Multi-Frequency Phase Synchronization and Class Averaging for Three-Dimensional Cryo-Electron Microscopy

Tingran Gao

Committee on Computational and Applied Mathematics (CCAM)
Department of Statistics
University of Chicago

Geometric Data Analysis
University of Chicago, Chicago IL

May 23, 2019

Outline

Motivation

- Synchronization

Multi-Frequency Phase Synchronization

- Phase Synchronization
- Multi-Frequency Formulation
- Theory

Application to Multi-Frequency Class Averaging

Joint work with **Yifeng Fan (UIUC)** & **Zhizhen Zhao (UIUC)**

Outline

Motivation

- **Synchronization**

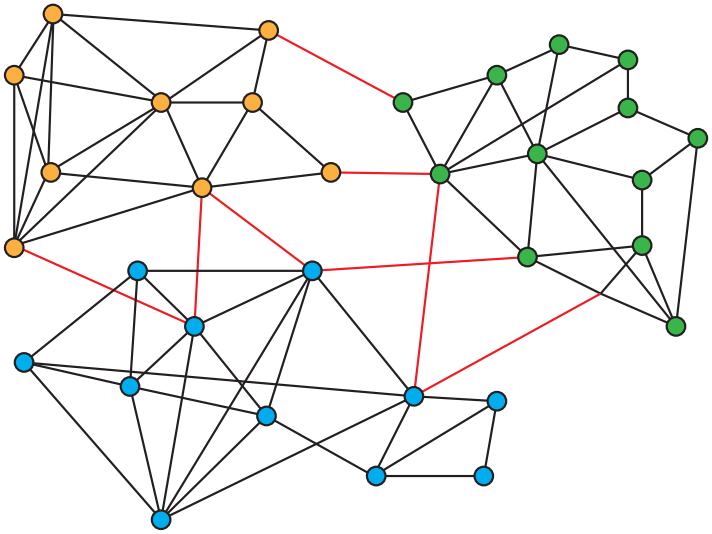
Multi-Frequency Phase Synchronization

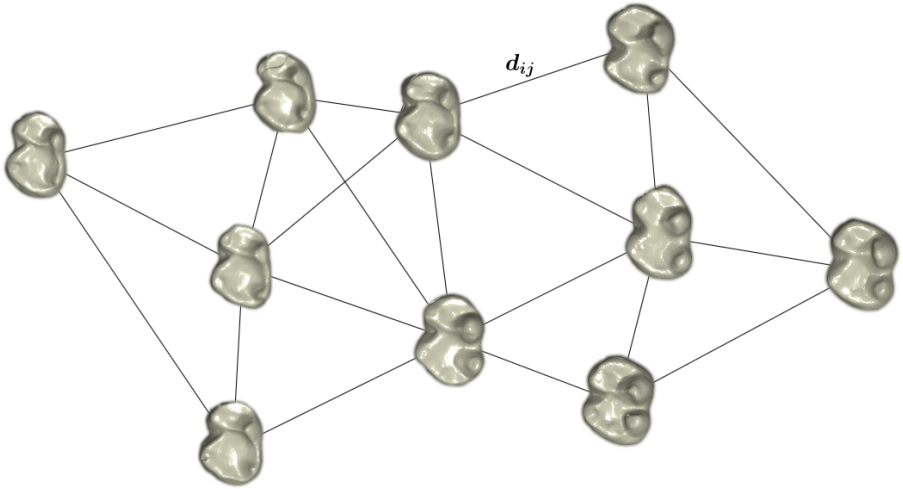
- Phase Synchronization
- Multi-Frequency Formulation
- Theory

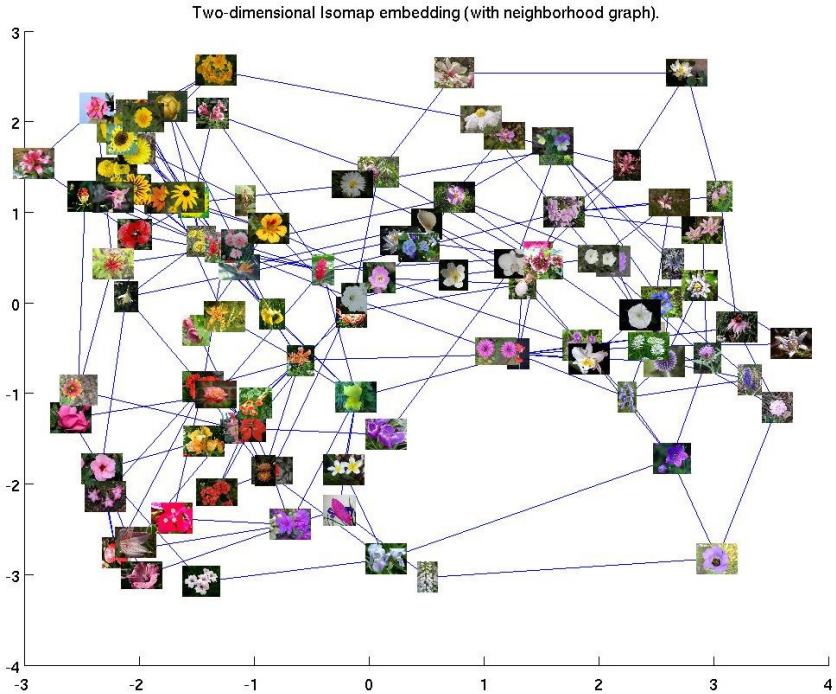
Application to Multi-Frequency Class Averaging

