# SUPPLEMENT TO "EXACT MINIMAX OPTIMALITY OF SPECTRAL METHODS IN PHASE SYNCHRONIZATION AND ORTHOGONAL GROUP SYNCHRONIZATION"

BY Anderson Ye Zhang

University of Pennsylvania

## APPENDIX A: PROOFS OF AUXILIARY LEMMAS OF SECTION 5

PROOF OF LEMMA 5.1. Let $\widetilde{\lambda}_1 \geq \widetilde{\lambda}_2 \geq \ldots \geq \widetilde{\lambda}_d$ be eigenvalues of $\widetilde{X}$. By Weyl's inequality, we have $\|\widetilde{\lambda}_{r+1} - \lambda_{r+1}\| \leq \|X - \widetilde{X}\|$. Under the assumption $\|X - \widetilde{X}\| < (\lambda_r - \lambda_{r+1})/4$, we have

$$\lambda_r - \widetilde{\lambda}_{r+1} = \lambda_r - \lambda_{r+1} + \lambda_{r+1} - \widetilde{\lambda}_{r+1} \geq \lambda_r - \lambda_{r+1} - \left\|X - \widetilde{X}\right\| > \frac{3}{4}(\lambda_r - \lambda_{r+1}) > 0.$$

Define

$$\Theta(U, \widetilde{U}) := \mathrm{diag}(\cos^{-1}\sigma_1, \ldots, \cos^{-1}\sigma_r) \in \mathbb{R}^{r\times r},$$

where $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_r$ are singular values of $U^{\mathrm{H}}\widetilde{U}$. Since $\lambda_r - \widetilde{\lambda}_{r+1} > 0$, by Davis-Kahan Theorem [13], we have

$$\left\|\sin\Theta(U, \widetilde{U})\right\| \leq \frac{\left\|X - \widetilde{X}\right\|}{\lambda_r - \widetilde{\lambda}_{r+1}} \leq \frac{4\left\|X - \widetilde{X}\right\|}{3(\lambda_r - \lambda_{r+1})}.$$

From page 10 of [13], we also have $\|\sin\Theta(U, \widetilde{U})\| = \|(I - UU^{\mathrm{H}})\widetilde{U}\|$. The proof is complete. □

PROOF OF LEMMA 5.2. Since both $x$ and $y$ are unit vectors, we have

$$\|x - yb\|^2 = 2 - x^{\mathrm{H}}yb - (yb)^{\mathrm{H}}x = 2 - 2\mathrm{Re}(x^{\mathrm{H}}yb), \forall b \in \mathbb{C}_1. \tag{47}$$

Therefore, when $x^{\mathrm{H}}y = 0$, we have $\|x - yb\| = \sqrt{2}$ independent of $b$. In this case, we also have $\|(I_n - xx^{\mathrm{H}})y\| = \|y\| = 1$. This proves the statement in the lemma for the $x^{\mathrm{H}}y = 0$ case. When $x^{\mathrm{H}}y \neq 0$, the infimum over $b$ in (47) is achieved when $b = y^{\mathrm{H}}x/|y^{\mathrm{H}}x|$. We then have

$$\begin{aligned}
\inf_{b\in\mathbb{C}_1} \|x - yb\|^2 &= \left\|y - \frac{x^{\mathrm{H}}y}{|x^{\mathrm{H}}y|}x\right\|^2 = \left\|y - xx^{\mathrm{H}}y + xx^{\mathrm{H}}y - \frac{x^{\mathrm{H}}y}{|x^{\mathrm{H}}y|}x\right\|^2 \\
&= \|y - xx^{\mathrm{H}}y\|^2 + \left\|\left(1 - \frac{1}{|x^{\mathrm{H}}y|}\right)(x^{\mathrm{H}}y)x\right\|^2 \\
&= \|y - xx^{\mathrm{H}}y\|^2 + \left|1 - \frac{1}{|x^{\mathrm{H}}y|}\right|^2 |x^{\mathrm{H}}y|^2 \\
&= \|y - xx^{\mathrm{H}}y\|^2 + |1 - |x^{\mathrm{H}}y||^2,
\end{aligned}$$

where we use the orthogonality between $(I_d - xx^{\mathrm{H}})y$ and $x$. With $\|y - xx^{\mathrm{H}}y\|^2 = 1 + \|xx^{\mathrm{H}}y\|^2 - 2y^{\mathrm{H}}xx^{\mathrm{H}}y = 1 - |x^{\mathrm{H}}y|^2 \geq (1 - |x^{\mathrm{H}}y|)^2$, where the last inequality is due to $0 \leq |x^{\mathrm{H}}y| \leq 1$, the proof is complete. □

PROOF OF LEMMA 5.3. Note that $\mathbb{E}A = pJ_n - pI_n$. Note that $(\mathbb{1}_n/\sqrt{n})^{\mathrm{T}}\mathbb{E}A(\mathbb{1}_n/\sqrt{n}) = (n-1)p$ and for any unit vector $u \in \mathbb{R}^n$ that is orthogonal to $\mathbb{1}_n/\sqrt{n}$, we have $u^{\mathrm{T}}\mathbb{E}Au = 0 - p\|u\|^2 = -p$. Hence, $(n-1)p$ is the largest eigenvalue with $\mathbb{1}_n/\sqrt{n}$ being the corresponding eigenvector, and $-p$ is another eigenvalue with multiplicity $n-1$.

