Multi-Representation Manifold Learning on Fibre Bundles

Tingran Gao

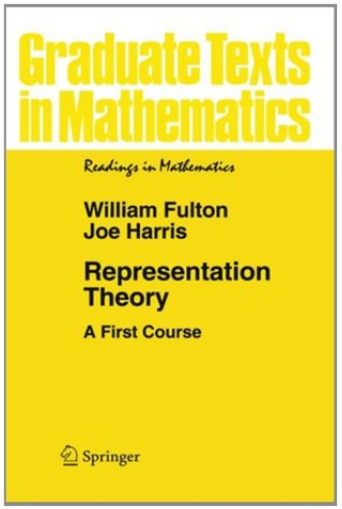
Committee on Computational and Applied Mathematics
Department of Statistics
University of Chicago

Computational Harmonic Analysis and Data Science
Casa Matemática Oaxaca
Oaxaca, Mexico

October 28, 2019

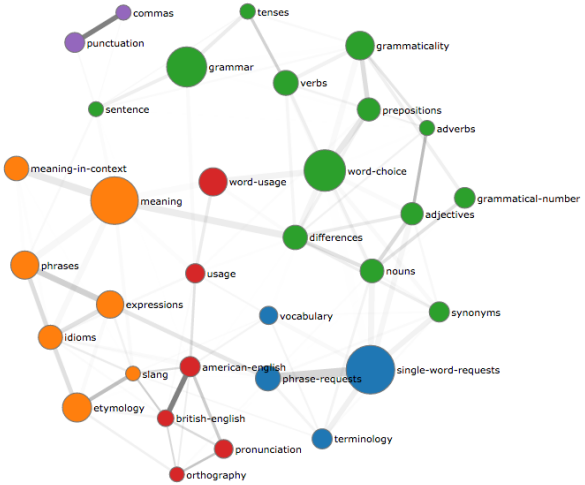
Outline

Representation?