Joint work with **Yifeng Fan (UIUC)** & **Zhizhen Zhao (UIUC)**

Graph-Based Data Analysis



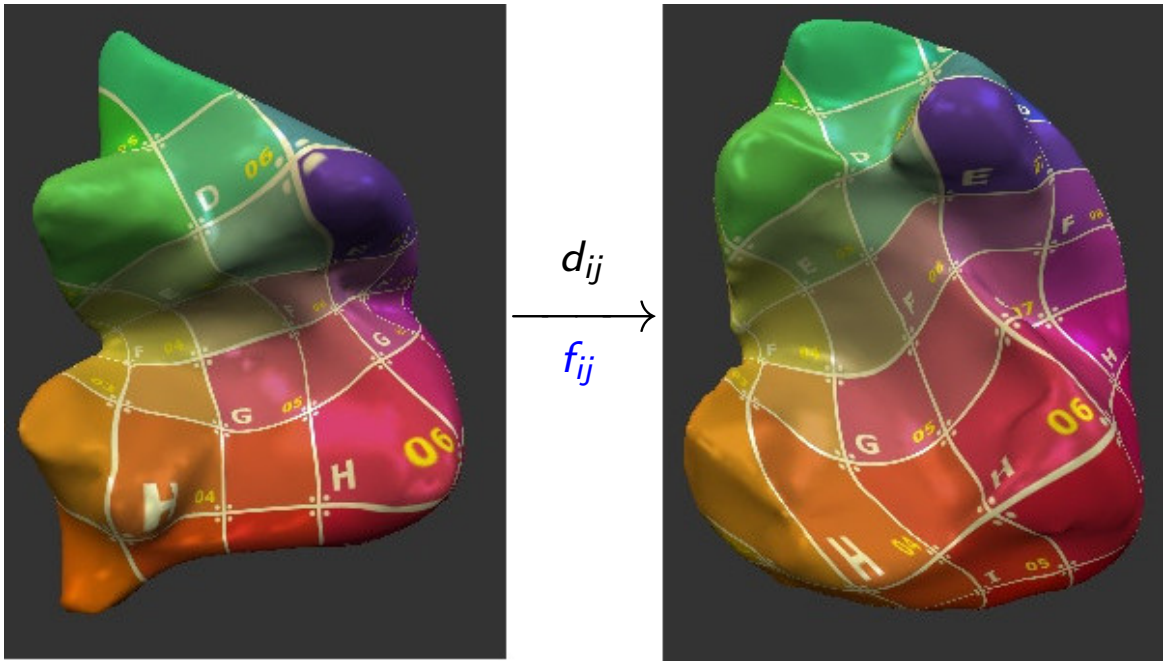




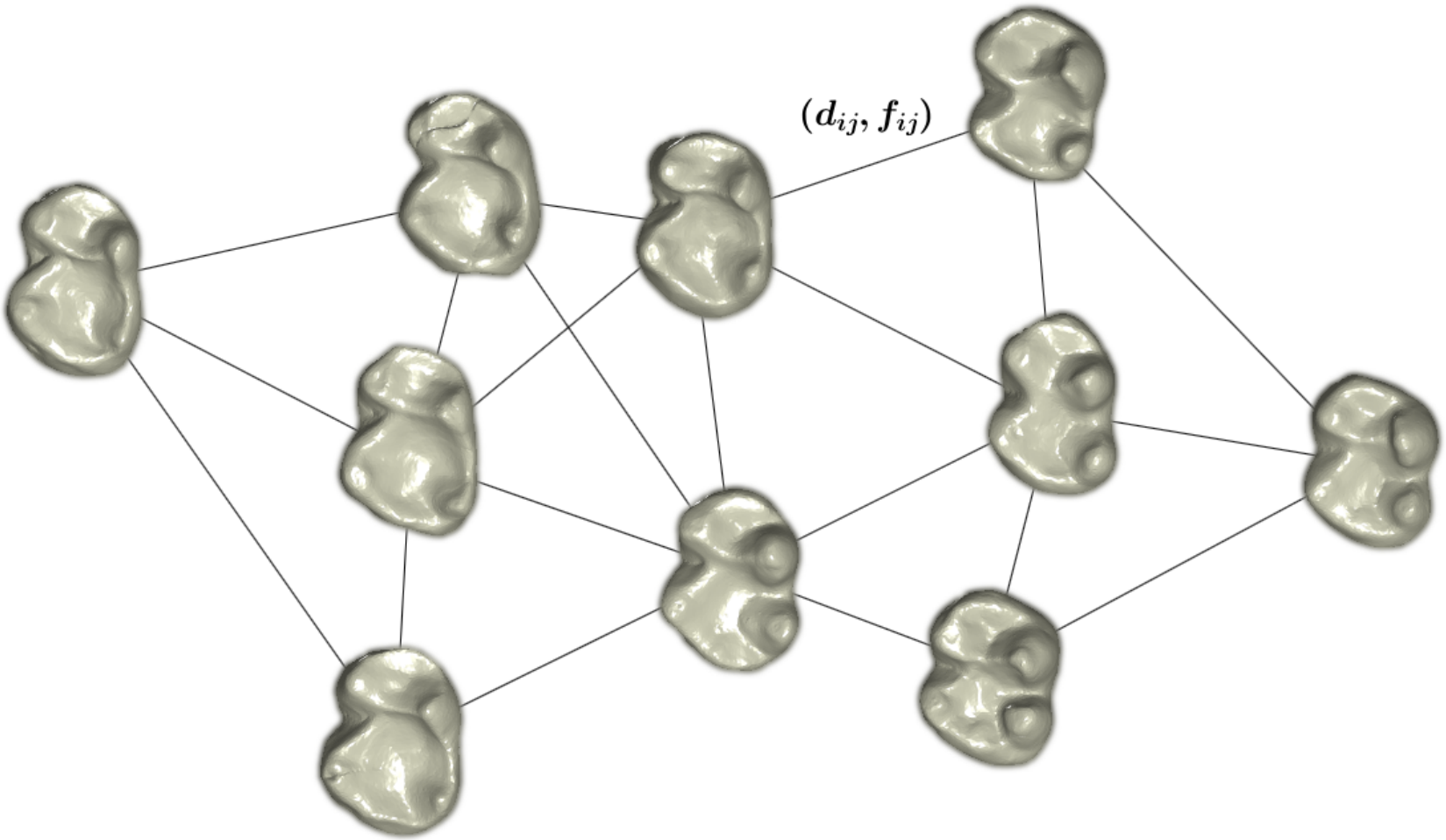
Evolutionary Biology

$$d_{\mathrm{cP}}\left(S_i, S_j\right) = \inf_{\mathcal{C}\in\mathcal{A}\left(S_i,S_j\right)} \inf_{R\in\mathbb{E}(3)} \left(\int_{S_i} \left\|R\left(x\right) - \mathcal{C}\left(x\right)\right\|^2 d\mathrm{vol}_{S_i}\left(x\right)\right)^{\frac{1}{2}}$$

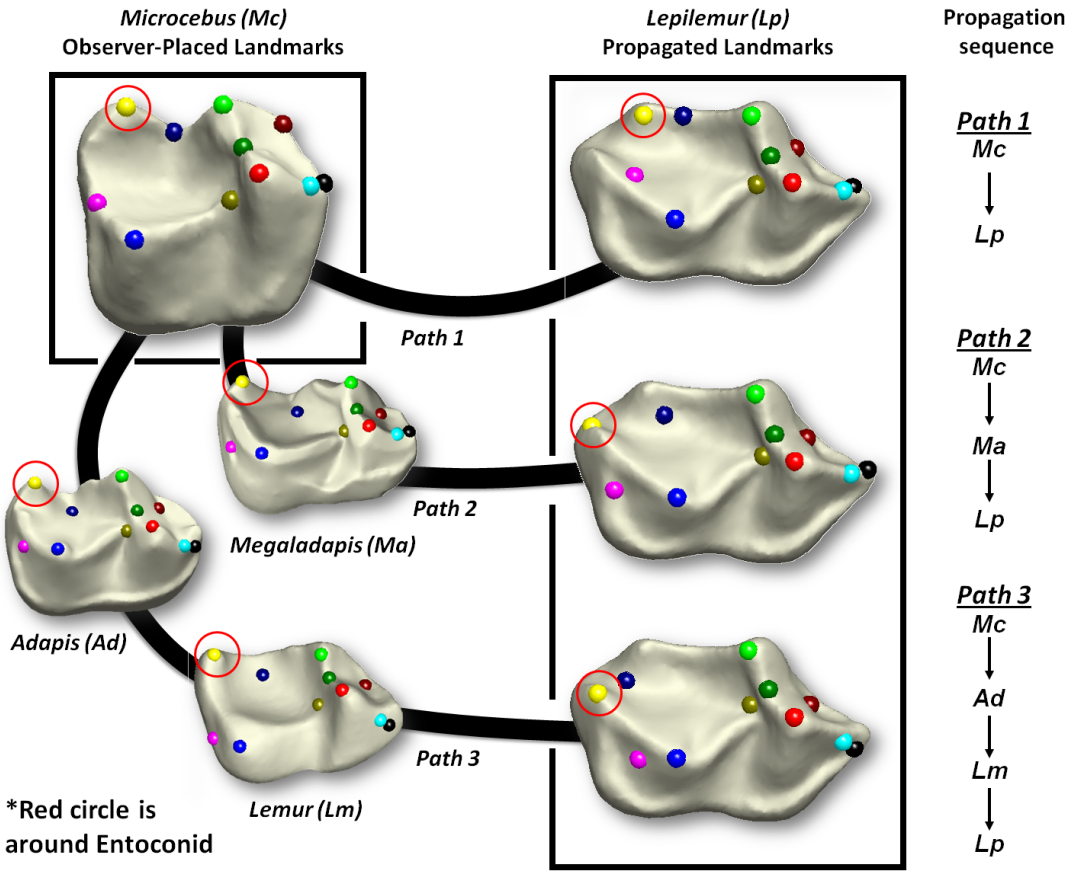
S_i



S_j

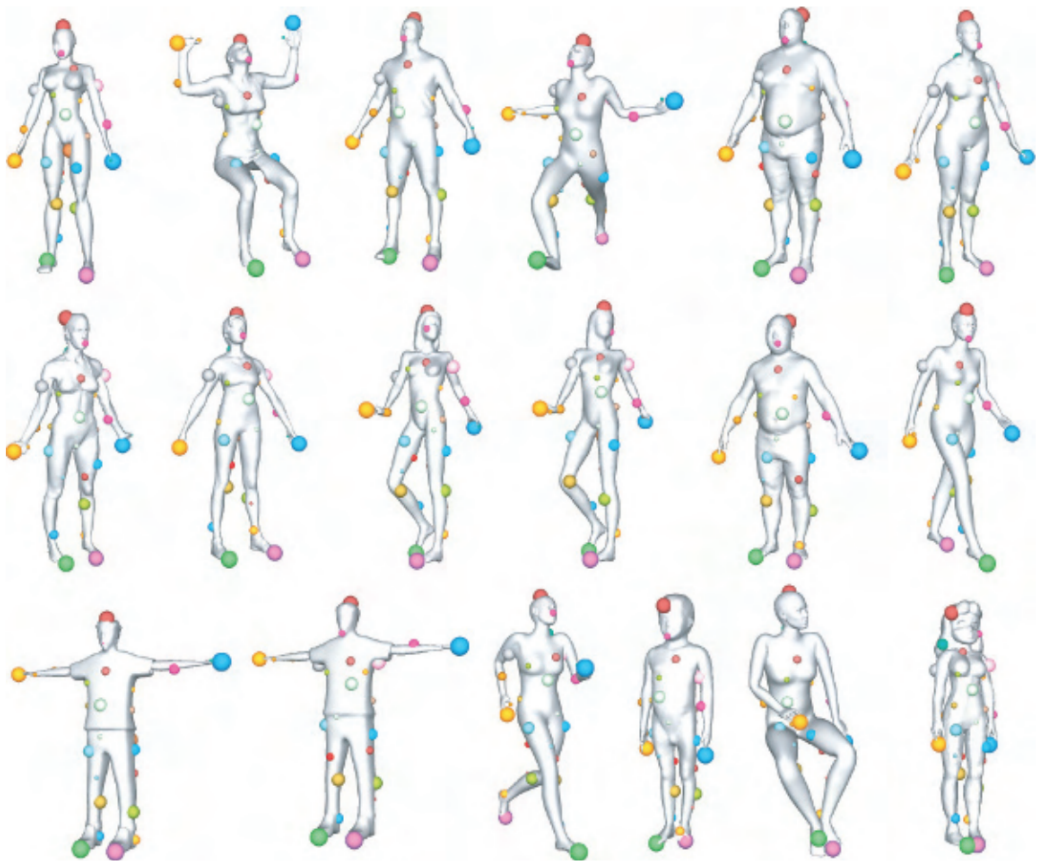

(d_{ij}, f_{ij})

