Exact Minimax Optimality of Spectral Methods in Phase Synchronization and Orthogonal Group Synchronization

Anderson Ye Zhang

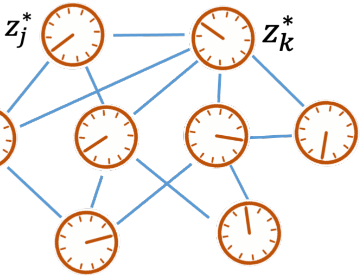
Department of Statistics and Data Science
University of Pennsylvania

Phase Synchronization

Problem Setup:

- n unit complex numbers $z_1^*, \ldots, z_n^* \in \mathbb{C}$, each one corresponds to a phase / angle in $(0, 2\pi]$

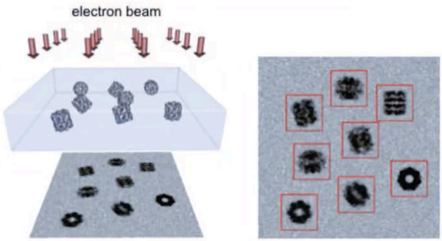
- We want to estimate them from their incomplete and noisy pairwise comparisons



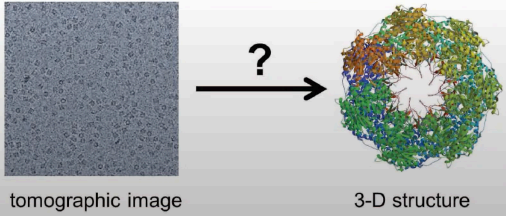
If not missing, $X_{jk} =$ noisy version of $z_j^* \overline{z_k^*}$

Motivation: Single Particle Cryo-EM

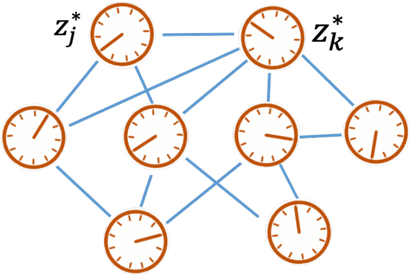
Schematic drawing of the imaging process:



The standard cryo-EM reconstruction problem:



Model



For $1 \leq j < k \leq n$,

$$X_{jk} := \begin{cases} z_j^* \overline{z_k^*} + \sigma W_{jk}, & \text{if } A_{jk} = 1, \\ 0, & \text{if } A_{jk} = 0, \end{cases}$$

where $A_{jk} \sim \text{Bernoulli}(p)$ and $W_{jk} \sim \mathcal{CN}(0, 1)$.

Matrix Form: Let $z^* = (z_1^*, \ldots, z_n^*)^T$. Then
$X = A \circ (z^* z^{*\mathsf{H}} + \sigma W) = A \circ (z^* z^{*\mathsf{H}}) + \sigma A \circ W$

Spectral Method (aka Eigenvector Method)

Motivation: $\mathbb{E}X = pz^*z^{*\mathsf{H}} - pI_n$. Its leading eigenvector is $z^*/\sqrt{n}$.

Step 1: Let u be the leading eigenvector of X.

Step 2: The spectral estimator $\hat{z}$ is defined as

$$\hat{z}_j = \begin{cases} \frac{u_j}{|u_j|}, & \text{if } u_j \neq 0, \\ 1, & \text{if } u_j = 0. \end{cases}$$

Eigendecomposition + Normalization

To measure its performance:

$$\ell(\hat{z}, z^*) := \frac{1}{n} \min_{a \in \mathbb{C}: |a|=1} \|\hat{z} - z^* a\|^2 = \frac{1}{n} \min_{a \in \mathbb{C}: |a|=1} \sum_{j=1}^{n} \left|\hat{z}_j - z_j^* a\right|^2$$

Existing Results

With high probability, if $\frac{np}{\log n} \to \infty$, then

$$\ell(\hat{z}, z^*) \leq C \left(\frac{\sigma^2}{np} + \frac{1}{np} \right).$$

Two sources of errors:

1. $\frac{\sigma^2}{np}$: from additive Gaussian noises
2. $\frac{1}{np}$: from missing data

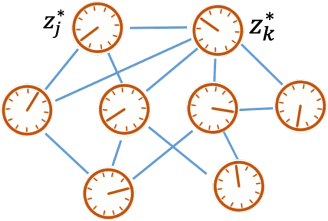
However, the minimax risk is

$$\inf_{z \in \mathbb{C}^n} \sup_{z^*} \mathbb{E}\ell(z, z^*) \geq (1 - o(1)) \frac{1}{2} \frac{\sigma^2}{np}.$$

(If we consider all possible methods, how small the error can be?)

