Parcellating connectivity in spatial maps

Christopher Baldassano[1], Diane M. Beck[2], Li Fei-Fei[1]

[1]Department of Computer Science, Stanford University

[2]Psychology Department and Beckman Institute, University of Illinois Urbana-Champaign

Summary

- **Goal:** Understand the spatial structure of interaction networks
- **Problem:** Existing methods provide fast approximate spatial parcellations, but fail under real-world high noise situations
- **Solution:** Fit a generative model using many passes over the connectivity data

Problem Definition

- Given a set of elements with spatial adjacency, and a connectivity matrix **D** describing interactions or flows
- Divide elements into spatially-contiguous parcels, such that the parcel x parcel connectivity matrix **A** approximates **D**
- We are interested in both the choice of parcels (revealing spatial sources in the data) and the summary connectivity matrix **A**

Previous Approaches

- Local Similarity: Define parcel borders by high spatial gradients in connectivity
- Normalized Cut: Cut adjacency graph based on local differences
- Region Growing: Iteratively grow seeds from high similarity regions
- Ward Clustering: Hierarchical greedy agglomerative clustering

Generative Parcellation Model

1. Sample a spatially-contiguous parcellation using the distance-dependent CRP

$$\mathbf{c} \sim \text{dd-CRP}(\alpha, \mathbf{f})$$

2. Sample connectivity strength and variance for each pair of parcels

$$\mathbf{A}_{mn}, \sigma^2_{mn} \sim \text{Normal-Inv-}\chi^2(\mu_0, \kappa_0, \sigma^2_0, \nu_0)$$

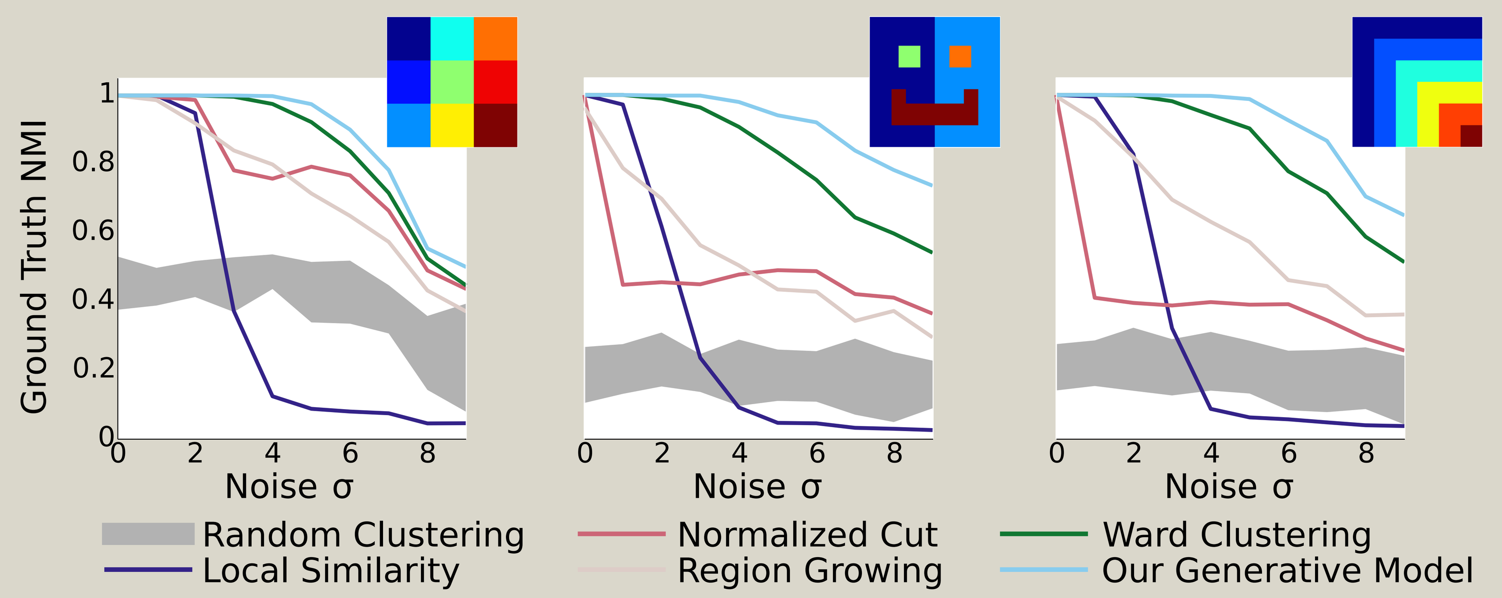
3. Sample noisy connectivity value between each pair of elements in parcels

$$\mathbf{D}_{ij}|\mathbf{z}(\mathbf{c}) \sim \text{Normal}(\mathbf{A}_{z(c)_i z(c)_j}, \sigma^2_{z(c)_i z(c)_j})$$

- Given **D**, perform collapsed Gibbs sampling on **c** to optimize parcellation

