Irreducible Representations and Multi-Frequency Angular Synchronization

Tingran Gao

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

2019 Joint Statistical Meetings (ASA)
Denver, CO

July 28, 2019

Outline

Motivation

- Class Averaging and Phase Synchronization

Multi-Frequency Phase Synchronization

- Multi-Frequency Formulation
- Theory

Application to Multi-Frequency Class Averaging

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

Outline

Motivation

- **Class Averaging and Phase Synchronization**

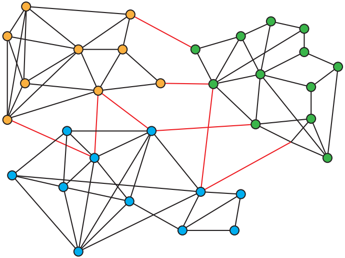
Multi-Frequency Phase Synchronization

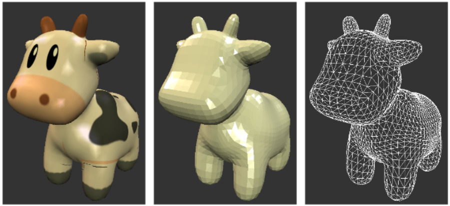
- Multi-Frequency Formulation
- Theory

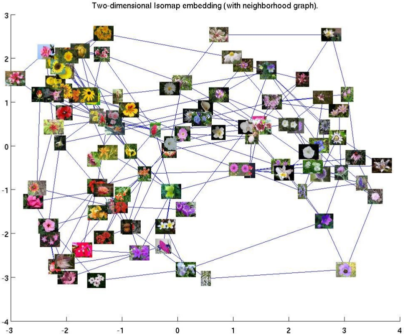
Application to Multi-Frequency Class Averaging

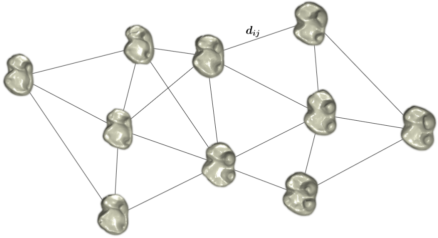
Joint work with **Yifeng Fan (UIUC)** & **Zhizhen Zhao (UIUC)**