Interpretability Issue



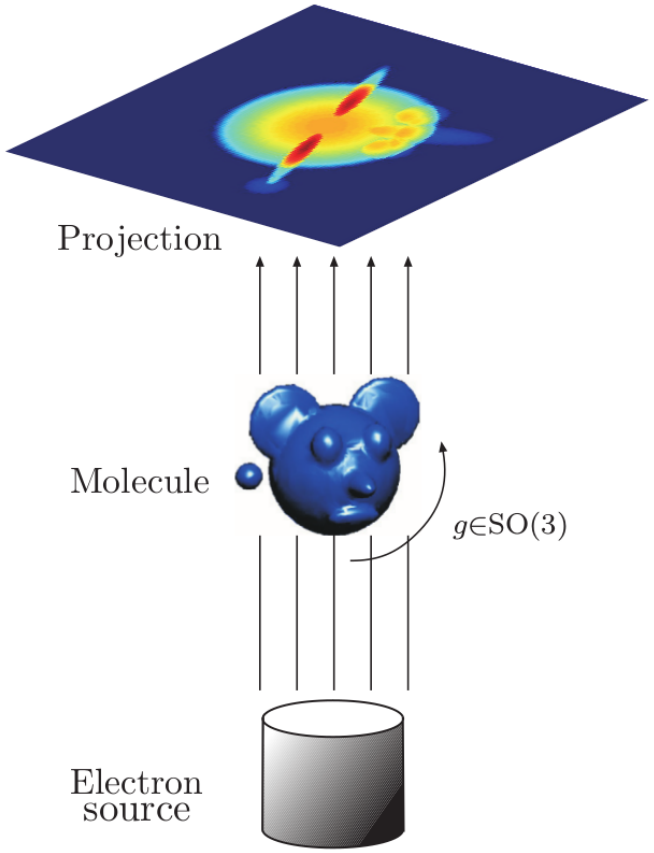
• Boyer et al. "Algorithms to Automatically Quantify the Geometric Similarity of Anatomical Surfaces." *Proceedings of the National Academy of Sciences* 108.45 (2011): 18221–18226.

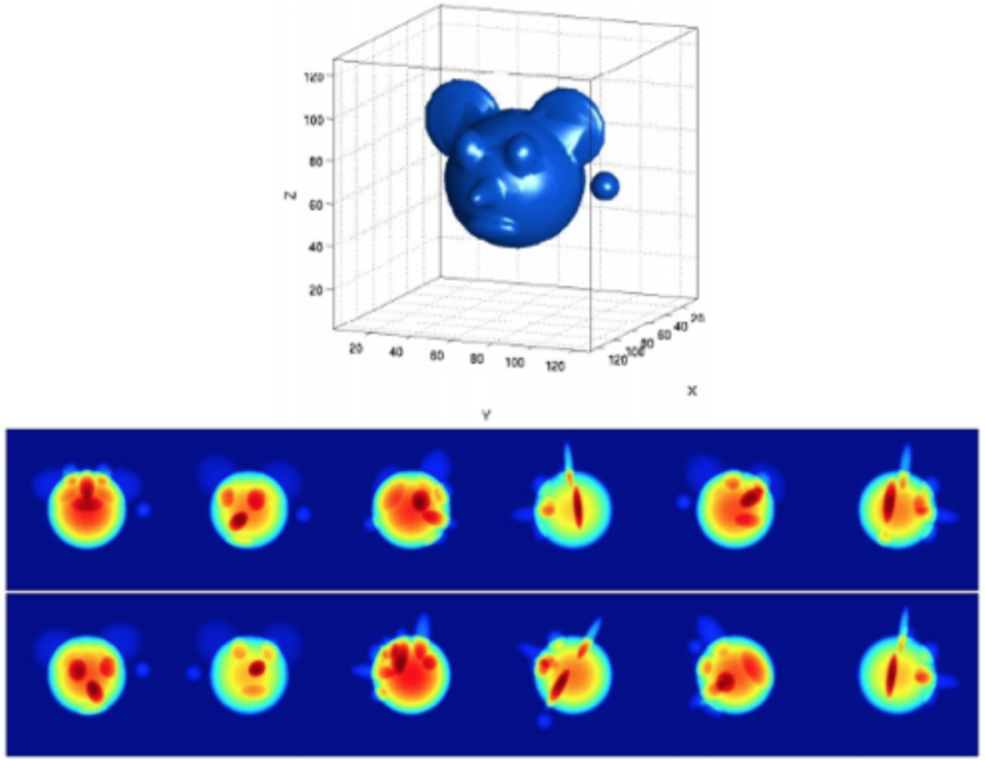
Collection Shape Analysis



• Q. Huang & L. Guibas. "Consistent Shape Maps via Semidefinite Programming", *Computer Graphics Forum*, vol. 32, no. 5, pp. 177–186 (2013).

Cryo-Electron Microscopy





• Singer et al. "Viewing Angle Classification of Cryo-Electron Microscopy Images using Eigenvectors", *SIAM Journal on Imaging Sciences*, 4 (2), pp. 543–572 (2011).

Cryo-Electron Microscopy: Real Challenge is Low SNR

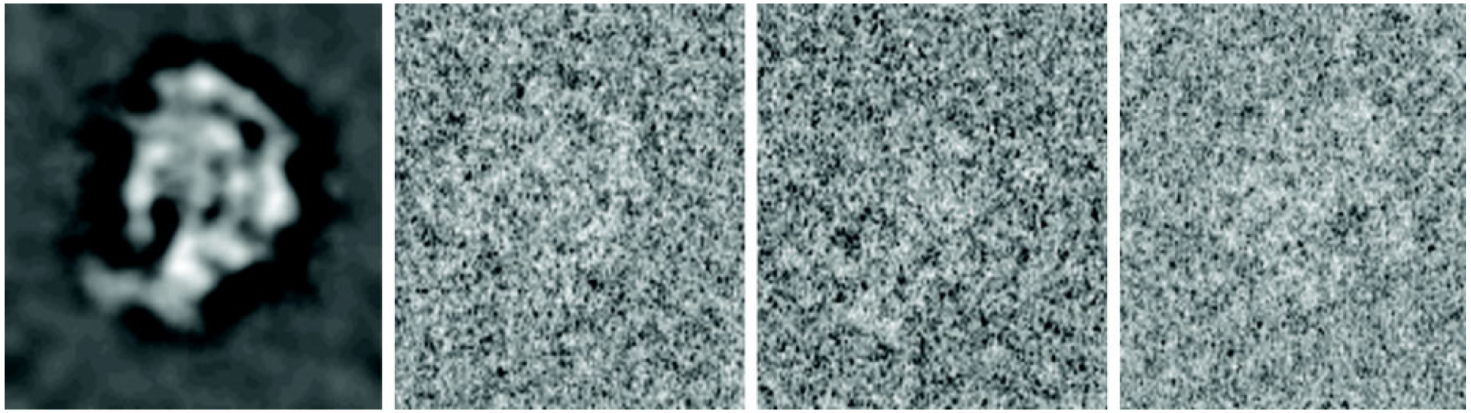
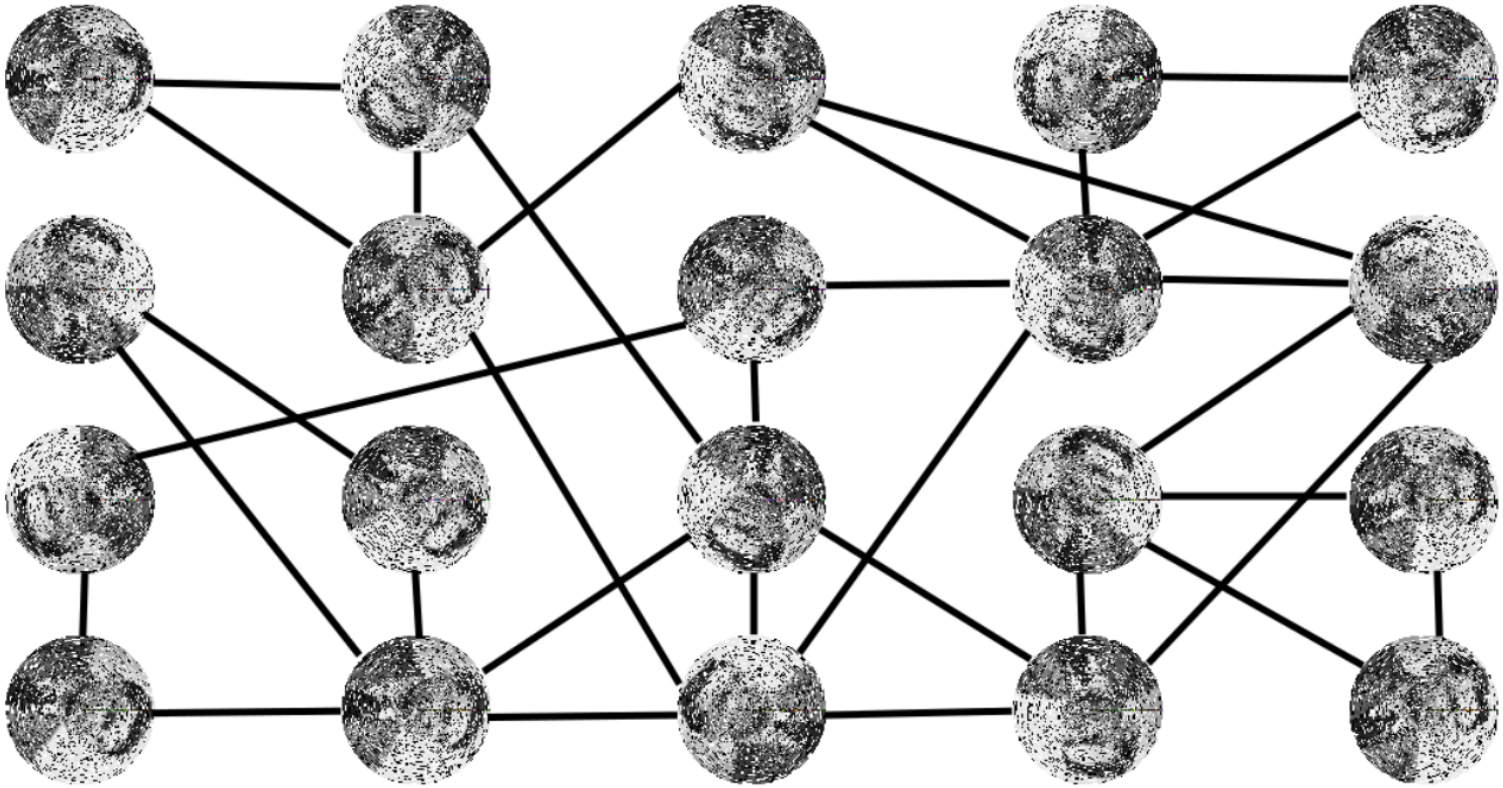


