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

Committee on Computational and Applied Mathematics
Department of Statistics
University of Chicago

IPAM Geometry and Learning Culminating Workshop
Lake Arrowhead CA

June 11, 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**

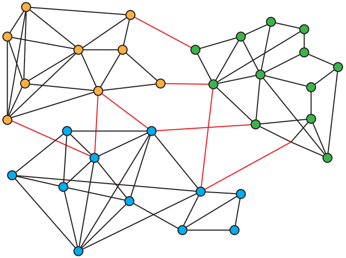
Multi-Frequency Phase Synchronization

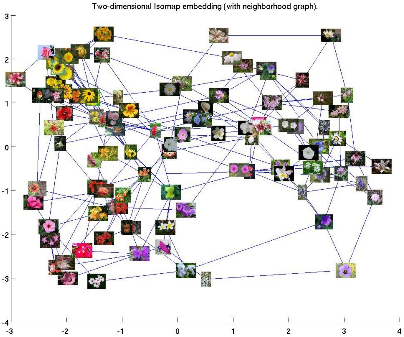
- Phase Synchronization
- Multi-Frequency Formulation
- Theory

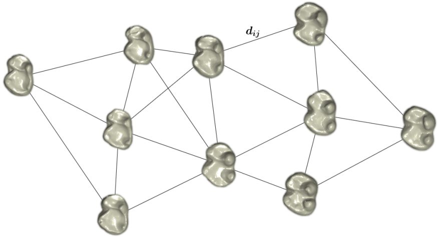
Application to Multi-Frequency Class Averaging

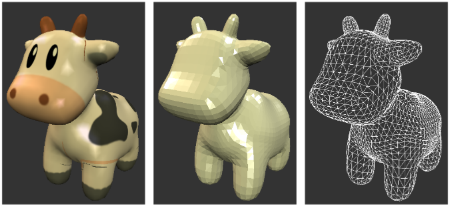
Joint work with **Yifeng Fan (UIUC)** & **Zhizhen Zhao (UIUC)**