Graph-Based Data Analysis





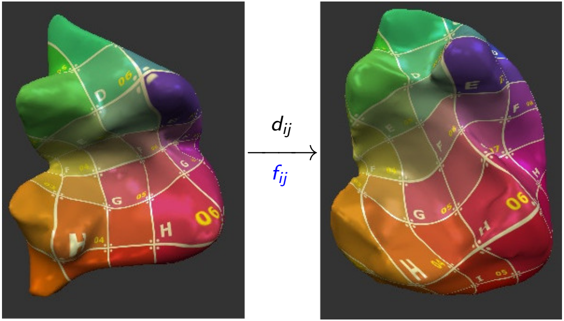




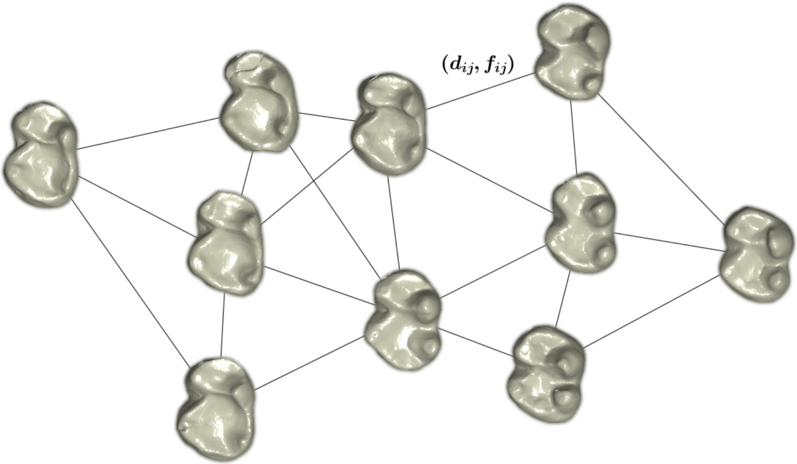
Non-Scalar Edge Weights

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

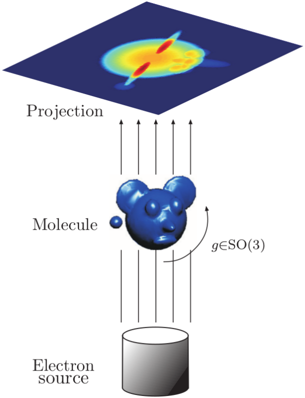
S_i

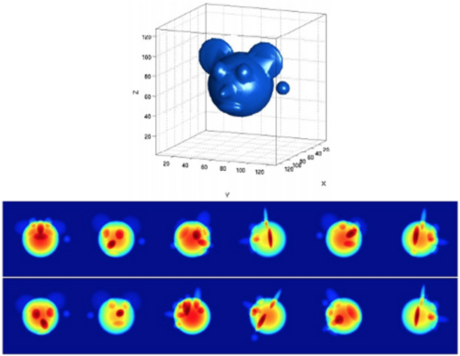


S_j


(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

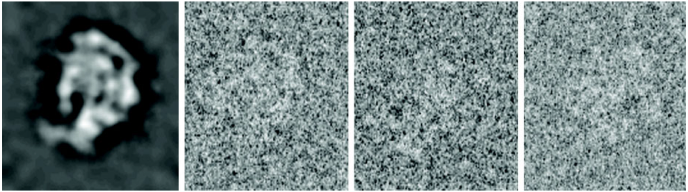


Fig. 3 The left most image is a clean simulated projection image of the E.coli 50S ribosomal subunit. The other three images are real electron microscope images of the same subunit

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).

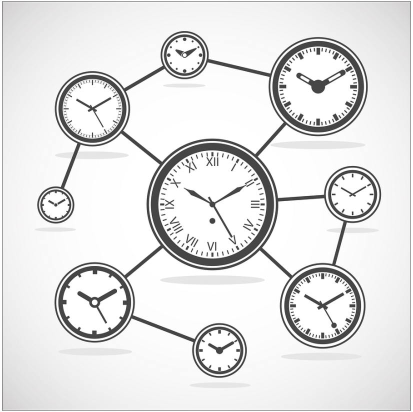
Phase Synchronization

- **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

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

  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)

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

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

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

  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

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

Outline

Motivation

- Class Averaging and Phase Synchronization

Multi-Frequency 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: Main Idea

- The rank-1 recovery formulation
$$\hat{x} := \arg\min_{x \in \mathbb{C}_1^n} \|xx^* - H\|_{\mathrm{F}}^2 \quad \Leftrightarrow \quad \hat{x} := \arg\max_{x \in \mathbb{C}_1^n} x^* H x$$
does not fully exploit that entries of x and H are **phases**!

Multi-Frequency Phase Synchronization: Main Idea

- The rank-1 recovery formulation
$$\hat{x} := \arg\min_{x \in \mathbb{C}_1^n} \|xx^* - H\|_{\mathrm{F}}^2 \quad \Leftrightarrow \quad \hat{x} := \arg\max_{x \in \mathbb{C}_1^n} x^* H x$$
does not fully exploit that entries of x and H are **phases**!
- **Key Observation:** Raising a phase to any power yields another phase! $e^{\iota\theta} \longmapsto e^{\iota k\theta}, \quad k = 1, 2, \ldots$