Fig. 3 The left most image is a clean simulated projection image of the E.coli 50S ribosomal subunit. The other three images are real electron microscope images of the same subunit

• Hadani & Singer. "Representation Theoretic Patterns in Three-Dimensional Cryo-Electron Microscopy II – The Class Averaging Problem," *Foundations of Computational Mathematics*, 11 (5), pp. 589–616 (2011).

$$y_i = R_i x + \xi_i$$
$$R_i \in O(d), \quad \xi_i \sim \text{i.i.d. noise}$$



- Afonso S. Bandeira. "Ten Lectures and Forty-Two Open Problems in the Mathematics of Data Science." (2015).

$$\boxed{\begin{array}{c} y_i = R_i x + \xi_i \\ R_i \in O(d), \quad \xi_i \sim \text{i.i.d. noise} \end{array}}$$

$$\begin{aligned} \text{Measurement:} \quad & R_{ij} \approx R_i^\top R_j \\ \text{Recover:} \quad & R_1, R_2, \cdots \end{aligned}$$

$$\boxed{\begin{array}{c} y_i = R_i x + \xi_i \\ R_i \in O(d), \quad \xi_i \sim \text{i.i.d. noise} \end{array}}$$

$$\begin{aligned} \text{Measurement:} \quad & R_{ij} \approx R_i^\top R_j \\ \text{Recover:} \quad & R_1, R_2, \cdots \end{aligned}$$

$\Rightarrow$ Solve the minimization problem

$$\min_{R_1, R_2, \cdots \in O(d)} \quad \sum_{(i,j) \in E} w_{ij} \left\| R_{ij} - R_i^\top R_j \right\|_{\mathrm{F}}^2$$

$$\frac{1}{2} \sum_{1 \leq i,j \leq N} w_{ij} \left\| R_{ij} - R_i^\top R_j \right\|_{\mathrm{F}}^2$$

$$
\begin{aligned}
&\frac{1}{2}\sum_{1\le i,j\le N} w_{ij}\left\|R_{ij}-R_i^\top R_j\right\|_{\mathrm{F}}^2 \\
&\quad=\frac{1}{2}\sum_{1\le i,j\le N} w_{ij}\mathrm{Tr}\left[\left(R_i^\top-R_{ij}R_j^\top\right)\left(R_i-R_jR_{ij}^\top\right)\right] \\
&\quad=\sum_{1\le i,j\le N} w_{ij}\mathrm{Tr}\left(I_{d\times d}-R_iR_{ij}R_j^\top\right)
\end{aligned}
$$

$$
\begin{aligned}
&\frac{1}{2}\sum_{1\le i,j\le N} w_{ij}\left\|R_{ij}-R_i^\top R_j\right\|_{\mathrm{F}}^2 \\
&\quad=\frac{1}{2}\sum_{1\le i,j\le N} w_{ij}\mathrm{Tr}\left[\left(R_i^\top-R_{ij}R_j^\top\right)\left(R_i-R_jR_{ij}^\top\right)\right] \\
&\quad=\sum_{1\le i,j\le N} w_{ij}\mathrm{Tr}\left(I_{d\times d}-R_iR_{ij}R_j^\top\right) \\
&\quad=\mathrm{Tr}\left(\left[R_1,\cdots,R_N\right]\left(\mathscr{D}-\mathscr{W}\right)\begin{bmatrix}R_1^\top\\ \vdots\\ R_N^\top\end{bmatrix}\right)
\end{aligned}
$$

where

$$
\mathscr{D}=\begin{bmatrix} d_1 I_{d\times d} & & \\ & \ddots & \\ & & d_N I_{d\times d}\end{bmatrix},\mathscr{W}=\begin{bmatrix} w_{11}R_{11} & \cdots & w_{1N}R_{1N}\\ \vdots & & \vdots\\ w_{N1}R_{N1} & \cdots & w_{NN}R_{NN}\end{bmatrix}
$$

QO-OC

$$\min_{R_1,R_2,\cdots\in O(d)} \quad \sum_{(i,j)\in E} w_{ij} \left\| R_{ij} - R_i^\top R_j \right\|_{\mathrm{F}}^2$$

$$\Leftrightarrow \quad \min_{R_1,\cdots,R_N\in\mathbb{R}^{d\times d}} \mathrm{Tr}\left(\left[R_1,\cdots,R_N\right] \left(\mathscr{D}-\mathscr{W}\right) \begin{bmatrix} R_1^\top \\ \vdots \\ R_N^\top \end{bmatrix} \right)$$

$$\text{s.t.} \qquad R_1^\top R_1 = \cdots = R_N^\top R_N = I_{d\times d}$$

• A. Nemirovski, "Sums of Random Symmetric Matrices and Quadratic Optimization Under Orthogonality Constraints", *Math. Programming*, 109 (2007), pp. 283–317.

Convex Relaxations (I)

$$\min_{R_1,\cdots,R_N \in \mathbb{R}^{d\times d}} \operatorname{Tr}\left(\left[R_1, \cdots, R_N\right] \left(\mathscr{D} - \mathscr{W}\right) \begin{bmatrix} R_1^\top \\ \vdots \\ R_N^\top \end{bmatrix} \right)$$
$$\text{s.t.} \qquad R_1^\top R_1 = \cdots = R_N^\top R_N = I_{d\times d}$$

- Spectral Relaxation

$$\boxed{\begin{aligned} &\min_{\mathcal{R} \in \mathbb{R}^{d\times Nd}} \operatorname{Tr}\left(\mathcal{R}\left(\mathscr{D} - \mathscr{W}\right)\mathcal{R}^\top\right) \\ &\text{s.t.} \qquad \mathcal{R}\mathcal{R}^\top = N \cdot I_{d\times d} \end{aligned}}$$

• Chaudhury et al. "Global Registration of Multiple Point Clouds using Semidefinite Programming", *SIAM Journal on Optimization*, 25(1):468–501, 2015.

