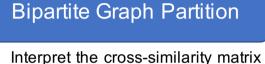




Divide-and-conquer based Landmark Selection Divide the landmark selection into small •

Approximate Similarity Matrix Construction

Find the *K*-nearest landmarks



sub-optimization problems

Construct cross-similarity matrix between data points and K -

as a bipartite graph

Find a low dimensional embedding

via a small eigen-problem

Apply k-means on the embedding

to obtain clustering results.

Use light-k-means for large subset Use *k*-means for large subset nearestlandmarks Set the subset centers as landmarks