Multi-Frequency Phase Synchronization: Main Idea

- The rank-1 recovery formulation
$$\hat{x} := \arg\min_{x\in\mathbb{C}_1^n} \|xx^* - H\|_{\mathrm{F}}^2 \quad \Leftrightarrow \quad \hat{x} := \arg\max_{x\in\mathbb{C}_1^n} x^* H x$$
does not fully exploit that entries of x and H are **phases**!
- **Key Observation:** Raising a phase to any power yields another phase! $e^{\iota\theta} \longmapsto e^{\iota k\theta}, \quad k = 1, 2, \ldots$
- Solve a family of coupled matrix factorization problems jointly
$$\hat{x}^{(k)} := \arg\max_{x\in\mathbb{C}_1^n} (x^k)^* H^{(k)} x^k, \qquad k = 1, 2, \ldots, k_{\max}$$
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 $H_{ij}^{(k)} := H_{ij}^k$, and then "stitch up" the individual estimates $\hat{x}^{(1)}, \ldots, \hat{x}^{(k_{\max})}$ to recover $\hat{x}$

Multi-Frequency Phase Synchronization: Main Idea

- The rank-1 recovery formulation
$$\hat{x} := \arg\min_{x \in \mathbb{C}_1^n} \|xx^* - H\|_{\mathrm{F}}^2 \quad \Leftrightarrow \quad \hat{x} := \arg\max_{x \in \mathbb{C}_1^n} x^* H x$$
does not fully exploit that entries of x and H are **phases**!
- **Key Observation:** Raising a phase to any power yields another phase! $e^{\iota\theta} \longmapsto e^{\iota k\theta}, \quad k = 1, 2, \ldots$
- Solve a family of coupled matrix factorization problems jointly
$$\hat{x}^{(k)} := \arg\max_{x \in \mathbb{C}_1^n} (x^k)^* H^{(k)} x^k, \qquad k = 1, 2, \ldots, k_{\max}$$
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 $H_{ij}^{(k)} := H_{ij}^k$, and then "stitch up" the individual estimates $\hat{x}^{(1)}, \ldots, \hat{x}^{(k_{\max})}$ to recover $\hat{x}$
- The last step strives to recover θ from (noisy) measured phases $e^{\iota\theta}, \ldots, e^{\iota k_{\max}\theta}$, which is a simplified version of the **harmonic retrieval** problem

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

Initialization: Inspired by Harmonic Retrieval

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

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

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

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

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

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

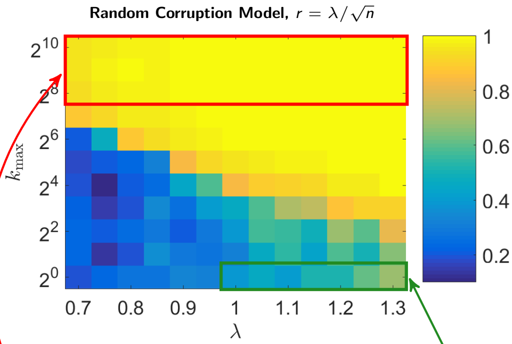
$$u^{(k)} := \arg\max_{v \in \mathbb{C}_1^n} \; v^* H^{(k)} v$$

  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

- Class Averaging and Phase Synchronization

Multi-Frequency Phase Synchronization

- Multi-Frequency Formulation
- **Theory**

Application to Multi-Frequency Class Averaging

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

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.

Why does it work?

- By a perturbation analysis, after solving each subproblem

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

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

- 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

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

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

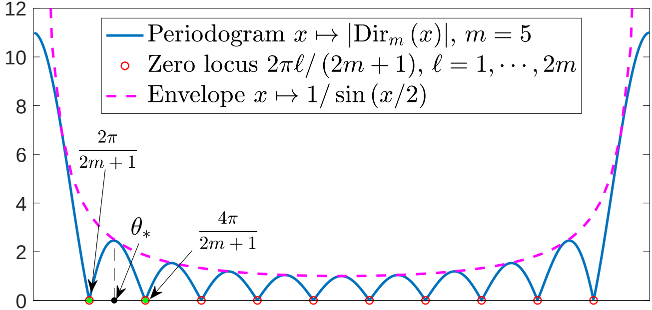
Landscape Analysis of the Dirichlet Kernel

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

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 and Zhizhen Zhao, "Multi-Frequency Phase Synchronization." Proceedings of the 36th International Conference on Machine Learning, PMLR 97:2132–2141, 2019.