Convex Relaxations (II)

$$\min_{R_1,\cdots,R_N \in \mathbb{R}^{d\times d}} \operatorname{Tr}\left((\mathscr{D} - \mathscr{W}) \begin{bmatrix} R_1^\top \\ \vdots \\ R_N^\top \end{bmatrix} \begin{bmatrix} R_1, \cdots, R_N \end{bmatrix} \right)$$
$$\text{s.t.} \qquad R_1^\top R_1 = \cdots = R_N^\top R_N = I_{d\times d}$$

-

$$\boxed{\begin{aligned} \min_{\mathcal{G} \in \mathbb{R}^{Nd\times Nd}} \quad & \operatorname{Tr}\left((\mathscr{D} - \mathscr{W})\, \mathcal{G} \right) \\ \text{s.t.} \quad & \mathcal{G} \succeq 0,\ \mathcal{G} = \mathcal{G}^\top \\ & \mathcal{G}_{ii} = I_{d\times d},\ 1 \le i \le N \\ & \operatorname{rank}(\mathcal{G}) = d \end{aligned}}$$

• Chaudhury et al. "Global Registration of Multiple Point Clouds using Semidefinite Programming", *SIAM Journal on Optimization*, 25(1):468–501, 2015.

Convex Relaxations (II)

$$\min_{R_1,\cdots,R_N \in \mathbb{R}^{d\times d}} \operatorname{Tr}\left((\mathscr{D} - \mathscr{W}) \begin{bmatrix} R_1^\top \\ \vdots \\ R_N^\top \end{bmatrix} \begin{bmatrix} R_1, \cdots, R_N \end{bmatrix} \right)$$
$$\text{s.t.} \qquad R_1^\top R_1 = \cdots = R_N^\top R_N = I_{d\times d}$$

- SDP Relaxation

$$\boxed{\begin{aligned} \min_{\mathcal{G} \in \mathbb{R}^{Nd\times Nd}} \quad & \operatorname{Tr}\left((\mathscr{D} - \mathscr{W})\, \mathcal{G} \right) \\ \text{s.t.} \quad & \mathcal{G} \succeq 0,\ \mathcal{G} = \mathcal{G}^\top \\ & \mathcal{G}_{ii} = I_{d\times d},\ 1 \le i \le N \\ & \cancel{\operatorname{rank}(\mathcal{G}) = d} \end{aligned}}$$

• Chaudhury et al. "Global Registration of Multiple Point Clouds using Semidefinite Programming", *SIAM Journal on Optimization*, 25(1):468–501, 2015.

Outline

Motivation

- Synchronization

Multi-Frequency Phase Synchronization

- **Phase Synchronization**
- Multi-Frequency Formulation
- Theory

Application to Multi-Frequency Class Averaging

Joint work with **Yifeng Fan (UIUC)** & **Zhizhen Zhao (UIUC)**

Phase Synchronization: Synchronization Over $\mathrm{U}(1)$

- Assume the underlying graph Γ is complete
- **Ground truth**: phase vector $z = (\exp \iota\theta_1, \ldots, \exp \iota\theta_n)^\top \in \mathbb{C}_1^n$
- **Measurements**: *noisy* pairwise phase offsets $H_{ij} \approx z_i \bar{z}_j$, or in matrix form

$$H := zz^* + \mathrm{noise}$$

- **Goal**: recover the true phase vector z (up to a global multiplicative factor)
- **Spectral Relaxation**: solve for the top eigenvector of H, denoted as $\tilde{x}$ (scaled to $\|\tilde{x}\|_2 = \sqrt{n}$), then define $\hat{x} \in \mathbb{C}_1^n$ by

$$\hat{x}_i := \tilde{x}_i / |\tilde{x}_i|$$

Noise Models

- **Additive Gaussian Noise Model**

$$H = zz^* + \sigma W$$

where W is a standard complex Wigner matrix with i.i.d. standard complex Gaussian entries above the diagonal

- **Random Corruption Model**

$$H_{ij} = \begin{cases} z_i \bar{z}_j & \text{with prob. } r \in [0,1] \\ \mathrm{Unif}\,(\mathrm{U}(1)) & \text{with prob. } 1-r \end{cases}$$

Optimization Problem Formulation

$$\min_{x \in \mathbb{C}_1^n} \|xx^* - H\|_{\mathrm{F}}^2 \quad \Leftrightarrow \quad \max_{x \in \mathbb{C}_1^n} x^* H x$$

- Nonconvex problem, can be attacked using convex (spectral or SDP) relaxation
- For additive Gaussian noise, convex relaxation almost recovers the ground truth $z \in \mathbb{C}_1^n$ with high probability for noise level up to $\sigma = O\left(\sqrt{n/\log n}\right)$ **[Zhong & Boumal (2018)]**; non-rigorous statistical-physics-based methods predict the same holds up to $\sigma = O\left(\sqrt{n}\right)$ **[Javanmard et al. (2016)]**
- For random corruption model, **[Singer (2011)]** argues that $|\mathrm{Corr}\,(\tilde{x}, z)| > 1/\sqrt{n}$ with high probability if $r > 1/\sqrt{n}$

Outline

Motivation

- Synchronization
- Geometric Model of Fibre Bundles

Multi-Frequency Phase Synchronization

- Phase Synchronization
- **Multi-Frequency Formulation**
- Theory

Application to Multi-Frequency Class Averaging

Joint work with **Yifeng Fan (UIUC)** & **Zhizhen Zhao (UIUC)**

Multi-Frequency Phase Synchronization

- Assume random corruption model
- Instead of solving

$$\max_{x \in \mathbb{C}_1^n} x^* H x$$

  let us solve

$$\boxed{\max_{x \in \mathbb{C}_1^n} \sum_{k=1}^{k_{\max}} (x^k)^* H^{(k)} x^k}$$

  where $x^k := \left(x_1^k, \ldots, x_n^k\right)^\top \in \mathbb{C}_1^n$, and $H^{(k)}$ is the n-by-n Hermitian matrix with $H_{ij}^{(k)} := H_{ij}^k$
- Solving this nonconvex optimization by a popular two-stage strategy: (i) produce a good initialization (ii) perform local methods e.g. gradient descent or (generalized) power iteration

• Tingran Gao & Zhizhen Zhao. "Multi-Frequency Phase Synchronization," *The 36th International Conference on Machine Learning* (ICML 2019). arXiv:1901.08235.

Multi-Frequency Phase Synchronization: Stage 1

- Fix $k_{\max} \geq 1$, build $H^{(2)}, \ldots, H^{(k_{\max})}$ out of $H = H^{(1)}$

Multi-Frequency Phase Synchronization: Stage 1

- Fix $k_{\max} \geq 1$, build $H^{(2)}, \ldots, H^{(k_{\max})}$ out of $H = H^{(1)}$
- For each $k = 1, \ldots, k_{\max}$, solve the subproblem

$$u^{(k)} := \arg\max_{v \in \mathbb{C}_1^n} \; v^* H^{(k)} v$$

using any convex relaxation, and set $W^{(k)} := u^{(k)} \left(u^{(k)}\right)^*$

Multi-Frequency Phase Synchronization: Stage 1

- Fix $k_{\max} \geq 1$, build $H^{(2)}, \ldots, H^{(k_{\max})}$ out of $H = H^{(1)}$
- For each $k = 1, \ldots, k_{\max}$, solve the subproblem

$$u^{(k)} := \arg\max_{v \in \mathbb{C}_1^n} \; v^* H^{(k)} v$$

  using any convex relaxation, and set $W^{(k)} := u^{(k)} \left(u^{(k)}\right)^*$
- For all $1 \leq i, j \leq n$, find the "peak location" of the spectrogram

$$\hat{\theta}_{ij} := \arg\max_{\phi \in [0, 2\pi]} \left| \frac{1}{2} \sum_{k=-k_{\max}}^{k_{\max}} W_{ij}^{(k)} e^{-\iota k \phi} \right|$$

Multi-Frequency Phase Synchronization: Stage 1

- Fix $k_{\max} \geq 1$, build $H^{(2)}, \ldots, H^{(k_{\max})}$ out of $H = H^{(1)}$
- For each $k = 1, \ldots, k_{\max}$, solve the subproblem

$$u^{(k)} := \arg\max_{v \in \mathbb{C}_1^n} \; v^* H^{(k)} v$$

  using any convex relaxation, and set $W^{(k)} := u^{(k)} \left(u^{(k)}\right)^*$
- For all $1 \leq i, j \leq n$, find the "peak location" of the spectrogram

$$\hat{\theta}_{ij} := \arg\max_{\phi \in [0, 2\pi]} \left| \frac{1}{2} \sum_{k=-k_{\max}}^{k_{\max}} W_{ij}^{(k)} e^{-\iota k \phi} \right|$$

- Construct Hermitian matrix $\widehat{H} \in \mathbb{C}_1^{n \times n}$ by $\widehat{H}_{ij} = \exp \iota \hat{\theta}_{ij}$, extract the top eigenvector $\tilde{x}$, and perform entrywise normalization to get $\hat{x} \in \mathbb{C}_1^n$

Multi-Frequency Phase Synchronization: Stage 2

Basically do generalized power iteration, but with a little extra work to ensure the $k_{\max}$ frequency channels "march in formation"

- For iteration counter $t = 0$, set $z^{(k,0)} := \hat{x}^k$ for $k = 1, \ldots, k_{\max}$
- Until convergence $(t = 1, \ldots, T)$,
  - $u^{(k,t)} := H^{(k)} z^{(k,t-1)}$ for $k = 1, \ldots, k_{\max}$
  - For each entry $1 \leq i \leq n$, soft-threshold the spectrogram

$$h_i^{(t)} := \frac{1}{2} \sum_{k=-k_{\max}}^{k_{\max}} u_i^{(k,t)} e^{\iota k \theta}$$

    then convert back

$$\hat{u}_i^{(k,t)} := \int_0^{2\pi} \mathrm{SoftThresh}\left(h_i^{(t)}(\theta)\right) \mathrm{d}\theta$$

    and finally get $z_i^{(k,t)} \in \mathbb{C}_1^n$ by entrywise normalization
- Output $\hat{z} := z^{(1,T)}$

How does it work? Check correlation $|\mathrm{Corr}\,(\hat{z}, z)|$.....