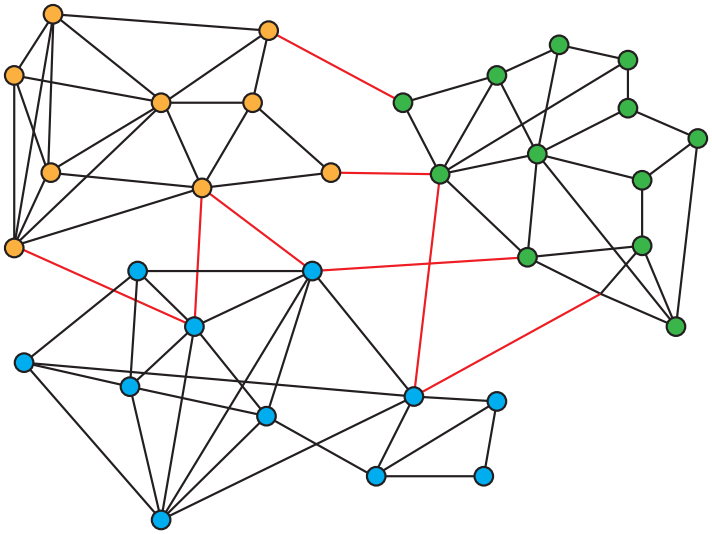
representation theory

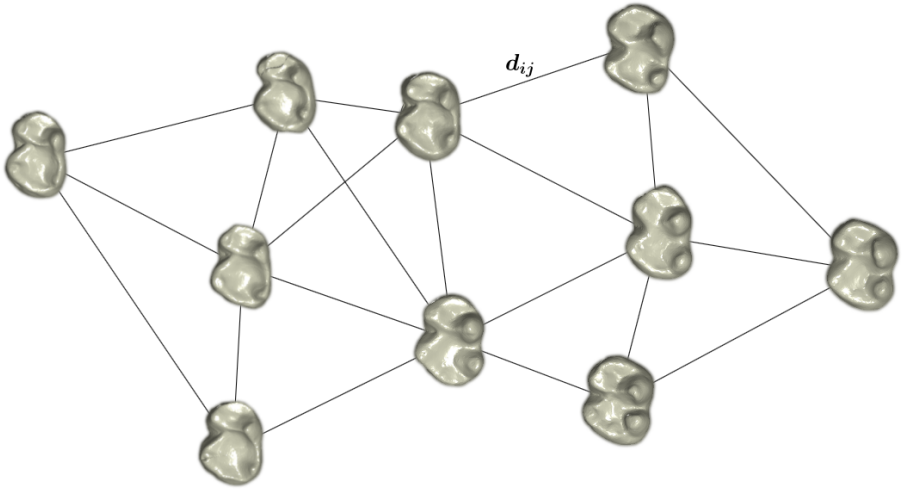
representationism

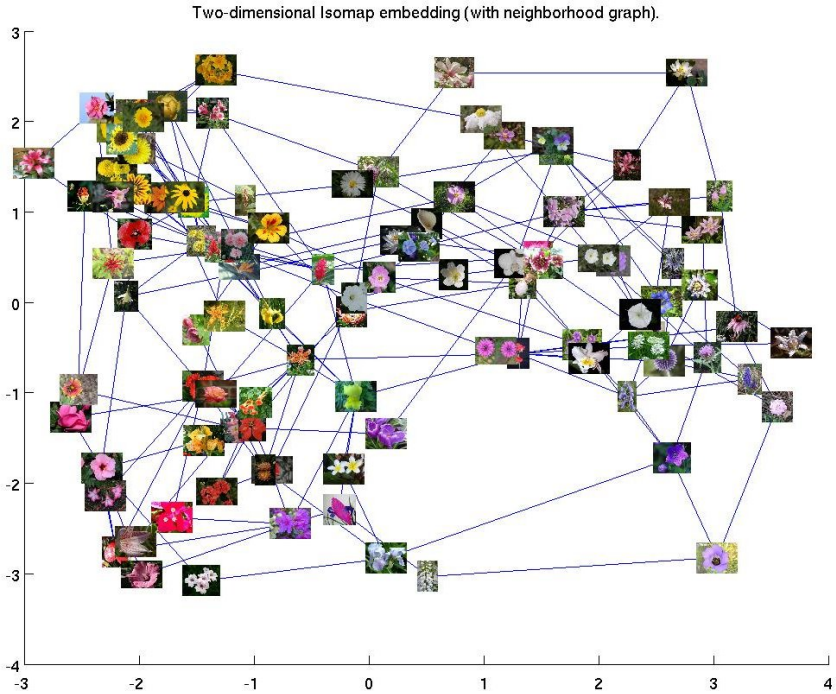


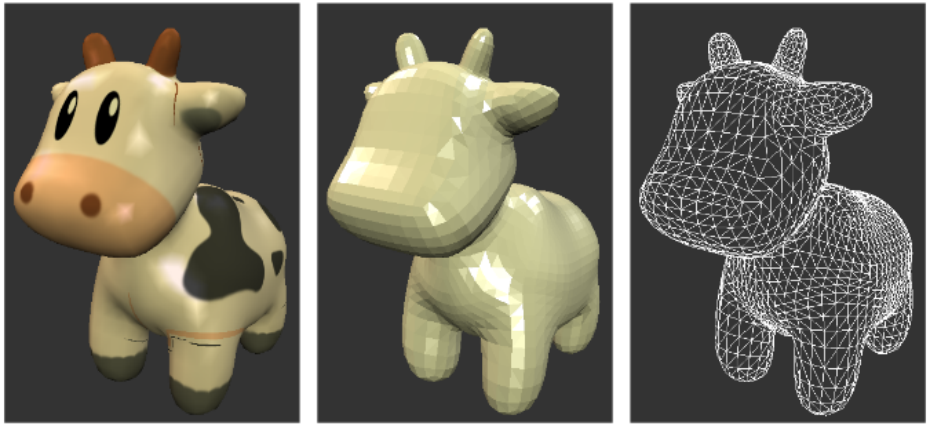
representation learning

Graph: A Flexible Data Representation





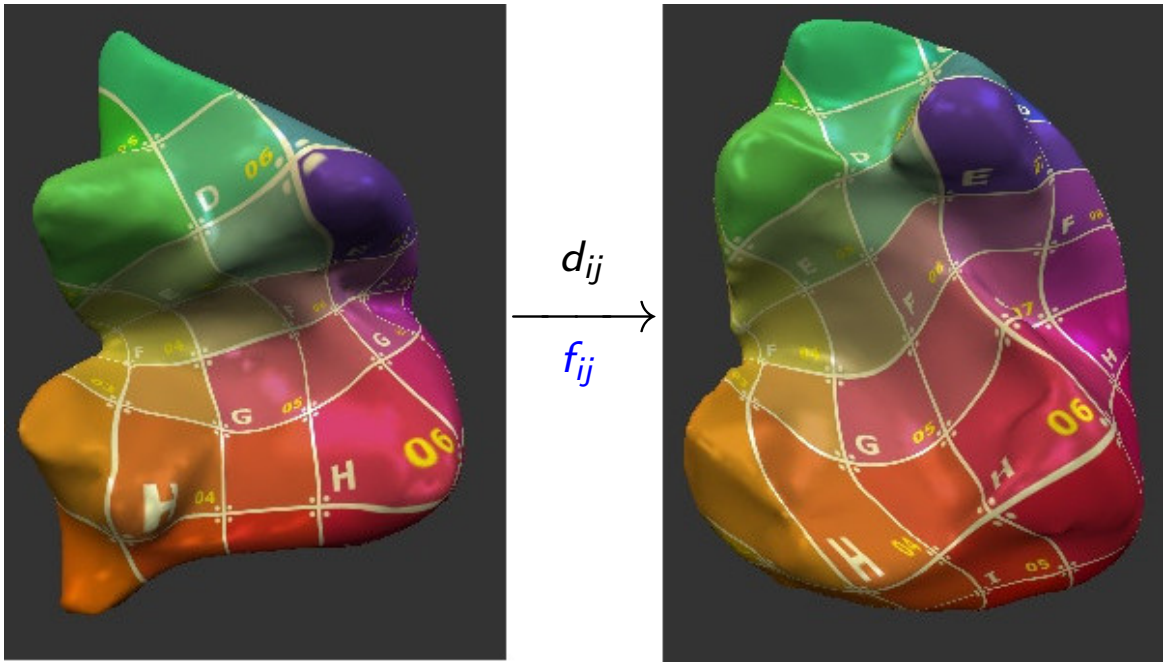




Non-Scalar Edge Weights

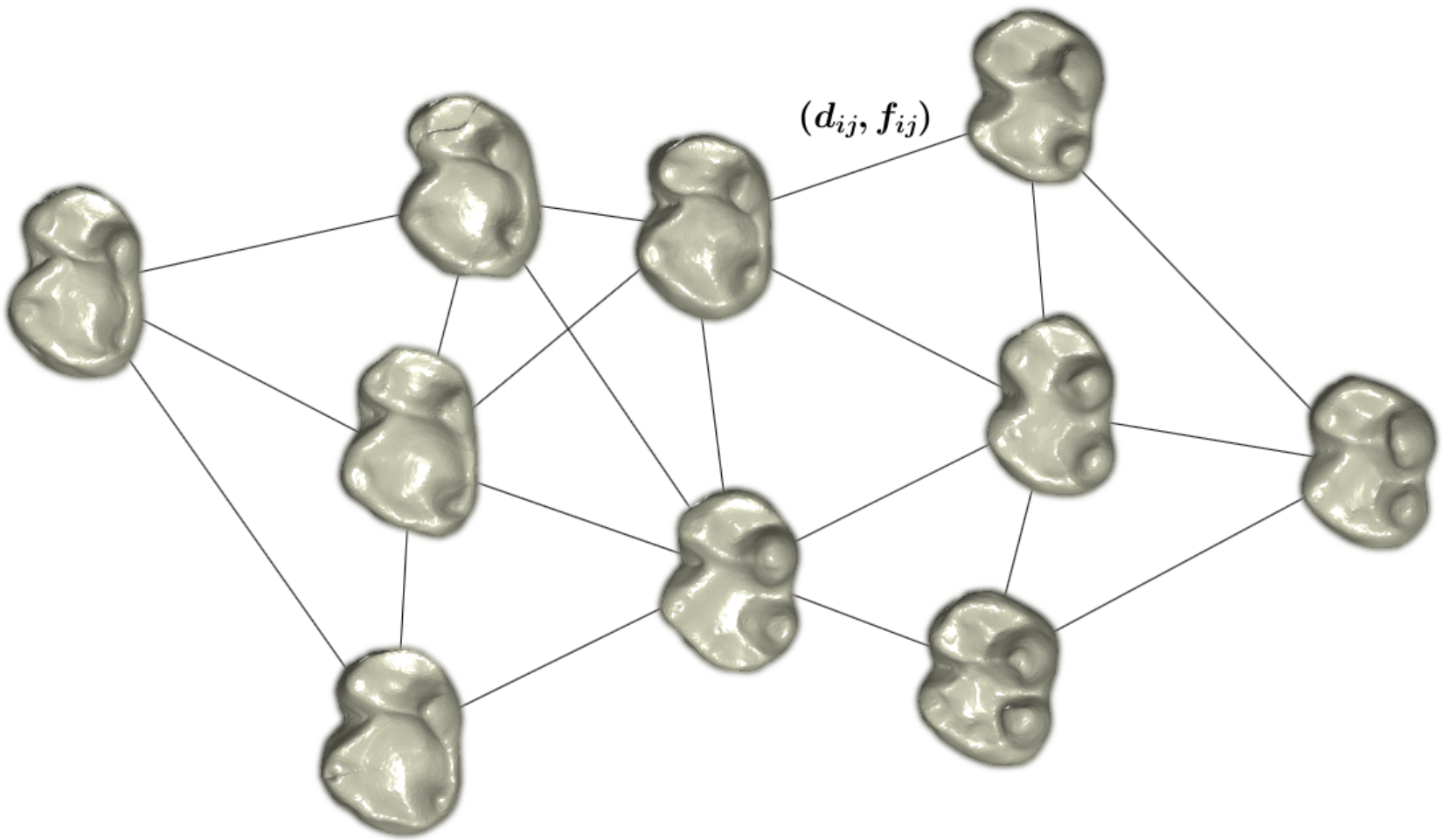
$$d_{\mathrm{cP}}\left(S_i, S_j\right) = \inf_{\mathcal{C}\in\mathcal{A}\left(S_i,S_j\right)} \inf_{R\in\mathbb{E}(3)} \left(\int_{S_i} \left\|R\left(x\right) - \mathcal{C}\left(x\right)\right\|^2 d\mathrm{vol}_{S_i}\left(x\right)\right)^{\frac{1}{2}}$$

S_i



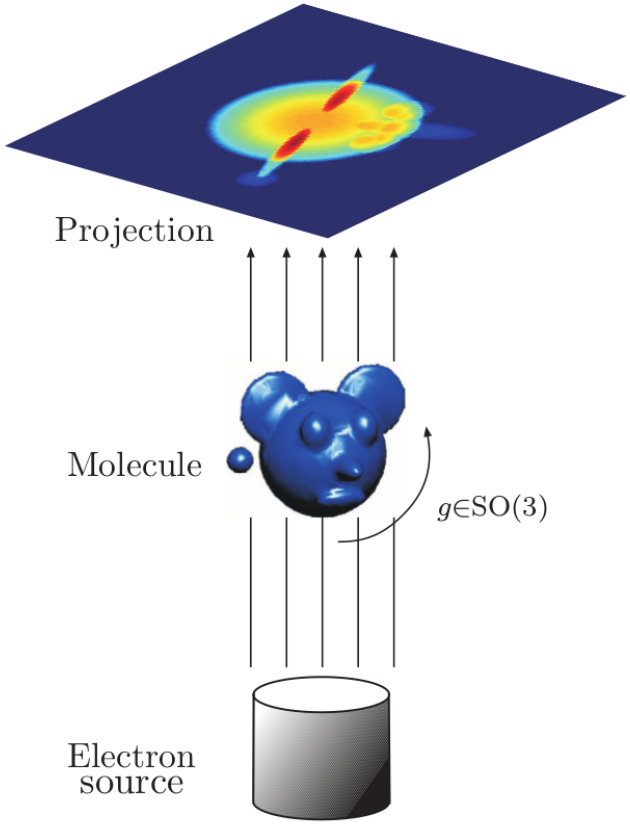
S_j

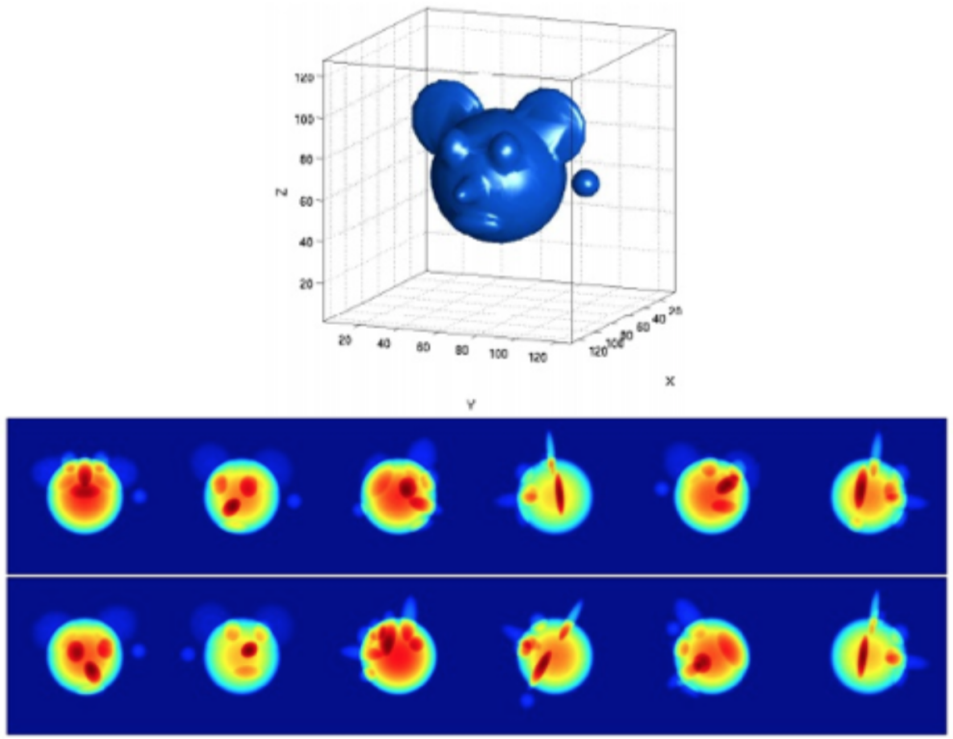
Do Graph Representations Have Enough Expressive Power?