By Weyl's inequality, we have $|\lambda' - (n-1)p|, \max_{2\leq j\leq n} |\lambda'_j - (-p)| \leq \|A - \mathbb{E}A\|$, which leads to (33) after rearrangement. This completes the proof, with $\lambda^* = \lambda'$ and $\lambda^*_2 = \lambda'_2$ by Lemma 2.1. □

PROOF OF LEMMA 5.4. The first two inequalities stem from Lemma 5 and Lemma 6 of [17], respectively. The third inequality is derived from Lemma 7 and (29) in [17]. □

PROOF OF LEMMA 5.7. It is proved in (31) of [17]. □

## APPENDIX B: PROOFS FOR ORTHOGONAL GROUP SYNCHRONIZATION

**B.1. Proof of Lemma 3.2.** Before the proof, we first state a technical lemma that is analogous to Lemma 5.2.

LEMMA B.1. *For any two matrices $U, V \in \mathcal{O}(d_1, d_2)$, we have*

$$\|(I_{d_1} - VV^{\mathrm{T}})U\| \leq \inf_{O\in\mathcal{O}(d_2)} \|V - UO\| \leq \sqrt{2}\,\|(I_{d_1} - VV^{\mathrm{T}})U\|.$$

PROOF. Let $V_\perp \in \mathbb{R}^{d_1\times(d_1-d_2)}$ be the complement of $V$ such that $(V, V_\perp) \in \mathcal{O}(d_1)$. From Lemma 2.5 and Lemma 2.6 of [11], we have $\|U^{\mathrm{T}}V_\perp\| \leq \inf_{O\in\mathcal{O}(d_2)} \|V - UO\| \leq \sqrt{2}\|U^{\mathrm{T}}V_\perp\|$. The proof is complete with $\|U^{\mathrm{T}}V_\perp\| = \|V_\perp V_\perp^{\mathrm{T}}U\| = \|(I_{d_1} - VV^{\mathrm{T}})U\|$. □

PROOF OF LEMMA 3.2. We first give an explicit expression for the first-order approximation $\widetilde{V}$. Denote $\mu_1 \geq \ldots \geq \mu_n$ as the eigenvalues of $Y$. Let $YV^* = GDN^{\mathrm{T}}$ be its SVD where $G \in \mathcal{O}(n,d)$, $N \in \mathcal{O}(d)$, and $D \in \mathbb{R}^{d\times d}$ is a diagonal matrix with singular values. Define $M^* = \mathrm{diag}(\mu^*_1, \ldots, \mu^*_d) \in \mathbb{R}^{d\times d}$. Since

$$YV^* = Y^*V^* + (Y - Y^*)V^* = V^*M^* + (Y - Y^*)V^*, \tag{48}$$

we have

$$\max_{i\in[d]} |D_{ii} - \mu^*_i| \leq \|(Y - Y^*)V^*\| \leq \|Y - Y^*\|, \tag{49}$$

by Weyl's inequality. Under the assumption that $\|Y - Y^*\| \leq \min\{\mu^*_d - \mu^*_{d+1}, \mu^*_d\}/4$, we have $\{D_{ii}\}_{i\in[d]}$ all being positive. Note that

$$\begin{aligned}\widetilde{V} &= \operatorname*{argmin}_{V'\in\mathcal{O}(n,d)} \left\|V' - YV^*\right\|_{\mathrm{F}}^2 = \operatorname*{argmax}_{V\in\mathcal{O}(n,d)} \left\langle V', YV^*\right\rangle \\ &= \operatorname*{argmax}_{V'\in\mathcal{O}(n,d)} \mathrm{tr}\left(V'^{\mathrm{T}}GDN^{\mathrm{T}}\right) = \operatorname*{argmax}_{V'\in\mathcal{O}(n,d)} \left\langle G^{\mathrm{T}}V'N, D\right\rangle.\end{aligned}$$

Due to the fact that $G, V' \in \mathcal{O}(n,d)$, $N \in \mathcal{O}(d)$, and the diagonal entries of $D$ are all positive, the maximum is achieved when $G^{\mathrm{T}}V'N = I_d$. This gives $\widetilde{V} = GN^{\mathrm{T}}$ which can also be written as

$$\widetilde{V} = YV^*S, \tag{50}$$

where

$$S := ND^{-1}N^{\mathrm{T}} \in \mathbb{R}^{d\times d} \tag{51}$$

can be seen as a linear operator and plays a similar role as $1/\|Xu^*\|$ for $\widetilde{u} = Xu^*/\|Xu^*\|$ in (9).

Define $M := \mathrm{diag}(\mu_1, \mu_2, \ldots, \mu_d) \in \mathbb{R}^{d\times d}$. Then we have

$$VM = YV,$$
$$\widetilde{V}M = YV^*SM,$$

and consequently,

$$(V - \widetilde{V})M = Y(V - V^*SM) = Y(V - \widetilde{V}) + Y(\widetilde{V} - V^*SM).$$

After rearranging, we have

$$Y\widetilde{V} - \widetilde{V}M = Y(\widetilde{V} - V^*SM).$$

Multiplying $(I - VV^{\mathrm{T}})$ on both sides, we have

$$\begin{aligned} Y(I - VV^{\mathrm{T}})\widetilde{V} - (I - VV^{\mathrm{T}})\widetilde{V}M &= (I - VV^{\mathrm{T}})Y\widetilde{V} - (I - VV^{\mathrm{T}})\widetilde{V}M \\ &= (I - VV^{\mathrm{T}})Y(\widetilde{V} - V^*SM), \end{aligned}$$

where the first equation is due to $Y(I - VV^{\mathrm{T}}) = (I - VV^{\mathrm{T}})Y$ as $V$ is the leading eigenspace of $Y$. Note that for any $x \in \mathrm{span}(I - VV^{\mathrm{T}})$ and for any $i \in [d]$, we have $\|Yx - \mu_i x\| \geq (\mu_i - \mu_{d+1})\|x\|$. Then we have

$$\left\|Y(I - VV^{\mathrm{T}})\widetilde{V} - (I - VV^{\mathrm{T}})\widetilde{V}M\right\| \geq (\mu_d - \mu_{d+1})\left\|(I - VV^{\mathrm{T}})\widetilde{V}\right\|.$$