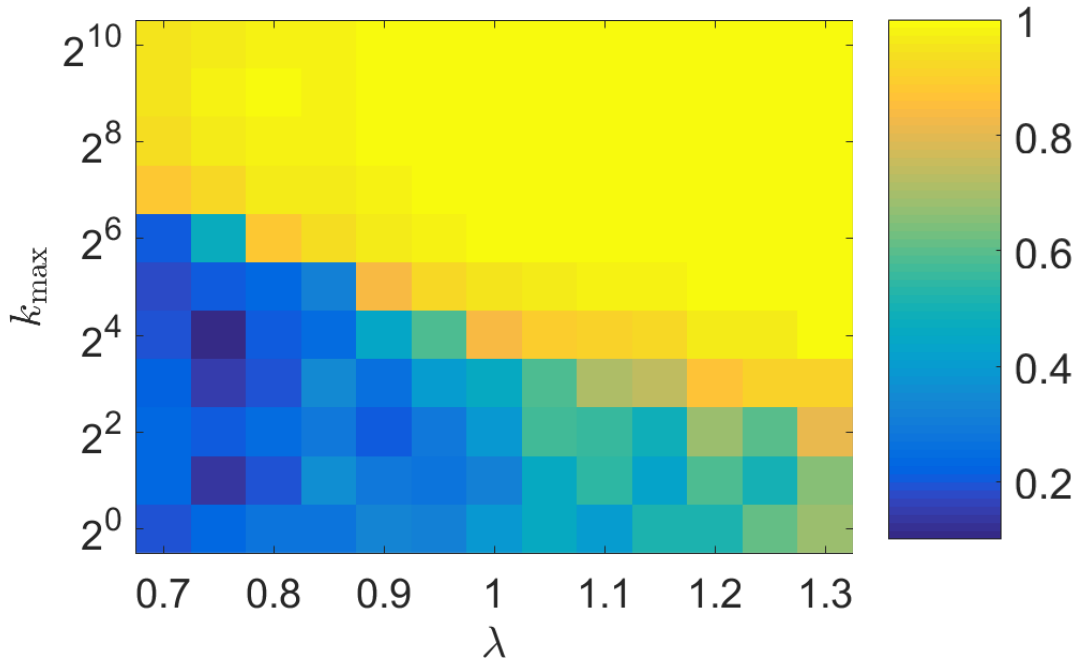


Figure: Random corruption model, $r = \lambda/\sqrt{n}$

- **[Singer (2011)]** only asserts $|\mathrm{Corr}\,(\hat{z}, z)| > 1/\sqrt{n}$ for $\lambda > 1$
- Even for $\lambda < 1$, our method attains $|\mathrm{Corr}\,(\hat{z}, z)| \to 1$ provided that $k_{\max} \gg 1$!

Outline

Motivation

- Synchronization
- Geometric Model of Fibre Bundles

Multi-Frequency Phase Synchronization

- Phase Synchronization
- Multi-Frequency Formulation
- **Theory**

Application to Multi-Frequency Class Averaging

Joint work with **Yifeng Fan (UIUC)** & **Zhizhen Zhao (UIUC)**

Focusing on Stage 1

- Fix $k_{\max} \geq 1$, build $H^{(2)}, \ldots, H^{(k_{\max})}$ out of $H = H^{(1)}$
- For each $k = 1, \ldots, k_{\max}$, solve the subproblem

$$u^{(k)} := \arg\max_{v \in \mathbb{C}_1^n} \; v^* H^{(k)} v$$

  using any convex relaxation, and set $W^{(k)} := u^{(k)} \left(u^{(k)}\right)^*$
- For all $1 \leq i, j \leq n$, find the "peak location" of the spectrogram

$$\hat{\theta}_{ij} := \arg\max_{\phi \in [0, 2\pi]} \left| \frac{1}{2} \sum_{k=-k_{\max}}^{k_{\max}} W_{ij}^{(k)} e^{-\iota k \phi} \right|$$

- Construct Hermitian matrix $\widehat{H} \in \mathbb{C}_1^{n \times n}$ by $\widehat{H}_{ij} = \exp \iota \hat{\theta}_{ij}$, extract the top eigenvector $\tilde{x}$, and perform entrywise normalization to get $\hat{x} \in \mathbb{C}_1^n$

Why does it work?

- By a perturbation analysis, we expect

$$\widehat{W}^{(k)} = u^{(k)}(u^{(k)})^* = z^k(z^k)^* + E^{(k)} \approx z^k(z^k)^*$$

from solving each subproblem

$$u^{(k)} := \arg\max_{v \in \mathbb{C}_1^n} \; v^* H^{(k)} v$$

- The peak finding step is expected to ensure

$$\begin{aligned}\hat{\theta}_{ij} &= \arg\max_{\phi \in [0,2\pi]} \left| \frac{1}{2} \sum_{k=1}^{k_{\max}} W_{ij}^{(k)} e^{-\iota k \phi} \right| \\ &\approx \arg\max_{\phi \in [0,2\pi]} \left| \frac{1}{2} \sum_{k=1}^{k_{\max}} z_i^k \bar{z}_j^k e^{-\iota k \phi} \right| = \theta_i - \theta_j\end{aligned}$$

provided that $E^{(k)}$ does not "perturb away" the maximum!

Landscape Analysis of the Dirichlet Kernel

$$\sum_{k=-k_{\max}}^{k_{\max}} W_{ij}^{(k)} e^{-\iota k \phi}$$

$$= \sum_{k=-k_{\max}}^{k_{\max}} e^{\iota k \left(\theta_i - \theta_j\right)} e^{-\iota k \phi} + \sum_{k=-k_{\max}}^{k_{\max}} E_{ij}^{(k)} e^{-\iota k \phi}$$

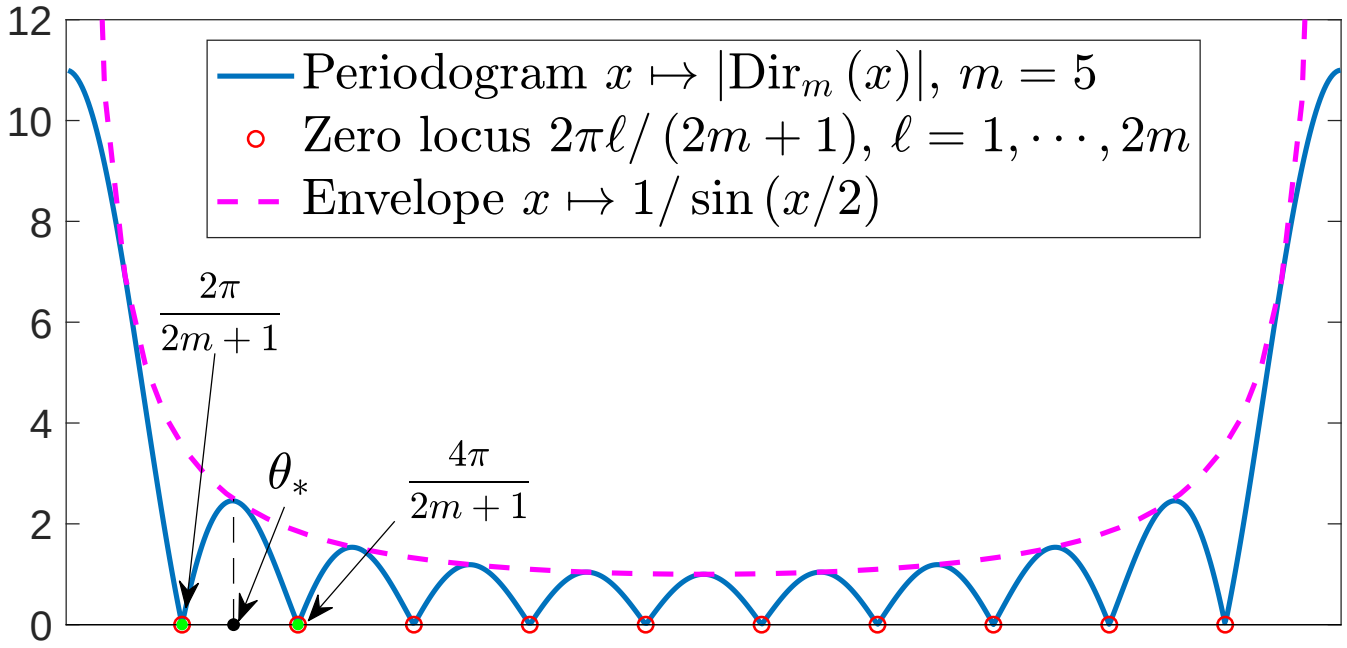
Landscape Analysis of the Dirichlet Kernel