Outline

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

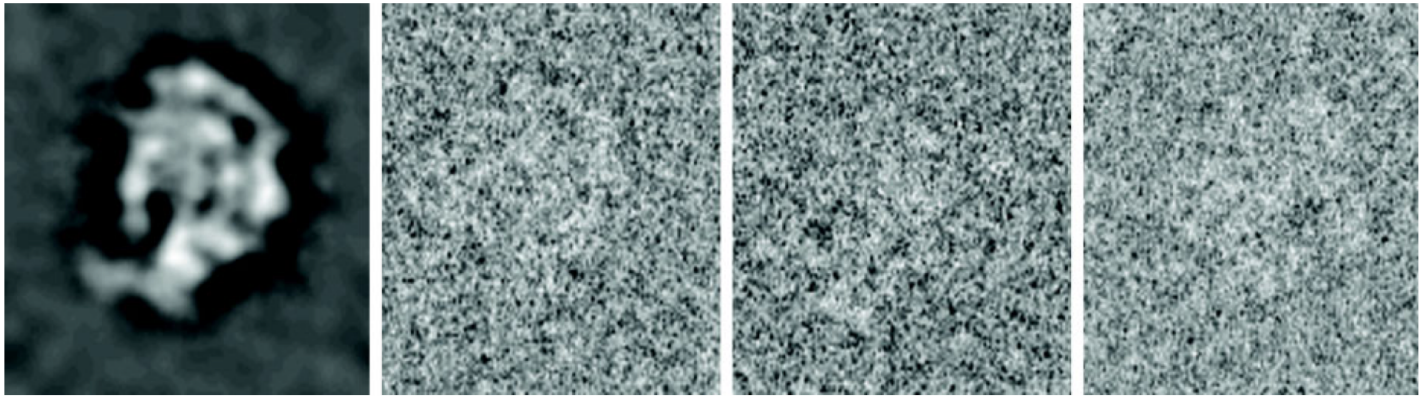


Fig. 3 The left most image is a clean simulated projection image of the E.coli 50S ribosomal subunit. The other three images are real electron microscope images of the same subunit

Apply **Class Averaging** to improve SNR!

- For each image, identify nearest neighbors in terms of similar viewing directions
- Average out the image with the identified neighbor images (with respect to the correct pairwise rotations)

• 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).

Class Averaging

- 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

Class Averaging

- 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

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

Class Averaging

- 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 $\mathbb{C}^3$ by

$$I_i \longmapsto \Psi(I_i) := \frac{(\psi_1(i), \psi_2(i), \psi_3(i))}{\|(\psi_1(i), \psi_2(i), \psi_3(i))\|}$$

Class Averaging

- 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 $\mathbb{C}^3$ by

$$I_i \longmapsto \Psi(I_i) := \frac{(\psi_1(i), \psi_2(i), \psi_3(i))}{\|(\psi_1(i), \psi_2(i), \psi_3(i))\|}$$

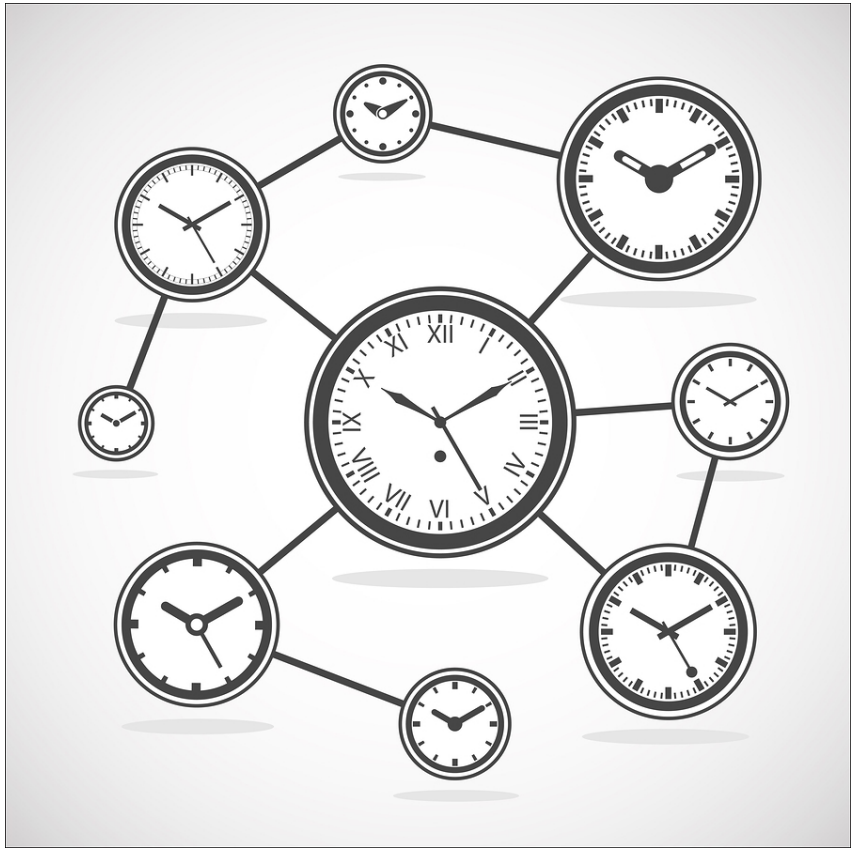
- Use correlation in the embedded $\mathbb{C}^3$ space to determine the closeness between viewing directions

Phase Synchronization (Singer'11)

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

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

  for some or all pairs of (i, j)
- **Examples:** Class averaging in cryo-EM image analysis, shape registration and community detection



Phase Synchronization (Singer'11)

- **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

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

  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

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

Phase Synchronization: Existing Methods

- Convex Relaxations: Singer'11, Chaudhury et al.'15, Bandeira et al.'16, Bandeira et al.'17
- Nonconvex Methods: Boumal'16, Zhong & Boumal'17
- Non-Unique Games: Bandeira et al.'15
- Approximate Message Passing: Perry et al.'18

• Singer, A. Angular synchronization by eigenvectors and semidefinite programming. *Applied and Computational Harmonic Analysis*, 30(1), 20–36 (2011).
• Chaudhury, K.N., Khoo, Y., Singer, A. Global registration of multiple point clouds using semidefinite programming. *SIAM Journal on Optimization*, 25(1), 468–501 (2015)
• Bandeira, A.S., Kennedy, C., Singer, A. Approximating the little Grothendieck problem over the orthogonal and unitary groups. *Mathematical Programming*, 160(1-2) pp. 433–475 (2016).
• Bandeira, A.S., Boumal, N., Singer, A. Tightness of the maximum likelihood semidenite relaxation for angular synchronization. *Mathematical Programming*, 163(1-2) pp. 145–167 (2017).
• Boumal, N. Nonconvex phase synchronization. *SIAM Journal on Optimization*, 26(4):2355–2377, 2016.
• Zhong, Y., Boumal, N. Near-optimal bounds for phase synchronization. *SIAM Journal on Optimization*, 28(2):989-1016, 2018
• Bandeira, A., Chen, Y., Singer, A. Non-unique games over compact groups and orientation estimation in cryo-EM. arXiv preprint arXiv:1505.03840, 2015.
• Perry, A., Wein, A. S., Bandeira, A. S., Moitra, A. Message-passing algorithms for synchronization problems over compact groups. *Communications on Pure and Applied Mathematics*, 2018.

