EXACT MINIMAX OPTIMALITY OF SPECTRAL METHODS IN PHASE SYNCHRONIZATION AND ORTHOGONAL GROUP SYNCHRONIZATION

BY ANDERSON YE ZHANG[a]

Department of Statistics and Data Science, University of Pennsylvania, [a]*ayz@wharton.upenn.edu*

We study the performance of the spectral method for the phase synchronization problem with additive Gaussian noises and incomplete data. The spectral method utilizes the leading eigenvector of the data matrix followed by a normalization step. We prove that it achieves the minimax lower bound of the problem with a matching leading constant under a squared ℓ_2 loss. This shows that the spectral method has the same performance as more sophisticated procedures including maximum likelihood estimation, generalized power method, and semidefinite programming, as long as consistent parameter estimation is possible. To establish our result, we first have a novel choice of the population eigenvector, which enables us to establish the exact recovery of the spectral method when there is no additive noise. We then develop a new perturbation analysis toolkit for the leading eigenvector and show it can be well-approximated by its first-order approximation with a small ℓ_2 error. We further extend our analysis to establish the exact minimax optimality of the spectral method for the orthogonal group synchronization.

1. Introduction. We consider the phase synchronization problem with additive Gaussian noises and incomplete data [2, 5, 17, 39]. Let $z_1^*, \ldots, z_n^* \in \mathbb{C}_1$ where $\mathbb{C}_1 := \{x \in \mathbb{C} : |z| = 1\}$, the set of all unit complex numbers. Then each z_j^* can be written equivalently as $e^{i\theta_j^*}$ for some phase (or angle) $\theta_j^* \in [0, 2\pi)$. For each $1 \le j < k \le n$, the observation $X_{jk} \in \mathbb{C}$ is missing at random. Let $A_{jk} \in \{0, 1\}$ and X_{jk} satisfy

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

where $A_{jk} \sim \text{Bernoulli}(p)$ and $W_{jk} \sim \mathcal{CN}(0, 1)$. That is, each X_{jk} is missing with probability $1 - p$ and is denoted as 0. If it is not missing, it is equal to $z_j^* \overline{z_k^*}$ with an additive noise σW_{jk} where W_{jk} follows the standard complex Gaussian distribution: $\text{Re}(W_{jk}), \text{Im}(W_{jk}) \sim \mathcal{N}(0, 1/2)$ independently. Each A_{jk} is the indicator of whether X_{jk} is observed or not. We assume all random variables $\{A_{jk}\}_{1 \le j < k \le n}$, $\{W_{j,k}\}_{1 \le j < k \le n}$ are independent of each other. The goal is to estimate the phase vector $z^* := (z_1^*, \ldots, z_n^*) \in \mathbb{C}_1^n$ from $\{A_{jk}\}_{1 \le j < k \le n}$ and $\{X_{jk}\}_{1 \le j < k \le n}$.

The observations can be seen as entries of a matrix $X \in \mathbb{C}^{n \times n}$ with $X_{jj} := 0$ and $X_{kj} := \overline{X_{jk}}$ for any $1 \le j < k \le n$. Define $A_{jj} := 0$ and $A_{kj} := A_{jk}$ for all $1 \le j < k \le n$. Then the matrix $A \in \{0, 1\}^{n \times n}$ can be interpreted as the adjacency matrix of an Erdős–Rényi random graph with edge probability p. Define $W \in \mathbb{C}^{n \times n}$ such that $W_{jj} := 0$ and $W_{kj} := \overline{W_{jk}}$ for all $1 \le j < k \le n$. Then all the matrices A, W, X are Hermitian and X can be written equivalently as

$$
(2) \qquad X = A \circ (z^* z^{*\mathrm{H}} + \sigma W) = A \circ (z^* z^{*\mathrm{H}}) + \sigma A \circ W.
$$

Note that X can be seen as a noisy version of

$$\mathbb{E}X = pz^*z^{*\mathrm{H}} - pI_n \tag{3}$$

whose leading eigenvector is $z^*/\sqrt{n}$. This motivates the following spectral method [12, 17, 35]. Let $u \in \mathbb{C}^n$ be the leading eigenvector of X. Then the spectral estimator $\widehat{z} \in \mathbb{C}_1^n$ is defined as

$$\widehat{z}_j := \begin{cases} \dfrac{u_j}{|u_j|} & \text{if } u_j \neq 0, \\ 1 & \text{if } u_j = 0, \end{cases} \tag{4}$$

for each $j \in [n]$, where each u_j is normalized so that $\widehat{z}_j \in \mathbb{C}_1$. The performance of the spectral estimator can be quantified by a normalized squared ℓ_2 loss

$$\ell(\widehat{z}, z^*) := \min_{a \in \mathbb{C}_1} \frac{1}{n} \sum_{j=1}^{n} |\widehat{z}_j - z_j^* a|^2, \tag{5}$$

where the minimum over $\mathbb{C}_1$ is due to the fact that $z_1^*, \ldots, z_n^*$ are identifiable only up to a phase.

The spectral estimator $\widehat{z}$ is simple and easy to implement. Regarding its theoretical performance, it was suggested in [17] that an upper bound $\ell(\widehat{z}, z^*) \leq C(\sigma^2 + 1)/(np)$ holds with high probability for some constant C. However, the minimax risk of the phase synchronization was established in [17] and has the following lower bound:

$$\inf_{z \in \mathbb{C}^n} \sup_{z^* \in \mathbb{C}_1^n} \mathbb{E}\ell(z, z^*) \geq (1 - o(1)) \frac{\sigma^2}{2np}. \tag{6}$$

To provably achieve the minimax risk, the spectral method is often used as an initialization for some more sophisticated procedures. For example, it was used to initialize a generalized power method (GPM) [7, 28, 30] in [17]. Nevertheless, numerically the performance of the spectral method is already very good and the improvement from GPM is often marginal. This raises the following questions about the performance of the spectral method: Can we derive a sharp upper bound? Does the spectral method already achieve the minimax risk or not?

In this paper, we provide complete answers to these questions. We carry out a sharp ℓ_2 analysis of the performance of the spectral estimator $\widehat{z}$ and further show it achieves the minimax risk with the correct constant. Our main result is summarized below in Theorem 1.1 in asymptotic form. Its nonasymptotic version will be given in Theorem 2.1 that only requires $\frac{np}{\sigma^2}$, $\frac{np}{\log n}$ to be greater than a certain constant. We note that in this paper, p and σ^2 are not constants but functions of n. This dependence can be more explicitly represented as p_n and σ_n^2. However, for simplicity of notation and readability, we choose to denote these as p and σ^2 throughout the paper.

THEOREM 1.1. *Assume $\frac{np}{\sigma^2} \to \infty$ and $\frac{np}{\log n} \to \infty$. There exists some $\delta = o(1)$ such that with high probability,*

$$\ell(\widehat{z}, z^*) \leq (1 + \delta) \frac{\sigma^2}{2np}. \tag{7}$$

As a consequence, when $\sigma = 0$ (i.e., there is no additive noise), the spectral method recovers z^* exactly (up to a phase) with high probability as long as $\frac{np}{\log n} \to \infty$.

Theorem 1.1 shows that $\widehat{z}$ is not only rate-optimal but also achieves the exact minimax risk with the correct leading constant in front of the optimal rate. The conditions needed in

Theorem 1.1 are necessary for consistent estimation of z^* in the phase synchronization problem. The condition $\frac{np}{\sigma^2} \to \infty$ is needed so that z^* can be estimated with a vanishing error according to the minimax lower bound (6). The condition $\frac{np}{\log n} \to \infty$ allows p to decrease as n grows and is close to the $\frac{np}{\log n} > (1 + \epsilon)$ condition required for A to be connected. If A has disjoint subgraphs, it is impossible to estimate z^* with a global phase. These two conditions are also needed in [17, 18] to establish the optimality of MLE (maximum likelihood estimation), GPM and SDP (semidefinite programming). Under these two conditions, [17] used the spectral method as an initialization for the GPM and shows GPM achieves the minimax risk after $\log(1/\sigma^2)$ iterations. On the contrary, Theorem 1.1 shows that the spectral method already achieves the minimax risk. This means that the spectral estimator is minimax optimal and achieves the correct leading constant whenever consistent estimation is possible, and in this parameter regime, it is as good as MLE, GPM and SDP.

There are two key novel components toward establishing Theorem 1.1. The first is a new idea regarding the choice of the "population eigenvector" as u can be viewed as its sample counterpart obtained from data. Due to (3) and the fact $pz^*z^{*H} = np(z^*/\sqrt{n})(z^*/\sqrt{n})^H$ is rank-one with the eigenvector $z^*/\sqrt{n}$, existing literature such as [17, 19, 25] treated $z^*/\sqrt{n}$ as the population eigenvector and study the perturbation of u with respect to $z^*/\sqrt{n}$. This seems natural but turns out to be unappealing as it fails to explain why the spectral estimator is able to recover all phases exactly when $\sigma = 0$, that is, when there is no additive noise. Instead, we denote $\breve{u} \in \mathbb{R}^n$ as the leading eigenvector of A and regard $u^* \in \mathbb{C}^n$, defined as

$$(8) \qquad u^* := z^* \circ \breve{u},$$

that is, $u_j^* = z_j^* \breve{u}_j$ for each $j \in [n]$, as the population eigenvector. Note that u^* is random as it depends on the graph A. A careful analysis of u^* reveals that it is the leading eigenvector of $A \circ z^* z^{*H}$ (see Lemma 2.1). In addition, Proposition 2.1 shows that with high probability, $u_j^*/|u_j^*| = z_j^*$ for each $j \in [n]$, up to a global phase. Since u equals u^* when $\sigma = 0$, it successfully explains the exact recovery of the spectral method in the no-additive-noise case. Another advantage of viewing u^* as the population eigenvector, instead of $z^*/\sqrt{n}$, is that intuitively u^* is closer to u than $z^*/\sqrt{n}$ is. This is because $A \circ z^* z^{*H}$ is closer to the data matrix X than $pz^* z^{*H}$ is.

The second key component is a novel perturbation analysis for u. Classical matrix perturbation theory such as Davis–Kahan theorem focuses on analyzing $\inf_{b \in \mathbb{C}_1} \|u - u^*b\|$. We go beyond it and show u can be well-approximated by its first-order approximation $\widetilde{u}$ defined as

$$(9) \qquad \widetilde{u} := \frac{Xu^*}{\|Xu^*\|},$$

in the sense that the difference between these two vectors (up to a phase) has a small ℓ_2 norm. This means that when $np \gtrsim \log n$ and $np \gtrsim \sigma^2$, we have

$$(10) \qquad \inf_{b \in \mathbb{C}_1} \|u - \widetilde{u}b\| \lesssim \frac{\sigma^2 + \sigma}{np},$$

with high probability (see Proposition 2.2). In fact, our perturbation analysis extends beyond the phase synchronization problem. What we establish is a general perturbation theory that can be applied to two arbitrary Hermitian matrices (see Lemma 2.2), which might be of independent interest.

With the help of these two key components, we then carry out an entrywise analysis for each $\widehat{z}_j = u_j/|u_j|$. Note that u_j can be decomposed into $\widetilde{u}_j$ and the difference between u_j and $\widetilde{u}_j$ (up to some global phase). We can decompose the error of $\widehat{z}_j$ into two parts, one is related to the estimation error of $\widetilde{u}_j/|\widetilde{u}_j|$, and the other is related to the magnitude of the difference between u_j and $\widetilde{u}_j$ (up to some global phase). Summing over all coordinates,

the first part eventually leads to the minimax risk $(1 + o(1))\sigma^2/2np$ and the second part is essentially negligible due to (10), which leads to the exact minimax optimality of the spectral estimator.

Orthogonal group synchronization. The above analysis for the phase synchronization can be extended to quantify the performance of the spectral method for orthogonal group synchronization, which is about orthogonal matrices instead of phases. Let $d > 0$ be an integer. Define

$$(11) \qquad \mathcal{O}(d) := \{U \in \mathbb{R}^{d \times d} : UU^\mathsf{T} = U^\mathsf{T}U = I_d\}$$

to include all orthogonal matrices in $\mathbb{R}^{d \times d}$. Let $Z_1^*, \ldots, Z_n^* \in \mathcal{O}(d)$. Analogous to (1), we consider the problem with additive Gaussian noise and incomplete data [19, 25, 26]. For each $1 \le j < k \le n$, we observe $\mathcal{X}_{jk} := Z_j^* Z_k^{*\mathsf{T}} + \sigma \mathcal{W}_{jk} \in \mathbb{R}^{d \times d}$ when $A_{jk} = 1$, where $\mathcal{W}_{jk}$ follows the standard matrix Gaussian distribution. The goal is to recover $Z_1^*, \ldots, Z_n^*$ from $\{\mathcal{X}_{jk}\}_{1 \le j < k \le n}$ and $\{A_{j,k}\}_{1 \le j < k \le n}$. This is known as the orthogonal group synchronization (or $\mathcal{O}(d)$ synchronization).

The observations $\{\mathcal{X}_{jk}\}_{1 \le j < k \le n}$ can be seen as submatrices of an $nd \times nd$ matrix $\mathcal{X}$ with $\mathcal{X}_{jj} := 0$ and $\mathcal{X}_{kj} := \mathcal{X}_{jk}^\mathsf{T}$ for any $1 \le j < k \le n$. Then $\mathcal{X}$ is symmetric and can be seen as a noisy version of

$$(12) \qquad \mathbb{E}\mathcal{X} = pZ^* Z^{*\mathsf{T}} - pI_{nd}$$

whose leading eigenspace is $Z^*/\sqrt{n}$, where $Z^* \in \mathcal{O}(d)^n$ is an $nd \times d$ matrix such that its jth submatrix is Z_j^*. Similar to the phase synchronization, we have the following spectral method. Let $\lambda_1 \ge \cdots \ge \lambda_d$ be the largest d eigenvalues of $\mathcal{X}$ and $u_1, \ldots, u_d \in \mathbb{R}^{nd}$ be their corresponding eigenvectors. Denote $U := (u_1, \ldots, u_d) \in \mathbb{R}^{nd \times d}$ as the eigenspace that includes the top d eigenvectors of $\mathcal{X}$. For each $j \in [n]$, denote $U_j \in \mathbb{R}^{d \times d}$ as its jth submatrix. Then the spectral estimator $\widehat{Z}_j \in \mathcal{O}(d)$ is defined as

$$(13) \qquad \widehat{Z}_j := \begin{cases} \mathcal{P}(U_j) & \text{if } \det(U_j) \ne 0, \\ I_d & \text{if } \det(U_j) = 0, \end{cases}$$

for each $j \in [n]$. Here the mapping $\mathcal{P} : \mathbb{R}^{d \times d} \to \mathcal{O}(d)$ is derived from the polar decomposition and serves as a normalization step for each U_j such that $\widehat{Z}_j \in \mathcal{O}(d)$. Let $\widehat{Z} \in \mathcal{O}(d)^n$ be an $nd \times d$ matrix such that $\widehat{Z}_j$ is its jth submatrix for each $j \in [n]$. Then the performance of $\widehat{Z}$ can be quantified by a loss function $\ell^{\mathrm{od}}(\widehat{Z}, Z^*)$ that is analogous to (5). The detailed definitions of $\mathcal{P}$ and ℓ^{od} are deferred to Section 3.