$$
\begin{aligned}
&\sum_{k=-k_{\max}}^{k_{\max}} W_{ij}^{(k)} e^{-\iota k \phi} \\
&= \sum_{k=-k_{\max}}^{k_{\max}} e^{\iota k \left(\theta_i - \theta_j\right)} e^{-\iota k \phi} + \sum_{k=-k_{\max}}^{k_{\max}} E_{ij}^{(k)} e^{-\iota k \phi} \\
&= \frac{\sin\left[\left(k_{\max} + \frac{1}{2}\right)(\theta_i - \theta_j - \phi)\right]}{\sin\left[\frac{1}{2}(\theta_i - \theta_j - \phi)\right]} + \sum_{k=-k_{\max}}^{k_{\max}} E_{ij}^{(k)} e^{-\iota k \phi} \\
&=: \mathrm{Dir}_{k_{\max}}(\theta_i - \theta_j - \phi) + R_{k_{\max}}(\phi)
\end{aligned}
$$

Landscape Analysis of the Dirichlet Kernel

Key Observation: $\left|\hat{\theta}_{ij} - (\theta_i - \theta_j)\right| \leq \theta_* < \frac{4\pi}{2k_{\max}+1}$ whenever

$$2k_{\max} + 1 - \|R_{k_{\max}}\|_\infty > \frac{1}{\sin(\theta_*/2)} + \|R_{k_{\max}}\|_\infty$$



Landscape Analysis of the Dirichlet Kernel

Key Observation: $\left|\hat{\theta}_{ij} - (\theta_i - \theta_j)\right| \leq \theta_* < \frac{4\pi}{2k_{\max}+1}$ whenever

$$2k_{\max} + 1 - \|R_{k_{\max}}\|_\infty > \frac{1}{\sin(\theta_*/2)} + \|R_{k_{\max}}\|_\infty$$

which is satisfied if

$$2k_{\max} + 1 - \|R_{k_{\max}}\|_\infty > \frac{1}{\sin\left(\frac{\pi}{2k_{\max}+1}\right)} + \|R_{k_{\max}}\|_\infty$$

$$\Leftarrow \left\|E_{ij}^{(k)}\right\|_\infty \sim \frac{1}{k_{\max}} \|R_{k_{\max}}\|_\infty < 1 - \frac{1}{2k_{\max}\sin\left(\frac{\pi}{2k_{\max}+1}\right)}$$

We just need a uniform upper bound for the $E_{ij}^{(k)}$'s!

Eigenvector perturbation analysis in ℓ_∞-norm!

- Standard perturbation bounds use ℓ_2-norm (e.g. Davis–Kahan), but most application scenarios of spectral methods require bounding the ℓ_∞-norm
- Active research area in recent years, e.g. **[Eldridge et al. (2017)]; [Abbe et al. (2017)]; [Fan et al. (2018)]; [Zhong & Boumal (2018)]**
- Sharpest results to date use a "leave-one-out" trick popularized by statisticians

Lemma. For any $0 < \epsilon \leq 2$, with probability at least $1 - O\left(n^{-(2+\epsilon)}\right)$, there exists absolute constant $C > 0$ such that

$$\left\| E_{ij}^{(k)} \right\|_\infty \leq C \sqrt{\frac{\log n}{n}}.$$

• Tingran Gao & Zhizhen Zhao. "Multi-Frequency Phase Synchronization," *The 36th International Conference on Machine Learning* (ICML 2019). arXiv:1901.08235.

Putting Everything Together

Theorem (Gao & Zhao 2019). With all (mild) assumptions satisfied, with high probability the multi-frequency phase synchronization algorithm produces an estimate $\hat{z}$ satisfying

$$\mathrm{Corr}\left(\hat{z}, z\right) \geq 1 - \frac{C'}{k_{\max}^2}$$

provided that

$$k_{\max} > \max\left\{5, \frac{1}{\sqrt{2}\,\pi\left(1 - 4C_2\sigma\sqrt{\log n/n}\right) - 2}\right\}.$$

In particular, $\mathrm{Corr}\left(\hat{z}, z\right) \to 1$ as $k_{\max} \to \infty$.

• Tingran Gao & Zhizhen Zhao. "Multi-Frequency Phase Synchronization," *The 36th International Conference on Machine Learning* (ICML 2019). arXiv:1901.08235.

Outline

Motivation

- Synchronization

Multi-Frequency Phase Synchronization

- Phase Synchronization
- Multi-Frequency Formulation
- Theory

Application to Multi-Frequency Class Averaging

Joint work with **Yifeng Fan (UIUC)** & **Zhizhen Zhao (UIUC)**

Class Averaging and the Intrinsic Model

- **Goal:** Classify cryo-EM images $I_1, I_2, \ldots$ according to their viewing directions
- **Procedure:**
  - Compute the *rotation-invariant distance* between all pairs of images $d_{\mathrm{RID}}(I_i, I_j) := \min_{\alpha \in [0,2\pi]} \|I_i - e^{\iota\alpha} I_j\|_{\mathrm{F}}$, and denote α_{ij} for the optimal alignment angle
  - Fix threshold $\epsilon > 0$ and define Hermitian $W \in \mathbb{C}^{n \times n}$ by

$$W_{ij} := \begin{cases} \exp(\iota\alpha_{ij}) & \text{if } d_{\mathrm{RID}}(I_i, I_j) < \epsilon \\ 0 & \text{otherwise} \end{cases}$$

  - Solve for the top 3 eigenvectors ψ_1, ψ_2, ψ_3 of W, which embeds $I_1, I_2, \ldots$ into $\mathbb{R}^3$
  - Use correlation in the embedded $\mathbb{R}^3$ space to determine the closeness between viewing directions

Multi-Frequency Class Averaging

- Idea directly descended from the multi-frequency phase synchronization, but the underlying fibre bundle is non-trivial!
- Fix threshold $\epsilon > 0$ and define for each $k = 1, \ldots, k_{\max}$ Hermitian $W^{(k)} \in \mathbb{C}^{n \times n}$ by

$$W_{ij}^{(k)} := \begin{cases} \exp\left(\iota k \alpha_{ij}\right) & \text{if } d_{\mathrm{RID}}\left(I_i, I_j\right) < \epsilon \\ 0 & \text{otherwise} \end{cases}$$

- Solve for the top $2k+1$ eigenvectors $\psi_1^{(k)}, \cdots, \psi_{2k+1}^{(k)}$ of $W^{(k)}$, which embeds $I_1, I_2, \ldots$ into $\mathbb{R}^{2k+1}$
- Use all correlations in the embedded $\mathbb{R}^{2k+1}$ $(k = 1, \ldots, k_{\max})$ spaces to determine the closeness between viewing directions

• Yifeng Fan & Zhizhen Zhao. "Cryo-Electron Microscopy Image Analysis Using Multi-Frequency Vector Diffusion Maps," *preprint.* arXiv:1904.07772.

Why $2k+1$? Some Representation Theoretic Patterns

- **(G., Fan, Zhao 2019)** For sufficiently large sample size n and appropriately small $\epsilon > 0$, the top eigenspace of $W^{(k)}$ is $(2k+1)$-dimensional, and the spectral gap grows linearly in k:

$$\lambda_k^{(k)} - \lambda_{k+1}^{(k)} \sim \frac{1+k}{4}\epsilon^2$$

- Larger $k \Rightarrow$ larger spectral gap $\Rightarrow$ better numerical stability!

• Tingran Gao, Yifeng Fan, Zhizhen Zhao. "Representation Theoretic Patterns in Multi-Frequency Class Averaging for Three-Dimensional Cryo-Electron Microscopy," *preprint*.

More Representation Theoretic Patterns.....