Graph-Based Data Analysis





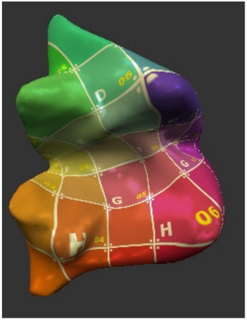




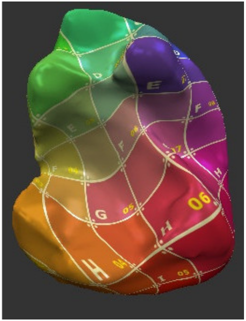
Non-Scalar Edge Weights

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

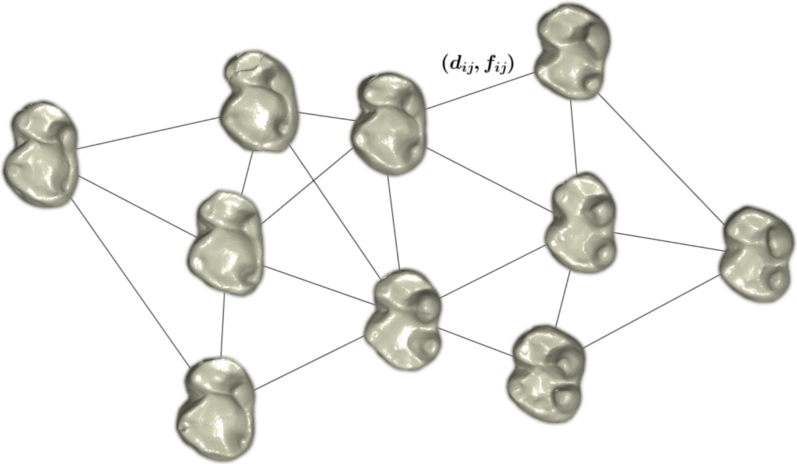
S_i



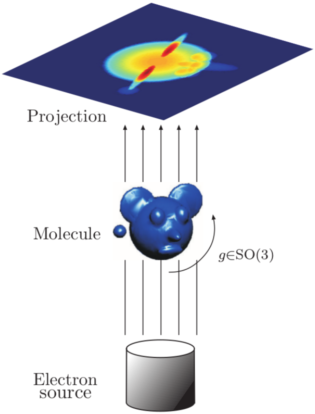
$\xrightarrow[f_{ij}]{d_{ij}}$

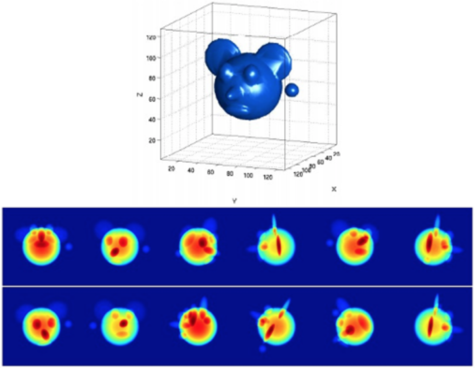


S_j


(d_{ij}, f_{ij})

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

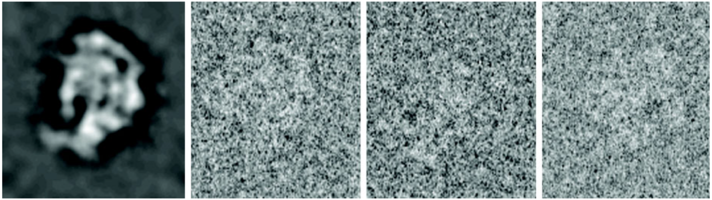


Fig. 3 The left most image is a clean simulated projection image of the E.coli 50S ribosomal subunit. The other three images are real electron microscope images of the same subunit

It is imperative to apply **class averaging** in preprocessing!

• 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).

Synchronization Problems

- Data:
  - graph $\Gamma = (V, E)$
  - topological group G, equipped with a norm $\|\cdot\|$, and a G-module F
  - **edge potential** $\rho : E \to G$ satisfying $\rho_{ij} = \rho_{ji}^{-1}, \ \forall\,(i,j) \in E$
- Goal:
  - find a **vertex potential** $f : V \to G$ or F such that

$$f_i = \rho_{ij} f_j, \quad \forall\,(i,j) \in E$$

- The goal can be achieved if and only if $\rho_{ij} = f_i f_j^{-1}$
- **Not** always feasible!

Synchronization Problems

- Data:
  - graph $\Gamma = (V, E)$
  - topological group G, equipped with a norm $\|\cdot\|$, and a G-module F
  - **edge potential** $\rho : E \to G$ satisfying $\rho_{ij} = \rho_{ji}^{-1}, \ \forall (i,j) \in E$
- Goal:
  - find a **vertex potential** $f : V \to G$ or F such that

$$f_i = \rho_{ij} f_j, \quad \forall (i,j) \in E$$

- The goal can be achieved if and only if $\rho_{ij} = f_i f_j^{-1}$
- **Not** always feasible!
- If infeasible, find the "closest solution" in the sense of

$$\min_{\substack{f:V\to G \\ \|f\|\neq 0}} \frac{1}{2} \frac{\sum_{i,j\in V} \|f_i - \rho_{ij} f_j\|^2}{\sum_{i\in V} \|f_i\|^2} (=: \eta(f))$$

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)**

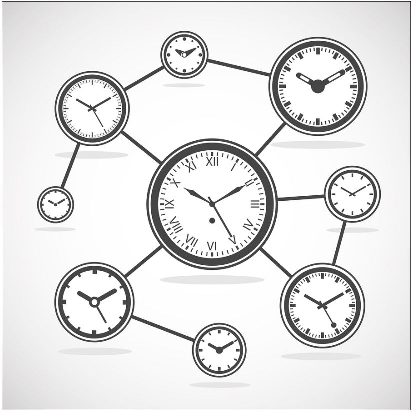
Phase Synchronization

- **Problem:** Recover rotation angles $\theta_1, \ldots, \theta_n \in [0, 2\pi]$ from noisy measurements of their pairwise offsets

$$\theta_{ij} = \theta_i - \theta_j + \text{noise}$$

for some or all pairs of (i, j)

- **Examples:** Class averaging in cryo-EM image analysis, shape registration and community detection