The spectral method $\widehat{Z}$ was used as an initialization in [19] for a variant of GPM to achieve the exact minimax risk $(1 + o(1))\frac{d(d-1)\sigma^2}{2np}$ for $d = O(1)$. To conduct a sharp analysis of its statistical performance, we extend our novel perturbation analysis from analyzing the leading eigenvector to the leading eigenspace. Recall $\breve{u}$ is the leading eigenvector of A. Analogous to (8), we have a novel choice of the population eigenspace U^*, defined as

$$(14) \qquad U^* := Z^* \circ (\breve{u} \otimes \mathbb{1}_d),$$

and view U as its sample counterpart. This is different from existing literature [19, 40] which uses $Z^*/\sqrt{n}$ as the population eigenspace. Our choice of U^* enables the establishment of the exact recovery of the spectral method when there is no additive noise (i.e., $\sigma = 0$), as seen in Proposition 3.1, and is closer to U than $Z^*/\sqrt{n}$ is.

The first-order approximation of U is a matrix determined by $\mathcal{X}U^*$ whose explicit expression will be given later in (22). We then show U can be well-approximated by its first-order approximation, analogous to (9), with a remainder term of a small ℓ_2 norm (see Proposition 3.2). This is a consequence of a more general eigenspace perturbation theory (see

Lemma 3.2) for two arbitrary Hermitian matrices. Using the first-order approximation, we then carry out an entrywise analysis for $\widehat{Z}$. Our main result for the spectral method in the $\mathcal{O}(d)$ synchronization is summarized below in Theorem 1.2. The nonasymptotic version will be given in Theorem 3.1.

THEOREM 1.2. *Assume* $\frac{np}{\sigma^2} \to \infty$, $\frac{np}{\log n} \to \infty$ *and* $2 \leq d = O(1)$. *There exists some* $\delta = o(1)$ *such that with high probability*,

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

As a consequence, when $\sigma = 0$ (*i.e., there is no additive noise*), *the spectral method recovers* Z^* *exactly* (*up to an orthogonal matrix*) *with high probability as long as* $\frac{np}{\log n} \to \infty$.

Theorem 1.2 shows that the spectral method $\widehat{Z}$ achieves exact minimax optimality as it matches the minimax lower bound $(1+o(1))\frac{d(d-1)\sigma^2}{2np}$ established in [19]. Similar to the phase synchronization, the two conditions needed in Theorem 1.2 so that consistent estimation of Z^* is possible. They are also needed in [19] to achieve the minimax risk by a variant of GPM initialized by the spectral method. On the contrary, Theorem 1.2 shows that in this parameter regime, the spectral method is already minimax optimal with the correct leading constant.

Related literature. Synchronization is a fundamental problem in applied math and statistics. Various methods have been studied for both phase synchronization and $\mathcal{O}(d)$ synchronization, including MLE [17, 39], GPM [17, 19, 26, 28, 33, 39], SDP [4, 15, 18, 21, 27, 36, 37], spectral methods [4, 8, 16, 29, 32, 36], and message passing [24, 30, 34]. The theoretical performance of spectral methods was investigated in [17, 19, 25] and crude error bounds under ℓ_2 or Frobenius norm were obtained. An ℓ_∞-type error bound for spectral methods was also given in [25].

Fine-grained perturbation analysis of eigenvectors has gained increasing attention in recent years for various low-rank matrix problems in machine learning and statistics. Existing literature has mostly focused on establishing ℓ_∞ bounds for eigenvectors [1, 11, 14] or $\ell_{2,\infty}$ bounds on eigenspaces [3, 9, 10, 22]. For instance, [1] developed ℓ_∞-type bounds for the difference between eigenvectors (or eigenspaces) and their first-order approximations. In this paper, we focus on developing sharp ℓ_2-type perturbation bounds, where direct applications of existing ℓ_∞-type results will result in extra logarithm factors.

For the phase synchronization problem, [16, 20, 31] investigated variants of spectral methods based on Laplacian matrices. Instead of using the leading eigenvector of X as in this paper, they utilize the eigenvector corresponding to the smallest eigenvalue of $D - X$ or $I_n - D^{-\frac{1}{2}}XD^{-\frac{1}{2}}$, where $D \in \mathbb{R}^{n \times n}$ is the degree matrix of A with diagonal entries $D_{jj} := \sum_{k \neq j} A_{jk}$ and off-diagonal entries set to zero. These studies have established upper bounds for the performance of their spectral methods applicable to general graphs A and additive noise W. Our focus, however, is on Erdős–Rényi random graphs with Gaussian noise. Under our setting, their results imply an upper bound of $\frac{C\sigma^2}{np}$, where C is a constant significantly greater than 1. In contrast, our work establishes a sharp upper bound with the correct leading constant $1/2$. Our analytical approach could potentially be extended to their methods to achieve the correct constant $1/2$.

The existing literature [17–19] explored the exact minimax risk in synchronization problems, focusing primarily on minimax lower bounds and analyzing MLE, GPM and SDP. While our study shares a thematic resemblance with these prior efforts, it fundamentally diverges in both analysis and proof techniques. Previous studies hinge on contraction properties of the generalized power iteration (GPI), demonstrating the iterative reduction in GPM error

until an optimal error is achieved. This approach further interprets MLE as a GPI fixed point and SDP as an extension of GPI in a higher-dimensional space, thereby establishing their optimality. In contrast, this paper employs a novel strategy specifically tailored for the spectral method. Instead of relying on the GPI framework, which proves inadequate for spectral analysis, we introduce a new perturbation toolkit designed for eigenvector analysis. This toolkit provides precise characterization of eigenvector perturbation and leads to the optimality of the spectral method. It opens new avenues for research and application beyond synchronization problems.

Organization. We study the phase synchronization in Section 2. We first establish the exact recovery of the spectral method in the no-additive-noise case in Section 2.1. Then in Section 2.2, we present our main technical tool for quantifying the distance between the leading eigenvector and its first-order approximation. We then carry out an entrywise analysis of the spectral method and obtain nonasymptotic sharp upper bounds in Section 2.3. Finally, we consider the extension to the orthogonal group synchronization in Section 3. Proofs of results for the phase synchronization are given in Section 5. Due to the page limit, we include proofs of results for the orthogonal group synchronization in the Supplementary Material [38].

Notation. For any positive integer n, we write $[n] := \{1, 2, \ldots, n\}$ and $\mathbb{1}_n := (1, 1, \ldots, 1)^{\mathrm{T}} \in \mathbb{R}^n$. Denote I_n as the $n \times n$ identity matrix and $J_n := \mathbb{1}_n \mathbb{1}_n^{\mathrm{T}} \in \mathbb{R}^{n \times n}$ as the $n \times n$ matrix with all entries being one. Given $a, b \in \mathbb{R}$, we denote $a \wedge b := \min\{a, b\}$ and $a \vee b := \max\{a, b\}$. For a complex number $x \in \mathbb{C}$, we use $\bar{x}$ for its complex conjugate, $\mathrm{Re}(x)$ for its real part, $\mathrm{Im}(x)$ for its imaginary part, and $|x|$ for its modulus. Denote $\mathbb{S}_n := \{x \in \mathbb{C}^n : \|x\| = 1\}$ as including all unit vectors in $\mathbb{C}^n$. For a complex vector $x = (x_j) \in \mathbb{C}^d$, we denote $\|x\| := (\sum_{j=1}^d |x_j|^2)^{1/2}$ as its Euclidean norm and $\|x\|_\infty := \max_{j \in [d]} |x_j|$ as its ℓ_∞ norm. For a complex matrix $B = (B_{jk}) \in \mathbb{C}^{d_1 \times d_2}$, we use $B^{\mathrm{H}} \in \mathbb{C}^{d_2 \times d_1}$ for its conjugate transpose such that $(B^{\mathrm{H}})_{jk} = \overline{B_{kj}}$. The Frobenius norm and the operator norm of B are defined by $\|B\|_{\mathrm{F}} := (\sum_{j=1}^{d_1} \sum_{k=1}^{d_2} |B_{jk}|^2)^{1/2}$ and $\|B\| := \sup_{u \in \mathbb{C}^{d_1}, v \in \mathbb{C}^{d_2} : \|u\| = \|v\| = 1} u^{\mathrm{H}} B v$. We use $\mathrm{Tr}(B)$ for the trace of a squared matrix B and define $\|B\|_{2 \to \infty} := \max_{j \in [d_1]} (\sum_{k=1}^{d_2} |B_{jk}|^2)^{1/2}$. The notation $\det(\cdot)$ and $\otimes$ are used for determinant and Kronecker product. For $U, V \in \mathbb{C}^{d_1 \times d_2}$, $U \circ V \in \mathbb{R}^{d_1 \times d_2}$ is the Hadamard product $(U \circ V)_{jk} = U_{jk} V_{jk}$. For any $B \in \mathbb{R}^{d_1 \times d_2}$, we denote $s_{\min}(B)$ as its smallest singular value. For two positive sequences $\{a_n\}$ and $\{b_n\}$, $a_n \lesssim b_n$ and $a_n = O(b_n)$ both mean $a_n \leq C b_n$ for some constant $C > 0$ independent of n. We also write $a_n = o(b_n)$ or $\frac{b_n}{a_n} \to \infty$ when $\limsup_n \frac{a_n}{b_n} = 0$. We use $\mathbb{I}\{\cdot\}$ as the indicator function. Define $\mathcal{O}(d_1, d_2) := \{V \in \mathbb{R}^{d_1 \times d_2} : V^{\mathrm{T}} V = I_{d_2}\}$ to include all $d_1 \times d_2$ real matrices that have orthonormal columns.

2. Phase synchronization.

2.1. *No-additive-noise case.*

We first study a special case where there is no additive noise (i.e., $\sigma = 0$). In this setting, the data matrix $X = A \circ z^* z^{*\mathrm{H}}$. Despite the data still being missing at random, we are going to show the spectral method is able to recover z^* exactly, up to a phase.

Recall that $\breve{u}$ is the leading eigenvector of A and u^* is defined in (8). The following lemma points out the connection between u^* and $A \circ z^* z^{*\mathrm{H}}$ as well as the connection between eigenvalues of A and those of $A \circ z^* z^{*\mathrm{H}}$.

LEMMA 2.1. *The unit vector u^* is the leading eigenvector of $A \circ z^* z^{*\mathrm{H}}$. That is, with λ^* denoting as the largest eigenvalue of $A \circ z^* z^{*\mathrm{H}}$, we have*

$$(15) \qquad (A \circ z^* z^{*\mathrm{H}}) u^* = \lambda^* u^*.$$

In addition, all the eigenvalues of A are also eigenvalues of $A \circ z^* z^{*\mathrm{H}}$, and vice versa.

Since $X = A \circ z^* z^{*H}$ in the no-additive-noise case, we have $u = u^*$. Note that $\widehat{z}_j = u_j / |u_j| = u_j^* / |u_j^*| = z_j^* \breve{u}_j / |z_j^* \breve{u}_j|$ for each $j \in [n]$. If $\breve{u}_j > 0$, we have $\widehat{z}_j = z_j^*$. If $\widehat{u}_j < 0$ instead, then $\widehat{z}_j = -z_j^*$. If all the coordinates of $\breve{u}$ are positive (or negative), we have $\widehat{z}$ being equal to z^* (or $-z^*$) exactly. The following proposition provides an ℓ_∞ control for the difference between $\breve{u}$ and $\mathbb{1}_n / \sqrt{n}$, which are eigenvectors of A and $\mathbb{E}A$, respectively. The proof of (16) follows proofs of results in [1]. When the right-hand side of (16) is smaller than $1/\sqrt{n}$, it immediately establishes the exact recovery of $\widehat{z}$.

PROPOSITION 2.1. *There exist some constants $C_1, C_2 > 0$ such that if $\frac{np}{\log n} > C_1$, we have*

$$(16) \qquad \min_{b \in \{1, -1\}} \max_{j \in [n]} \left| \breve{u}_j - \frac{1}{\sqrt{n}} b \right| \leq C_2 \left(\sqrt{\frac{\log n}{np}} + \frac{1}{\log(np)} \right) \frac{1}{\sqrt{n}},$$

with probability at least $1 - 8n^{-10}$. As a result, if $\frac{np}{\log n} > \max\{C_1, 2C_2^2\}$, we have $\ell(\widehat{z}, z^*) = 0$ with probability at least $1 - 8n^{-10}$.

Lemma 2.1 and Proposition 2.1 together establish the exact recovery of $\widehat{z}$ for the special case where $\sigma = 0$, through studying u^*. This provides a starting point for our analysis of the general case where $\sigma \neq 0$. From (2), the data matrix X is a noisy version of $A \circ z^* z^{*H}$ with additive noise $\sigma A \circ W$ that scales with σ. As a result, in the following sections, we view u^* as the population eigenvector and u as its sample counterpart, studying the performance of the spectral method.

2.2. *First-order approximation of the leading eigenvector.* In this section, we provide a fine-grained perturbation analysis for the eigenvector u. Classical matrix perturbation theory, such as Davis–Kahan theorem, can only give a crude upper bound for $\inf_{b \in \mathbb{C}_1} \|u - u^* b\|$, which turns out to be insufficient to derive a sharp bound for $\ell(\widehat{z}, z^*)$. Instead, we develop a more powerful tool for perturbation analysis of u using its first-order approximation $\widetilde{u}$ defined in (9). In fact, our tool goes beyond the phase synchronization problem and can be applied to arbitrary Hermitian matrices.

LEMMA 2.2. *Consider two Hermitian matrices $Y, Y^* \in \mathbb{C}^{n \times n}$. Let $\mu_1^* \geq \mu_2^* \geq \cdots \geq \mu_n^*$ 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*

$$\inf_{b \in \mathbb{C}_1} \left\| v - \frac{Yv^*}{\|Yv^*\|} b \right\| \leq \frac{40\sqrt{2}}{9(\mu_1^* - \mu_2^*)} \left(\left(\frac{4}{\mu_1^* - \mu_2^*} + \frac{2}{\mu_1^*} \right) \|Y - Y^*\|^2 \right.$$
$$\left. + \frac{\max\{|\mu_2^*|, |\mu_n^*|\}}{\mu_1^*} \|Y - Y^*\| \right).$$

In Lemma 2.2, there are two matrices Y, Y^* whose leading eigenvectors are v, v^*, respectively. It studies the ℓ_2 difference between v and $Yv^*/\|Yv^*\|$ up to a phase. Let μ_1 be the largest eigenvalue of Y. The unit vector $Yv^*/\|Yv^*\|$ is interpreted as the first-order approximation of v, as v can be decomposed into $v = Yv/\mu_1 = Yv^*/\mu_1 + Y(v - v^*)/\mu_1$ where the first term Yv^*/μ_1 is proportional to $Yv^*/\|Yv^*\|$. If Y^* is rank-one, meaning $\mu_2^* = \mu_n^* = 0$, the upper bound in Lemma 2.2 becomes $80\sqrt{2}\|Y - Y^*\|^2/(3\mu_1^{*2})$. Lemma 2.2 itself might be of independent interest and be useful in other low-rank matrix problems.