Putting Everything Together

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

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

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

• Tingran Gao and Zhizhen Zhao, "Multi-Frequency Phase Synchronization." Proceedings of the 36th International Conference on Machine Learning, PMLR 97:2132–2141, 2019.

Detour: Multi-Frequency Synchronization over $\mathrm{SO}(3)$

Peter–Weyl: $f(g) = \sum_{k=0}^{\infty} d_k \mathrm{Tr}\left[\hat{f}(k)\,\rho_k(g)\right], \ \forall f \in L^2(\mathrm{SO}(3))$

- Random corruption model:

$$g_{ij} = \begin{cases} g_i g_j^{-1} & \text{with probability } r \\ \mathrm{Unif}\left(\mathrm{SO}(3)\right) & \text{with probability } 1-r \end{cases}$$

- Use spectral methods to estimate $\rho_1(g_{ij}), \ldots, \rho_{k_{\max}}(g_{ij})$, denote $\widehat{H}_{ij}^{(k)}$ for the estimator of $\rho_k(g_{ij})$
- Solve a generalized harmonic retrieval problem on $\mathrm{SO}(3)$:

$$\hat{g}_{ij} = \operatorname*{arg\,max}_{g \in \mathrm{SO}(3)} \sum_{k=1}^{k_{\max}} d_k \mathrm{Tr}\left[\widehat{H}_{ij}^{(k)} \rho_k^*(g)\right]$$

- Works fantastic in practice, but no theory yet!

Outline

Motivation

- Class Averaging and Phase Synchronization

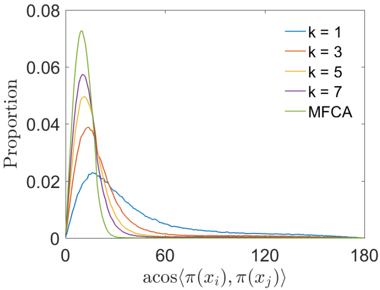
Multi-Frequency Phase Synchronization

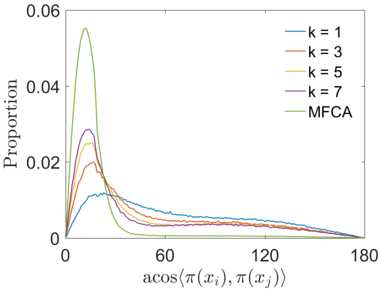
- Multi-Frequency Formulation
- Theory

Application to Multi-Frequency Class Averaging

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

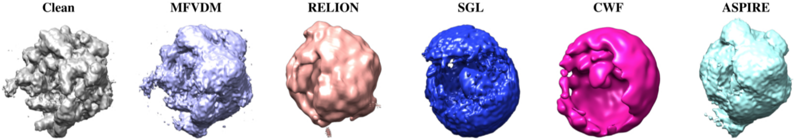
Multi-Frequency Information Improves Class Averaging





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

Multi-Frequency Information Improves Class Averaging

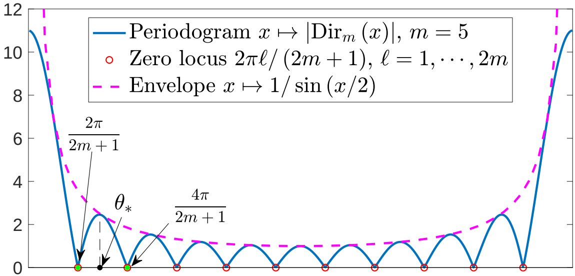


- Yifeng Fan & Zhizhen Zhao. "Cryo-Electron Microscopy Image Analysis Using Multi-Frequency Vector Diffusion Maps," *preprint.* arXiv:1904.07772.
- Tingran Gao, Yifeng Fan, Zhizhen Zhao. "Representation Theoretic Patterns in Multi-Frequency Class Averaging for Three-Dimensional Cryo-Electron Microscopy,"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



Center for Data and Computing

- 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," *submitted.* arxiv:1906.01082.