Outline

Initialization: Inspired by Harmonic Retrieval

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

Initialization: Inspired by Harmonic Retrieval

- Fix $k_{\max} \geq 1$, build $H^{(2)}, \ldots, H^{(k_{\max})}$ out of $H = H^{(1)}$ by taking entrywise powers of H
- For each $k = 1, \ldots, k_{\max}$, find a reasonably good symmetric rank-1 approximation

$$W^{(k)} := \underset{\substack{Y = Y^\top \\ \mathrm{rank}(Y) = 1}}{\arg\max} \left\| H^{(k)} - Y \right\|_F^2$$

using e.g. spectral method

Initialization: Inspired by Harmonic Retrieval

- Fix $k_{\max} \geq 1$, build $H^{(2)}, \ldots, H^{(k_{\max})}$ out of $H = H^{(1)}$ by taking entrywise powers of H
- For each $k = 1, \ldots, k_{\max}$, find a reasonably good symmetric rank-1 approximation

$$W^{(k)} := \underset{\substack{Y = Y^\top \\ \operatorname{rank}(Y)=1}}{\arg\max} \left\| H^{(k)} - Y \right\|_F^2$$

  using e.g. spectral method
- 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 $H = H^{(1)}$ by taking entrywise powers of H
- For each $k = 1, \ldots, k_{\max}$, find a reasonably good symmetric rank-1 approximation

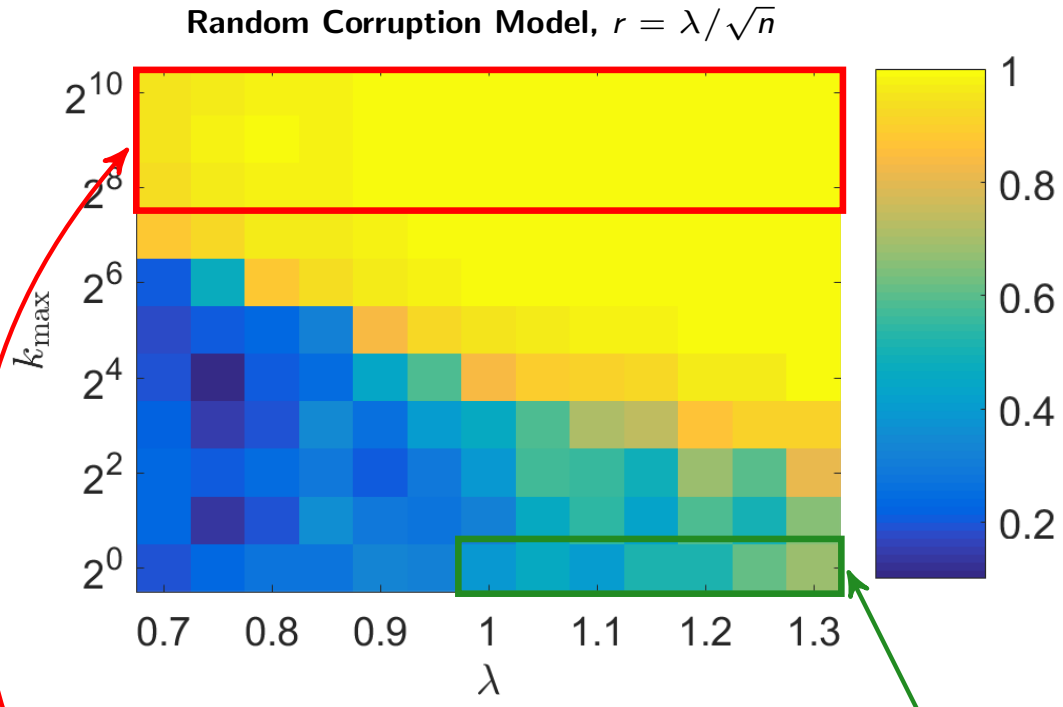
$$W^{(k)} := \underset{\substack{Y = Y^\top \\ \operatorname{rank}(Y)=1}}{\arg\max} \left\| H^{(k)} - Y \right\|_F^2$$

  using e.g. spectral method
- For all $1 \leq i, j \leq n$, find the "peak location" of the spectrogram

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

- Apply the spectral method yet again to the Hermitian matrix $\widehat{H}$ to get $\hat{x} \in \mathbb{C}_1^n$, where $\widehat{H}_{ij} = e^{\iota \hat{\theta}_{ij}}$

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

New Theory: Strong Recovery

Theorem (G.–Zhao 2019). Under (mild) assumptions, with high probability, the multi-frequency phase synchronization algorithm produces an estimate $\hat{x}$ satisfying

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

provided that

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

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

Why does it work?

- By a perturbation analysis, after solving each subproblem

$$W^{(k)} := \underset{\substack{Y=Y^\top \\ \mathrm{rank}(Y)=1}}{\arg\max} \left\| H^{(k)} - Y \right\|_F^2$$

  we expect $W^{(k)} = z^k (z^k)^* + E^{(k)} \approx z^k (z^k)^*$, where $z_i = e^{\iota\theta_i}$

- The peak finding step is expected to ensure

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

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

Landscape Analysis of the Dirichlet Kernel

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

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

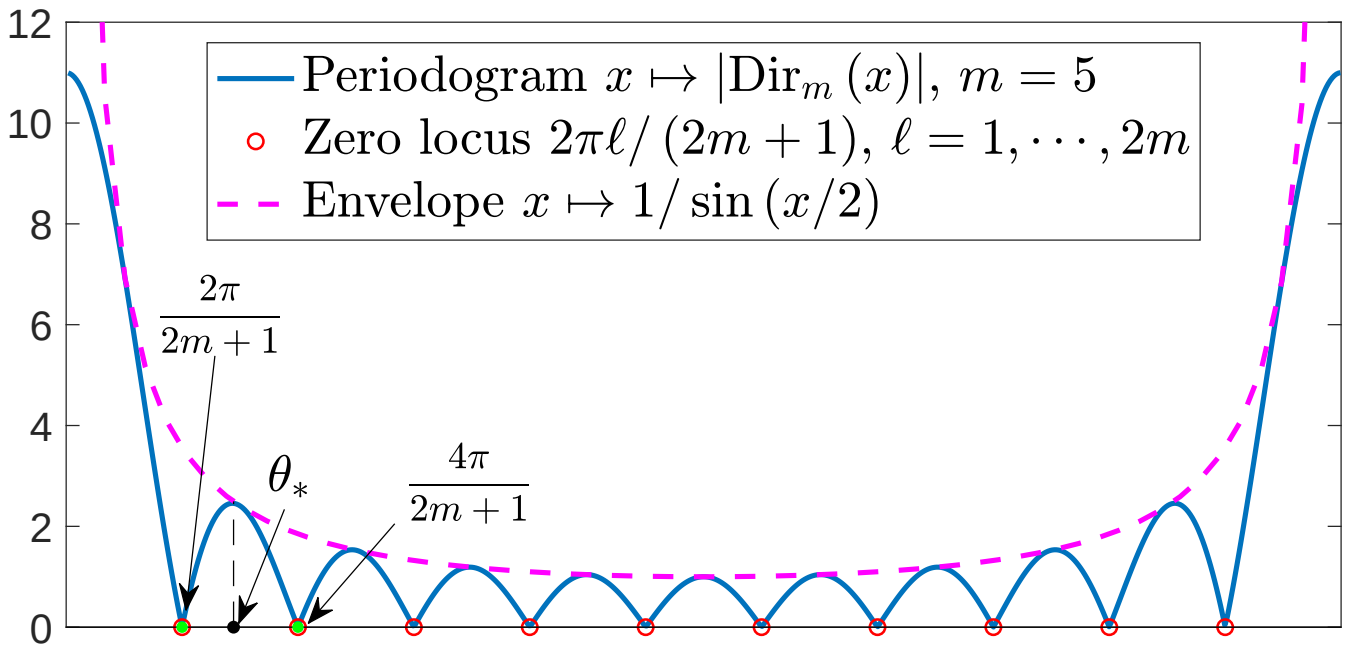
Landscape Analysis of the Dirichlet Kernel

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

A Proof by Picture (Notation: $\Delta\theta_{ij} := \hat{\theta}_{ij} - (\theta_i - \theta_j)$)

Key Observation: $|\Delta\theta_{ij}| \leq \theta_* < \frac{4\pi}{2k_{\max}+1}$ whenever