The key to Lemma 2.2 is the following equation. Since $\mu_1 v = Yv$ and $\|Yv^*\| \frac{Yv^*}{\|Yv^*\|} = Yv^*$, we can derive (see (29) in the proof of Lemma 2.2):

$$\mu_1^{-1} \|Yv^*\| (\mu_1 I_n - Y)\left(v - \frac{Yv^*}{\|Yv^*\|}\right) = Y(\mu_1^{-1} Yv^* - v^*).$$

Its left-hand side can be shown to be related to $\inf_{b \in \mathbb{C}_1} \|v - Yv^*b/\|Yv^*\|\|$. By carefully studying and upper bounding its right-hand side, which does not involve v, we derive Lemma 2.2.

Lemma 2.2 requires that the perturbation between Y and Y^* is not only small compared to the eigengap $\mu_1^* - \mu_2^*$, but also small compared to the leading eigenvalue μ_1^*. A similar requirement is also needed in [1] to establish ℓ_∞ bounds for the difference between the eigenvector and its first-order approximation. In contrast, classical theory such as Davis–Kahan theorem (see Lemma 5.1) only needs the perturbation to be small compared to the eigengap to bound $\inf_{b \in \mathbb{C}_1} \|v - v^*b\|$. A natural question is whether the bound in Lemma 2.2 can be modified to depend on eigenvalues only through the eigengaps. It turns out this is not feasible, as it deals with the distance between v and its first-order approximation $Yv^*/\|Yv^*\|$, not the distance between v and v^* as in Davis–Kahan theorem. To illustrate it, consider the following counterexample. Let $e_1, \ldots, e_n$ be the canonical basis of $\mathbb{R}^n$. Let $\delta > 0$. Define

$$(17) \quad Y^* := \operatorname{diag}(0, -1, -1, \ldots, -1) \in \mathbb{R}^{n \times n}, \quad \text{and} \quad Y := Y^* + \delta(e_1 + e_2)(e_1 + e_2)^{\mathsf{T}}/2.$$

Then $\mu_1^* = 0$, $\mu_2^* = -1$, $\mu_1^* - \mu_2^* = 1$, $v^* = e_1$, $\|Y - Y^*\| = \delta$ and $Yv^*/\|Yv^*\| = (e_1 + e_2)/\sqrt{2}$. We can show v has the following explicit expression (see Appendix C in the Supplementary Material for detailed calculation):

$$(18) \qquad v = \sqrt{\frac{1}{2}\left(1 + \frac{1}{\sqrt{1+\delta^2}}\right)} e_1 + \sqrt{\frac{1}{2}\left(1 - \frac{1}{\sqrt{1+\delta^2}}\right)} e_2.$$

When δ is sufficiently close to 0, we have $v \approx v^*$. This is not surprising as it is consistent with the bound from Davis–Kahan theorem as the ratio between the perturbation and eigengap is $\|Y - Y^*\|/(\mu_1^* - \mu_2^*) = \delta \approx 0$. On the other hand, $\|v - Yv^*/\|Yv^*\|\| \approx \|e_1 - (e_1 + e_2)/\sqrt{2}\| = 2 - \sqrt{2} > 0$ no matter how small δ may be. As a result, in this counterexample, $Yv^*/\|Yv^*\|$ is not a good approximation of v despite the sufficiently small perturbation.

Applying Lemma 2.2 to the phase synchronization, we have the following result.

PROPOSITION 2.2.   *There exist constants $C_1, C_2, C_3 > 0$ such that if $\frac{np}{\log n} > C_1$ and $\frac{np}{\sigma^2} > C_2$, we have*

$$\inf_{b \in \mathbb{C}_1} \|u - \widetilde{u}b\| \le C_3 \frac{\sigma^2 + \sigma}{np},$$

with probability at least $1 - 3n^{-10}$.

Proposition 2.2 shows that u is well-approximated by its first-order approximation $\widetilde{u}$ (up to a phase) with an approximation error that is at most in the order of $(\sigma^2 + \sigma)/np$. Note that we can show $\inf_{b \in \mathbb{C}_1} \|u - u^*b\|$ is of order $\sigma/\sqrt{np}$ by using Davis–Kahan theorem. This is much larger than the upper bound derived in Proposition 2.2, particularly when np/σ^2 is large. As a result, $\widetilde{u}$ provides a precise characterization of u with negligible ℓ_2 error.

2.3. *Sharp ℓ_2 analysis of the spectral estimator.* In this section, we will conduct a sharp analysis of the performance of the spectral estimator $\widehat{z}$ using the first-order approximation $\widetilde{u}$ of the eigenvector u. According to Proposition 2.2, u is close to $\widetilde{u}$ (up to a phase) with a small difference. Then intuitively, $\widehat{z}$ should be close to its counterpart that uses $\widetilde{u}$ instead of u in (4), up to a global phase. For each $j \in [n]$, the distance of $\widetilde{u}_j/|\widetilde{u}_j|$ from z_j^* is essentially determined by $\overline{z_j^*}\widetilde{u}_j$. By the definition in (9), $\widetilde{u}_j$ is proportional to $[Xu^*]_j$, the jth coordinate of Xu^*. With (2), it leads to $\overline{z_j^*}\widetilde{u}_j \propto \lambda^* \overline{z_j^*}u_j^* + \sigma \sum_{k \neq j} A_{jk}W_{jk}\overline{z_j^*}u_k^*$. Here the first term $\lambda^* \overline{z_j^*}u_j^*$ can be interpreted as the signal as it is related to the population quantity u_j^*, which gives the exact recovery of the spectral method in the no-additive-noise case in Proposition 2.1. As u^* is close to $z^*/\sqrt{n}$, the second term is approximately equal to $n^{-1/2}\sum_{k \neq j} A_{jk}W_{jk}\overline{z_j^*}z_k^*$. Its contribution toward the estimation error is essentially determined by its imaginary part $n^{-1/2}\mathrm{Im}(\sum_{k \neq j} A_{jk}W_{jk}\overline{z_j^*}z_k^*)$, which can be interpreted as the main error term. Summing over all $j \in [n]$, the signals and the main error terms together lead to the minimax risk $\sigma^2/(2np)$. At the same time, contributions of approximation errors such as $\inf_{b \in \mathbb{C}_1} \|u - \widetilde{u}b\|$ turn out to be negligible. This leads to the following theorem on the performance of the spectral estimator.

THEOREM 2.1. *There exist constants $C_1, C_2, C_3 > 0$ such that if $\frac{np}{\log n} > C_1$ and $\frac{np}{\sigma^2} > C_2$, we have*

$$\ell(\widehat{z}, z^*) \leq \left(1 + C_3\left(\left(\frac{\sigma^2}{np}\right)^{\frac{1}{4}} + \sqrt{\frac{\log n}{np}} + \frac{1}{\log(np)}\right)\right)\frac{\sigma^2}{2np},$$

with probability at least $1 - n^{-9} - \exp(-\frac{1}{32}(\frac{np}{\sigma^2})^{\frac{1}{4}})$.

Theorem 2.1 is nonasymptotic and its asymptotic version is presented in Theorem 1.1. It covers the no-additive-noise case (i.e., Proposition 2.1), as it implies that $\ell(\widehat{z}, z^*) = 0$ with high probability when $\sigma = 0$. Theorem 2.1 shows that $\ell(\widehat{z}, z^*)$ is equal to $\sigma^2/(2np)$ up to a factor that is determined by $(\sigma^2/(np))^{1/4}$, $\sqrt{\log n/(np)}$ and $1/\log(np)$. The first term is related to various approximation errors including the one from Proposition 2.2. The second and third terms are derived from (16).

We can make a comparison between Theorem 2.1 and the existing result $\ell(\widehat{z}, z^*) \lesssim (\sigma^2 + 1)/np$ in [17]. There are two main improvements. First, we obtain the exact constant $1/2$ for the error term $\frac{\sigma^2}{np}$, which gives a more accurate characterization of the performance of the spectral estimator. Second, the $1/np$ error term in $(\sigma^2 + 1)/np$ no longer exists in Theorem 2.1. We further compare Theorem 2.1 with the minimax lower bound for the phase synchronization problem. The paper [17] proved that there exist constants $C_4, C_5 > 0$ such that if $\frac{np}{\sigma^2} \geq C_4$, we have

$$(19) \qquad \inf_{z \in \mathbb{C}^n} \sup_{z^* \in \mathbb{C}_1^n} \mathbb{E}\ell(z, z^*) \geq \left(1 - C_5\left(\frac{\sigma^2}{np} + \frac{1}{n}\right)\right)\frac{\sigma^2}{2np}.$$

Compared with (19), the spectral estimator $\widehat{z}$ is exact minimax optimal as it not only achieves the correct rate $\sigma^2/(np)$ but also the correct constant $1/2$. Under the parameter regime as in Theorem 2.1, [17, 18] showed that MLE, GPM (if properly initialized), and SDP achieve the exact minimax risk. Theorem 2.1 points out that the spectral method is as good as these methods.

3. Orthogonal group synchronization. In this section, we will extend our analysis to matrix synchronizations where the quantities of interest are orthogonal matrices instead of phases. The orthogonal group synchronization problem has been briefly introduced in Section 1. Here we provide more details about the problem.

Let $d > 0$ be an integer. Recall the definition of $\mathcal{O}(d)$ in (11) and that $Z_1^*, \ldots, Z_n^* \in \mathcal{O}(d)$. For each $1 \leq j < k \leq n$, the observation $\mathcal{X}_{jk} \in \mathbb{R}^{d \times d}$ is given by

$$(20) \qquad \mathcal{X}_{jk} := \begin{cases} Z_j^* Z_k^{*\mathrm{T}} + \sigma \mathcal{W}_{jk} & \text{if } A_{jk} = 1, \\ 0 & \text{if } A_{jk} = 0, \end{cases}$$

where $A_{jk} \sim \text{Bernoulli}(p)$ and $\mathcal{W}_{jk} \sim \mathcal{MN}(0, I_d, I_d)$, that is, the standard matrix Gaussian distribution.[1] We assume $\{A_{jk}\}_{1 \leq j < k \leq n}$, $\{\mathcal{W}_{j,k}\}_{1 \leq j < k \leq n}$ are all independent of each other. Similar to the phase synchronization problem, the observations are missing at random with additive Gaussian noises. The goal is to recover $Z_1^*, \ldots, Z_n^*$ from $\{\mathcal{X}_{jk}\}_{1 \leq j < k \leq n}$ and $\{A_{j,k}\}_{1 \leq j < k \leq n}$.

The data matrix $\mathcal{X} \in \mathbb{R}^{nd \times nd}$ can be written equivalently in a way that is analogous to (2). Define $A_{jj} := 0$ and $A_{kj} := A_{jk}$ for all $1 \leq j < k \leq n$. Define $\mathcal{W} \in \mathbb{C}^{nd \times nd}$ such that $\mathcal{W}_{jj} := 0_{d \times d}$ and $\mathcal{W}_{kj} := \mathcal{W}_{jk}^{\mathrm{T}}$ for all $1 \leq j < k \leq n$. Then we have the expression

$$(21) \qquad \mathcal{X} = (A \otimes J_d) \circ \left(Z^* Z^{*\mathrm{T}} + \sigma \mathcal{W} \right) = (A \otimes J_d) \circ Z^* Z^{*\mathrm{T}} + \sigma(A \otimes J_d) \circ \mathcal{W}.$$

From (12), the data matrix $\mathcal{X}$ can be seen as a noisy version of $pZ^* Z^{*\mathrm{T}} - pI_{nd}$. Since the columns of Z^* are orthogonal to each other, we have the following eigendecomposition: $pZ^* Z^{*\mathrm{T}} = np(Z^*/\sqrt{n})(Z^*/\sqrt{n})^{\mathrm{T}}$ where $Z^*/\sqrt{n} \in \mathcal{O}(nd, d)$. That is, np is the only nonzero eigenvalue of $pZ^* Z^{*\mathrm{T}}$ with multiplicity d.

The definition of the spectral estimator $\widehat{Z}_1, \ldots, \widehat{Z}_n$ is given in (13). The mapping $\mathcal{P} : \mathbb{R}^{d \times d} \to \mathcal{O}(d)$ is from the polar decomposition and is defined as follows. For any matrix $B \in \mathbb{R}^{d \times d}$ that is full-rank, it admits a singular value decomposition (SVD): $B = MDV^{\mathrm{T}}$ with $M, V \in \mathcal{O}(d)$ and D a diagonal matrix. Then its polar decomposition is $B = (MV^{\mathrm{T}})(VDV^{\mathrm{T}})$ and $\mathcal{P}(B) := MV^{\mathrm{T}}$ is defined as its first factor.

Recall that $\breve{u}$ is the leading eigenvector of A and the population eigenspace U^* is defined in (14). That is, $U^* \in \mathbb{R}^{nd \times d}$ and its jth submatrix is $U_j^* = \breve{u}_j Z_j^* \in \mathbb{R}^{d \times d}$ for each $j \in [n]$. Following the proof of Lemma 2.1, we can show U^* is the leading eigenspace of $(A \otimes J_d) \circ Z^* Z^{*\mathrm{T}}$:

LEMMA 3.1. *Denote* $\lambda_1^* \geq \lambda_2^* \geq \cdots \geq \lambda_{nd}^*$ *as the eigenvalues of* $(A \otimes J_d) \circ Z^* Z^{*\mathrm{T}}$*. Then* $\lambda_1^* = \lambda_2^* = \cdots = \lambda_d^*$*, all equal the leading eigenvalue of* A*. In addition,* λ_{d+1}^* *is equal to the second largest eigenvalue of* A*. Furthermore,* U^* *is the eigenspace of* $(A \otimes J_d) \circ Z^* Z^{*\mathrm{T}}$ *corresponding to* λ_1^**, that is,*

$$\left((A \otimes J_d) \circ Z^* Z^{*\mathrm{T}} \right) U^* = \lambda_1^* U^*.$$

Following the proof of Proposition 2.1, particularly using (16), we can further establish the exact recovery of $\widehat{Z}$, up to an orthogonal matrix, in the no-additive-noise case.

PROPOSITION 3.1. *Consider the no-additive-noise case where* $\sigma = 0$*. There exists some constant* $C_1 > 0$ *such that if* $\frac{np}{\log n} > C_1$*, we have* $\ell(\widehat{z}, z^*) = 0$ *with probability at least* $1 - 7n^{-10}$*.*

---

[1] A random matrix X follows a matrix Gaussian distribution $\mathcal{MN}(M, \Sigma, \Omega)$ if its density function is proportional to $\exp(-\frac{1}{2}\mathrm{Tr}(\Omega^{-1}(X - M)^{\mathrm{T}}\Sigma^{-1}(X - M)))$.

Similar to the phase synchronization, we can study the first-order approximation of the eigenspace U. Denote $\Lambda := \operatorname{diag}(\lambda_1, \ldots, \lambda_d) \in \mathbb{R}^{d \times d}$ as the diagonal matrix of the d largest eigenvalues of $\mathcal{X}$. Then U can be expressed as $U = \mathcal{X} U \Lambda^{-1}$. Define

$$\widetilde{U} := \underset{U' \in \mathcal{O}(nd, d)}{\operatorname{argmin}} \|U' - \mathcal{X} U^*\|_{\mathrm{F}}^2. \tag{22}$$

Then $\widetilde{U}$ is the projection of $\mathcal{X} U^*$ onto $\mathcal{O}(nd, d)$. This is similar to the definition of $\widetilde{u}$ in (9) for the phase synchronization, where $\widetilde{u}$ is the projection of $X u^*$ onto the unit sphere. As a result, $\widetilde{U}$ can be regarded as the first-order approximation of U.

