Fast Computation of the Barycenter Graph in the Spectral Domain

Olivia Courtney
ocourtne@ucsc.edu
Applied Mathematics, University of California, Santa Cruz

François G. Meyer
Applied Mathematics, University of Colorado, Boulder

Objective and Contributions

- ❖ Given a set of graph-valued data, how do we **estimate a mean graph** which inherits structural properties?
  - **Essential location parameter** of many ML algorithms working on networks
  - Challenge of capturing **topological information on various scales**
- ❖ Main contributions of our algorithm
  - Mean graph inherits large-scale topological information (community structure)
  - Independent of graph size → can compare **graphs of different sizes**
  - **Computationally efficient** (parametric approach)
  - Supported by **theoretical results**

Code implementation at https://github.com/ocourtney/frechet_mean.git

Algorithm Description

The Sample Fréchet Mean

Notation:

- n: graph size
- $\boldsymbol{A}$: adjacency matrix (unweighted, symmetric, $n \times n$)
- $\boldsymbol{\lambda}$: vector of ordered eigenvalues, $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$

❖ We replace the algebraic notion of a mean with a minimization problem

❖ Finds the most central object, or 'center of mass'

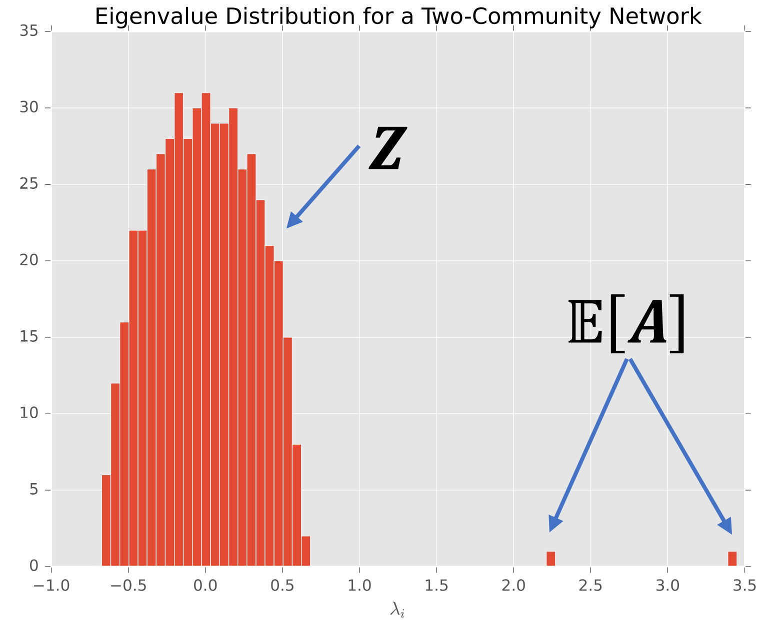
Given a sample of N graphs $\{G^{(1)}, G^{(2)}, \ldots, G^{(N)}\}$ and distance d, the **sample Fréchet mean** is the solution to

$$\widehat{\mathbb{E}}_N[G] = \arg\min_{G \in \mathcal{G}} \frac{1}{N} \sum_{k=1}^{N} d^2\left(G, G^{(k)}\right),$$

where $\mathcal{G}$ is an unweighted, undirected graph.

Spectral Pseudo-Distance

- Common pairwise distances: Edit, Hamming, Frobenius Norm
    - Good for fine-scale features, but poorly recover large-scale features (i.e. community structure)
- The spectrum contains multi-scale information, so we use this distance



$$\boldsymbol{A} = \mathbb{E}[\boldsymbol{A}] + \boldsymbol{Z},$$

Given graphs G and G' with respective spectra $\boldsymbol{\lambda}(\boldsymbol{A})$ and $\boldsymbol{\lambda}'(\boldsymbol{A})$, the <u>adjacency spectral distance</u> between the two graphs is defined as

$$d(G, G') = \sqrt{\sum_{j=1}^{n} \left(\lambda_j - \lambda_j'\right)^2}$$

Problem Statement

- Use spectral distance in Fréchet sample mean definition
- Assume that the mean graph has a Stochastic Block Model (SBM($n, \boldsymbol{P}$)) structure
  - Random generative model to construct graphs with community structure
  - Has universal approximating properties
  - k communities $C_1, \ldots, C_k,\ a_{ij} \sim Bernoulli(p_{ij})$
  - Edge probabilities stored in $\boldsymbol{P}$
- We prove that
  - $\lambda_i(\mathbb{E}[\boldsymbol{A}])$ depend linearly on the density probabilities
  - The error in $\lambda_i(\boldsymbol{A}) - \lambda_i(\mathbb{E}[\boldsymbol{A}])$ is bounded
- So we can effectively optimize this difference as a function of $\boldsymbol{P}$