As a result, we have

$$\left\|(I - VV^{\mathrm{T}})\widetilde{V}\right\| \leq \frac{1}{\mu_d - \mu_{d+1}}\left\|(I - VV^{\mathrm{T}})Y(\widetilde{V} - V^*SM)\right\|, \tag{52}$$

which is analogous to (31) in the proof of Lemma 3.2. By Lemma B.1, we have

(53)
$$\inf_{O\in\mathcal{O}(d)}\left\|V - \widetilde{V}O\right\| \leq \sqrt{2}\left\|(I - VV^{\mathrm{T}})\widetilde{V}\right\| \leq \frac{\sqrt{2}}{\mu_d - \mu_{d+1}}\left\|(I - VV^{\mathrm{T}})Y(\widetilde{V} - V^*SM)\right\|.$$

In the next, we are going to analyze $(I - VV^{\mathrm{T}})Y(\widetilde{V} - V^*SM)$. Using (50), we have

$$\begin{aligned} &(I - VV^{\mathrm{T}})Y(\widetilde{V} - V^*SM) \\ &= (I - VV^{\mathrm{T}})Y\left(YV^*S - V^*SM\right) \\ &= (I - VV^{\mathrm{T}})Y\left(V^*M^*S + (Y - Y^*)V^*S - V^*SM\right) \\ &= (I - VV^{\mathrm{T}})YV^*\left(M^*S - SM\right) + (I - VV^{\mathrm{T}})Y(Y - Y^*)V^*S \\ &= (I - VV^{\mathrm{T}})\left(V^*M^* + (Y - Y^*)V^*\right)\left(M^*S - SM\right) \\ &\quad + (I - VV^{\mathrm{T}})V^*M^*V^{*\mathrm{T}}(Y - Y^*)V^*S \\ &\quad + (I - VV^{\mathrm{T}})(Y^* - V^*M^*V^{*\mathrm{T}})(Y - Y^*)V^*S + (I - VV^{\mathrm{T}})(Y - Y^*)(Y - Y^*)V^*S \\ &= (I - VV^{\mathrm{T}})V^*M^*\left((M^*S - SM) + V^{*\mathrm{T}}(Y - Y^*)V^*S\right) \\ &\quad + (I - VV^{\mathrm{T}})(Y - Y^*)V^*\left(M^*S - SM\right) \\ &\quad + (I - VV^{\mathrm{T}})(Y^* - V^*M^*V^{*\mathrm{T}})(Y - Y^*)V^*S + (I - VV^{\mathrm{T}})(Y - Y^*)(Y - Y^*)V^*S, \end{aligned}$$

where in the second to last equation, we use (48) and the decomposition $Y = V^*M^*V^{*\top} + (Y^* - V^*M^*V^{*\top}) + (Y - Y^*)$. Hence, with $\|Y^* - V^*M^*V^{*\top}\| = \max\{|\mu^*_{d+1}|, |\mu^*_n|\}$, we have

$$\begin{aligned}
&\left\|(I - VV^\top)Y(\widetilde{V} - V^*SM)\right\| \\
&\leq \mu_1^* \left\|(I - VV^\top)V^*\right\| \left(\|M^*S - SM\| + \|Y - Y^*\| \|S\|\right) \\
&\quad + \|Y - Y^*\| \|M^*S - SM\| + \max\{|\mu^*_{d+1}|, |\mu^*_n|\} \|Y - Y^*\| \|S\| + \|Y - Y^*\|^2 \|S\|.
\end{aligned}$$

Then from (53), we have

$$\begin{aligned}
\inf_{O \in \mathcal{O}(d)} \left\|V - \widetilde{V}O\right\| \leq \frac{\sqrt{2}}{\mu_d - \mu_{d+1}} \Bigg( &\mu_1^* \left\|(I - VV^\top)V^*\right\| \left(\|M^*S - SM\| + \|Y - Y^*\| \|S\|\right) \\
&+ \|Y - Y^*\| \|M^*S - SM\| + \max\{|\mu^*_{d+1}|, |\mu^*_n|\} \|Y - Y^*\| \|S\| \\
&+ \|Y - Y^*\|^2 \|S\| \Bigg).
\end{aligned}$$

In the rest of the proof, we are going to simplify the display above. By Weyl's inequality, we have

$$\max_{i \in [n]} |\mu_i - \mu_i^*| \leq \|Y - Y^*\|. \tag{54}$$

Since $\|Y - Y^*\| \leq (\mu^*_d - \mu^*_{d+1})/4$ is assumed, we have

$$\mu_d - \mu_{d+1} \geq \frac{\mu^*_d - \mu^*_{d+1}}{2}.$$

By this assumption and Lemma 5.1, we have

$$\left\|(I - VV^\top)V^*\right\| \leq \frac{2\|Y - Y^*\|}{\mu^*_d - \mu^*_{d+1}}.$$

By (49) and the definition of $S$ in (51), we have

$$\|S\| = \left\|D^{-1}\right\| \leq \frac{1}{\mu^*_d - \|Y - Y^*\|} \leq \frac{4}{3\mu^*_d}.$$

