

Hierarchical Latent Space Models for Social Network Analysis

Alex Loewi^{1,4}, Francisco Ralston², Octavio Mesner^{2,3,4}

¹Public Policy, ²Engineering & Public Policy, ³Statistics & Data Science, ⁴Machine Learning

Introduction

Multigraphs: “Parallel” network models sharing identical nodes but different edges connecting nodes, used to model data where actors (nodes) maintain varying types of relationships.

Social Network Analysis (SNA), where nodes represent individual people and edges represent differing relationships, such as friendship, professional colleague, or facebook connected, for example.

Problem: Social science and network analysis rely on intuitive and interpretable models. Visualizing the similarities between multiple related graphs is challenging.

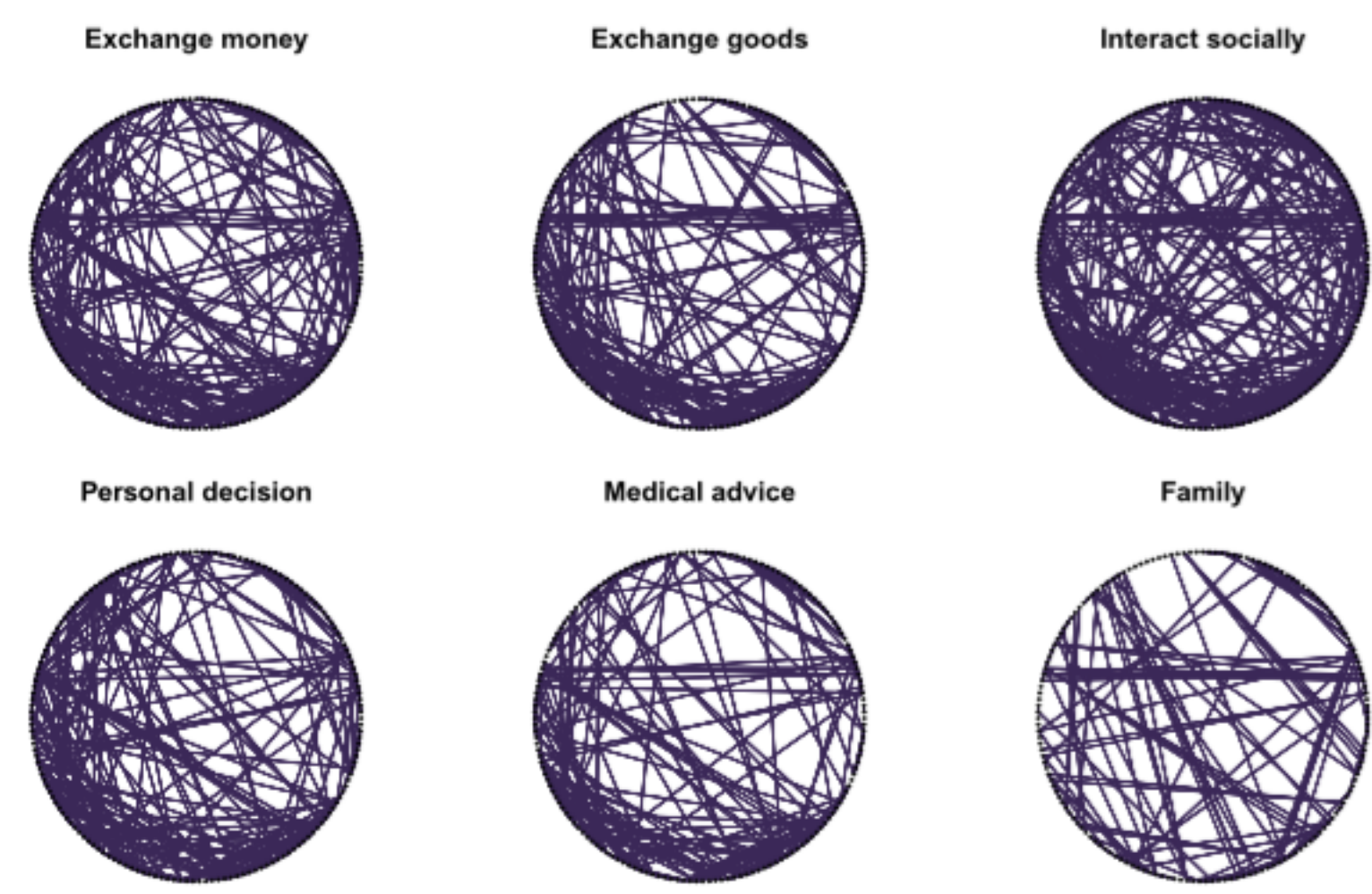


Figure 1: (Source: Salter-Townshend, Michael, and Tyler H. McCormick)

Solution: Using a LASSO penalty, we aim to force similar graphs to be displayed in comparable, visualizable, ways

Employee Relationships Data

A recently conducted study at a charter school in West Baltimore collected a five-layer valued network between the approximately eighty employees of the school:

- Frequency of interaction
- Discussing Academics
- Discussing Behavior
- Social interaction
- Professionally Helpful Relationship

While empirically these layers are found to have very high correlations, modeling them in a way that demonstrates those correlations is challenging.