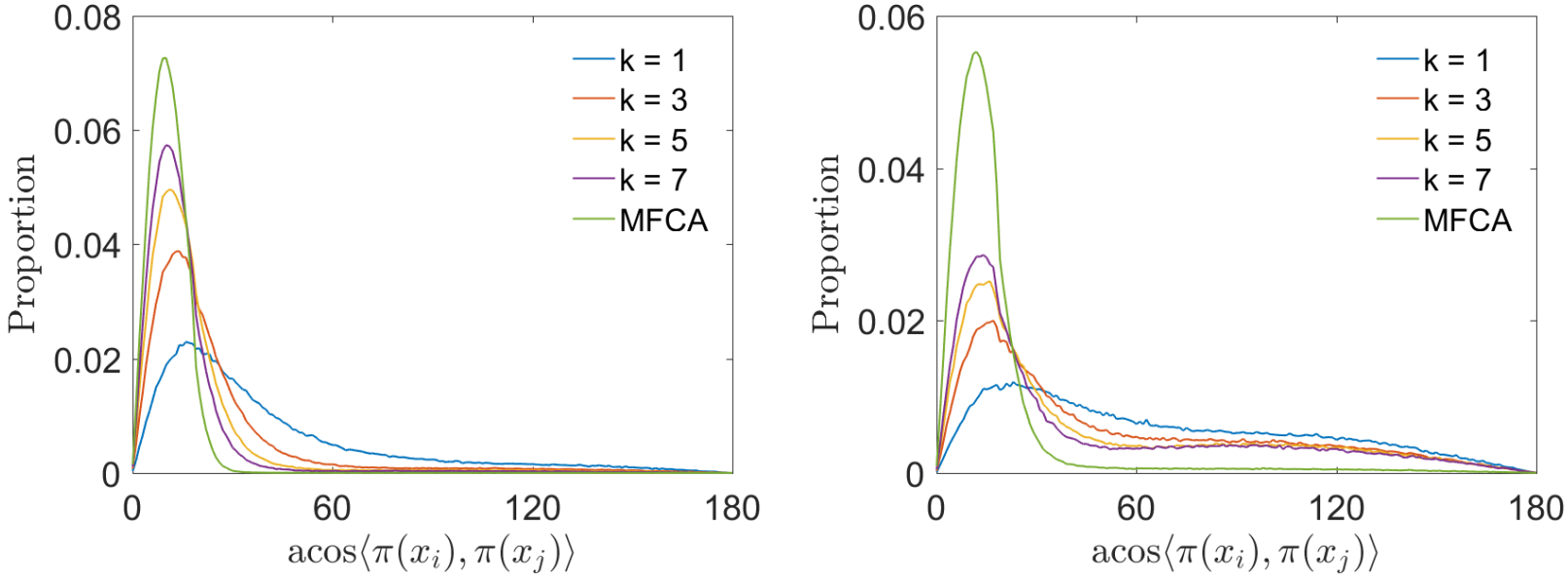
- **(G., Fan, Zhao 2019)** The viewing angle θ_{ij} between image i and image j satisfies

$$\left|\sum_{\ell=1}^{2k+1} \psi_\ell^{(k)}(i)\,\overline{\psi_\ell^{(k)}(j)}\right| = \left(\frac{1+\cos\theta_{ij}}{2}\right)^k$$

- Larger $k \Rightarrow$ easier thresholding
- Can also jointly use $k = 1, \cdots, k_{\max}$ to construct polynomial filters for $\cos\theta_{ij}$

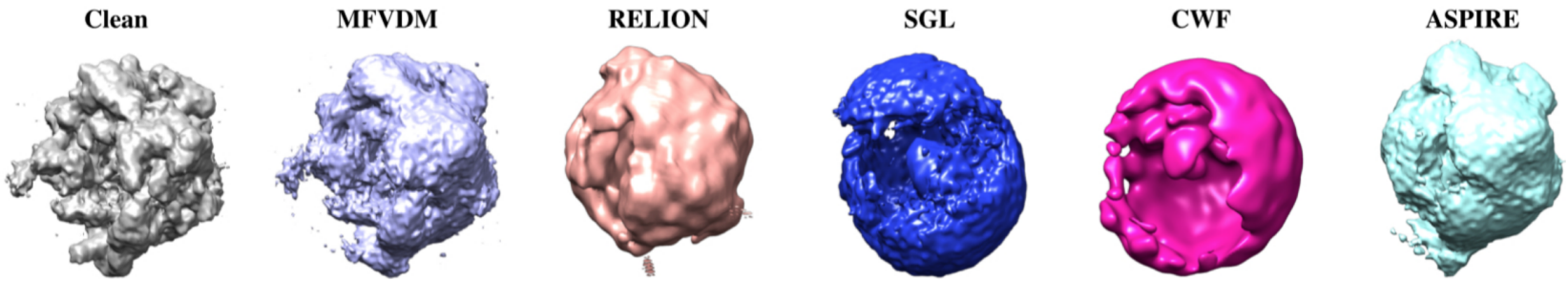
• Tingran Gao, Yifeng Fan, Zhizhen Zhao. "Representation Theoretic Patterns in Multi-Frequency Class Averaging for Three-Dimensional Cryo-Electron Microscopy," *preprint*.

Multi-Frequency Information Improves Class Averaging



• Tingran Gao, Yifeng Fan, Zhizhen Zhao. "Representation Theoretic Patterns in Multi-Frequency Class Averaging for Three-Dimensional Cryo-Electron Microscopy," *preprint*.

Multi-Frequency Information Improves Class Averaging

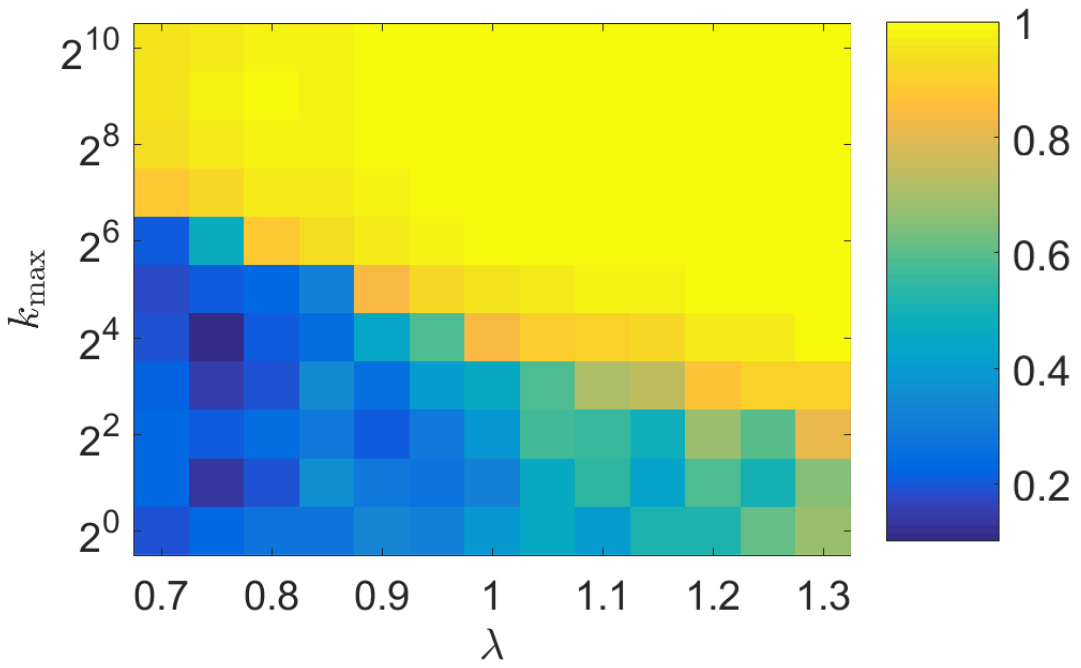


• Yifeng Fan & Zhizhen Zhao. "Cryo-Electron Microscopy Image Analysis Using Multi-Frequency Vector Diffusion Maps," *preprint*. arXiv:1904.07772.
• Tingran Gao, Yifeng Fan, Zhizhen Zhao. "Representation Theoretic Patterns in Multi-Frequency Class Averaging for Three-Dimensional Cryo-Electron Microscopy," *preprint*.

Open Problems

- Landscape and convergence analysis for the Stage 2 algorithm?
- Algorithm works for $\mathrm{SO}(3)$ seamlessly, but theory?
- Possible extension to other synchronization and multireference alignment problems over compact/noncompact Lie groups?
- Mult-frequency vector diffusion maps?
- A learning paradigm on sheaves?

Thank You!



• Tingran Gao & Zhizhen Zhao. "Multi-Frequency Phase Synchronization," *The 36th International Conference on Machine Learning* (ICML 2019). arXiv:1901.08235.
• Tingran Gao, Yifeng Fan, Zhizhen Zhao. "Representation Theoretic Patterns in Multi-Frequency Class Averaging for Three-Dimensional Cryo-Electron Microscopy," *preprint*.