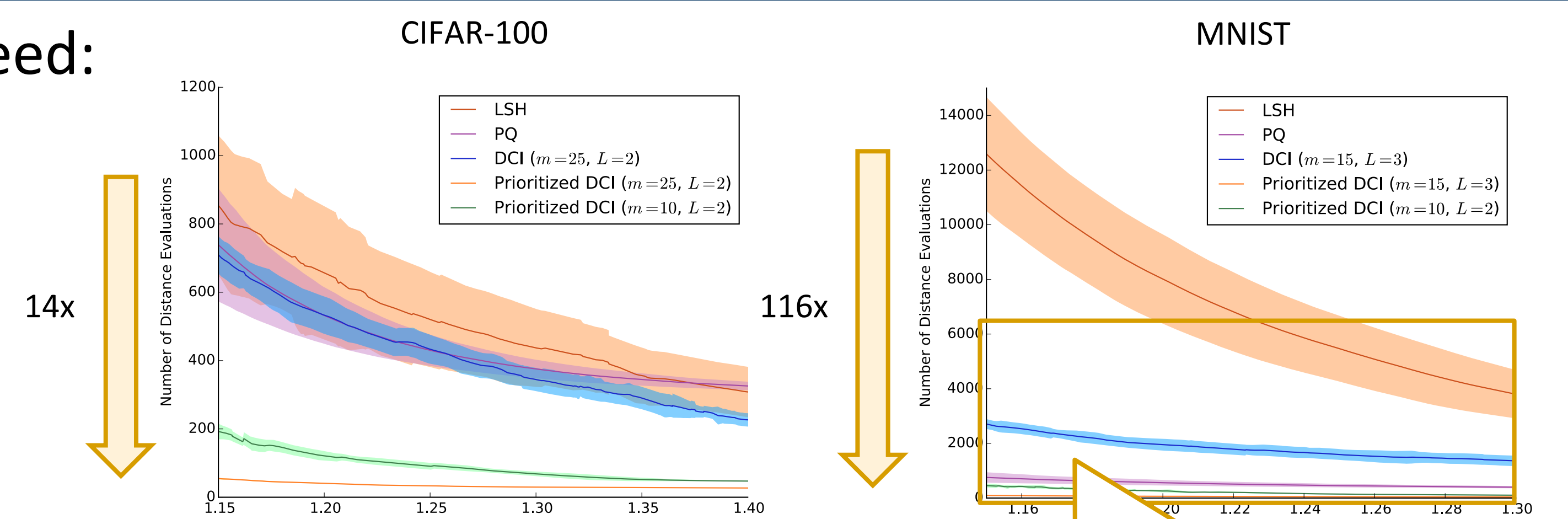
Complexity

- Intuition:
 - Points are added to the candidate set in the order of maximum projected distance to the query.
 - Maximum projected distance is a lower bound on the true distance.
 - As # of projection directions $\rightarrow \infty$, this \rightarrow true distance.
- Lower Bound for Related Problem:
 - For k -nearest neighbour search along a direction (Chazelle, 1989): $\Omega(n^{1-1/d'} / \log n)$

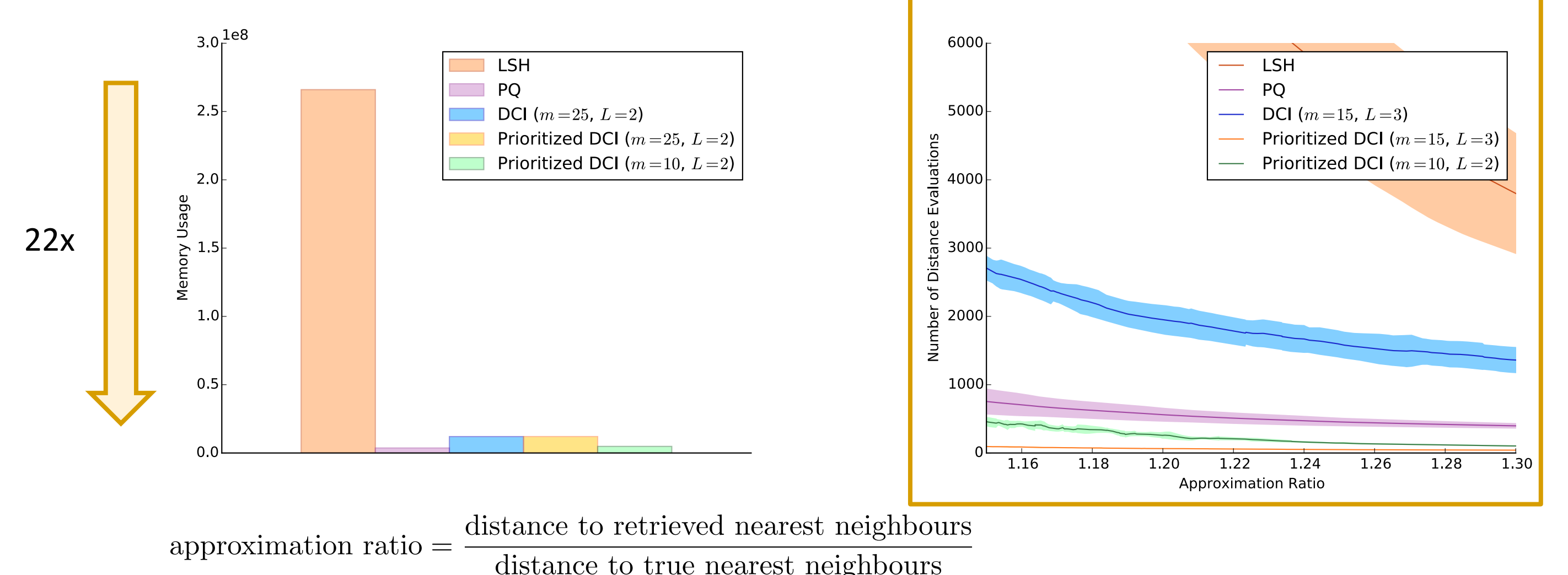


Experiments

Speed:



Memory Usage:



Code available at <https://people.eecs.berkeley.edu/~ke.li/>