Latent Space Projection

- Map nodes, $i \mapsto z_i \in \mathbb{R}^n$ where proximity, $\|z_i - z_j\|_2 = d_{ij} < 1$, indicates nodes are connected (and not connected otherwise)
- Intuitively captures reciprocity ($j \rightarrow i \Rightarrow i \rightarrow j$) and transitivity ($i \rightarrow j, j \rightarrow k \Rightarrow i \rightarrow k$)
- Edge probability:
 $\sigma_{ij} = \mathbb{P}(Y_{ij} = 1 | z_i, z_j) = \text{logit}^{-1}(\alpha + \|z_i - z_j\|_2^2)$
where $Y_{ij} = 1$ indicates i, j are connected in the data, find z
- Likelihood: $\prod_{i < j} \sigma_{ij}^{y_{ij}} (1 - \sigma_{ij})^{1 - y_{ij}}$

Hierarchical Models

Goal: Collapse similar networks indicated by k
We include a LASSO penalty on the log-likelihood:

$$\sum_k \sum_{i < j} [y_{ijk} \ln \sigma_{ijk} + (1 - y_{ijk}) \ln(1 - \sigma_{ijk})] + \lambda \sum_i \sum_k \|\epsilon_{ik}\|_1$$

where $z_{ik} = b_i + \epsilon_{ik}$ for initialization points, b_i

Optimization Approaches:

- 1 Proximal Gradient Descent
- 2 Coordinate-Wise Optimization
- 3 Hamiltonian Monte Carlo methods