Synthetic Data Experiments

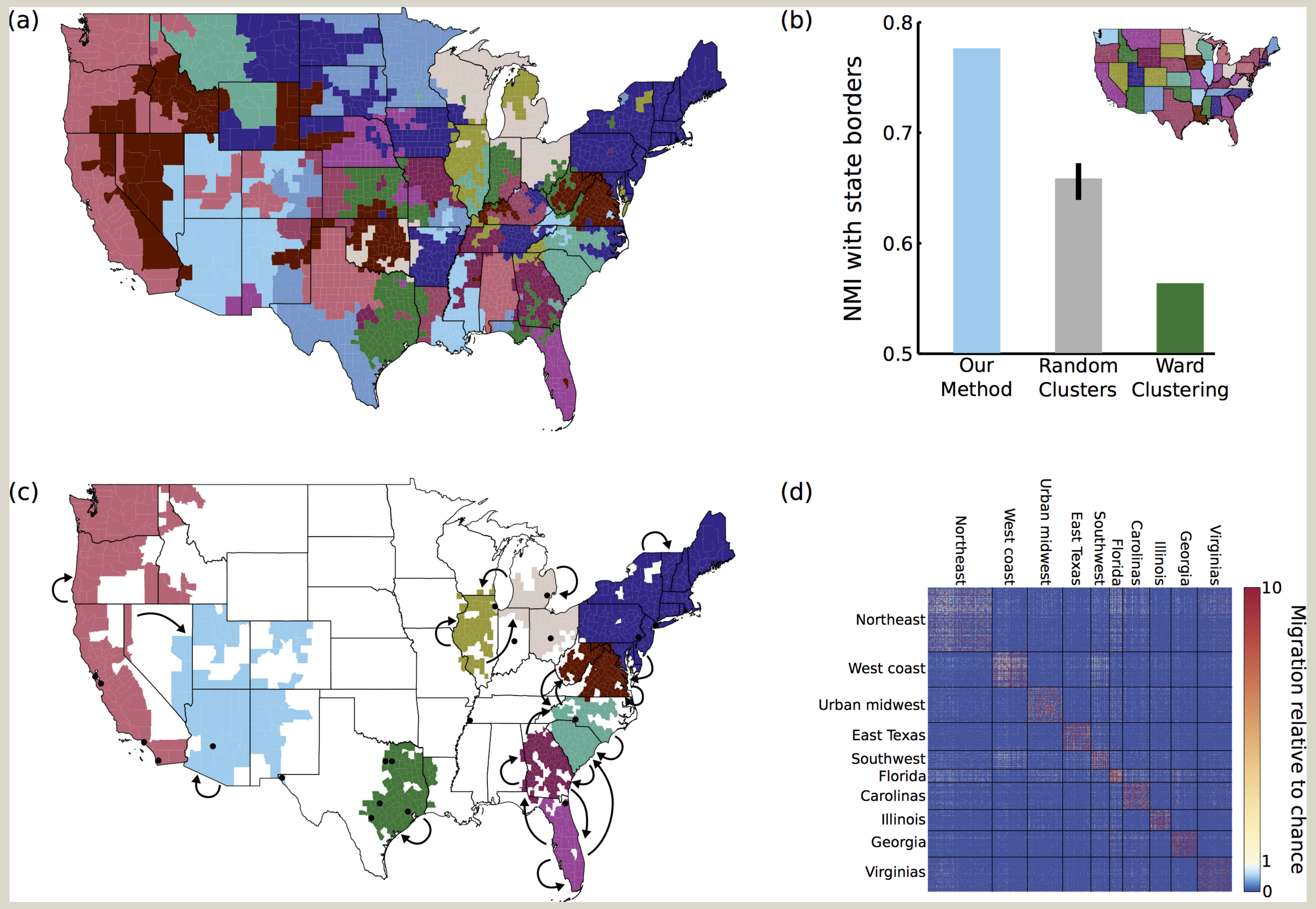
- Generated 18 x 18 grid of elements with 3 different ground truth patterns and random connectivity patterns
- Tried to recover ground truth under varying levels of noise



- Our model outperforms greedy methods, especially when parcels are nonuniform

US Migration Data

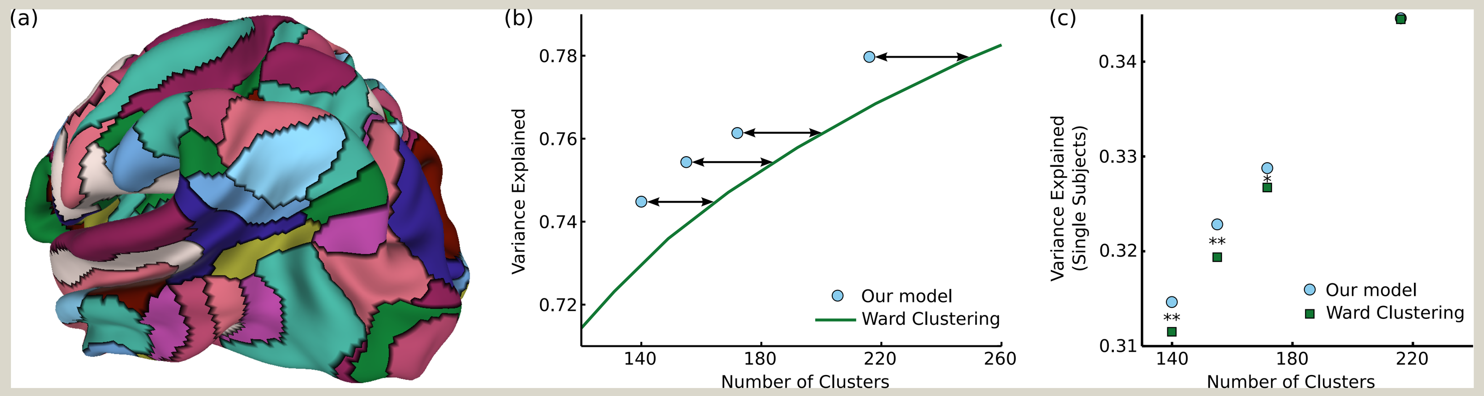
- Can we explain county-level human migration patterns using a small number of regions?



- Find parcels often related to state borders, with migration among most populous parcels well summarized by a 10x10 block matrix

Brain Functional Connectivity Data

- Can we explain brain connectivity in terms of interactions between local regions?



- Find parcels that explain larger proportion of the full 60k x 60k connectivity matrix, and generalize better to individual subjects, than greedy Ward clustering