In addition,

$$\begin{aligned}
\|M^*S - SM\| &\leq \|M^*S - SM^*\| + \|S(M - M^*)\| \\
&\leq \|(M^* - \mu^*_d I_d)S + S(\mu^*_d I_d - M^*)\| + \|S\| \|M - M^*\| \\
&\leq \|S\| \left(2\|M^* - \mu^*_d I_d\| + \|M - M^*\|\right) \\
&\leq \frac{4}{3\mu^*_d} \left(2(\mu^*_1 - \mu^*_d) + \|Y - Y^*\|\right),
\end{aligned}$$

where in the last inequality we use the fact $\|M - M^*\| = \max_{i \in [d]} |\mu_i - \mu_i^*|$ and (54). Combining all the results together, we have

$$\begin{aligned}
&\inf_{O \in \mathcal{O}(d)} \left\|V - \widetilde{V}O\right\| \\
&\leq \frac{2\sqrt{2}}{\mu^*_d - \mu^*_{d+1}} \left( \mu^*_1 \frac{2\|Y - Y^*\|}{\mu^*_d - \mu^*_{d+1}} \left( \frac{4\left(2(\mu^*_1 - \mu^*_d) + \|Y - Y^*\|\right)}{3\mu^*_d} + \frac{4\|Y - Y^*\|}{3\mu^*_d} \right) \right)
\end{aligned}$$

$$
\begin{aligned}
&+ \frac{4}{3\mu_d^*}\left(2(\mu_1^* - \mu_d^*) + \|Y - Y^*\|\right)\|Y - Y^*\| + \frac{4\max\{|\mu_{d+1}^*|, |\mu_n^*|\}\,\|Y - Y^*\|}{3\mu_d^*} \\
&+ \left.\frac{4\,\|Y - Y^*\|^2}{3\mu_d^*}\right) \\
\le\ & \frac{16\sqrt{2}}{3\left(\mu_d^* - \mu_{d+1}^*\right)\mu_d^*}\left(\frac{2\mu_1^*}{3(\mu_d^* - \mu_{d+1}^*)} + 1\right)\|Y - Y^*\|^2 \\
&+ \frac{8\sqrt{2}}{3\left(\mu_d^* - \mu_{d+1}^*\right)\mu_d^*}\left(\frac{4\mu_1^*\left(\mu_1^* - \mu_d^*\right)}{\mu_d^* - \mu_{d+1}^*} + 2(\mu_1^* - \mu_d^*) + \max\{|\mu_{d+1}^*|, |\mu_n^*|\}\right)\|Y - Y^*\|.
\end{aligned}
$$

□

### B.2. Proofs of Lemma 3.1, Proposition 3.1, and Proposition 3.2.

PROOF OF LEMMA 3.1. Similar to the proof of Lemma 2.1, we can show each eigenvalue of $A$ is also an eigenvalue of $(A \otimes J_d) \circ Z^* Z^{*\mathrm{T}}$ with multiplicity $d$. At the same time, each eigenvalue of $(A \otimes J_d) \circ Z^* Z^{*\mathrm{T}}$ must be an eigenvalue of $A$. The proof is omitted here. □

PROOF OF PROPOSITION 3.1. Since $\sigma = 0$, we have $U = U^*$. Then $\widehat{Z}_j = \mathcal{P}(U_j) = \mathcal{P}(U_j^*) = \mathcal{P}(Z_j^* \breve{u}_j)$. Since $Z_j^*$ is an orthogonal matrix, we have $\widehat{Z}_j = Z_j^* \mathrm{sign}(\breve{u}_j)$. Then by (16), the proposition is proved by the same argument used to prove Proposition 2.1. □

Before proving Proposition 3.2, we state some properties of $A$ and $\mathcal{W}$. The following lemma can be seen as an analog of Lemma 5.4.

LEMMA B.2. *There exist constants $C_1, C_2 > 0$ such that if $\frac{np}{\log n} > C_1$, then we have*

$$\|(A \otimes J_d) \circ \mathcal{W}\| \le C_2\sqrt{dnp},$$

$$\sum_{i=1}^n \left\|\sum_{j\in[n]\setminus\{i\}} A_{ij}\left(Z_i^{*\mathrm{T}}\mathcal{W}_{ij}Z_j^* - Z_j^{*\mathrm{T}}\mathcal{W}_{ji}Z_i^*\right)\right\|_{\mathrm{F}}^2 \le 2d(d-1)n^2p\left(1 + C_2\sqrt{\frac{\log n}{n}}\right),$$

$$\sum_{i=1}^n \left\|\sum_{j\in[n]\setminus\{i\}} A_{ij}\mathcal{W}_{ij}Z_j^*\right\|_{\mathrm{F}}^2 \le d^2n^2p\left(1 + C_2\sqrt{\frac{\log n}{n}}\right),$$

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

PROOF. The first inequality is from Lemma 4.2 of [19]. The second and third inequalities are from (59) and (60), together with Lemma 4.3, of [19], respectively. □

PROOF OF PROPOSITION 3.2. By Lemma 5.4 and Lemma B.2, there exist constants $c_1, c_2 > 0$ such that when $\frac{np}{\log n} > c_1$, we have $\|A - \mathbb{E}A\| \le c_2\sqrt{np}$ and $\|(A \otimes J_d) \circ \mathcal{W}\| \le c_2\sqrt{dnp}$ with probability at least $1 - 6n^{-10}$. By Lemma 3.1 and Lemma 5.3, we have $\lambda_1^* = \lambda_d^* \ge (n-1)p - c_2\sqrt{np}$, $\max\{|\lambda_{d+1}^*|, |\lambda_n^*|\} \le p + c_2\sqrt{np}$, and $\lambda_d^* - \lambda_{d+1}^* \ge np - 2c_2\sqrt{np}$. Note that $d$ is a constant. When $\frac{np}{\log n}$ and $\frac{np}{\sigma^2}$ are greater than some sufficiently large constant, we have $4\sigma\left\|(A\otimes J_d)\circ\mathcal{W}\right\|\leq np/2\leq\min\{\lambda_d^*,\lambda_d^*-\lambda_{d+1}^*\}$ satisfied. Since $\mathcal{X}-(A\otimes J_d)\circ Z^*Z^{*\mathrm{H}}=\sigma(A\otimes J_d)\circ\mathcal{W}$, a direct application of Lemma 3.2 leads to