Objective Functions

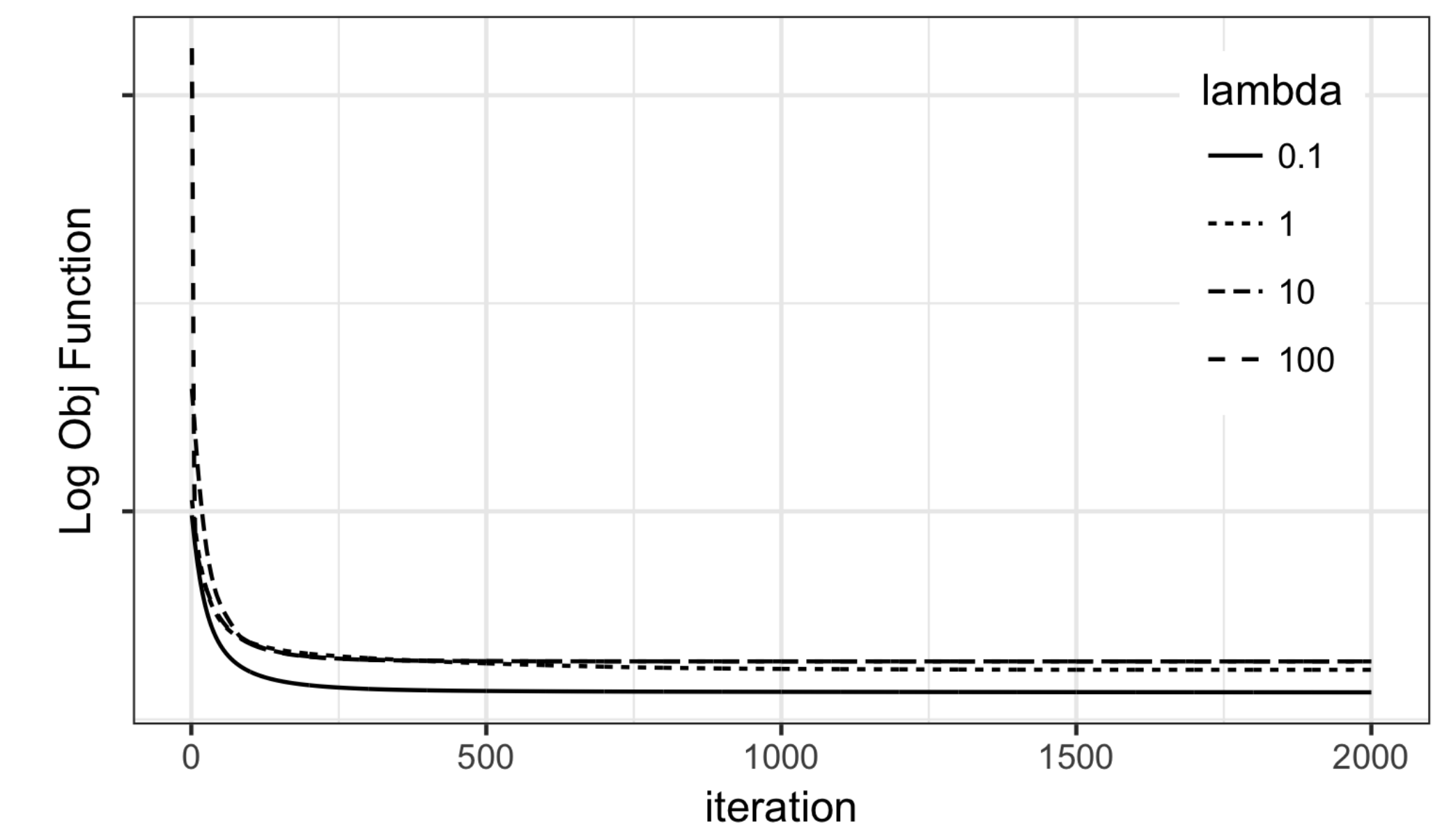


Figure 3: Objective function for different values of λ

Results

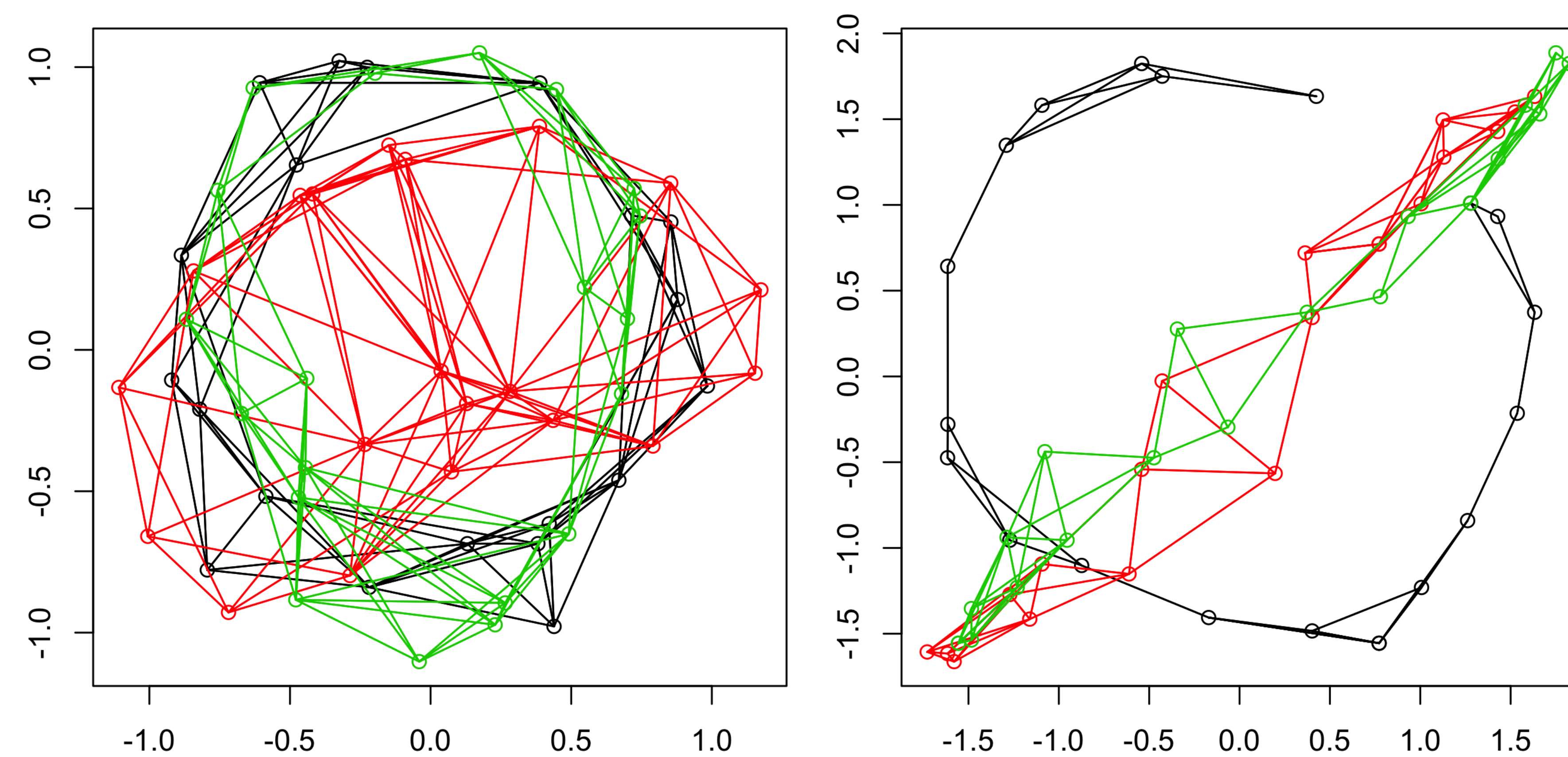


Figure 2: Initial and final positions using $\lambda = 1$

Challenges

- Negative log-likelihood is not convex in the positions
Fix: We started from carefully chosen initializations
- Distances between positions will yield same likelihood under translations and rotation
Fix: Fit coordinate-wise, not allowing for spin, unlike the standard sampling approaches

Conclusion

We found that proximal gradient methods are an effective tool for refining an HLSM fit, given a good starting point.

References

- [1]
- [2]
- [3]
- [4]
- [5]
- [6]
- [7]
- [8]
- [9]
- [10]