$$\mathbb{E}[\mathbf{A}] = \begin{bmatrix} p_1 & \cdots & p_1 & q & \cdots & q \\ \vdots & & \vdots & \vdots & & \vdots \\ p_1 & \cdots & p_1 & q & \cdots & q \\ q & \cdots & q & p_2 & \cdots & p_2 \\ \vdots & & \vdots & \vdots & & \vdots \\ q & \cdots & q & p_2 & \cdots & p_2 \end{bmatrix}$$

$$\boxed{\hat{G} = \arg \min_{\boldsymbol{P} \in [0,1]^k} \frac{1}{N} \sum_{i=1}^{N} \sum_{j=1}^{k} \left(\lambda_j(\mathbb{E}[\boldsymbol{A}]_{\hat{\boldsymbol{P}}}) - \lambda_j(\boldsymbol{A}^i) \right)^2}$$

The Sample Fréchet Mean Algorithm

- Input: Sample networks $\boldsymbol{G}_1, \boldsymbol{G}_2, \dots, \boldsymbol{G}_N$ (or simply their dominant eigenvalues)
- Output: Probability matrix from which $\hat{G}$ can be generated
- Objective function: $F(\boldsymbol{P}) = \frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{k} ||\lambda_j\left(\mathbb{E}[\widehat{\boldsymbol{A}}]_{\boldsymbol{P}}\right) - \lambda_j(\boldsymbol{A}^{(i)})||^2$

Algorithm 1 Approximation of the Sample Fréchet Mean

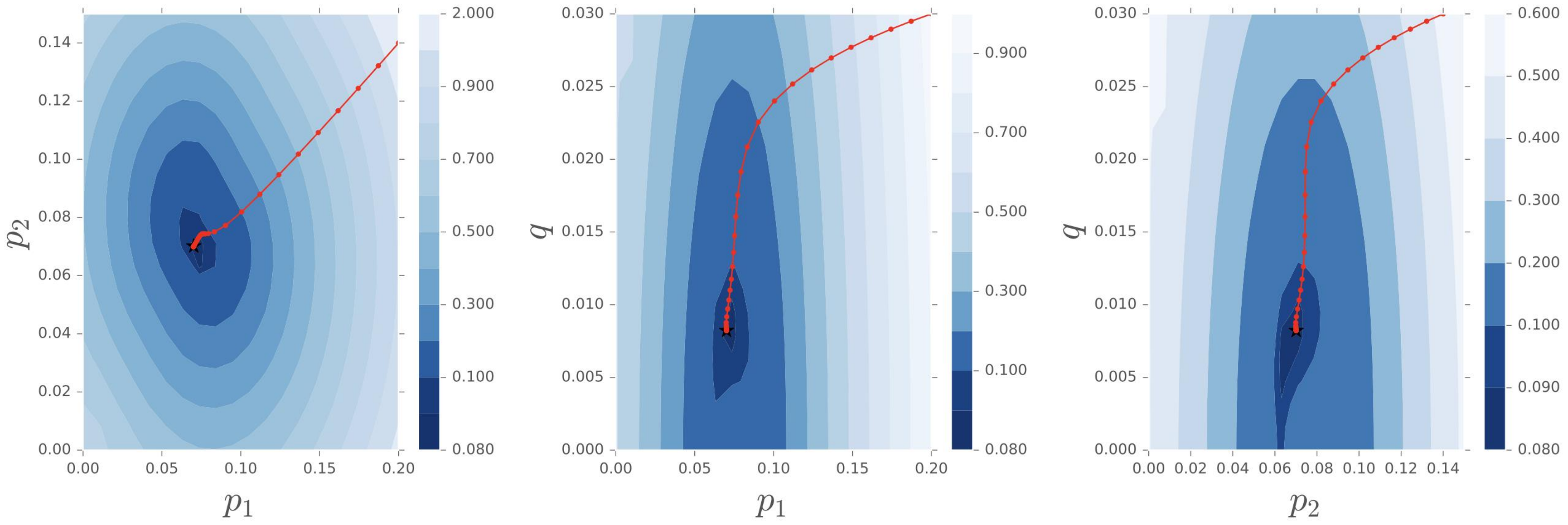
Require: A set of graphs $G = \{G^{(i)}\}_{i=1}^{N}$.
1: Set k and initialize $\mathbf{P} = \mathbf{P_0}$ via the desired method **.
2: For $i = 1, \dots, N$, compute $\boldsymbol{\lambda}^i = [\lambda_1(\mathbf{A}^i), \dots, \lambda_k(\mathbf{A}^i)]$.
3: **while** The relative change in $\mathbf{P}$ is greater than the tolerance **do**
4: Estimate the gradient of $F(\mathbf{P})$ (given by (4)) via finite differences.
5: Update $\mathbf{P}$ via a projected gradient step.
6: **end while**
7: **return** $\mathbf{P}$.