$$\begin{aligned}
&\inf_{O\in\mathcal{O}(d)}\left\|U-\widetilde{U}O\right\| \\
&\leq\frac{8\sqrt{2}}{3(\lambda_1^*-\lambda_{d+1}^*)}\Bigg(\left(\frac{4}{3(\lambda_1^*-\lambda_{d+1}^*)}+\frac{2}{\lambda_1^*}\right)\sigma^2\left\|(A\otimes J_d)\circ\mathcal{W}\right\|^2 \\
&\quad+\frac{\max\{|\lambda_{d+1}^*|,|\lambda_n^*|\}}{\lambda_1^*}\sigma\left\|(A\otimes J_d)\circ\mathcal{W}\right\|\Bigg) \\
&=\frac{8\sqrt{2}}{3(np/2)}\left(\left(\frac{4}{3(np/2)}+\frac{2}{np/2}\right)\sigma^2c_2^2dnp+\frac{p+c_2\sqrt{np}}{np/2}\sigma c_2\sqrt{dnp}\right) \\
&\leq c_3\frac{\sigma^2d+\sigma\sqrt{d}}{np},
\end{aligned}$$

for some constant $c_3>0$. □

**B.3. Proof of Theorem 3.1.** We first state useful technical lemmas. They are analogs of Lemma 5.7 and Lemma 5.8, respectively. Lemma B.3 is proved in (31) of [19].

LEMMA B.3. *There exists some constant $C>0$ such that for any $\rho$ that satisfies $\frac{\rho^2np}{d^2\sigma^2}\geq C$, we*

$$\sum_{i=1}^n\mathbb{I}\left\{\frac{2\sigma}{np}\left\|\sum_{j\in[n]\setminus\{i\}}A_{ij}\mathcal{W}_{ij}Z_j^*\right\|>\rho\right\}\leq\frac{\sigma^2}{\rho^2p}\exp\left(-\sqrt{\frac{\rho^2np}{\sigma^2}}\right),$$

*with probability at least* $1-\exp\left(-\sqrt{\frac{\rho^2np}{\sigma^2}}\right)$.

LEMMA B.4 (Lemma 2.1 of [19]). *Let $X,\widetilde{X}\in\mathbb{R}^{d\times d}$ be two matrices of full rank. Then,*

$$\left\|\mathcal{P}(X)-\mathcal{P}(\widetilde{X})\right\|_{\mathrm{F}}\leq\frac{2}{s_{\min}(X)+s_{\min}(\widetilde{X})}\left\|X-\widetilde{X}\right\|_{\mathrm{F}}.$$

PROOF OF THEOREM 3.1. Let $O\in\mathcal{O}(d)$ satisfy $\|U-\widetilde{U}O\|=\inf_{O'\in\mathcal{O}(d)}\|U-\widetilde{U}O'\|$. Define $\Delta:=U-\widetilde{U}O\in\mathbb{R}^{nd\times d}$. Recall $\breve{u}$ is the leading eigenvector of $A$. From Proposition 2.1, Proposition 3.2, Lemma 5.4, and Lemma B.2, there exist constants $c_1,c_2>0$ such that if $\frac{np}{\log n},\frac{np}{\sigma^2}>c_1$, we have

$$\|\Delta\|\leq c_2\frac{\sigma^2d+\sigma\sqrt{d}}{np}, \tag{55}$$

$$\max_{j\in[n]}\left|\breve{u}_j-\frac{1}{\sqrt{n}}b_2\right|\leq c_2\left(\sqrt{\frac{\log n}{np}}+\frac{1}{\log(np)}\right)\frac{1}{\sqrt{n}}, \tag{56}$$

$$\|A-\mathbb{E}A\|\leq c_2\sqrt{np}, \tag{57}$$

$$\|(A\otimes J_d)\circ\mathcal{W}\|\leq c_2\sqrt{npd}, \tag{58}$$

$$\sum_{i=1}^{n}\left\|\sum_{j\in[n]\setminus\{i\}} A_{ij}\left(Z_i^{*\mathrm{T}}\mathcal{W}_{ij}Z_j^*-Z_j^{*\mathrm{T}}\mathcal{W}_{ji}Z_i^*\right)\right\|_{\mathrm{F}}^2 \leq 2d(d-1)n^2p\left(1+c_2\sqrt{\frac{\log n}{n}}\right), \tag{59}$$

$$\sum_{i=1}^{n}\left\|\sum_{j\in[n]\setminus\{i\}} A_{ij}\mathcal{W}_{ij}Z_j^*\right\|_{\mathrm{F}}^2 \leq d^2n^2p\left(1+c_2\sqrt{\frac{\log n}{n}}\right), \tag{60}$$

with probability at least $1-n^{-9}$, for some $b_2\in\{-1,1\}$. By Lemma 3.1 and Lemma 5.3, we have $\lambda_1^*=\lambda_d^*$, $|\lambda_d^*-(n-1)p|\leq c_2\sqrt{np}$, $\left|\lambda_{d+1}^*\right|\leq p+c_2\sqrt{np}$, and $\lambda_d^*-\lambda_{d+1}^*\geq np-2c_2\sqrt{np}$.

