



Supplementary Fig. 1. Density estimates for the S-RNase-w (A) and SFB-w (B) data sets based on the simulated data, for the number of clusters, average cluster size, maximum cluster size, and average distance between words within clusters (average jumps). Red line - value observed in the empirical data set.