Numerical Results

1. Stehlé J, Voirin N, Barrat A, Cattuto C, Isella L, Pinton J-F, et al. (2011) High-Resolution Measurements of Face-to-Face Contact Patterns in a Primary School. PLoS ONE 6(8): e23176. https://doi.org/10.1371/journal.pone.0023176

Convergence of gradient descent

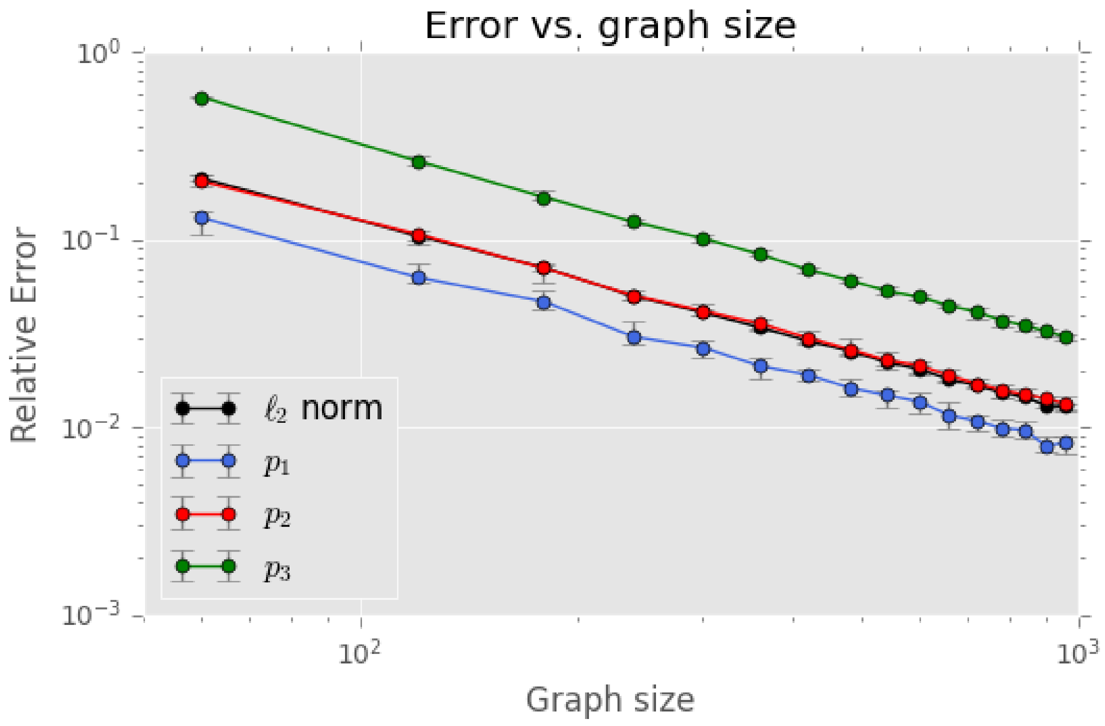
- Sample: real network data[1]
  - Temporal social contact networks of a primary school
  - $n = 242$
  - We see convergence over the objective function



Convergence in the spectral domain

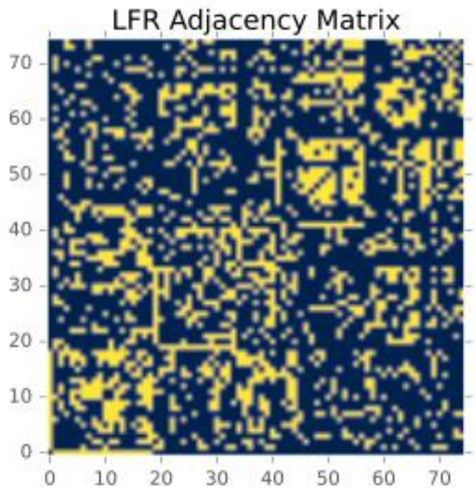
- Sample: SBM networks
  - $n \in [50, 1000], \mathrm{N} = 30$
  - $k = 3$, with distinct sizes that are relative to n
  - $\boldsymbol{P} =$ constant, but distinct p densities
  - Median of 20 trials is plotted
- Spectral error: $\|\hat{\boldsymbol{\lambda}} - \boldsymbol{\lambda}\|_2$

❖ We see the algorithm converges to the true mean network as the graph size (n) grows.