Using the same argument as (50) and (51) in the proof of Lemma 3.2, we can have an explicit expression for $\widetilde{U}$. Recall the definition of $\widetilde{U}$ in (22). Let $\mathcal{X}U^*=GDN^{\mathrm{T}}$ be its SVD where $G\in\mathcal{O}(nd,d)$, $N\in\mathcal{O}(d)$, and $D\in\mathbb{R}^{d\times d}$ is a diagonal matrix with singular values. By the decomposition (21), we have

$$\mathcal{X}U^*=((A\otimes J_d)\circ Z^*Z^{*\mathrm{T}})U^*+\sigma((A\otimes J_d)\circ\mathcal{W})U^*=\lambda_1^*U^*+\sigma((A\otimes J_d)\circ\mathcal{W})U^*. \tag{61}$$

Since the diagonal entries of $D$ correspond to the leading singular values of $\mathcal{X}U^*$, Weyl's inequality leads to $\max_{j\in[d]}|D_{jj}-\lambda_1^*|\leq\sigma\left\|(A\otimes J_d)\circ\mathcal{W}\right\|\leq c_2\sigma\sqrt{dnp}$. Denote

$$t:=p+c_2\sqrt{np}+c_2\sigma\sqrt{dnp}. \tag{62}$$

We then have

$$\max_{j\in[d]}|D_{jj}-np|\leq t. \tag{63}$$

When $\frac{np}{\log n},\frac{np}{d\sigma^2}$ are greater than some sufficiently large constant, we have $np/2\leq\lambda_1^*$ and $np/2\leq D_{jj}\leq 3np/2$ for all $j\in[d]$. As a consequence, all the diagonal entries of $D$ are positive. Then $\widetilde{U}$ can be written as

$$\widetilde{U}=\mathcal{X}U^*S,$$

where

$$S:=ND^{-1}N^{\mathrm{T}}\in\mathbb{R}^{d\times d}. \tag{64}$$

Then (63) leads to

$$\left\|\frac{1}{np}I_d-S\right\|=\left\|\frac{1}{np}I_d-D^{-1}\right\|\leq\frac{1}{np-t}-\frac{1}{np}\leq\frac{2t}{(np)^2}, \tag{65}$$

and

$$\|S\|=\left\|D^{-1}\right\|\leq\frac{2}{np}. \tag{66}$$

Using (61), we have the following decomposition for $U$:

$$U=\widetilde{U}O+\Delta=\mathcal{X}U^*SO+\Delta=\left(\lambda_1^*U^*+\sigma((A\otimes J_d)\circ\mathcal{W})U^*\right)SO+\Delta.$$

Recall the definition of $U^*$ in (14). Define $\Delta^*:=U^*-\frac{1}{\sqrt{n}}Z^*b_2$. When $\frac{np}{\log n}\geq 2c_2^*$, by the same argument used to derive (39) as in the proof of Theorem 2.1, we have

$$\|\Delta^*\|=\left\|Z^*\circ\left(\breve{u}\otimes\mathbb{1}_d-\frac{1}{\sqrt{n}}\mathbb{1}_n\otimes\mathbb{1}_d b_2\right)\right\|=\left\|\breve{u}\otimes\mathbb{1}_d-\frac{1}{\sqrt{n}}\mathbb{1}_n\otimes\mathbb{1}_d\right\|=\sqrt{d}\left\|\breve{u}-\frac{1}{\sqrt{n}}\mathbb{1}_n b_2\right\|$$

$$\leq\frac{2c_2\sqrt{np}+2p}{np}\sqrt{d}. \tag{67}$$

Then $U$ can be further decomposed into

$$U=\left(\lambda_1^*U^*+\sigma((A\otimes J_d)\circ\mathcal{W})\left(\frac{1}{\sqrt{n}}Z^*b_2+\Delta^*\right)\right)SO+\Delta.$$

For any $j\in[n]$, denote $[(A\otimes J_d)\circ\mathcal{W}]_{j\cdot}\in\mathbb{R}^{d\times nd}$ as the submatrix corresponding to its rows from the $((j-1)d+1)$th to the $(jd)$th. Note that $SO\in\mathbb{R}^{d\times d}$. Then $U_j$ has an expression:

$$\begin{aligned}U_j&=\left(\lambda_1^*U_j^*+\frac{\sigma}{\sqrt{n}}[(A\otimes J_d)\circ\mathcal{W}]_{j\cdot}Z^*b_2+\sigma[(A\otimes J_d)\circ\mathcal{W}]_{j\cdot}\Delta^*\right)SO+\Delta_j\\&=\left(\lambda_1^*Z_j^*\breve{u}_j+\frac{\sigma}{\sqrt{n}}\sum_{k\neq j}A_{jk}\mathcal{W}_{jk}Z_k^*b_2+\sigma[(A\otimes J_d)\circ\mathcal{W}]_{j\cdot}\Delta^*\right)SO+\Delta_j,\end{aligned}$$

where $\Delta_j\in\mathbb{R}^{d\times d}$ is denoted as the $j$th submatrix of $\Delta$.