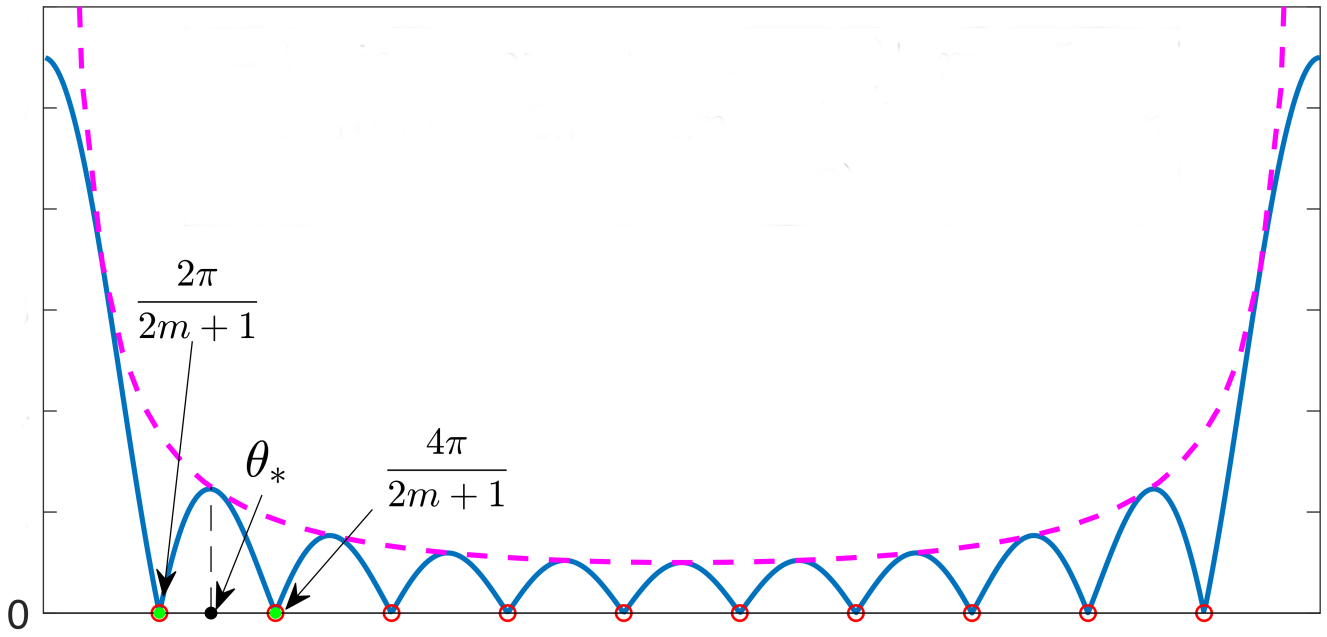
$$2k_{\max} + 1 - \|R_{k_{\max}}\|_\infty > \frac{1}{\sin(\theta_*/2)} + \|R_{k_{\max}}\|_\infty$$



A Proof by Picture (Notation: $\Delta\theta_{ij} := \hat{\theta}_{ij} - (\theta_i - \theta_j)$)

Key Observation: $|\Delta\theta_{ij}| \leq \theta_* < \frac{4\pi}{2k_{\max}+1}$ whenever

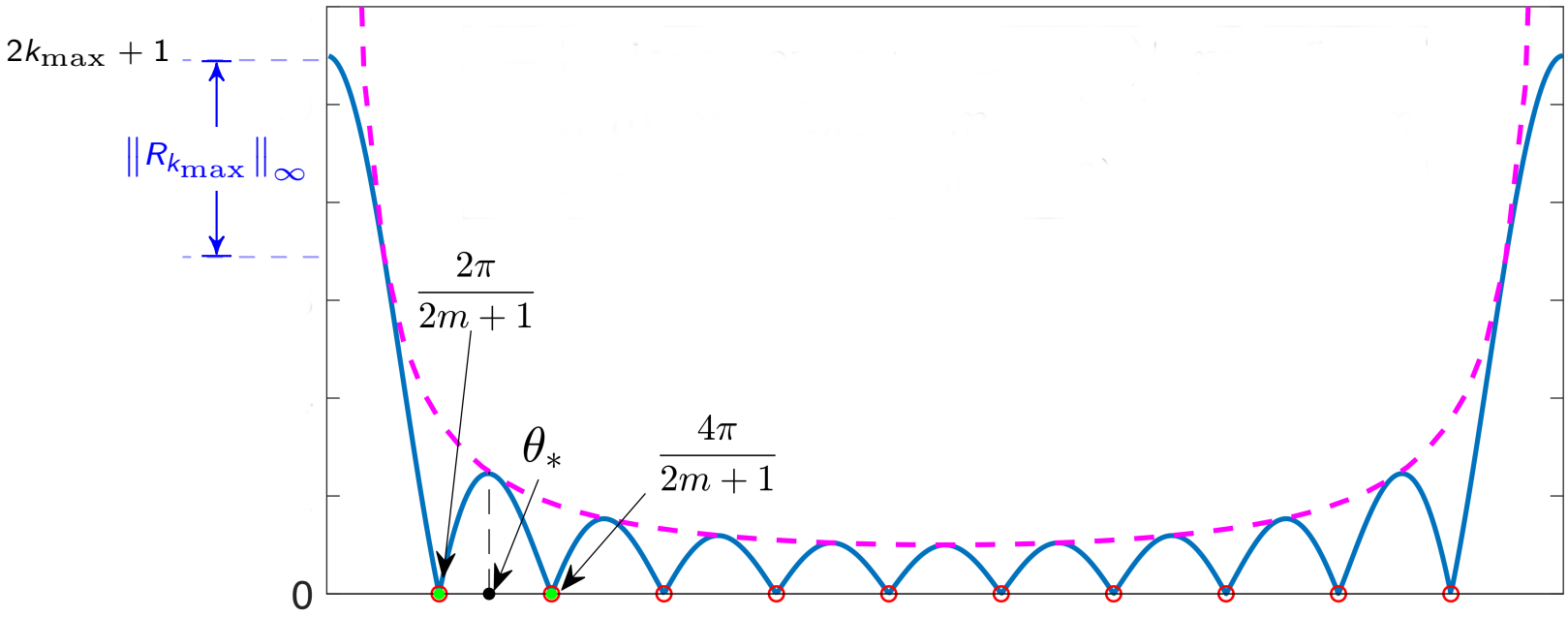
$$2k_{\max} + 1 - \|R_{k_{\max}}\|_\infty > \frac{1}{\sin(\theta_*/2)} + \|R_{k_{\max}}\|_\infty$$



A Proof by Picture (Notation: $\Delta\theta_{ij} := \hat{\theta}_{ij} - (\theta_i - \theta_j)$)

Key Observation: $|\Delta\theta_{ij}| \leq \theta_* < \frac{4\pi}{2k_{\max}+1}$ whenever

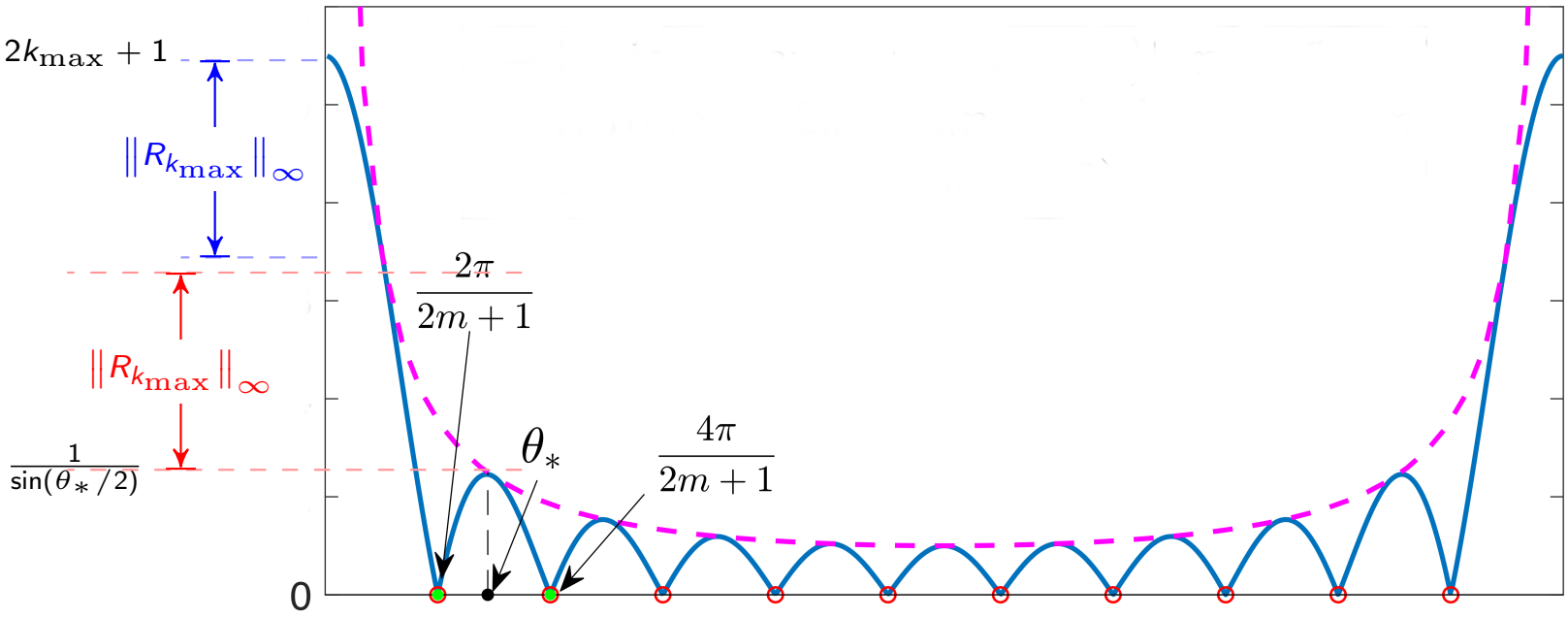
$$2k_{\max} + 1 - \|R_{k_{\max}}\|_\infty > \frac{1}{\sin(\theta_*/2)} + \|R_{k_{\max}}\|_\infty$$



A Proof by Picture (Notation: $\Delta\theta_{ij} := \hat{\theta}_{ij} - (\theta_i - \theta_j)$)