New Result 1: Exact Recovery for No-additive-noise Case

When $\sigma = 0$:



$$X_{jk} = \begin{cases} z_j^* \overline{z_k^*}, & \text{if } A_{jk} = 1, \\ 0, & \text{if } A_{jk} = 0. \end{cases}$$

Matrix form: $X = A \circ (z^* z^{*\mathsf{H}})$

Lemma

If $\sigma = 0$ and $\frac{np}{\log n} \to \infty$. With high probability, $\ell(\hat{z}, z^*) = 0$, i.e., the spectral method achieves the exact recovery.

New Result 2: Exact Minimax Optimality

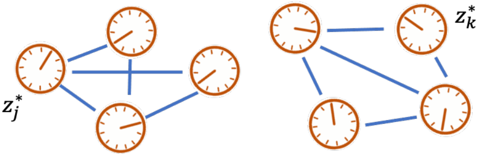
Theorem

Assume $\frac{np}{\sigma^2} \to \infty$ *and* $\frac{np}{\log n} \to \infty$. *With high probability*

$$\ell(\hat{z}, z^*) \leq (1 + o(1)) \frac{1}{2} \frac{\sigma^2}{np}.$$

Remarks:

- Achieves the exact minimax risk
- $\frac{np}{\sigma^2} \to \infty$ is for consistency
- $\frac{np}{\log n} \gtrsim 1$ is for the comparison graph $A \sim$ Erdös-Rényi(n, p) to be connected



New Result 2: Exact Minimax Optimality

Theorem

Assume $\frac{np}{\sigma^2} \to \infty$ *and* $\frac{np}{\log n} \to \infty$. *With high probability*

$$\ell(\hat{z}, z^*) \leq (1 + o(1)) \frac{1}{2} \frac{\sigma^2}{np}.$$

Remarks:

- As good as more sophisticated procedures including maximum likelihood estimation (MLE), generalized power method (GPM), and semidefinite programming (SDP), under this parameter regime.

Technical Tools: Eigenvector Perturbation Analysis

Novelty 1: Choice of the "population matrix"

- In literature, X is viewed as a perturbation of $\mathbb{E}X$

	Sample		Population
Matrix	X	$\longrightarrow$	$\mathbb{E}X$
		perturbation	
Eigenvector	u	$\longrightarrow$	$z^*/\sqrt{n}$

- Consequently, u is viewed as a perturbation of $z^*/\sqrt{n}$, the leading eigenvector of $\mathbb{E}X$.
- The distance between u and $z^*/\sqrt{n}$ can be upper bounded by the Davis-Kahan Theorem, which leads to the existing loose bound.

Technical Tools: Eigenvector Perturbation Analysis

Novelty 1: Choice of the "population matrix"

- In our analysis, recall $X = A \circ (z^* z^{*\mathsf{H}}) + \sigma A \circ W$. We view X as a perturbation of $A \circ (z^* z^{*\mathsf{H}})$, ie., $\mathbb{E}(X|A)$.

	Sample		Population
Matrix	X	$\longrightarrow$	$A \circ (z^* z^{*\mathsf{H}})$
		perturbation	
Eigenvector	u	$\longrightarrow$	u^*

This is the data matrix if there was no additive noise

No explicit formula but can be shown to be well-behaved

- Consequently, we view u as a perturbation of u^*, the leading eigenvector of $A \circ (z^* z^{*\mathsf{H}})$.
- u is closer to u^* than to $z^*/\sqrt{n}$.

Technical Tools: Eigenvector Perturbation Analysis

Novelty 2: Approximating eigenvectors by their first-order approximations

- Classical matrix perturbation theory such as Davis-Kahan Theorem focuses on analyzing $\inf_{b\in\mathbb{C}:|b|=1} \|u - u^* b\|$.
- We show u can be well-approximated by its first-order approximation $\tilde{u}$ defined as

$$\tilde{u} := \frac{X u^*}{\|X u^*\|},$$

- $\inf_{b\in\mathbb{C}:|b|=1} \|u - \tilde{u} b\|$ is much smaller than $\inf_{b\in\mathbb{C}:|b|=1} \|u - u^* b\|$, meaning u is closer to $\tilde{u}$ than to u^*.
- We study $\tilde{u}$ to understand behavior of u and the performance of the spectral method.