Note that we have following properties for the mapping $\mathcal{P}$. For any $B\in\mathbb{R}^{d\times d}$ of full rank and any $F\in\mathcal{O}(d)$, we have $\mathcal{P}(BF)=\mathcal{P}(B)F$. In addition, if $B$ is positive-definite, $\mathcal{P}(B)=I_d$. Since we have shown the diagonal entries of $D$ are all lower bounded by $np/2$, (64) leads to $\mathcal{P}(S)=I_d$. Then

$$\left\|\widehat{Z}_j-Z_j^*Ob_2\right\|_{\mathrm{F}}=\left\|\mathcal{P}(U_j)-Z_j^*Ob_2\right\|_{\mathrm{F}}=\left\|\mathcal{P}(Z_j^{*\mathrm{T}}U_jO^{\mathrm{T}}b_2)-I_d\right\|_{\mathrm{F}}.$$

We have

$$Z_j^{*\mathrm{T}}U_jO^{\mathrm{T}}b_2=\left(\lambda_1^*\breve{u}_jb_2I_d+\frac{\sigma}{\sqrt{n}}\Xi_j+\sigma b_2Z_j^{*\mathrm{T}}[(A\otimes J_d)\circ\mathcal{W}]_{j\cdot}\Delta^*\right)S+Z_j^{*\mathrm{T}}\Delta_jO^{\mathrm{T}}b_2$$

where

$$\Xi_j:=\sum_{k\neq j}A_{jk}Z_j^{*\mathrm{T}}\mathcal{W}_{jk}Z_k^*.$$

Note that from (56), we have

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

As long as $\frac{np}{\log n}$ is greater than some sufficiently large constant, we have $b_2\breve{u}_j\geq\frac{1}{2\sqrt{n}}$. Since $\lambda_1^*$ is also positive, we have

$$\frac{Z_j^{*\mathrm{T}}U_jO^{\mathrm{T}}b_2}{\lambda_1^*\breve{u}_jb_2}=S+T_j \tag{68}$$

where $T_j$ is defined as

$$\begin{aligned}T_j&:=\frac{1}{\lambda_1^*\breve{u}_jb_2}\left(\left(\frac{\sigma}{\sqrt{n}}\Xi_j+\sigma b_2Z_j^{*\mathrm{T}}[(A\otimes J_d)\circ\mathcal{W}]_{j\cdot}\Delta^*\right)S+Z_j^{*\mathrm{T}}\Delta_jO^{\mathrm{T}}b_2\right)\\&=\frac{1}{\lambda_1^*\breve{u}_jb_2}\frac{\sigma}{\sqrt{n}}\Xi_jS+\frac{\sigma b_2Z_j^{*\mathrm{T}}[(A\otimes J_d)\circ\mathcal{W}]_{j\cdot}\Delta^*S}{\lambda_1^*\breve{u}_jb_2}+\frac{Z_j^{*\mathrm{T}}\Delta_jO^{\mathrm{T}}b_2}{\lambda_1^*\breve{u}_jb_2}.\end{aligned}$$

As a consequence, when $\det(U_j)\neq0$, we have

$$\left\|\widehat{Z}_j-Z_j^*Ob_2\right\|_{\mathrm{F}}=\left\|\mathcal{P}\left(\frac{Z_j^{*\mathrm{T}}U_jO^{\mathrm{T}}b_2}{\lambda_1^*\breve{u}_jb_2}\right)-I_d\right\|_{\mathrm{F}}=\left\|\mathcal{P}\left(S+T_j\right)-I_d\right\|_{\mathrm{F}}. \tag{69}$$

Let $0<\gamma,\rho<1/8$ whose values will be determined later. To simplify $\|\widehat{Z}_j - Z_j^* O b_2\|_{\rm F}$, consider the following two cases.

(1) If

$$\left\|\frac{1}{\lambda_1^*\breve{u}_j b_2}\frac{\sigma}{\sqrt{n}}\Xi_j S\right\| \leq \frac{\gamma}{np} \tag{70}$$

$$\left\|\frac{\sigma b_2 Z_j^{*\rm T}[(A\otimes J_d)\circ\mathcal{W}]_{j\cdot}\Delta^* S}{\lambda_1^*\breve{u}_j b_2}\right\| \leq \frac{\rho}{np}$$

$$\left\|\frac{Z_j^{*\rm T}\Delta_j O^{\rm T} b_2}{\lambda_1^*\breve{u}_j b_2}\right\| \leq \frac{\rho}{np} \tag{71}$$

all hold, then

$$\begin{aligned} s_{\min}(S+T_j) &\geq s_{\min}(S) - \|T_j\| = s_{\min}(D^{-1}) - \|T_j\| = D_{11}^{-1} - \|T_j\| \\ &\geq D_{11}^{-1} - \frac{\gamma+2\rho}{np}, \end{aligned}$$

which is greater than 0 by (63). Together with (68), we have $\det(U_j)\neq 0$. The same lower bound holds for $s_{\min}(S+(T_j+T_j^{\rm T})/2)$. Since $S$ is positive-definite, we have $\mathcal{P}(S+(T_j+T_j^{\rm T})/2) = I_d$. By Lemma B.4 and (69), we have

$$\begin{aligned} &\left\|\widehat{Z}_j - Z_j^* O b_2\right\|_{\rm F} \\ &= \left\|\mathcal{P}\left(S+T_j\right) - \mathcal{P}\left(S+\frac{T_j+T_j^{\rm T}}{2}\right)\right\|_{\rm F} \\ &\leq \frac{1}{\left(D_{11}^{-1}-\frac{\gamma+2\rho}{np}\right)}\left\|\frac{T_j-T_j^{\rm T}}{2}\right\|_{\rm F} \\ &\leq \frac{1}{\lambda_1^*\breve{u}_j b_2}\frac{1}{2\left(D_{11}^{-1}-\frac{\gamma+2\rho}{np}\right)}\left(\frac{\sigma}{\sqrt{n}}\left\|\Xi_j S - S^{\rm T}\Xi_j^{\rm T}\right\|_{\rm F} + 2\left\|\sigma b_2 Z_j^{*\rm T}[(A\otimes J_d)\circ\mathcal{W}]_{j\cdot}\Delta^* S\right\|_{\rm F}\right. \\ &\qquad \left. + 2\left\|Z_j^{*\rm T}\Delta_j O^{\rm T} b_2\right\|_{\rm F}\right). \end{aligned}$$