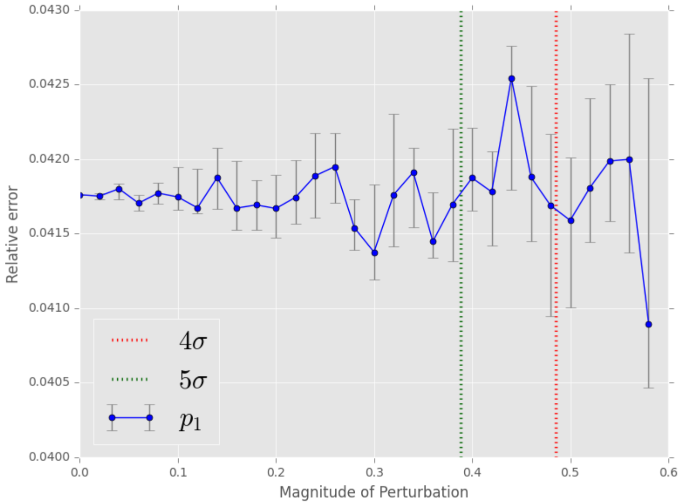


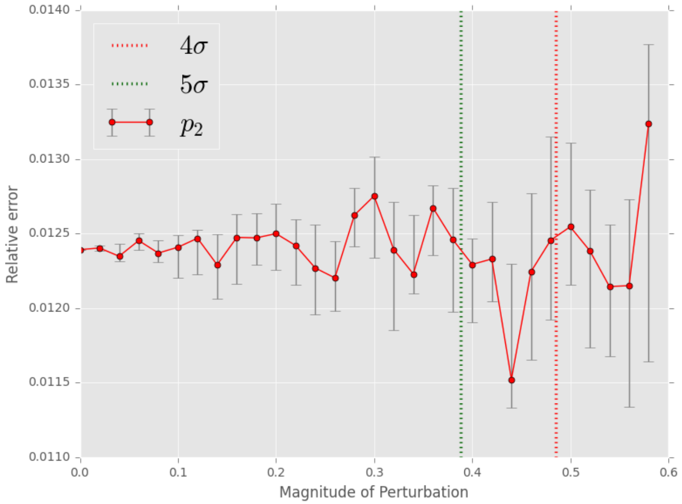
2. A. Lancichinetti, S. Fortunato, and F. Radicchi.(2008) Benchmark graphs for testing community detection algorithms. Physical Review E, 78.

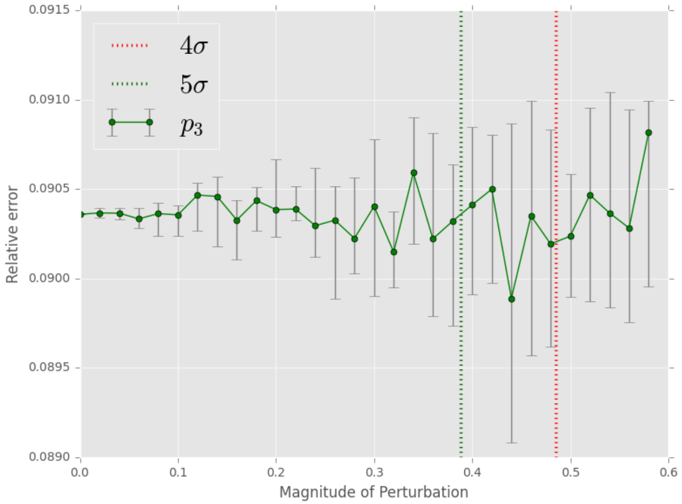
Real Network



- True mean: LFR-Benchmark[2] network
  - $n = 75$, unweighted, undirected
  - $k = 3$, with distinct sizes, non-overlapping
  - Median of 20 trials is plotted
- Recovery of the true mean network given a noisy sample
- Density error: $\left\|\widehat{\boldsymbol{P}} - \boldsymbol{P}\right\|_2$
- Algorithm will on expectation return true mean graph, acts as an unbiased estimator







Conclusion

- Given a sample of networks, we can **approximate the mean** using a parameterized approach in the spectral domain

- Using the eigenvalues to measure proximity between graphs allows for the **recovery of community structure**

- **Limitation**: algorithms relies of our ability to approximate the mean graph with a SBM

- **Future work**:
    - Modular analysis of errors from SBM approximation and from algorithm design
    - Experiments on graphs with no community structure

My email: ocourtne@ucsc.edu