The following lemma provides an upper bound for a leading eigenspace and its first-order approximation of two arbitrary Hermitian matrices. It is an extension of Lemma 2.2 which is only about the perturbation of a leading eigenvector. The proof of Lemma 3.2 follows that of Lemma 2.2 but is more involved, as it needs to deal with matrix multiplication which is not commutative.

LEMMA 3.2.   *Consider two symmetric matrices $Y, Y^* \in \mathbb{R}^{n \times n}$. Let $\mu_1^* \geq \mu_2^* \geq \cdots \geq \mu_n^*$ be the eigenvalues of Y^*. Let $V^* \in \mathbb{R}^{n \times d}$ (resp. V) be the leading eigenspace of Y^* (resp. Y) corresponding to its d largest eigenvalues. Define $\widetilde{V} := \operatorname{argmin}_{V' \in \mathcal{O}(n,d)} \|V' - Y V^*\|_{\mathrm{F}}^2$. If $\|Y - Y^*\| \leq \min\{\mu_d^* - \mu_{d+1}^*, \mu_d^*\}/4$, we have*

$$\inf_{O \in \mathcal{O}(d)} \|V - \widetilde{V} O\|$$

$$\leq \frac{16\sqrt{2}}{3(\mu_d^* - \mu_{d+1}^*)\mu_d^*} \left(\frac{2\mu_1^*}{3(\mu_d^* - \mu_{d+1}^*)} + 1 \right) \|Y - Y^*\|^2$$

$$+ \frac{8\sqrt{2}}{3(\mu_d^* - \mu_{d+1}^*)\mu_d^*} \left(\frac{4\mu_1^*(\mu_1^* - \mu_d^*)}{\mu_d^* - \mu_{d+1}^*} + 2(\mu_1^* - \mu_d^*) + \max\{|\mu_{d+1}^*|, |\mu_n^*|\} \right) \|Y - Y^*\|.$$

Lemma 3.2 includes Lemma 2.2 as a special case when $d = 1$. For $d > 1$, if $\mu_1^* = \mu_d^*$, that is, the largest d eigenvalues of Y^* are all equal, the upper bound in Lemma 3.2 simplifies to

$$\inf_{O \in \mathcal{O}(d)} \|V - \widetilde{V} O\| \lesssim \frac{1}{\mu_d^* - \mu_{d+1}^*} \left(\left(\frac{1}{\mu_d^* - \mu_{d+1}^*} + \frac{1}{\mu_d^*} \right) \|Y - Y^*\|^2 \right.$$

$$\left. + \frac{\max\{|\mu_{d+1}^*|, |\mu_n^*|\}}{\mu_d^*} \|Y - Y^*\| \right),$$

which is similar in form to the upper bound in Lemma 2.2. This expression can be used in the $\mathcal{O}(d)$ synchronization problem as λ_1^* is shown to be equal to λ_d^* in Lemma 3.1. A direct application of this expression leads to the following proposition regarding the perturbation between U and $\widetilde{U}$.

PROPOSITION 3.2.   *Assume $2 \leq d \leq C_0$ for some constant $C_0 > 0$. There exist constants $C_1, C_2, C_3 > 0$ such that if $\frac{np}{\log n} > C_1$ and $\frac{np}{\sigma^2} > C_2$, we have*

$$\inf_{O \in \mathcal{O}(d)} \|U - \widetilde{U} O\| \leq C_3 \frac{\sigma^2 d + \sigma\sqrt{d}}{np},$$

with probability at least $1 - 6n^{-10}$.

When $d = 1$, Proposition 3.2 reduces to Proposition 2.2. With Proposition 3.2, we can carry out a sharp ℓ_2 analysis of the performance of the spectral estimator $\widehat{Z}$ using $\widetilde{U}$. The loss function is defined analogously to (5) as

$$\ell^{\mathrm{od}}(\widehat{Z}, Z^*) := \min_{O \in \mathcal{O}(d)} \frac{1}{n} \|\widehat{Z}_j - Z_j^* O\|_{\mathrm{F}}^2.$$

In this way, we have the following theorem which is similar to Theorem 2.1. Its asymptotic version is given in Theorem 1.2. The proof of Theorem 3.1 follows that of Theorem 2.1 but is more complicated due to the existence of the mapping $\mathcal{P}$ in the definition of the spectral method. To prove Theorem 3.1, note that for each $j \in [n]$, $\|\widehat{Z}_j - Z_j^*\|_{\mathrm{F}} = \|\mathcal{P}(U_j) - Z_j^*\|_{\mathrm{F}} = \|\mathcal{P}(Z_j^{*\mathrm{T}} U_j) - I_d\|_{\mathrm{F}}$ where $Z_j^{*\mathrm{T}} U_j$ can be approximated by $Z_j^{*\mathrm{T}} \widetilde{U}_j$ according to Proposition (3.2). The term $Z_j^{*\mathrm{T}} \widetilde{U}_j$ can be further expanded using (21) and Lemma 3.1, leading to $\sum_{k \neq j} A_{jk} Z_j^{*\mathrm{T}} \mathcal{W}_{jk} Z_k^*$ and several approximation error terms. Careful analysis of $\sum_{k \neq j} A_{jk} Z_j^{*\mathrm{T}} \mathcal{W}_{jk} Z_k^*$ eventually leads to the minimax risk $d(d-1)\sigma^2/(2np)$ and all the other error terms turn out to be negligible.

THEOREM 3.1. *Assume $2 \leq d \leq C_0$ for some constant $C_0 > 0$. There exist constants C_1, C_2, C_3 such that if $\frac{np}{\log n} > C_1$ and $\frac{np}{\sigma^2} > C_2$, we have*

$$\ell^{od}(\widehat{Z}, Z^*) \leq \left(1 + C_3 \left(\left(\frac{\sigma^2}{np}\right)^{\frac{1}{4}} + \sqrt{\frac{\log n}{np}} + \frac{1}{\log(np)}\right)\right) \frac{d(d-1)\sigma^2}{2np}$$

holds with probability at least $1 - n^{-9} - \exp(-\frac{1}{32}(\frac{np}{\sigma^2})^{\frac{1}{4}})$.

We can compare the upper bound in Theorem 3.1 to existing results for the $\mathcal{O}(d)$ synchronization. [19] derived an upper bound for the spectral method: $\ell(\widehat{Z}, Z^*) \lesssim d^4(1 + \sigma^2 d)/(np)$ with high probability. In comparison, our upper bound has a smaller factor of $d(d-1)/2$ for σ^2/np. In addition, it does not have the d^4/np error term. The paper [19] also established the minimax lower bound: when $2 \leq d \leq C_0$, there exist constants $C_4, C_5 > 0$ such that if $\frac{np}{\sigma^2} > C_4$, we have

$$\inf_{Z \in \mathcal{O}(d)^n} \sup_{Z^* \in \mathcal{O}(d)^n} \ell^{\mathrm{od}}(\widehat{Z}, Z^*) \geq \left(1 - C_5 \left(\frac{1}{n} + \frac{\sigma^2}{np}\right)\right) \frac{d(d-1)\sigma^2}{2np}.$$

Compared to the lower bound, the spectral estimator $\widehat{Z}$ is exact minimax optimal as it achieves the correct rate with the correct constant $d(d-1)/2$ in front of the optimal rate σ^2/np.

4. Discussions.

4.1. *Comparison of spectral method and other methods.* In synchronization problems, the spectral method offers computational advantages over alternative methods such as MLE, SDP and GPM. According to Theorem 1.1, the spectral method attains statistical optimality in the limit as $\frac{np}{\sigma^2} \to \infty$, achieving the minimum possible risk. The performance of the spectral method in scenarios where $\frac{np}{\sigma^2}$ does not approach infinity, however, remains less understood.

Previous studies [21, 23] have explored the PCA method in Bayesian settings for synchronization problems with $p = 1$. Unlike the spectral method, as defined in (4), PCA does not involve entrywise normalization but scales the leading eigenvector u to minimize the mean square error (MSE). These studies offer a comprehensive asymptotic analysis of PCA's MSE

and that of the Bayes-optimal estimator, demonstrating both methods' ability to achieve substantial accuracy when σ^2 is below a specific threshold. However, PCA tends to exhibit a higher MSE compared to the Bayes-optimal estimator. Furthermore, [21] indicates that the MSE of SDP falls between that of PCA and the Bayes-optimal estimator, leaning more towards the latter.

While Theorem 1.1 addresses the regime where $\frac{np}{\sigma^2} \to \infty$, Theorem 2.1 establishes an upper bound in scenarios where $\frac{np}{\sigma^2}$ exceeds a certain constant. This suggests a complex interplay between the performance of the spectral method and the ratio $\frac{\sigma^2}{np}$ in the constant $\frac{np}{\sigma^2}$ regime. To better understand this relationship, we conducted numerical experiments using the spectral method, GPM and SDP under various σ^2 levels. The GPM, initialized with the spectral estimator $\widehat{z}_{\mathrm{GPM}}^{(0)} := \widehat{z}$, iteratively updates $\widehat{z}_{\mathrm{GPM}}^{(t)} := f(X\widehat{z}_{\mathrm{GPM}}^{(t-1)})$ for $t \geq 1$, where $f : \mathbb{C}^n \to \mathbb{C}_1^n$ is an entrywise normalization function defined as $[f(x)]_i := x_i/|x_i|\mathbb{I}\{x_i \neq 0\} + \mathbb{I}\{x_i = 0\}$ for any $x \in \mathbb{C}^n$. The SDP, a convex optimization problem, maximizes $\max_{Z \in \mathbb{C}^{n \times n}: Z = Z^{\mathrm{H}}, \mathrm{diag}(Z) = I_n, Z \succeq 0} \mathrm{Tr}(XZ)$ over complex positive-semidefinite Hermitian matrices with unit diagonal entries and can be initialized using the spectral method. We assessed their performances using the normalized squared ℓ_2 loss (5).

Figure 1 summarizes the comparative performances of these methods. For low σ^2 values, the error rates of all methods approximate $\frac{\sigma^2}{2np}$. The left panel of the figure shows that as σ^2 increases, their error rates rise more steeply than $\frac{\sigma^2}{2np}$. As σ^2 continues to increase, the spectral method exhibits higher error rates, as expected, since the other two methods use the spectral method for initialization and enhance it through more complex procedures. For a deeper insight into the numerical performance differences, we compare the high-order terms in their errors. Specifically, the normalized squared ℓ_2 loss for each method can be expressed as $(1+\delta)\frac{\sigma^2}{2np}$, where δ represents the high-order term. The right panel of Figure 1 compares δ for these three methods. It reveals that even at small σ^2 values, the spectral method's performance diverges from those of the other methods. This suggests that while δ diminishes to 0 for all three methods as σ^2 decreases (thus achieving exact minimax optimality), the spectral method's δ diminishes more slowly than those of the other two methods.

Deriving explicit expressions for these error rates would be insightful, yet it falls outside the scope of this paper and presents an avenue for future research.
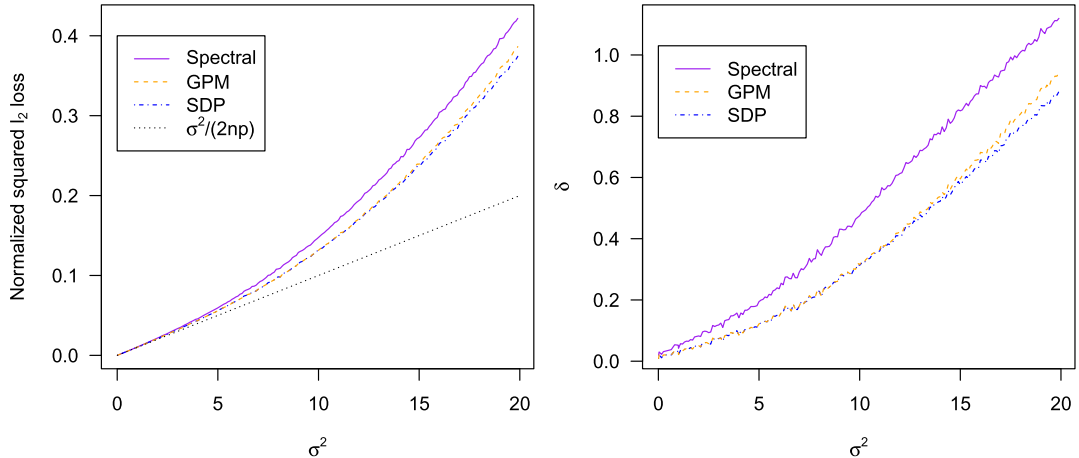

FIG. 1. *Numerical results for the spectral method, GPM and SDP in phase synchronization, with $n = 100$, $p = 0.5$ and σ^2 varying within $[0, 20]$. Left: Error comparison measured by the normalized squared ℓ_2 loss. Right: Comparison of the high-order term in their errors.*

4.2. *Condition on p*.   In the phase synchronization problem (1), observations are missing at random, forming an Erdős–Rényi random graph A with edge probability p. The value of p cannot be excessively small, as this could result in A being disconnected, thereby making accurate estimation of z^* under a global phase impossible. Theorem 1.1 assumes $\frac{np}{\log n} \to \infty$ to establish the exact minimax optimality of the spectral method. A less stringent condition, where $\frac{np}{\log n}$ exceeds a certain constant, is considered in Theorem 2.1. However, it is known that A is connected with high probability when $\frac{np}{\log n} > 1 + \epsilon$ for any constant $\epsilon > 0$. This raises the question of how the spectral method performs when $\frac{np}{\log n}$ is a small constant.

Our analysis requires $\frac{np}{\log n}$ to be greater than a certain constant for several technical reasons. This condition ensures desired bounds hold for critical quantities such as $\|A - \mathbb{E}A\|$ and $\|A \circ W\|$, which are essential for the ℓ_∞ analysis in Proposition 2.1 and the ℓ_2 analysis of the first order approximation in Proposition 2.2. Moreover, the proof of Theorem 2.1 leverages the ℓ_∞ results from Proposition 2.1, leading to the inclusion of the $\sqrt{\frac{\log n}{np}}$ factor in the theorem's upper bound. This requires $\frac{np}{\log n}$ to approach infinity for the upper bound to asymptotically match the exact minimax risk. Obtaining precise bounds for the performance of the spectral method when $\frac{np}{\log n}$ is a small constant would require an extension beyond our current analytic framework, a task we leave for future research.

4.3. *Other low-rank problems*.   The synchronization problems investigated in this manuscript are part of a broader category of problems characterized by low-rank matrix structures disrupted by additive noise and incomplete data. The methodologies developed herein are applicable to a variety of related problems, such as matrix completion, principal component analysis, factor models, mixture models, and ranking from pairwise comparison data. A key observation is that many of these problems encompass multiple sources of randomness, such as that arising from missing data and additive noise. An effective approach, as demonstrated in this study, is to isolate these sources and evaluate their individual contributions to the overall estimation error. This strategy is exemplified in our analysis of synchronization problems, where we introduce a novel population eigenvector and eigenspace. Furthermore, Lemma 2.2 and Lemma 3.2 offer a general framework for the perturbation analysis of eigenvectors and eigenspaces.