Technical Tools: Eigenvector Perturbation Analysis

Novelty 2: Approximating eigenvectors by their first-order approximations

A general perturbation result:

Lemma

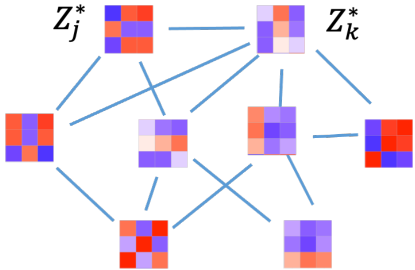
Consider two Hermitian matrices $Y, Y^* \in \mathbb{C}^{n\times n}$*. Let* $\mu_1^* \geq \mu_2^* \geq \ldots \geq \mu_n^*$ *be the eigenvalues of* Y^**. Let* v *(resp.* v^**) be the eigenvector of* Y *(resp.* Y^**) corresponding to its largest eigenvalue. If* $\|Y - Y^*\| \leq \min\{\mu_1^* - \mu_2^*, \mu_1^*\}/4$*, we have*

$$\inf_{b\in\mathbb{C}:|b|=1} \left\| v - \frac{Yv^*}{\|Yv^*\|} b \right\| \leq \frac{40\sqrt{2}}{9(\mu_1^* - \mu_2^*)} \left(\left(\frac{4}{\mu_1^* - \mu_2^*} + \frac{2}{\mu_1^*} \right) \|Y - Y^*\|^2 + \frac{\max\{|\mu_2^*|, |\mu_n^*|\}}{\mu_1^*} \|Y - Y^*\| \right).$$

If Y^* is rank-one, it gives $\left\|Y - Y^*\right\|^2 / (\mu_1^*)^2$ vs. $\left\|Y - Y^*\right\| / \mu_1^*$ from Davis-Kahan.

Generalization to Orthogonal Group Synchronization

$Z_1^*, \ldots, Z_n^* \in \mathcal{O}(d)$ are $d \times d$ orthogonal matrices



For $1 \leq j < k \leq n$,

$$X_{jk} := \begin{cases} Z_j^*(Z_k^*)^T + \sigma W_{jk}, & \text{if } A_{jk} = 1, \\ 0, & \text{if } A_{jk} = 0, \end{cases}$$

where $A_{jk} \sim \text{Bernoulli}(p)$ and $W_{jk} \sim \mathcal{MN}(0, I_d, I_d)$.

Generalization to Orthogonal Group Synchronization

Spectral Method:

Step 1: $U = (u_1, \ldots, u_d) \in \mathbb{R}^{nd \times d}$ to include the leading d eigenvectors of X. Write

$$U = \begin{pmatrix} U_1 \\ U_2 \\ \ldots \\ U_n \end{pmatrix}$$

such that $U_j \in \mathbb{R}^{d \times d}$ is its jth block.

Step 2:

$$\hat{Z}_j := \begin{cases} \mathcal{P}(U_j), & \text{if } \det(U_j) \neq 0, \\ I_d, & \text{if } \det(U_j) = 0, \end{cases}$$

Here the mapping $\mathcal{P} : \mathbb{R}^{d \times d} \to \mathcal{O}(d)$ is from the polar decomposition.

Exact Minimax Optimality in Orthogonal Group Synchronization

Theorem

Assume $d = O(1)$. Assume $\frac{np}{\sigma^2} \to \infty$ and $\frac{np}{\log n} \to \infty$. With high probability

$$\ell^{\text{od}}(\hat{Z}, Z^*) \leq (1 + o(1))\frac{d(d-1)\sigma^2}{2np}.$$

The minimax risk is

$$\inf_{Z \in \mathbb{R}^{nd \times d}} \sup_{Z^* \in \mathcal{O}(d)^n} \mathbb{E}\ell^{\text{od}}(Z, Z^*) \geq (1 - o(1))\frac{d(d-1)\sigma^2}{2np}.$$

Reference:
Anderson Ye Zhang. Exact Minimax Optimality of Spectral Methods in Phase Synchronization and Orthogonal Group Synchronization. *Annals of Statistics.* 2024

Thank You