Phase Synchronization

- **Setup**: Phase vector $z = \left(e^{\iota\theta_1}, \ldots, e^{\iota\theta_n}\right)^\top \in \mathbb{C}_1^n$, noisy pairwise measurements in n-by-n Hermitian matrix

$$H_{ij} = \begin{cases} e^{\iota\left(\theta_i - \theta_j\right)} = z_i \bar{z}_j & \text{with prob. } r \in [0,1] \\ \mathrm{Uniform}\left(\mathbb{C}_1\right) & \text{with prob. } 1 - r \end{cases}$$

  and $H_{ij} = \overline{H_{ji}}$. This is known as a random corruption model.
- **Goal**: recover the true phase vector z (up to a global multiplicative factor)
- **Existing method**: Rank-1 recovery (e.g. convex relaxations)

$$\hat{x} := \operatorname*{arg\,min}_{x \in \mathbb{C}_1^n} \|xx^* - H\|_{\mathrm{F}}^2 \quad \Leftrightarrow \quad \hat{x} := \operatorname*{arg\,max}_{x \in \mathbb{C}_1^n} x^* H x$$

Phase Synchronization

- **Setup**: Phase vector $z = \left(e^{\iota\theta_1}, \dots, e^{\iota\theta_n}\right)^\top \in \mathbb{C}_1^n$, noisy pairwise measurements in n-by-n Hermitian matrix

$$H_{ij} = \begin{cases} e^{\iota\left(\theta_i - \theta_j\right)} = z_i \bar{z}_j & \text{with prob. } r \in [0,1] \\ \mathrm{Uniform}\left(\mathbb{C}_1\right) & \text{with prob. } 1 - r \end{cases}$$

  and $H_{ij} = \overline{H_{ji}}$. This is known as a random corruption model.
- **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

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

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)**

Multi-Frequency Phase Synchronization

- **Multi-Frequency Formulation:**

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

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 $H_{ij}^{(k)} := H_{ij}^k$

- **Intuition:** Matching higher trigonometric moments
- **Two-stage Algorithm:** (i) Good initialization (ii) Local methods e.g. gradient descent or (generalized) power iteration

Initialization: Inspired by Harmonic Retrieval

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

Initialization: Inspired by Harmonic Retrieval

- 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

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

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

Initialization: Inspired by Harmonic Retrieval

- 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

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

  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

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

Initialization: Inspired by Harmonic Retrieval

- 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

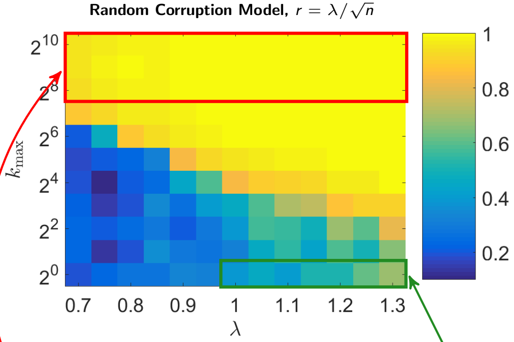
$$u^{(k)} := \arg\max_{v \in \mathbb{C}_1^n} \; v^* H^{(k)} v$$

  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

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

- Entrywise normalize the top eigenvector $\tilde{x}$ of Hermitian matrix $\widehat{H}$, defined by $\widehat{H}_{ij} = e^{\iota \hat{\theta}_{ij}}$, to get $\hat{x} \in \mathbb{C}_1^n$

How well does it work? Evaluate correlation $|\mathrm{Corr}(\hat{x}, z)|$



Our Method: $|\mathrm{Corr}(\hat{x}, z)| \longrightarrow 1$ as $k_{\max} \gg 1$, **even for** $\lambda < 1$!

Previous Art: Only ensures $|\mathrm{Corr}(\hat{x}, z)| > \frac{1}{\sqrt{n}}$ **for** $\lambda > 1$

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)**

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**

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

Previous Theory — Weak Recovery

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

- Nonconvex problem, can be attacked using convex (spectral or SDP) relaxation or nonconvex methods
- 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 $r > 1/\sqrt{n}$

Theory Now — Strong Recovery

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

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

provided that

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

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

- Tingran Gao and Zhizhen Zhao, "Multi-Frequency Phase Synchronization." Proceedings of the 36th International Conference on Machine Learning, PMLR 97:2132–2141, 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)**

Class Averaging

- **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 $d_{\mathrm{RID}}(I_i, I_j) := \min_{\alpha \in [0,2\pi]} \left\| I_i - e^{\iota\alpha} I_j \right\|_{\mathrm{F}}$, and denote α_{ij} for the optimal alignment angle
  - Fix threshold $\epsilon > 0$ and define Hermitian $W \in \mathbb{C}^{n \times n}$ by

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

  - 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

• Tingran Gao, Yifeng Fan, Zhizhen Zhao. "Representation Theoretic Patterns in Multi-Frequency Class Averaging for Three-Dimensional Cryo-Electron Microscopy," *preprint.* arxiv:1906.01082.

