3D structures

To communities

Algorithm

1. xyz to weighted graphs via UMAP
2. Apply a community algorithm on each graph realization
3. Build the frequency matrix:
   a. p_{ij} = # (ij are in the same community) / #(of realizations)
4. NMF to simplify the P matrix
5. Community algorithm on P to call structures as communities on it?

Algorithm

- Points in 3d can be mapped to a graph whose edges encode proximities
- The graph can be partitioned into clusters of nodes, called communities, representing group of highly connected nodes
- An ensemble of graphs can then be used to study recurrent patterns of the communities

Analysis of community structures: frequency matrix

- Build a frequency matrix P:
    - P_{ij} = the numb of times nodes (i,j) occur in the same community in the ensemble of structure
- Consider P as an in-silico HiC matrix and used it to identify typical structures and their fluctuations (these are not accessible from the real HiC matrix)

Analysis of community structures: NMF

- Consider the triplet (i,j,s) of nodes/loci (i,j) in the s-th structure realization
- Construct a 2d array $M_{ij;s}$ with rows labeled by the pairs (i,j) and columns labeled by s
- Decompose M using non-negative matrix factorization:
    - $M_{ij;s} = \sum_k N_{ij;k} H_{k;s}$
    - Each column of N should be thought of as an basic state of the structures
    - Each row of H gives the weights to use in the mixture of basic states to get back M