On the other hand, synchronization problems are special in that their leading eigenvector or eigenspace is spread out. In the literature [11], the coherence of an eigenvector u is defined as $\max_{i \in [n]} |u_i|^2/n$, where $u_1, \ldots, u_n$ are its coordinates. In phase synchronization, the leading eigenvector of $\mathbb{E}X$ in (3) possesses uniformly equal magnitude $1/\sqrt{n}$, indicating maximal coherence. Contrastingly, in many low-rank problems, eigenvectors exhibit lower coherence, which naturally factors into theoretical analysis. Therefore, when extending the concepts and methodologies from this paper to other scenarios, it is crucial to monitor eigenvector coherence for more precise and insightful analysis.

5. Proofs for phase synchronization.

5.1. *Proof of Lemma 2.2*.   We first present a variant of Davis–Kahan theorem [13] and an inequality about $\inf_{b \in \mathbb{C}_1} \|x - yb\|$ and $\|(I_d - xx^H)y\|$ that will be used in the proof of Lemma 2.2.

LEMMA 5.1.   *Let $X, \widetilde{X} \in \mathbb{C}^{d \times d}$ be two Hermitian matrices. Let $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_d$ be the eigenvalues of X. Consider any $r \in [d]$. Let $U \in \mathbb{C}^{d \times r}$ (resp. $\widetilde{U}$) be the eigenspace of*

X (resp. $\widetilde{X}$) *that includes its leading r eigenvectors. Under the assumption that* $\|X - \widetilde{X}\| <$ $(\lambda_r - \lambda_{r+1})/4$, *we have*

$$\|(I - UU^{\mathrm{H}})\widetilde{U}\| \leq \frac{4\|X - \widetilde{X}\|}{3(\lambda_r - \lambda_{r+1})}.$$

LEMMA 5.2. *For any unit vectors* $x, y \in \mathbb{C}^d$, *we have* $\inf_{b \in \mathbb{C}_1} \|x - yb\| \leq \sqrt{2}\|(I_d - xx^{\mathrm{H}})y\|$.

PROOF OF LEMMA 2.2. Denote $\mu_1 \geq \cdots \geq \mu_n$ as the eigenvalues of Y. We first give some inequalities for the eigenvalues and $\|Yv^*\|$ that will be used later in the proof. By Weyl's inequality, we have

$$\max\{|\mu_1 - \mu_1^*|, |\mu_2 - \mu_2^*|\} \leq \|Y - Y^*\|.$$

Since $\|Y - Y^*\| \leq \min\{\mu_1^* - \mu_2^*, \mu_1^*\}/4$ is assumed, we have

$$(23) \qquad \frac{3}{4}\mu_1^* \leq \mu_1 \leq \frac{5}{4}\mu_1^*, \qquad \mu_1 - \mu_2 \geq \frac{\mu_1^* - \mu_2^*}{2},$$

and

$$(24) \qquad \left|\frac{\mu_1^*}{\mu_1} - 1\right| = \frac{|\mu_1^* - \mu_1|}{\mu_1} \leq \frac{\|Y - Y^*\|}{\mu_1^* - \|Y - Y^*\|} \leq \frac{4\|Y - Y^*\|}{3\mu_1^*}.$$

Regarding $\|Yv^*\|$, using the decomposition

$$Y = Y^* + (Y - Y^*) = \mu_1^* v^* v^{*\mathrm{H}} + (Y^* - \mu_1^* v^* v^{*\mathrm{H}}) + (Y - Y^*),$$

and its consequence

$$(25) \qquad Yv^* = Y^* v^* + (Y - Y^*)v^* = \mu_1^* v^* + (Y - Y^*)v^*,$$

we have

$$(26) \qquad \|Yv^*\| \geq \mu_1^* - \|Y - Y^*\| \geq \frac{3\mu_1^*}{4}.$$

We define $\breve{v} \in \mathbb{C}^n$ and $\widetilde{v} \in \mathbb{S}_n$ as

$$(27) \qquad \breve{v} := \frac{Yv^*}{\mu_1},$$

$$(28) \qquad \widetilde{v} := \frac{Yv^*}{\|Yv^*\|}.$$

Then $\widetilde{v}$ is the first-order approximation of v, written equivalently as $\widetilde{v} = \breve{v}/\|\breve{v}\|$. Note that with $\|Yv^*\| > 0$ as shown in (26), $\widetilde{v}$ is well defined.

Since v is the eigenvector of Y corresponding to μ_1, we have

$$\mu_1 v = Yv,$$

$$\mu_1 \widetilde{v} = Yv^*/\|\breve{v}\|.$$

Subtracting the second equation from the first one, we have

$$\mu_1(v - \widetilde{v}) = Y\left(v - \frac{v^*}{\|\breve{v}\|}\right) = Y(v - \widetilde{v}) + Y\left(\widetilde{v} - \frac{v^*}{\|\breve{v}\|}\right) = Y(v - \widetilde{v}) + \frac{1}{\|\breve{v}\|}Y(\breve{v} - v^*).$$

After rearranging, we have

$$(29) \qquad \|\breve{v}\|(\mu_1 I_n - Y)(v - \widetilde{v}) = Y(\breve{v} - v^*).$$

Since $(\mu_1 I_n - Y)v = 0$, we have $\text{span}(\mu_1 I_n - Y)$ being orthogonal to v. As a result, $\|\breve{v}\|(\mu_1 I_n - Y)(v - \widetilde{v}) = \|\breve{v}\|(\mu_1 I_n - Y)(I_n - vv^H)\widetilde{v}$. In addition, since the left-hand side of (29) belongs to $\text{span}(I_n - vv^H)$, its right-hand side must also belong to $\text{span}(I_n - vv^H)$. That is, $Y(\breve{v} - v^*) = (I_n - vv^H)Y(\breve{v} - v^*)$. Then (29) leads to

$$(30) \qquad \|\breve{v}\|(\mu_1 I_n - Y)(I_n - vv^H)\widetilde{v} = (I_n - vv^H)Y(\breve{v} - v^*).$$

Observe that $0 \leq \mu_1 - \mu_2 \leq \cdots \leq \mu_1 - \mu_n$ are the eigenvalues of $\mu_1 I_n - Y$. In particular, the eigenvector corresponding to 0 is v. Since $(I_n - vv^H)\widetilde{v}$ is orthogonal to v, from (30) we have

$$\|\breve{v}\|(\mu_1 - \mu_2)\|(I_n - vv^H)\widetilde{v}\| \leq \|\breve{v}\|\|(\mu_1 I_n - Y)(I_n - vv^H)\widetilde{v}\| = \|(I_n - vv^H)Y(\breve{v} - v^*)\|.$$

Hence,

$$(31) \qquad \|(I_n - vv^H)\widetilde{v}\| \leq \frac{1}{\|\breve{v}\|(\mu_1 - \mu_2)}\|(I_n - vv^H)Y(\breve{v} - v^*)\|.$$

From Lemma 5.2, we have $\inf_{b \in \mathbb{C}_1} \|v - \widetilde{v}b\| \leq \sqrt{2}\|(I_n - vv^H)\widetilde{v}\|$. With this, (31) leads to

$$(32) \qquad \inf_{b \in \mathbb{C}_1} \|v - \widetilde{v}b\| \leq \frac{\sqrt{2}}{\|\breve{v}\|(\mu_1 - \mu_2)}\|(I_n - vv^H)Y(\breve{v} - v^*)\|.$$

In the following, we are going to analyze $(I_n - vv^H)Y(\breve{v} - v^*)$. We have

$$(I_n - vv^H)Y(\breve{v} - v^*)$$

$$= (I_n - vv^H)Y\left(\frac{Yv^*}{\mu_1} - v^*\right)$$

$$= (I_n - vv^H)Y\left(\frac{\mu_1^*}{\mu_1} - 1\right)v^* + \frac{1}{\mu_1}(I_n - vv^H)Y(Y - Y^*)v^*$$

$$= \left(\frac{\mu_1^*}{\mu_1} - 1\right)(I_n - vv^H)\mu_1^*v^* + \left(\frac{\mu_1^*}{\mu_1} - 1\right)(I_n - vv^H)(Y - Y^*)v^*$$

$$\quad + \frac{1}{\mu_1}(I_n - vv^H)\mu_1^*v^*v^{*H}(Y - Y^*)v^* + \frac{1}{\mu_1}(I_n - vv^H)(Y^* - \mu_1^*v^*v^{*H})(Y - Y^*)v^*$$

$$\quad + \frac{1}{\mu_1}(I_n - vv^H)(Y - Y^*)(Y - Y^*)v^*$$

$$= \left(\left(\frac{\mu_1^*}{\mu_1} - 1\right) + \frac{1}{\mu_1}v^{*H}(Y - Y^*)v^*\right)\mu_1^*(I_n - vv^H)v^*$$

$$\quad + \left(\frac{\mu_1^*}{\mu_1} - 1\right)(I_n - vv^H)(Y - Y^*)v^* + \frac{1}{\mu_1}(I_n - vv^H)(Y^* - \mu_1^*v^*v^{*H})(Y - Y^*)v^*$$

$$\quad + \frac{1}{\mu_1}(I_n - vv^H)(Y - Y^*)(Y - Y^*)v^*.$$

Hence,

$$\|(I_n - vv^H)Y(\breve{v} - v^*)\|$$

$$\leq \left(\left|\frac{\mu_1^*}{\mu_1} - 1\right| + \frac{|v^{*H}(Y - Y^*)v^*|}{\mu_1}\right)\mu_1^*\|(I_n - vv^H)v^*\| + \left|\frac{\mu_1^*}{\mu_1} - 1\right|\|Y - Y^*\|$$

$$\quad + \frac{\|Y^* - \mu_1^*v^*v^{*H}\|\|Y - Y^*\|}{\mu_1} + \frac{\|Y - Y^*\|^2}{\mu_1}$$

$$\leq \left(\left| \frac{\mu_1^*}{\mu_1} - 1 \right| + \frac{\|Y - Y^*\|}{\mu_1} \right) \mu_1^* \|(I_n - vv^{\mathrm{H}})v^*\| + \left| \frac{\mu_1^*}{\mu_1} - 1 \right| \|Y - Y^*\|$$
$$+ \frac{\max\{|\mu_2^*|, |\mu_n^*|\}\|Y - Y^*\|}{\mu_1} + \frac{\|Y - Y^*\|^2}{\mu_1},$$

where we use the fact that $\|I_n - vv^{\mathrm{H}}\| = 1$ and $\|Y^* - \mu_1^* v^* v^{*\mathrm{H}}\| = \max\{|\mu_2^*|, |\mu_n^*|\}$. Then together with (32), we have

$$\inf_{b \in \mathbb{C}_1} \|v - \tilde{v}b\| \leq \frac{\sqrt{2}}{\|\tilde{v}\|(\mu_1 - \mu_2)} \left(\left(\left| \frac{\mu_1^*}{\mu_1} - 1 \right| + \frac{\|Y - Y^*\|}{\mu_1} \right) \mu_1^* \|(I_n - vv^{\mathrm{H}})v^*\| \right.$$
$$\left. + \left| \frac{\mu_1^*}{\mu_1} - 1 \right| \|Y - Y^*\| + \frac{\max\{|\mu_2^*|, |\mu_n^*|\}\|Y - Y^*\|}{\mu_1} + \frac{\|Y - Y^*\|^2}{\mu_1} \right).$$

In the rest of the proof, we are going to simplify the display above. From (23) and (26), we have

$$\|\tilde{v}\| = \frac{\|Yv^*\|}{\mu_1} \geq \frac{3}{5}.$$

Using Lemma 5.1 and the assumption $\|Y - Y^*\| \leq (\mu_1^* - \mu_2^*)/4$, we have

$$\|(I_n - vv^{\mathrm{H}})v^*\| \leq \frac{2\|Y - Y^*\|}{\mu_1^* - \mu_2^*}.$$

With the above results, together with (23) and (24), we have

$$\inf_{b \in \mathbb{C}_1} \|v - \tilde{v}b\|$$
$$\leq \frac{\sqrt{2}}{\frac{3}{5}\frac{\mu_1^* - \mu_2^*}{2}} \left(\left(\frac{4\|Y - Y^*\|}{3\mu_1^*} + \frac{\|Y - Y^*\|}{\frac{3}{4}\mu_1^*} \right) \mu_1^* \frac{2\|Y - Y^*\|}{\mu_1^* - \mu_2^*} + \frac{4\|Y - Y^*\|}{3\mu_1^*} \|Y - Y^*\| \right.$$
$$\left. + \frac{\max\{|\mu_2^*|, |\mu_n^*|\}\|Y - Y^*\|}{\frac{3}{4}\mu_1^*} + \frac{\|Y - Y^*\|^2}{\frac{3}{4}\mu_1^*} \right)$$
$$= \frac{10\sqrt{2}}{3(\mu_1^* - \mu_2^*)} \left(\left(\frac{16}{3(\mu_1^* - \mu_2^*)} + \frac{8}{3\mu_1^*} \right) \|Y - Y^*\|^2 + \frac{4\max\{|\mu_2^*|, |\mu_n^*|\}}{3\mu_1^*} \|Y - Y^*\| \right)$$
$$\leq \frac{40\sqrt{2}}{9(\mu_1^* - \mu_2^*)} \left(\left(\frac{4}{(\mu_1^* - \mu_2^*)} + \frac{2}{\mu_1^*} \right) \|Y - Y^*\|^2 + \frac{\max\{|\mu_2^*|, |\mu_n^*|\}}{\mu_1^*} \|Y - Y^*\| \right). \qquad \square$$

5.2. *Proofs of Lemma* 2.1, *Proposition* 2.1 *and Proposition* 2.2.

PROOF OF LEMMA 2.1.   Denote λ' as an eigenvalue of A with its corresponding eigenvector u'. Then we have $Au' = \lambda' u'$. This can be equivalently written as

$$\sum_{k \neq j} A_{jk} u_k' = \lambda' u_j' \quad \forall j \in [n].$$

Multiplying by z_j^* on both sides, we have

$$\sum_{k \neq j} A_{jk} z_j^* u_k' = \sum_{k \neq j} A_{jk} z_j^* \overline{z_k^*} (z_k^* u_k') = \lambda' z_j^* u_j' \quad \forall j \in [n].$$

That is, $(A \circ z^* z^{*\mathrm{H}})(z^* \circ u') = \lambda'(z^* \circ u')$. Hence, λ' is an eigenvalue of $A \circ z^* z^{*\mathrm{H}}$ with the corresponding eigenvector $z^* \circ u'$.

By the same argument, we can show each eigenvalue of $A \circ z^* z^{*H}$ is also an eigenvalue of A. As a result, since $\breve{u}$ is the leading eigenvector of A, $z^* \circ \breve{u}$ is the leading eigenvector of $A \circ z^* z^{*H}$. □

Before proving Proposition 2.1 and Proposition 2.2, we first state some technical lemmas related to A and W.