Multi-Frequency Class Averaging

- **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

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

  - 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

• Tingran Gao, Yifeng Fan, Zhizhen Zhao. "Representation Theoretic Patterns in Multi-Frequency Class Averaging for Three-Dimensional Cryo-Electron Microscopy," *preprint*. arxiv:1906.01082.

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*. arxiv:1906.01082.

More Representation Theoretic Patterns......

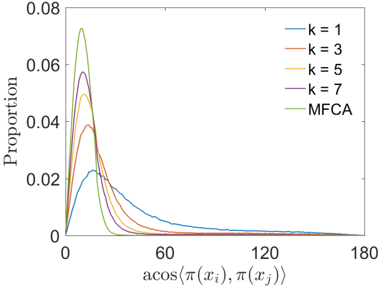
- **(G., Fan, Zhao 2019)** The viewing angle θ_{ij} between image i and image j satisfies

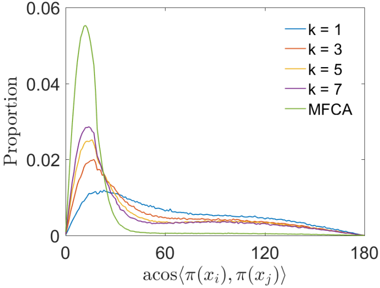
$$\left|\sum_{\ell=1}^{2k+1} \psi_\ell^{(k)}(i)\,\overline{\psi_\ell^{(k)}(j)}\right| = \left(\frac{1+\cos\theta_{ij}}{2}\right)^k$$

- 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*. arxiv:1906.01082.

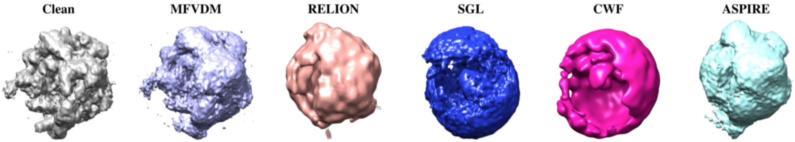
Multi-Frequency Information Improves Class Averaging





- Tingran Gao, Yifeng Fan, Zhizhen Zhao. "Representation Theoretic Patterns in Multi-Frequency Class Averaging for Three-Dimensional Cryo-Electron Microscopy," *preprint.* arxiv:1906.01082.

Multi-Frequency Information Improves Class Averaging

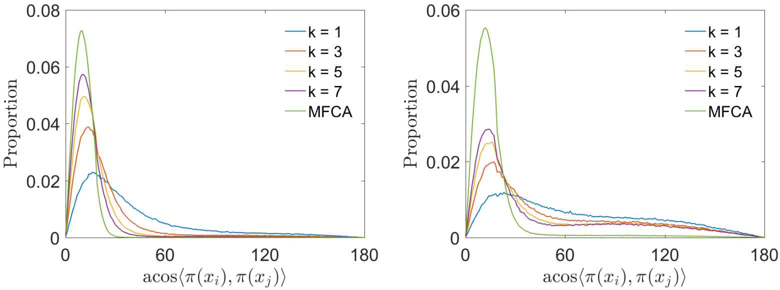


• Yifeng Fan & Zhizhen Zhao. "Cryo-Electron Microscopy Image Analysis Using Multi-Frequency Vector Diffusion Maps," *preprint.* arXiv:1904.07772.
• Tingran Gao, Yifeng Fan, Zhizhen Zhao. "Representation Theoretic Patterns in Multi-Frequency Class Averaging for Three-Dimensional Cryo-Electron Microscopy," *preprint.* arxiv:1906.01082.

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 and Zhizhen Zhao, "Multi-Frequency Phase Synchronization." Proceedings of the 36th International Conference on Machine Learning, PMLR 97:2132–2141, 2019.
- Tingran Gao, Yifeng Fan, and Zhizhen Zhao. "Representation Theoretic Patterns in Multi-Frequency Class Averaging for Three-Dimensional Cryo-Electron Microscopy," *arxiv:1906.01082*.