Key Observation: $|\Delta\theta_{ij}| \leq \theta_* < \frac{4\pi}{2k_{\max}+1}$ whenever

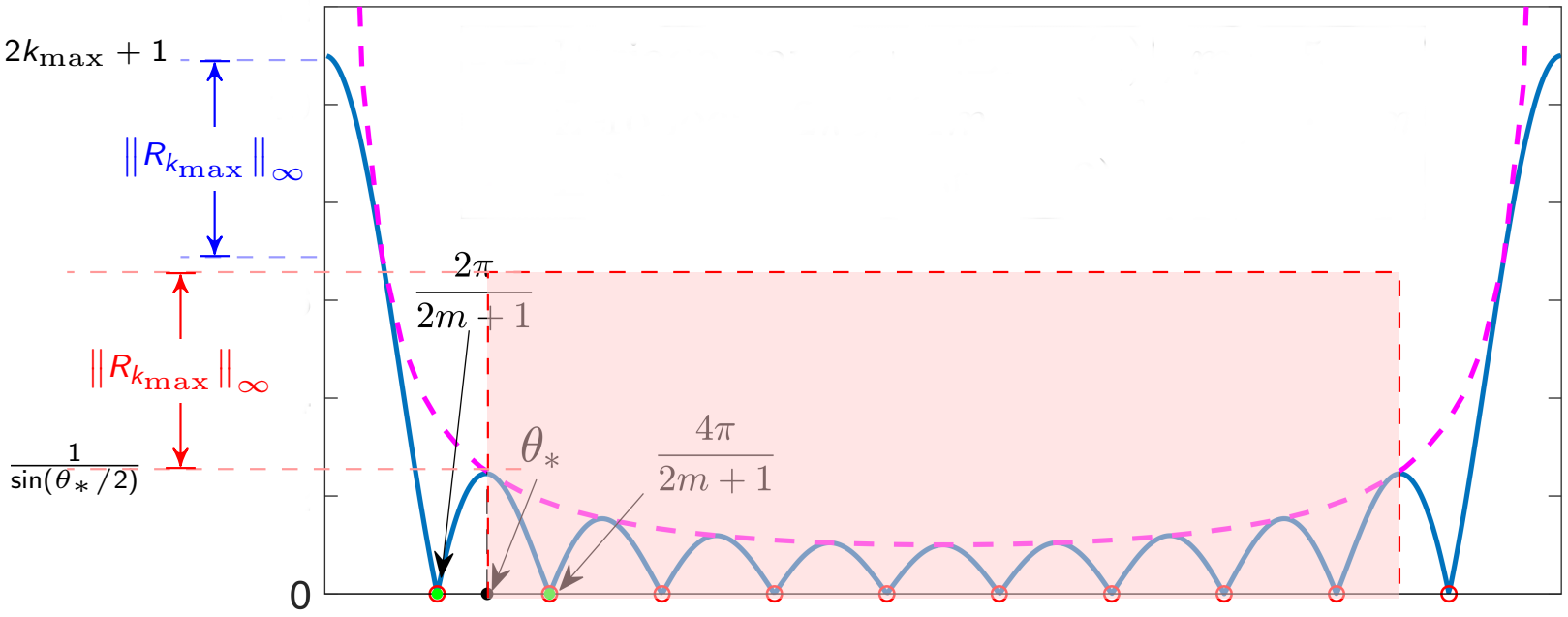
$$2k_{\max} + 1 - \|R_{k_{\max}}\|_\infty > \frac{1}{\sin(\theta_*/2)} + \|R_{k_{\max}}\|_\infty$$



A Proof by Picture (Notation: $\Delta\theta_{ij} := \hat{\theta}_{ij} - (\theta_i - \theta_j)$)

Key Observation: $|\Delta\theta_{ij}| \leq \theta_* < \frac{4\pi}{2k_{\max}+1}$ whenever

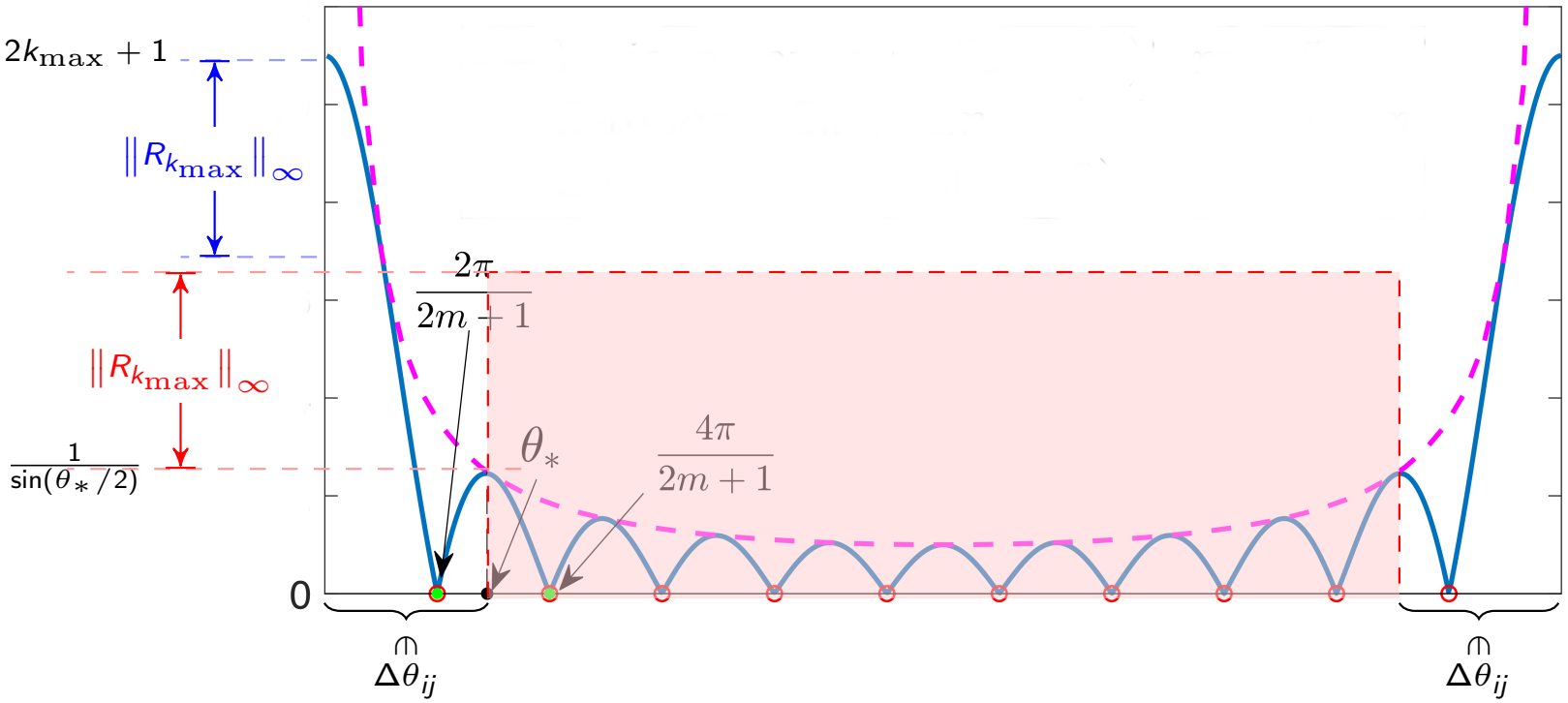
$$2k_{\max} + 1 - \|R_{k_{\max}}\|_\infty > \frac{1}{\sin(\theta_*/2)} + \|R_{k_{\max}}\|_\infty$$



A Proof by Picture (Notation: $\Delta\theta_{ij} := \hat{\theta}_{ij} - (\theta_i - \theta_j)$)

Key Observation: $|\Delta\theta_{ij}| \leq \theta_* < \frac{4\pi}{2k_{\max}+1}$ whenever

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



A Proof by Picture (Notation: $\Delta\theta_{ij} := \hat{\theta}_{ij} - (\theta_i - \theta_j)$)

Key Observation: $|\Delta\theta_{ij}| \leq \theta_* < \frac{4\pi}{2k_{\max}+1}$ whenever

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

A Proof by Picture (Notation: $\Delta\theta_{ij} := \hat{\theta}_{ij} - (\theta_i - \theta_j)$)

Key Observation: $|\Delta\theta_{ij}| \leq \theta_* < \frac{4\pi}{2k_{\max}+1}$ whenever

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

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

A Proof by Picture (Notation: $\Delta\theta_{ij} := \hat{\theta}_{ij} - (\theta_i - \theta_j)$)

Key Observation: $|\Delta\theta_{ij}| \leq \theta_* < \frac{4\pi}{2k_{\max}+1}$ whenever

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

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

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

What are the odds?
Can be estimated with 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 (G.–Zhao 2019). For any $0 < \epsilon \leq 2$, with probability at least $1 - O\left(n^{-(2+\epsilon)}\right)$, there exists absolute constant $C_2 > 0$ s.t.

$$\left\| E_{ij}^{(k)} \right\|_\infty \leq C_2 \sigma \sqrt{\frac{\log n}{n}}.$$

Putting Everything Together, with a Union Bound

With high probability, uniformly for all $i, j \in [n]$, $\hat{\theta}_{ij}$ is close to the true offset $\theta_i - \theta_j$ as long as

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

$$\Leftarrow \; 1 - \frac{1}{2k_{\max} \sin\left(\frac{\pi}{2k_{\max}+1}\right)} > C_2 \sigma \sqrt{\frac{\log n}{n}}$$

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

New Theory: Strong Recovery

Theorem (G.–Zhao 2019). Under (mild) assumptions, with high probability, the multi-frequency phase synchronization algorithm produces an estimate $\hat{x}$ satisfying

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

provided that

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

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

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:

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

- 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 $\mathrm{SO}(3)$:

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

- Works fantastic in practice, but no theory yet!