LEMMA 5.3.   *The largest eigenvalue of $\mathbb{E}A$ is $(n-1)p$ and the corresponding eigenvector is $\mathbb{1}_n / \sqrt{n}$. The remaining eigenvalues of $\mathbb{E}A$ are $-p$ with multiplicity $n-1$. Denote $\lambda' \geq \lambda'_2 \geq \cdots \geq \lambda'_n$ as the eigenvalues of A. We have*

$$(33) \quad \left| \lambda' - (n-1)p \right|, \max_{2 \leq j \leq n} \left| \lambda'_j + p \right| \leq \| A - \mathbb{E}A \|, \text{ and } \lambda' - \lambda'_2 \geq np - 2\| A - \mathbb{E}A \|.$$

LEMMA 5.4.   *There exist constants $C_1, C_2 > 0$ such that if $\frac{np}{\log n} > C_1$, then we have*

$$\| A - \mathbb{E}A \| \leq C_2 \sqrt{np},$$
$$\| A \circ W \| \leq C_2 \sqrt{np},$$
$$\sum_{j \in [n]} \left| \text{Im} \left(\sum_{k \neq j} A_{jk} W_{jk} \overline{z^*_j} z^*_k \right) \right|^2 \leq \frac{n^2 p}{2} \left(1 + C_2 \sqrt{\frac{\log n}{n}} \right),$$

with probability at least $1 - 3n^{-10}$.

The first part of Proposition 2.1 (i.e., (16)) can be proved using Theorem 2.1 of [1] which we include below for completeness. The statement of Theorem 2.1 in [1] is complicated as the theorem works for perturbation of eigenspaces. However, what we need to consider here is only the perturbation of the leading eigenvector. For easier reference, we present below a simpler version of the theorem.

LEMMA 5.5 (A simpler version of Theorem 2.1 of [1]).   *Consider two symmetric matrices $Y, Y^* \in \mathbb{R}^{n \times n}$. Let the eigenvalues of Y^* be $\mu^*_1 \geq \mu^*_2 \geq \cdots \geq \mu^*_n$. Define $\Delta^* := \min\{\mu^*_1 - \mu^*_2, \mu^*_1\}$ and $\kappa := \max\{|\mu^*_1|, |\mu^*_n|\}/\Delta^*$. Let the leading eigenvector of Y (resp. Y^*) be v (resp. v^*). Assume the following conditions are satisfied for some $\gamma \geq 0$ and some function $\phi : [0, +\infty) \to [0, +\infty)$:*

1. *$\| Y^* \|_{2 \to \infty} \leq \gamma \Delta^*$.*
2. *For any $m \in [n]$, $\{Y_{jk} : j = m \text{ or } k = m\}$ are independent of $\{Y_{jk} : j \neq m, k \neq m\}$.*
3. *$32\kappa \max\{\gamma, \phi(\gamma)\} \leq 1$ and for some $\delta_0 \in (0, 1)$, $\mathbb{P}(\| Y - Y^* \| \leq \gamma \Delta^*) \geq 1 - \delta_0$.*
4. *Suppose $\phi(x)$ is continuous and nondecreasing in $[0, +\infty)$ with $\phi(0) = 0$, $\phi(x)/x$ is nonincreasing in $[0, +\infty)$ and $\delta_1 \in (0, 1)$. For any $m \in [n]$ and $w \in \mathbb{R}^n$,*

$$\mathbb{P} \left(\left| [Y - Y^*]_{m.} w \right| \leq \Delta^* \| w \|_\infty \phi \left(\frac{\| w \|}{\sqrt{n} \| w \|_\infty} \right) \right) \geq 1 - \frac{\delta_1}{n}.$$

Then with probability at least $1 - \delta_0 - 2\delta_1$, there exists some constant $C > 0$ and some $b \in \{-1, 1\}$ such that

$$\| vb - Yv^*/\mu^*_1 \|_\infty \leq C \left(\kappa(\kappa + \phi(1))(\gamma + \phi(\gamma)) \| v^* \|_\infty + \gamma \| Y^* \|_{2 \to \infty}/\Delta^* \right).$$

The following Lemma 5.6 provides two Bernstein-type concentration inequalities to be used in the proof of Proposition 2.1. The first one is the classical Bernstein inequality; see Section 2.8 of [6] for its proof. The second one is proved in Lemma 7 of [1].

LEMMA 5.6. *Let $B_1, \ldots, B_n$ be real independent random variables such that* $\max_{j \in [n]} |B_j| \leq M$ *for some $M > 0$. Then*

$$\mathbb{P}\left(\left| \sum_{j \in [n]} (B_j - \mathbb{E}B_j) \right| \geq t \right) \leq 2 \exp\left(-\frac{\frac{1}{2}t^2}{\sum_{j \in [n]} \mathbb{E}(B_j - \mathbb{E}B_j)^2 + \frac{1}{3}Mt} \right).$$

Let $w \in \mathbb{R}^n$ be a fixed vector and $\alpha \geq 0$. If $\{B_j\}_{j \in [n]} \overset{i.i.d.}{\sim}$ Bernoulli(p), we have

$$\mathbb{P}\left(\left| \sum_{j \in [n]} w_j(B_j - p) \right| \geq \frac{(2+\alpha)np}{1 \vee \log(\frac{\sqrt{n}\|w\|_\infty}{\|w\|})} \|w\|_\infty \right) \leq 2 \exp(-\alpha np).$$

PROOF OF PROPOSITION 2.1. We use Lemma 5.5 to prove the first part of the proposition. Denote $\mu_1^* \geq \mu_2^* \geq \cdots \geq \mu_n^*$ as the eigenvalues of $\mathbb{E}A$. Define Δ^* and κ the same as in Lemma 5.5. From Lemma 5.3, we have $\Delta^* = (n-1)p$, $\kappa = 1$, with $\mathbb{1}_n/\sqrt{n}$ being the leading eigenvector of $\mathbb{E}A$. Since $\mathbb{E}A = pJ_n - pI_n$, we have $\|\mathbb{E}A\|_{2\to\infty} = \sqrt{(n-1)}p$. By Lemma 5.4, there exist constants $c_1, c_2 > 1$ such that if $\frac{np}{\log n} > c_1$, then $\|A - \mathbb{E}A\| \leq c_2\sqrt{np}$ with probability at least $1 - 3n^{-10}$. Define $\gamma := 2c_2/\sqrt{np}$, $\delta_0 := 2n^{-10}$ and $\phi(x) := 3(1 \vee \log(x^{-1}))^{-1}$. Then the first assumption of Lemma 5.5 is satisfied as long as $c_2 \geq 1$. When $\frac{np}{\log n}$ is greater than some sufficiently large constant, we have $\phi(\gamma) \leq 8/\log(np)$ and the third assumption is satisfied. We can also verify that the second assumption is also satisfied. For any $m \in [n]$, denote $[A - \mathbb{E}A]_m$. as the jth row of $A - \mathbb{E}A$. For any $w \in \mathbb{R}^n$, since $[A - \mathbb{E}A]_m \cdot w$ is a weighted average of centered Bernoulli random variables, the second inequality of Lemma 5.6 can be applied to have

$$\mathbb{P}\left(\left| [A - \mathbb{E}A]_m \cdot w \right| > \Delta^* \|w\|_\infty \phi\left(\frac{\|w\|}{\sqrt{n}\|w\|_\infty} \right) \right)$$

$$\leq \mathbb{P}\left(\left| [A - \mathbb{E}A]_m \cdot w \right| \geq \frac{2.5np}{1 \vee \log(\frac{\sqrt{n}\|w\|_\infty}{\|w\|})} \|w\|_\infty \right) \leq 2n^{-11},$$

when $\frac{np}{\log n} \geq 11$ is greater than some sufficiently large constant. Define $\delta_1 := 2n^{-10}$. Then the last assumption of Lemma 5.5 is satisfied. Then Lemma 5.5 leads to the conclusion that with probability at least $1 - 6n^{-10}$, there exists some constant $c_1 > 0$ and some $b \in \{-1, 1\}$ such that

$$\left\| \breve{u}b - \frac{1}{\mu_1^*\sqrt{n}} A\mathbb{1}_n \right\|_\infty \leq c_1\left(\kappa(\kappa + \phi(1))(\gamma + \phi(\gamma)) \left\| \frac{1}{\sqrt{n}}\mathbb{1}_n \right\|_\infty + \gamma \frac{\|\mathbb{E}A\|_{2\to\infty}}{\Delta^*} \right)$$

$$\leq c_1\left((1+3)\left(\frac{2c_2}{\sqrt{np}} + \frac{8}{\log(np)} \right) \frac{1}{\sqrt{n}} + \frac{2c_2}{\sqrt{np}} \frac{\sqrt{(n-1)}p}{(n-1)p} \right)$$

$$\leq \frac{c_2}{\log(np)} \frac{1}{\sqrt{n}},$$

for some constant $c_2 > 0$. Note that

$$\frac{1}{\mu_1^*\sqrt{n}} A\mathbb{1}_n = \frac{1}{\mu_1^*\sqrt{n}}\mathbb{E}A\mathbb{1}_n + \frac{1}{\mu_1^*\sqrt{n}}(A - \mathbb{E}A)\mathbb{1}_n = \frac{1}{\sqrt{n}}\mathbb{1}_n + \frac{1}{(n-1)p\sqrt{n}}(A - \mathbb{E}A)\mathbb{1}_n.$$

Then we have

$$\left\| \breve{u}b - \frac{1}{\sqrt{n}}\mathbb{1}_n \right\|_\infty \leq \frac{c_2}{\log(np)} \frac{1}{\sqrt{n}} + \frac{1}{(n-1)p} \left\| \frac{1}{\sqrt{n}}(A - \mathbb{E}A)\mathbb{1}_n \right\|_\infty.$$

For any $m \in [n]$, by the first inequality of Lemma 5.6, there exists some constant $c_3 > 0$ such that

$$\mathbb{P}\big(|[A - \mathbb{E}A]_{m \cdot} \mathbb{1}_n| \geq c_3 \sqrt{np \log n}\big) \leq 2 \exp\bigg(-\frac{\frac{c_3^2}{2} np \log n}{(n-1)p(1-p) + \frac{c_3}{3}\sqrt{np \log n}}\bigg)$$

$$\leq 2n^{-11}.$$

Together with a union bound, we have $\mathbb{P}(\|(A - \mathbb{E}A)\mathbb{1}_n\| \geq c_3\sqrt{np \log n}) \leq 2n^{-10}$. Hence,

$$\bigg\|\breve{u}b - \frac{1}{\sqrt{n}}\mathbb{1}_n\bigg\|_\infty \leq \frac{c_2}{\log(np)}\frac{1}{\sqrt{n}} + \frac{1}{(n-1)p}\frac{c_3\sqrt{np \log n}}{\sqrt{n}} \leq c_4\bigg(\sqrt{\frac{\log n}{np}} + \frac{1}{\log(np)}\bigg)\frac{1}{\sqrt{n}},$$

for some constant $c_4 > 0$ with probability at least $1 - 8n^{-10}$.

The second part of the proposition is an immediate consequence of the first part. If $\frac{np}{\log n} > \max\{C_1, 2C_2^2\}$, all the coordinates of $\breve{u}$ have the same sign according to (16). From Lemma 2.1, we have $u = u^*$ as u^* is the leading eigenvector of $A \circ z^* z^{*\mathrm{H}}$. If $\{\breve{u}_j\}_{j \in [n]}$ are all positive, we have

$$\widehat{z}_j = u_j^*/|u_j^*| = z_j^* \breve{u}_j/\breve{u}_j = z_j^*,$$

for each $j \in [n]$. That is, $\widehat{z} = z^*$. If $\{\breve{u}_j\}_{j \in [n]}$ are all negative, we then have $\widehat{z} = -z^*$. $\square$

PROOF OF PROPOSITION 2.2.   Recall λ^* is the largest eigenvalue of $A \circ z^* z^{*\mathrm{H}}$. From Lemma 2.1, u^* is the corresponding eigenvector. Denote $\lambda_2^* \geq \cdots \geq \lambda_n^*$ as its remaining eigenvalues. By Lemma 5.4, there exist constants $c_1, c_2 > 0$ such that when $\frac{np}{\log n} > c_1$, we have $\|A - \mathbb{E}A\| \leq c_2\sqrt{np}$ and $\|A \circ W\| \leq c_2\sqrt{np}$ with probability at least $1 - 3n^{-10}$. By Lemma 2.1 and Lemma 5.3, we have $\lambda^* \geq (n-1)p - c_2\sqrt{np}$, $\max\{|\lambda_2^*|, |\lambda_n^*|\} \leq p + c_2\sqrt{np}$, and $\lambda^* - \lambda_2^* \geq np - 2c_2\sqrt{np}$. When $\frac{np}{\log n}$ and $\frac{np}{\sigma^2}$ are greater than some sufficiently large constant, we have $4\sigma\|A \circ W\| \leq np/2 \leq \min\{\lambda_1^*, \lambda^* - \lambda_2^*\}$ satisfied. Since $X - A \circ z^* z^{*\mathrm{H}} = \sigma A \circ W$, a direct application of Lemma 2.2 leads to

$$\inf_{b \in \mathbb{C}_1} \|u - \widetilde{u}b\|$$

$$\leq \frac{40\sqrt{2}}{9(\lambda^* - \lambda_2^*)}\bigg(\bigg(\frac{4}{\lambda^* - \lambda_2^*} + \frac{2}{\lambda^*}\bigg)\sigma^2\|A \circ W\|^2 + \frac{\max\{|\lambda_2^*|, |\lambda_n^*|\}\sigma\|A \circ W\|}{\lambda^*}\bigg)$$

$$\leq \frac{40\sqrt{2}}{9np/2}\bigg(\bigg(\frac{4}{np/2} + \frac{2}{np/2}\bigg)c_2^2\sigma^2 np + \frac{(p + c_2\sqrt{np})c_2\sigma\sqrt{np}}{np/2}\bigg)$$

$$\leq c_3\frac{\sigma^2 + \sigma}{np},$$

for some constant $c_3 > 0$. $\square$

5.3. *Proof of Theorem* 2.1.   We first state some technical lemmas that will be used in the proof of Theorem 2.1.

LEMMA 5.7.   *There exists some constant $C_1 > 0$ such that for any γ satisfying $\frac{\gamma^2 np}{\sigma^2} \geq C_1$, we have*

$$\sum_{j \in [n]} \mathbb{I}\bigg\{\frac{2\sigma}{np}\bigg|\sum_{k \neq j} A_{jk}W_{jk}\overline{z_j^*}z_k^*\bigg| \geq \gamma\bigg\} \leq \frac{4\sigma^2}{\gamma^2 p}\exp\bigg(-\frac{1}{16}\sqrt{\frac{\gamma^2 np}{\sigma^2}}\bigg),$$

holds with probability at least $1 - \exp(-\frac{1}{32}\sqrt{\frac{\gamma^2 np}{\sigma^2}})$.

