

## **Spectral Clustering in the Gaussian Mixture Block Model**

**Shuangping Li**

**Department of Statistics, Stanford University, USA**

**Email: [fifalsp@stanford.edu](mailto:fifalsp@stanford.edu)**

Gaussian mixture block models are distributions over graphs that strive to model modern networks: to generate a graph from such a model, we associate each vertex with a latent feature vector sampled from a mixture of Gaussians, and we add edge if and only if the feature vectors are sufficiently similar. The different components of the Gaussian mixture represent the fact that there may be different types of nodes with different distributions over features—for example, in a social network each component represents the different attributes of a distinct community. Natural algorithmic tasks associated with these networks are embedding (recovering the latent feature vectors) and clustering (grouping nodes by their mixture component).

In this talk, we focus on clustering and embedding graphs sampled from high-dimensional Gaussian mixture block models, where the dimension of the latent feature vectors goes to infinity as the size of the network goes to infinity. This high-dimensional setting is most appropriate in the context of modern networks, in which we think of the latent feature space as being high-dimensional. We analyze the performance of canonical spectral clustering and embedding algorithms for such graphs in the case of 2-component spherical Gaussian mixtures and begin to sketch out the information-computation landscape for clustering and embedding in these models.

This is based on joint work with Tselil Schramm.