Outline

Class Averaging

- 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 $\mathbb{C}^3$ by

$$I_i \longmapsto \Psi(I_i) := \frac{(\psi_1(i), \psi_2(i), \psi_3(i))}{\|(\psi_1(i), \psi_2(i), \psi_3(i))\|}$$

- Use correlation in the embedded $\mathbb{C}^3$ space to determine the closeness between viewing directions

Multi-Frequency Class Averaging

- 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 $\mathbb{C}^{2k+1}$ by

$$I_i \longmapsto \Psi^{(k)}(I_i) := \frac{(\psi_1(i), \ldots, \psi_{2k+1}(i))}{\|(\psi_1(i), \ldots, \psi_{2k+1}(i))\|}$$

- Use all correlations in the embedded $\mathbb{C}^{2k+1}$ $(k = 1, \ldots, k_{\max})$ spaces to determine the closeness between viewing directions

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

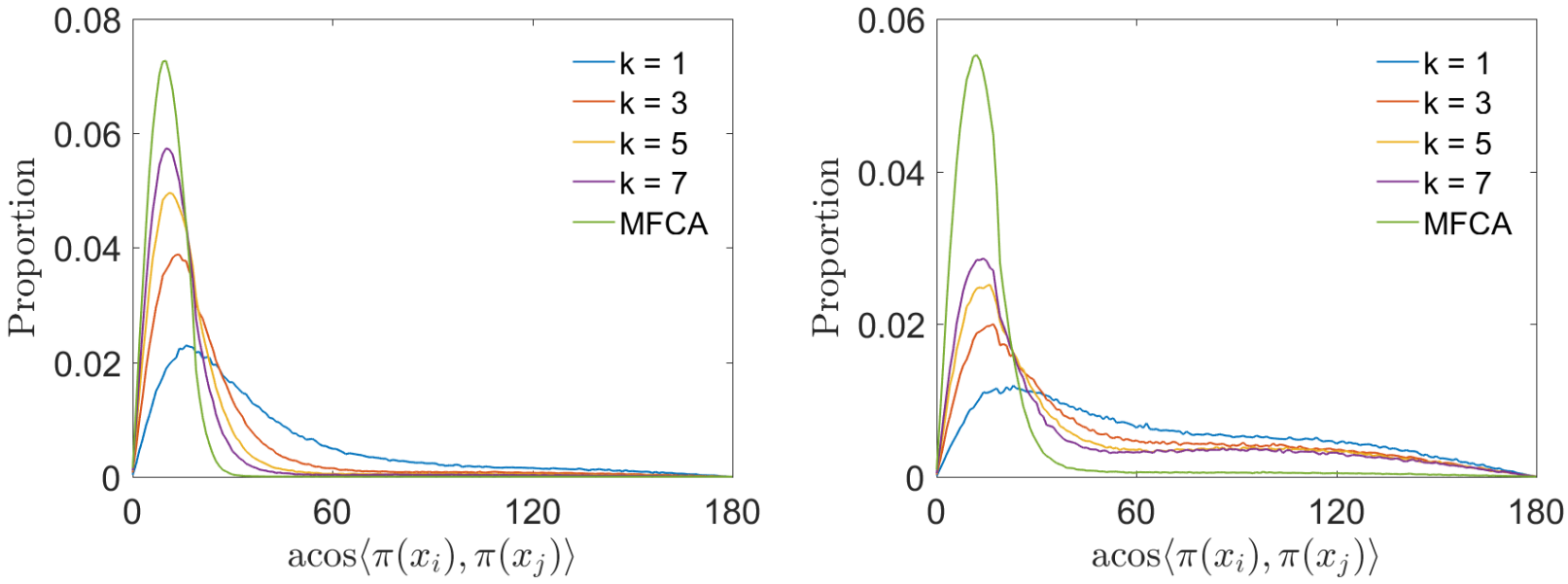
- **(G.–Fan–Zhao 2019)** The *extrinsic model* S^2 of viewing directions and the *intrinsic* model of the top eigenspace of $W^{(k)}$ are isomorphic, following (Hadani & Singer, 2011)
- For sufficiently large sample size n and appropriately small $\epsilon > 0$, the top eigenspace of $W^{(k)}$ is $(2k+1)$-dimensional, 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!

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

Multi-Frequency Information Improves Class Averaging



Histograms of true viewing angles between each image and its 50 nearest neighboring images

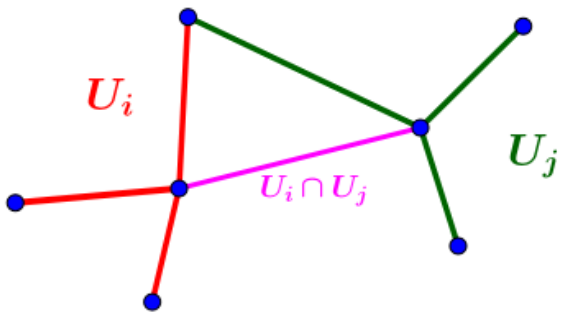
Outline

Synchronization Problems: Geometric Picture

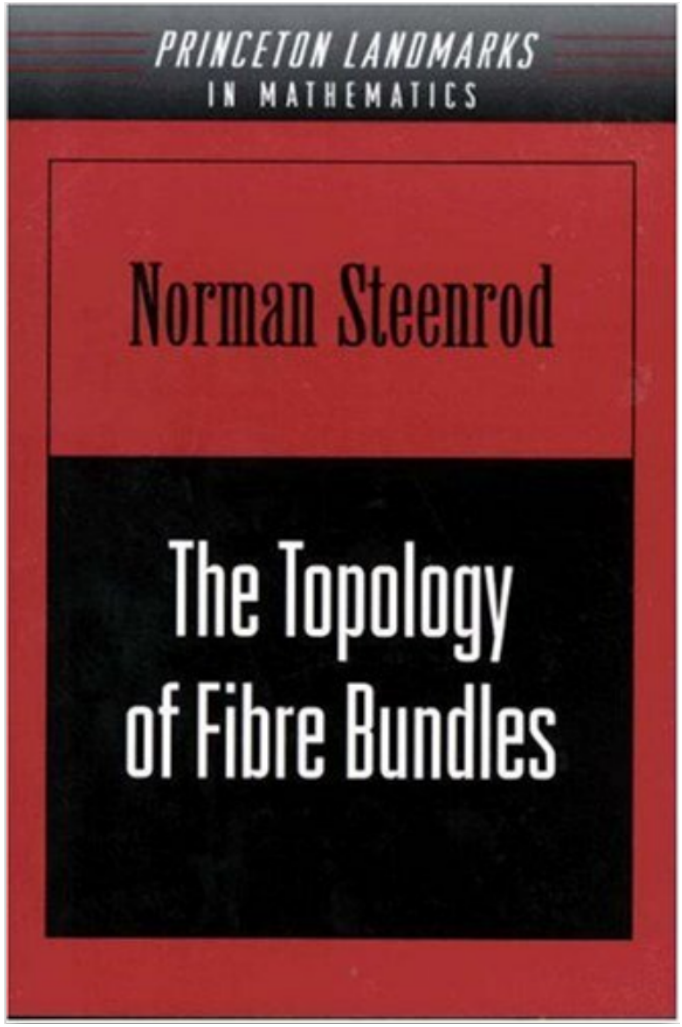
- Data:
  - graph $\Gamma = (V, E)$
  - topological group G, equipped with a norm $\|\cdot\|$, and a G-module F
  - **edge potential** $g : E \to G$ satisfying $g_{ij} = g_{ji}^{-1}$, $\forall (i,j) \in E$

Synchronization Problems: Geometric Picture

- Data:
  - graph $\Gamma = (V, E)$
  - topological group G, equipped with a norm $\|\cdot\|$, and a G-module F
  - **edge potential** $g : E \to G$ satisfying $g_{ij} = g_{ji}^{-1}$, $\forall (i,j) \in E$
- Flat Principal G-Bundle:
  - Let $\mathfrak{U} = \{U_i \mid 1 \leq i \leq |V|\}$ be an open cover of Γ (viewed as a 1-dimensional simplicial complex), where U_i is the *(open) star neighborhood* of vertex i.