LEMMA 5.8 (Lemma 10 and Lemma 11 of [17]).   *For any $x \in \mathbb{C}$ such that $\mathrm{Re}(x) > 0$,*
$|\frac{x}{|x|} - 1| \le |\frac{\mathrm{Im}(x)}{\mathrm{Re}(x)}|$. *For any $x \in \mathbb{C} \setminus \{0\}$ and any $y \in \mathbb{C}_1$, we have $|\frac{x}{|x|} - y| \le 2|x - y|$.*

PROOF OF THEOREM 2.1.   Let $b_1 \in \mathbb{C}_1$ satisfy $\|u - \widetilde{u}b_1\| = \inf_{a \in \mathbb{C}_1} \|u - \widetilde{u}a\|$. Denote
$\delta := u - \widetilde{u}b_1 \in \mathbb{C}^n$. Recall $\breve{u}$ is the leading eigenvector of A. From Proposition 2.1, Proposition 2.2 and Lemma 5.4, there exist constants $c_1, c_2 > 0$ such that if $\frac{np}{\log n}, \frac{np}{\sigma^2} > c_1$, we have

$$\|\delta\| \le c_2 \frac{\sigma^2 + \sigma}{np}, \tag{34}$$

$$\max_{j \in [n]} \left| \breve{u}_j - \frac{1}{\sqrt{n}}b_2 \right| \le c_2 \left(\sqrt{\frac{\log n}{np}} + \frac{1}{\log(np)} \right) \frac{1}{\sqrt{n}}, \tag{35}$$

$$\|A - \mathbb{E}A\| \le c_2\sqrt{np}, \tag{36}$$

$$\|A \circ W\| \le c_2\sqrt{np}, \tag{37}$$

$$\sum_{j \in [n]} \left| \mathrm{Im}\left(\sum_{k \ne j} A_{jk}W_{jk}\overline{z_j^*}z_k^* \right) \right|^2 \le \frac{n^2 p}{2}\left(1 + c_2\sqrt{\frac{\log n}{n}} \right), \tag{38}$$

with probability at least $1 - n^{-9}$, for some $b_2 \in \{-1, 1\}$.

From (35), when $\frac{np}{\log n} \ge 2c_2^2$, $\breve{u}$ is closer to $\mathbb{1}_n/\sqrt{n}b_2$ than to $-\mathbb{1}_n/\sqrt{n}b_2$ with respect to
ℓ_2 norm. From Lemma 5.3, $\mathbb{1}_n/\sqrt{n}$ is the leading eigenvector of $\mathbb{E}A$. By Lemma 5.1 and
Lemma 5.2, we have

$$\|\breve{u} - \mathbb{1}_n/\sqrt{n}b_2\| \le \sqrt{2}\|(I - \mathbb{1}_n\mathbb{1}_n^{\mathrm{T}}/n)\breve{u}\| \le \frac{2\|A - \mathbb{E}A\|}{np} \le \frac{2c_2}{\sqrt{np}}.$$

Recall that u^* is defined as $z^* \circ \breve{u}$ in (8). Define $\delta^* := u^* - \frac{1}{\sqrt{n}}z^*b_2$. This yields

$$\|\delta^*\| = \left\| z^* \circ \breve{u} - \frac{1}{\sqrt{n}}z^* \circ \mathbb{1}_n b_2 \right\| = \left\| z^* \circ \left(\breve{u} - \frac{1}{\sqrt{n}}\mathbb{1}_n b_2 \right) \right\|$$

$$= \left\| \breve{u} - \frac{1}{\sqrt{n}}\mathbb{1}_n b_2 \right\| \le \frac{2c_2\sqrt{np} + 2p}{np}. \tag{39}$$

By the definition of $\widetilde{u}$ in (9), we can decompose u into

$$u = \widetilde{u}b_1 + \delta = \frac{Xu^*}{\|Xu^*\|}b_1 + \delta = \frac{b_1}{\|Xu^*\|}\left((A \circ z^*z^{*\mathrm{H}})u^* + \sigma(A \circ W)u^*\right) + \delta$$

$$= \frac{b_1}{\|Xu^*\|}(\lambda^* u^* + \sigma(A \circ W)u^*) + \delta, \tag{40}$$

where we use the fact that u^* is the eigenvector of $A \circ z^*z^{*\mathrm{H}}$ corresponding to the eigenvalue
λ^* by Lemma 2.1. With the definition of u^* and also its approximation $\frac{1}{\sqrt{n}}z^*b_2$, (40) leads to

$$u = \frac{b_1}{\|Xu^*\|}\left(\lambda^*(z^* \circ \breve{u}) + \sigma(A \circ W)\left(\frac{1}{\sqrt{n}}z^*b_2 + \delta^* \right) \right) + \delta.$$

For any $j \in [n]$, denote $[A \circ W]_{j\cdot}$ as its jth row. From the display above, we can express u_j
as

$$u_j = \frac{b_1}{\|Xu^*\|}\left(\lambda^* z_j^* \breve{u}_j + \frac{\sigma}{\sqrt{n}} \sum_{k \ne j} A_{jk}W_{jk}z_k^*b_2 + \sigma[A \circ W]_{j\cdot}\delta^* \right) + \delta_j.$$

By (4), when $u_j \neq 0$, we have

$$|\widehat{z}_j - z_j^* b_1 b_2| = |b_2 \overline{b_1 z_j^*} \widehat{z}_j - 1| = \left| b_2 \overline{b_1 z_j^*} \frac{u_j}{|u_j|} - 1 \right| = \left| \frac{b_2 \overline{b_1 z_j^*} u_j}{|b_2 \overline{b_1 z_j^*} u_j|} - 1 \right|$$

(41)

$$= \left| \frac{\frac{\|Xu^*\|}{\lambda^*} b_2 \overline{b_1 z_j^*} u_j}{|\frac{\|Xu^*\|}{\lambda^*} b_2 \overline{b_1 z_j^*} u_j|} - 1 \right|$$

which is all about $\frac{\|Xu^*\|}{\lambda^*} b_2 \overline{b_1 z_j^*} u_j$. With

$$\xi_j := \sum_{k \neq j} A_{jk} W_{jk} \overline{z_j^*} z_k^*,$$

we have

(42) $$\frac{\|Xu^*\|}{\lambda^*} b_2 \overline{b_1 z_j^*} u_j = b_2 \breve{u}_j + \frac{\sigma}{\lambda^* \sqrt{n}} \xi_j + \frac{\sigma [A \circ W]_{j.} \delta^* b_2 \overline{z_j^*}}{\lambda^*} + \frac{\|Xu^*\|}{\lambda^*} \delta_j b_2 \overline{b_1 z_j^*}.$$

Note that from (35), we have

(43) $$b_2 \breve{u}_j \geq \left(1 - c_2 \left(\sqrt{\frac{\log n}{np}} + \frac{1}{\log(np)} \right) \right) \frac{1}{\sqrt{n}}.$$

Let $0 < \gamma, \rho < 1/8$ whose values will be given later. Consider the following two cases.
(1) If

(44) $$\left| \frac{\sigma}{\lambda^* \sqrt{n}} \xi_j \right| \leq \frac{\gamma}{\sqrt{n}},$$

(45) $$\left| \frac{\sigma [A \circ W]_{j.} \delta^*}{\lambda^*} \right| \leq \frac{\rho}{\sqrt{n}},$$

(46) $$\left| \frac{\|Xu^*\|}{\lambda^*} \delta_j \right| \leq \frac{\rho}{\sqrt{n}}$$

all hold, then from (42) and (43), we have

$$\text{Re} \left(\frac{\|Xu^*\|}{\lambda^*} b_2 \overline{b_1 z_j^*} u_j \right) \geq \left(1 - c_2 \left(\sqrt{\frac{\log n}{np}} + \frac{1}{\log(np)} \right) - \gamma - 2\rho \right) \frac{1}{\sqrt{n}},$$

which can be further lower bounded by $1/(2\sqrt{n})$ for sufficiently large $\frac{np}{\log n}$. Therefore, $u_j \neq 0$ in this case. Then by Lemma 5.8 and (41) we have

$$|\widehat{z}_j - z_j^* b_1 b_2|$$

$$\leq \frac{|\text{Im}(\frac{\sigma}{\lambda^* \sqrt{n}} \xi_j + \frac{\sigma [A \circ W]_{j.} \delta^* b_2 \overline{z_j^*}}{\lambda^*} + \frac{\|Xu^*\|}{\lambda^*} \delta_j b_2 \overline{b_1 z_j^*})|}{(1 - c_2(\sqrt{\frac{\log n}{np}} + \frac{1}{\log(np)}) - \gamma - 2\rho) \frac{1}{\sqrt{n}}}$$

$$\leq \frac{|\text{Im}(\frac{\sigma}{\lambda^* \sqrt{n}} \xi_j)|}{(1 - c_2(\sqrt{\frac{\log n}{np}} + \frac{1}{\log(np)}) - \gamma - 2\rho) \frac{1}{\sqrt{n}}} + \frac{|\frac{\sigma [A \circ W]_{j.} \delta^*}{\lambda^*}| + |\frac{\|Xu^*\|}{\lambda^*} \delta_j|}{\frac{1}{2\sqrt{n}}}$$

$$= \frac{\frac{\sigma}{\lambda^*} |\text{Im}(\xi_j)|}{(1 - c_2(\sqrt{\frac{\log n}{np}} + \frac{1}{\log(np)}) - \gamma - 2\rho)} + \frac{2\sqrt{n} \sigma}{\lambda^*} |[A \circ W]_{j.} \delta^*| + \frac{2\sqrt{n} \|Xu^*\|}{\lambda^*} |\delta_j|.$$

Note that for any $x, y \in \mathbb{R}$ and any $\eta > 0$, we have $(x + y)^2 = x^2 + 2(\eta^{1/2}x)(\eta^{-1/2}y) + y^2 \leq (1 + \eta)x^2 + (1 + \eta^{-1})y^2$. We have

$$|\widehat{z}_j - z_j^* b_1 b_2|^2 \leq \frac{(1 + \eta)\frac{\sigma^2}{\lambda^{*2}}|\mathrm{Im}(\xi_j)|^2}{(1 - c_2(\sqrt{\frac{\log n}{np}} + \frac{1}{\log(np)}) - \gamma - 2\rho)^2}$$
$$+ (1 + \eta^{-1})\frac{8n\sigma^2}{\lambda^{*2}}|[A \circ W]_{j\cdot}\delta^*|^2 + (1 + \eta^{-1})\frac{8n\|Xu^*\|^2}{\lambda^{*2}}|\delta_j|^2,$$

where the value of $\eta > 0$ will be given later.

(2) If any one of (44)–(46) does not hold, we simply upper bound $|\widehat{z}_j - z_j^* b_1 b_2|$ by 2. Then this case can be written as

$$|\widehat{z}_j - z_j^* b_1 b_2|^2$$
$$\leq 4\left(\mathbb{I}\left\{\left|\frac{\sigma}{\lambda^*\sqrt{n}}\xi_j\right| > \frac{\gamma}{\sqrt{n}}\right\} + \mathbb{I}\left\{\left|\frac{\sigma[A \circ W]_{j\cdot}\delta^*}{\lambda^*}\right| > \frac{\rho}{\sqrt{n}}\right\} + \mathbb{I}\left\{\left|\frac{\|Xu^*\|}{\lambda^*}\delta_j\right| > \frac{\rho}{\sqrt{n}}\right\}\right)$$
$$\leq 4\left(\mathbb{I}\{\sigma|\xi_j| \geq \gamma\lambda^*\} + \frac{\sigma^2 n|[A \circ W]_{j\cdot}\delta^*|^2}{\rho^2\lambda^{*2}} + \frac{n\|Xu^*\|^2|\delta_j|^2}{\rho^2\lambda^{*2}}\right),$$

where in the last inequality we use the fact $\mathbb{I}\{x \geq y\} \leq x^2/y^2$ for any $x, y > 0$.

Combining the above two cases together, we have

$$|\widehat{z}_j - z_j^* b_1 b_2|^2$$
$$\leq \frac{(1 + \eta)\frac{\sigma^2}{\lambda^{*2}}|\mathrm{Im}(\xi_j)|^2}{(1 - c_2(\sqrt{\frac{\log n}{np}} + \frac{1}{\log(np)}) - \gamma - 2\rho)^2}$$
$$+ (1 + \eta^{-1})\frac{8n\sigma^2}{\lambda^{*2}}|[A \circ W]_{j\cdot}\delta^*|^2 + (1 + \eta^{-1})\frac{8n\|Xu^*\|^2}{\lambda^{*2}}|\delta_j|^2$$
$$+ 4\left(\mathbb{I}\{\sigma|\xi_j| \geq \gamma\lambda^*\} + \frac{\sigma^2 n|[A \circ W]_{j\cdot}\delta^*|^2}{\rho^2\lambda^{*2}} + \frac{n\|Xu^*\|^2|\delta_j|^2}{\rho^2\lambda^{*2}}\right)$$
$$\leq \frac{(1 + \eta)\frac{\sigma^2}{\lambda^{*2}}|\mathrm{Im}(\xi_j)|^2}{(1 - c_2(\sqrt{\frac{\log n}{np}} + \frac{1}{\log(np)}) - \gamma - 2\rho)^2} + 4\mathbb{I}\{\sigma|\xi_j| \geq \gamma\lambda^*\}$$
$$+ 8(1 + \eta^{-1} + \rho^{-2})\frac{n\sigma^2}{\lambda^{*2}}|[A \circ W]_{j\cdot}\delta^*|^2 + 8(1 + \eta^{-1} + \rho^{-2})\frac{n\|Xu^*\|^2}{\lambda^{*2}}|\delta_j|^2.$$

The display above holds for each $j \in [n]$. Summing over j, we have

$$n\ell(\widehat{z}, z^*)$$
$$\leq \sum_{j \in [n]}|\widehat{z}_j - z_j^* b_1 b_2|^2$$
$$\leq \frac{(1 + \eta)\frac{\sigma^2}{\lambda^{*2}}}{(1 - c_2(\sqrt{\frac{\log n}{np}} + \frac{1}{\log(np)}) - \gamma - 2\rho)^2}\sum_{j \in [n]}|\mathrm{Im}(\xi_j)|^2 + 4\sum_{j \in [n]}\mathbb{I}\{\sigma|\xi_j| \geq \gamma\lambda^*\}$$
$$+ 8(1 + \eta^{-1} + \rho^{-2})\frac{n\sigma^2}{\lambda^{*2}}\sum_{j \in [n]}|[A \circ W]_{j\cdot}\delta^*|^2 + 8(1 + \eta^{-1} + \rho^{-2})\frac{n\|Xu^*\|^2}{\lambda^{*2}}\sum_{j \in [n]}|\delta_j|^2$$

$$\leq \frac{(1+\eta)\frac{\sigma^2}{\lambda^{*2}}}{(1-c_2(\sqrt{\frac{\log n}{np}}+\frac{1}{\log(np)})-\gamma-2\rho)^2} \sum_{j\in[n]} |\mathrm{Im}(\xi_j)|^2 + 4\sum_{j\in[n]} \mathbb{I}\{\sigma|\xi_j|\geq\gamma\lambda^*\}$$

$$+ 8(1+\eta^{-1}+\rho^{-2})\frac{n\sigma^2}{\lambda^{*2}}\|A\circ W\|^2\|\delta^*\|^2 + 8(1+\eta^{-1}+\rho^{-2})\frac{n\|Xu^*\|^2}{\lambda^{*2}}\|\delta\|^2,$$

where in the last inequality, we use $\sum_{j\in[n]}|[A\circ W]_j.\delta^*|^2 = \|(A\circ W)\delta^*\|^2 \leq \|A\circ W\|^2\|\delta^*\|^2$.