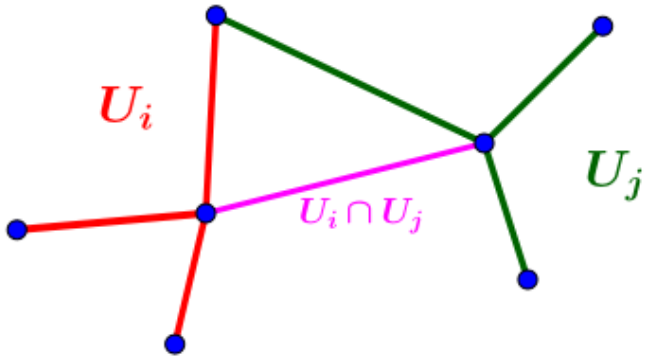
  - Triplet (g, G, Γ) defines a *flat principal G-bundle $\mathscr{B}_\rho$* over Γ



Theorem (Steenrod 1951, §2).
If Lie group G acts on Y,
$\mathfrak{U} = \{U_i\}$ is an open cover of X,
$\{g_{ij} \in G \mid U_i \cap U_j \neq \emptyset\}$ satisfies

$$\begin{aligned} g_{ii} &= e \in G \quad \text{for all } U_i \\ g_{ij} &= g_{ji}^{-1} \quad \text{if } U_i \cap U_j \neq \emptyset \\ g_{ij} g_{jk} &= g_{ik} \quad \text{if } U_i \cap U_j \cap U_k \neq \emptyset \end{aligned}$$

then there exists a fibre bundle $\mathscr{B}$ with base space X, fibre Y, group G, and bundle transformations $\{g_{ij}\}$.



No triple intersections!

Theorem (Steenrod 1951, §2).
If Lie group G acts on Y,
$\mathfrak{U} = \{U_i\}$ is an open cover of X,
$\{g_{ij} \in G \mid U_i \cap U_j \neq \emptyset\}$ satisfies

$$g_{ii} = e \in G \quad \text{for all } U_i$$

$$g_{ij} = g_{ji}^{-1} \quad \text{if } U_i \cap U_j \neq \emptyset$$

$$\cancel{g_{ij} g_{jk} = g_{ik} \quad \text{if } U_i \cap U_j \cap U_k \neq \emptyset}$$

then there exists a fibre bundle $\mathscr{B}$ with base space X, fibre Y, group G, and bundle transformations $\{g_{ij}\}$.

Representation and Associated Bundles

- If M is a principal G-bundle over B, $\rho : G \to \mathrm{Aut}\,(F)$ is a representation of G over a vector space F. Then ρ induces an **associated** F-bundle over B:

$$M \times_\rho F := M \times F \Big/ \sim$$

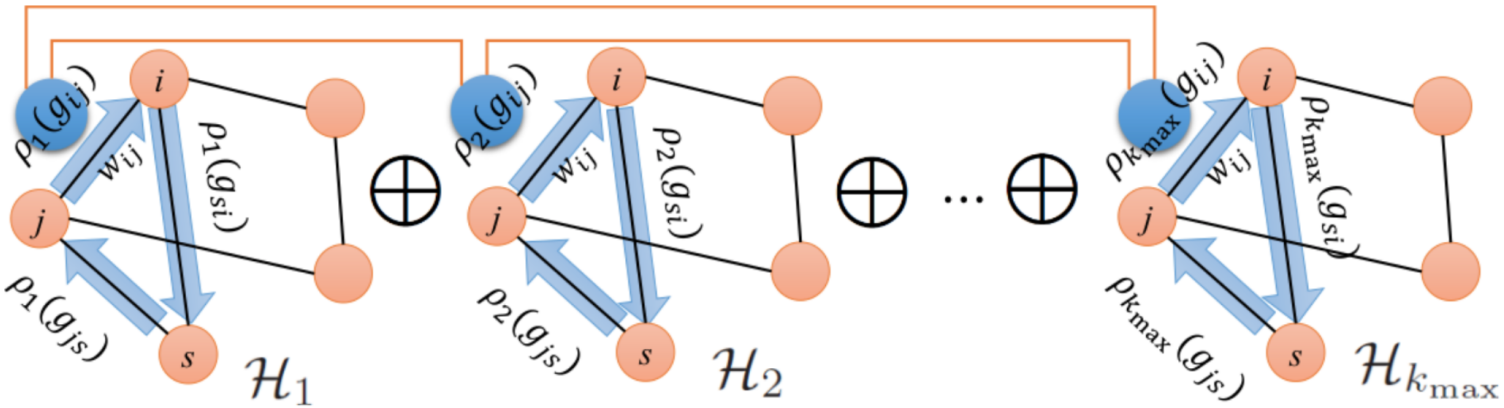
where the equivalence relation is defined by

$$(m \cdot g, v) \sim (m, \rho\,(g)\, v)$$

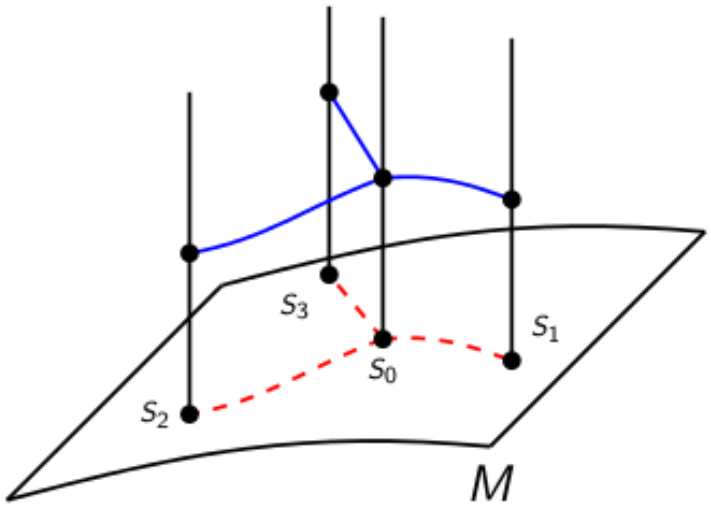
- Non-equivalent irreducible representations gives rise to distinct associated bundles

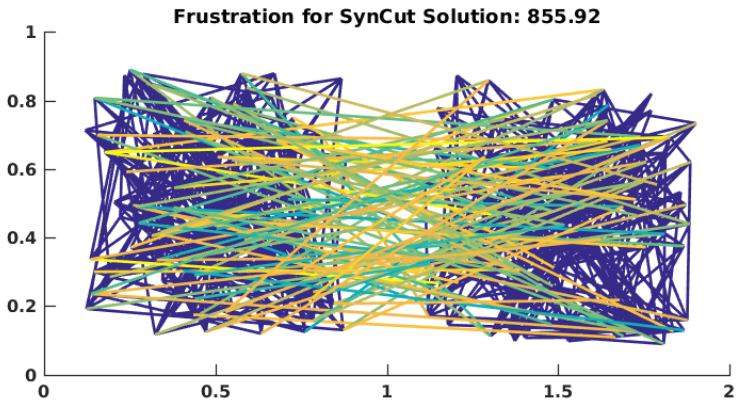
Multi-Frequency $\leftrightarrow$ Multiple Associated Bundles

- All irreducible representations of $\mathrm{U}(1)$: $\left\{\theta \to e^{\iota k\theta} \mid k \in \mathbb{Z}\right\}$
- Entrywise power: Inducing multiple irreducible representations, effectively creating many associated bundles
- Multi-frequency phase synchronization and class averaging both strive to distill features across multiple associate bundles (associated with the same principal bundles)



Diffusion and Community Detection on Fibre Bundles





• Tingran Gao. *The Diffusion Geometry of Fibre Bundles: Horizontal Diffusion Maps.* Applied and Computational Harmonic Analysis, online first, pp.1–69 (2019)
• Tingran Gao, Jacek Brodzki, and Sayan Mukherjee. *The Geometry of Synchronization Problems and Learning Group Actions*, Discrete & Computational Geometry, online first, pp.1–62 (2019)

Open Problems

- Extension to other synchronization and multireference alignment problems over compact/noncompact Lie groups?
- Fundamental statistical/computational limits in synchronizability-based community detection?
- **A learning paradigm on sheaves?**

Thank You!

- AMS–Simons Travel Grant
- UChicago CDAC Data Science Discovery Seed Grant
- NSF CDS&E-MSS DMS-1854831

Center for Data and Computing

• Tingran Gao and Zhizhen Zhao. *Multi-Frequency Phase Synchronization.* Proceedings of the 36th International Conference on Machine Learning (ICML 2019), 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.
• Tingran Gao. *The Diffusion Geometry of Fibre Bundles: Horizontal Diffusion Maps.* Applied and Computational Harmonic Analysis, online first, pp.1–69 (2019)
• Tingran Gao, Jacek Brodzki, and Sayan Mukherjee. *The Geometry of Synchronization Problems and Learning Group Actions*, Discrete & Computational Geometry, online first, pp.1–62 (2019)