We are going to simplify the display above. From (34), (37), (39) and (38), we have upper bounds for $\|\delta\|$, $\|A\circ W\|$, $\|\delta^*\|$ and $\sum_{j\in[n]}|\mathrm{Im}(\xi_j)|^2$. Using (36), Lemma 2.1 and Lemma 5.3, we have $\lambda^* \geq (n-1)p - c_2\sqrt{np}$ and a crude bound $np/2 \leq \lambda^* \leq 2np$ when $\frac{np}{\log n}$ is greater than some sufficiently large constant. Due to the decomposition $X = A\circ z^*z^{*\mathrm{H}} + \sigma A\circ W$ and that $(A\circ z^*z^{*\mathrm{H}})u^* = \lambda^*u^*$, we have

$$\|Xu^*\| = \|\lambda^*u^* + \sigma A\circ Wu^*\| \leq \lambda^* + \sigma\|A\circ W\| \leq np + c_2\sigma\sqrt{np}.$$

From Lemma 5.7, if γ satisfies $\frac{\gamma^2 np}{\sigma^2} > c_3$ for some constant $c_3 > 0$, we have

$$\sum_{j\in[n]} \mathbb{I}\{\sigma|\xi_j|\geq\gamma\lambda^*\} \leq \sum_{j\in[n]} \mathbb{I}\left\{\frac{2\sigma}{np}|\xi_j|\geq\gamma\right\} \leq \frac{4\sigma^2}{\gamma^2 p}\exp\left(-\frac{1}{16}\sqrt{\frac{\gamma^2 np}{\sigma^2}}\right),$$

holds with probability at least $1 - \exp(-\frac{1}{32}\sqrt{\frac{\gamma^2 np}{\sigma^2}})$. When c_3 is sufficiently large, we have

$$\frac{4\sigma^2}{\gamma^2 np}\exp\left(-\frac{1}{16}\sqrt{\frac{\gamma^2 np}{\sigma^2}}\right) \leq \left(\frac{\sigma^2}{\gamma^2 np}\right)^3,$$

which is due to the fact $4\exp(-\sqrt{x}/16) \leq 1/x^2$ when $x \geq x_0$ for some large $x_0 > 0$.

Combining the above results together, we have

$$\ell(\hat{z}, z^*) \leq \frac{(1+\eta)(\frac{1}{1-c_2\frac{1}{\sqrt{np}}-\frac{1}{n}})^2}{(1-c_2(\sqrt{\frac{\log n}{np}}+\frac{1}{\log(np)})-\gamma-2\rho)^2}\left(1+c_2\sqrt{\frac{\log n}{n}}\right)\frac{\sigma^2}{2np} + \left(\frac{\sigma^2}{\gamma^2 np}\right)^3$$

$$+ 32(1+\eta^{-1}+\rho^{-2})c_2^2\left(\frac{2c_2}{\sqrt{np}}\right)^2\frac{\sigma^2}{np}$$

$$+ 128(1+\eta^{-1}+\rho^{-2})\left(1+\frac{c_2^2\sigma^2}{np}\right)c_2^2\frac{\sigma^4+\sigma^2}{(np)^2}.$$

Note that $\frac{1}{(1-x)^2} \leq 1+16x$, $\forall 0\leq x\leq\frac{1}{2}$. We have $(1-c_2(\sqrt{\frac{\log n}{np}}+\frac{1}{\log(np)})-\gamma-2\rho)^{-2} \leq 1+16(c_2(\sqrt{\frac{\log n}{np}}+\frac{1}{\log(np)})+\gamma+2\rho)$ and $(1-c_2\frac{1}{\sqrt{np}}-\frac{1}{n})^{-2} \leq 1+16(c_2\frac{1}{\sqrt{np}}+\frac{1}{n})$ as long as $\frac{np}{\log n}$ is greater than some sufficiently large constant. After rearrangement, there exists some constant $c_5 > 0$ such that

$$\ell(\hat{z}, z^*) \leq \left(1 + c_5\left(\eta+\gamma+\rho+\sqrt{\frac{\log n}{np}}+\frac{1}{\log(np)}+\gamma^{-6}\left(\frac{\sigma^2}{np}\right)^2\right.\right.$$

$$\left.\left.+ (\eta^{-1}+\rho^{-2})\left(\frac{1+\sigma^2}{np}\right)\right)\right)\frac{\sigma^2}{2np}.$$

We can choose $\gamma^2 = \sqrt{\sigma^2/(np)}$ (then $\frac{\gamma^2 np}{\sigma^2} > c_3$ is guaranteed as long as $\frac{np}{\sigma^2} > c_3^2$). We also set $\rho^2 = \sqrt{(1+\sigma^2)/np}$ and let $\eta = \rho^2$. Then, there exists some constant $c_6 > 0$ such that

$$\ell(\widehat{z}, z^*) \leq \left(1 + c_6\left(\left(\frac{\sigma^2}{np}\right)^{\frac{1}{4}} + \sqrt{\frac{\log n}{np}} + \frac{1}{\log(np)}\right)\right)\frac{\sigma^2}{2np}.$$

This holds with probability at least $1 - n^{-9} - \exp(-\frac{1}{32}(\frac{np}{\sigma^2})^{\frac{1}{4}})$. $\quad\square$

SUPPLEMENTARY MATERIAL

Supplement to "Exact minimax optimality of spectral methods in phase synchronization and orthogonal group synchronization" (DOI: 10.1214/24-AOS2424SUPP; .pdf). In the supplement [38], we first prove the auxiliary lemmas of Section 5 in Appendix A. We then include proofs of results for the orthogonal group synchronization in Appendix B. A calculation to derive (18) is included in Appendix C.

REFERENCES

[1] ABBE, E., FAN, J., WANG, K. and ZHONG, Y. (2020). Entrywise eigenvector analysis of random matrices with low expected rank. *Ann. Statist.* **48** 1452–1474. MR4124330 https://doi.org/10.1214/19-AOS1854

[2] ABBE, E., MASSOULIÉ, L., MONTANARI, A., SLY, A. and SRIVASTAVA, N. (2018). Group synchronization on grids. *Math. Stat. Learn.* **1** 227–256. MR4059722 https://doi.org/10.4171/msl/6

[3] AGTERBERG, J., LUBBERTS, Z. and PRIEBE, C. E. (2022). Entrywise estimation of singular vectors of low-rank matrices with heteroskedasticity and dependence. *IEEE Trans. Inf. Theory* **68** 4618–4650. MR4449064 https://doi.org/10.1109/tit.2022.3159085

[4] ARIE-NACHIMSON, M., KOVALSKY, S. Z., KEMELMACHER-SHLIZERMAN, I., SINGER, A. and BASRI, R. (2012). Global motion estimation from point matches. In 2012 *Second International Conference on 3D Imaging, Modeling, Processing, Visualization & Transmission* 81–88. IEEE, New York.

[5] BANDEIRA, A. S., BOUMAL, N. and SINGER, A. (2017). Tightness of the maximum likelihood semidefinite relaxation for angular synchronization. *Math. Program.* **163** 145–167. MR3632977 https://doi.org/10.1007/s10107-016-1059-6

[6] BOUCHERON, S., LUGOSI, G. and BOUSQUET, O. (2003). Concentration inequalities. In *Summer School on Machine Learning* 208–240. Springer, Berlin.

[7] BOUMAL, N. (2016). Nonconvex phase synchronization. *SIAM J. Optim.* **26** 2355–2377. MR3566919 https://doi.org/10.1137/16M105808X

[8] BOUMAL, N., SINGER, A. and ABSIL, P.-A. (2013). Robust estimation of rotations from relative measurements by maximum likelihood. In 52*nd IEEE Conference on Decision and Control* 1156–1161. IEEE, New York.

[9] CAI, C., LI, G., CHI, Y., POOR, H. V. and CHEN, Y. (2021). Subspace estimation from unbalanced and incomplete data matrices: $\ell_{2,\infty}$ statistical guarantees. *Ann. Statist.* **49** 944–967. MR4255114 https://doi.org/10.1214/20-aos1986

[10] CAPE, J., TANG, M. and PRIEBE, C. E. (2019). The two-to-infinity norm and singular subspace geometry with applications to high-dimensional statistics. *Ann. Statist.* **47** 2405–2439. MR3988761 https://doi.org/10.1214/18-AOS1752

[11] CHEN, Y., CHI, Y., FAN, J., MA, C. et al. (2021). Spectral methods for data science: A statistical perspective. *Found. Trends Mach. Learn.* **14** 566–806.

[12] CUCURINGU, M. (2016). Sync-rank: Robust ranking, constrained ranking and rank aggregation via eigenvector and SDP synchronization. *IEEE Trans. Netw. Sci. Eng.* **3** 58–79. MR3477931 https://doi.org/10.1109/TNSE.2016.2523761

[13] DAVIS, C. and KAHAN, W. M. (1970). The rotation of eigenvectors by a perturbation. III. *SIAM J. Numer. Anal.* **7** 1–46. MR0264450 https://doi.org/10.1137/0707001

[14] FAN, J., WANG, W. and ZHONG, Y. (2017). An ℓ_∞ eigenvector perturbation bound and its application to robust covariance estimation. *J. Mach. Learn. Res.* **18** Paper No. 207, 42 pp. MR3827095

[15] FAN, Y., KHOO, Y. and ZHAO, Z. (2022). Joint community detection and rotational synchronization via semidefinite programming. *SIAM J. Math. Data Sci.* **4** 1052–1081. MR4467534 https://doi.org/10.1137/21M1419702

[16] FILBIR, F., KRAHMER, F. and MELNYK, O. (2021). On recovery guarantees for angular synchronization. *J. Fourier Anal. Appl.* **27** Paper No. 31, 26 pp. MR4240787 https://doi.org/10.1007/s00041-021-09834-1

[17] GAO, C. and ZHANG, A. Y. (2021). Exact minimax estimation for phase synchronization. *IEEE Trans. Inf. Theory* **67** 8236–8247. MR4346085 https://doi.org/10.1109/TIT.2021.3112712

[18] GAO, C. and ZHANG, A. Y. (2022). SDP achieves exact minimax optimality in phase synchronization. *IEEE Trans. Inf. Theory* **68** 5374–5390. MR4476402 https://doi.org/10.1109/tit.2022.3167603

[19] GAO, C. and ZHANG, A. Y. (2023). Optimal orthogonal group synchronization and rotation group synchronization. *Inf. Inference* **12** 591–632. MR4565748 https://doi.org/10.1093/imaiai/iaac022

[20] IWEN, M. A., PRESKITT, B., SAAB, R. and VISWANATHAN, A. (2020). Phase retrieval from local measurements: Improved robustness via eigenvector-based angular synchronization. *Appl. Comput. Harmon. Anal.* **48** 415–444. MR4016998 https://doi.org/10.1016/j.acha.2018.06.004

[21] JAVANMARD, A., MONTANARI, A. and RICCI-TERSENGHI, F. (2016). Phase transitions in semidefinite relaxations. *Proc. Natl. Acad. Sci. USA* **113** E2218–E2223. MR3494080 https://doi.org/10.1073/pnas.1523097113

[22] LEI, L. (2019). Unified $\ell_{2\to\infty}$ eigenspace perturbation theory for symmetric random matrices. Preprint. Available at arXiv:1909.04798.

[23] LELARGE, M. and MIOLANE, L. (2019). Fundamental limits of symmetric low-rank matrix estimation. *Probab. Theory Related Fields* **173** 859–929. MR3936148 https://doi.org/10.1007/s00440-018-0845-x

[24] LERMAN, G. and SHI, Y. (2022). Robust group synchronization via cycle-edge message passing. *Found. Comput. Math.* **22** 1665–1741. MR4509114 https://doi.org/10.1007/s10208-021-09532-w

[25] LING, S. (2022). Near-optimal performance bounds for orthogonal and permutation group synchronization via spectral methods. *Appl. Comput. Harmon. Anal.* **60** 20–52. MR4387245 https://doi.org/10.1016/j.acha.2022.02.003

[26] LING, S. (2022). Improved performance guarantees for orthogonal group synchronization via generalized power method. *SIAM J. Optim.* **32** 1018–1048. MR4425911 https://doi.org/10.1137/20M1389571

[27] LING, S. (2023). Solving orthogonal group synchronization via convex and low-rank optimization: Tightness and landscape analysis. *Math. Program.* **200** 589–628. MR4590243 https://doi.org/10.1007/s10107-022-01896-3

[28] LIU, H., YUE, M.-C. and SO, A. M.-C. (2017). On the estimation performance and convergence rate of the generalized power method for phase synchronization. *SIAM J. Optim.* **27** 2426–2446. MR3735300 https://doi.org/10.1137/16M110109X

[29] PERRY, A., WEIN, A. S., BANDEIRA, A. S. and MOITRA, A. (2016). Optimality and sub-optimality of PCA for spiked random matrices and synchronization. Preprint. Available at arXiv:1609.05573.

[30] PERRY, A., WEIN, A. S., BANDEIRA, A. S. and MOITRA, A. (2018). Message-passing algorithms for synchronization problems over compact groups. *Comm. Pure Appl. Math.* **71** 2275–2322. MR3862091 https://doi.org/10.1002/cpa.21750

[31] PRESKITT, B. P. (2018). Phase retrieval from locally supported measurements. University of California, San Diego. MR3908057

[32] ROMANOV, E. and GAVISH, M. (2020). The noise-sensitivity phase transition in spectral group synchronization over compact groups. *Appl. Comput. Harmon. Anal.* **49** 935–970. MR4135427 https://doi.org/10.1016/j.acha.2019.05.002

[33] SHEN, Y., HUANG, Q., SREBRO, N. and SANGHAVI, S. (2016). Normalized spectral map synchronization. *Adv. Neural Inf. Process. Syst.* **29**.

[34] SHI, Y. and LERMAN, G. (2020). Message passing least squares framework and its application to rotation synchronization. In *International Conference on Machine Learning* 8796–8806. PMLR.

[35] SINGER, A. (2011). Angular synchronization by eigenvectors and semidefinite programming. *Appl. Comput. Harmon. Anal.* **30** 20–36. MR2737931 https://doi.org/10.1016/j.acha.2010.02.001

[36] SINGER, A. and SHKOLNISKY, Y. (2011). Three-dimensional structure determination from common lines in cryo-EM by eigenvectors and semidefinite programming. *SIAM J. Imaging Sci.* **4** 543–572. MR2810897 https://doi.org/10.1137/090767777

[37] WANG, L. and SINGER, A. (2013). Exact and stable recovery of rotations for robust synchronization. *Inf. Inference* **2** 145–193. MR3311446 https://doi.org/10.1093/imaiai/iat005

[38] ZHANG, A. Y. (2024). Supplement to "Exact minimax optimality of spectral methods in phase synchroniza-
     tion and orthogonal group synchronization." https://doi.org/10.1214/24-AOS2424SUPP

[39] ZHONG, Y. and BOUMAL, N. (2018). Near-optimal bounds for phase synchronization. *SIAM J. Optim.* **28**
     989–1016. MR3782406 https://doi.org/10.1137/17M1122025

[40] ZHU, L., WANG, J. and SO, A. M.-C. (2021). Orthogonal group synchronization with incomplete measure-
     ments: Error bounds and linear convergence of the generalized power method. Preprint. Available at
     arXiv:2112.06556.