We can further simplify the first term in the display above. We have

$$\begin{aligned} \left\|\Xi_j S - S^{\rm T}\Xi_j^{\rm T}\right\|_{\rm F} &= \left\|\frac{1}{np}\left(\Xi_j - \Xi_j^{\rm T}\right) - \Xi_j\left(\frac{1}{np}I_d - S\right) + \left(\frac{1}{np}I_d - S^{\rm T}\right)\Xi_j^{\rm T}\right\|_{\rm F} \\ &\leq \frac{1}{np}\left\|\Xi_j - \Xi_j^{\rm T}\right\|_{\rm F} + 2\left\|\frac{1}{np}I_d - S\right\|\left\|\Xi_j\right\|_{\rm F}. \end{aligned}$$

Using (65) and (66), we have

$$\begin{aligned} \left\|\widehat{Z}_j - Z_j^* O b_2\right\|_{\rm F} &\leq \frac{1}{\lambda_1^*\breve{u}_j b_2}\frac{1}{2\left(D_{11}^{-1}-\frac{\gamma+2\rho}{np}\right)}\left(\frac{\sigma}{\sqrt{n}}\frac{1}{np}\left\|\Xi_j - \Xi_j^{\rm T}\right\|_{\rm F} + \frac{\sigma}{\sqrt{n}}\frac{t}{(np)^2}\left\|\Xi_j\right\|_{\rm F}\right. \\ &\qquad \left. + \frac{4}{np}\sigma\left\|[(A\otimes J_d)\circ\mathcal{W}]_{j\cdot}\Delta^*\right\|_{\rm F} + 2\left\|\Delta_j\right\|_{\rm F}\right). \end{aligned}$$

Using the lower bounds for $\lambda_1^*$, $\breve{u}_j b_2$, and $D_{11}^{-1}$, as given at the beginning of this proof, we have

$$\begin{aligned}&\left\|\widehat{Z}_j - Z_j^* O b_2\right\|_{\mathrm{F}}\\ &\leq \frac{1}{\left(np-p-c_2\sqrt{np}\right)\left(1-c_2\left(\sqrt{\frac{\log n}{np}}+\frac{1}{\log(np)}\right)\right)\left(\frac{1}{np+t}-\frac{\gamma+2\rho}{np}\right)}\frac{\sigma}{2np}\left\|\Xi_j-\Xi_j^{\mathrm{T}}\right\|_{\mathrm{F}}\\ &\quad+\frac{4\sigma t}{(np)^2}\left\|\Xi_j\right\|_{\mathrm{F}}+\frac{16\sigma\sqrt{n}}{np}\left\|[(A\otimes J_d)\circ\mathcal{W}]_{j\cdot}\Delta^*\right\|_{\mathrm{F}}+16\sqrt{n}\left\|\Delta_j\right\|_{\mathrm{F}}.\end{aligned}$$

Let $\eta>0$ whose value will be given later. By the same argument as used in the proof of Theorem 2.1, we have

$$\begin{aligned}&\left\|\widehat{Z}_j - Z_j^* O b_2\right\|_{\mathrm{F}}^2\\ &\leq \frac{1+\eta}{\left(np-p-c_2\sqrt{np}\right)^2\left(1-c_2\left(\sqrt{\frac{\log n}{np}}+\frac{1}{\log(np)}\right)\right)^2\left(\frac{1}{np+t}-\frac{\gamma+2\rho}{np}\right)^2}\frac{\sigma^2}{4(np)^2}\left\|\Xi_j-\Xi_j^{\mathrm{T}}\right\|_{\mathrm{F}}^2\\ &\quad+3(1+\eta^{-1})\frac{16\sigma^2t^2}{(np)^4}\left\|\Xi_j\right\|_{\mathrm{F}}^2+3(1+\eta^{-1})\frac{256\sigma^2 n}{(np)^2}\left\|[(A\otimes J_d)\circ\mathcal{W}]_{j\cdot}\Delta^*\right\|_{\mathrm{F}}^2\\ &\quad+3(1+\eta^{-1})64n\left\|\Delta_j\right\|_{\mathrm{F}}^2.\end{aligned}$$

(2) If any one of (70)-(71) does not hold, we simply upper bound $\|\widehat{Z}_j - Z_j^*\widetilde{Q}b_2\|_{\mathrm{F}}$ by $2\sqrt{d}$. Then this case can be written as

$$\begin{aligned}&\left\|\widehat{Z}_j - Z_j^* O b_2\right\|_{\mathrm{F}}^2\\ &\leq 4d\Bigg(\mathbb{I}\left\{\left\|\frac{1}{\lambda_1^*\breve{u}_j b_2}\frac{\sigma}{\sqrt{n}}\Xi_j S\right\|>\frac{\gamma}{np}\right\}+\mathbb{I}\left\{\left\|\frac{\sigma b_2 Z_j^{*\mathrm{T}}[(A\otimes J_d)\circ\mathcal{W}]_{j\cdot}\Delta^* S}{\lambda_1^*\breve{u}_j b_2}\right\|>\frac{\rho}{np}\right\}\\ &\quad+\mathbb{I}\left\{\left\|\frac{Z_j^{*\mathrm{T}}\Delta_j O^{\mathrm{T}} b_2}{\lambda_1^*\breve{u}_j b_2}\right\|>\frac{\rho}{np}\right\}\Bigg).\end{aligned}$$

Using (66), $\lambda_1^*\geq np/2$, and $\breve{u}_j b_2\geq 1/(2\sqrt{n})$, we have

$$\begin{aligned}&\left\|\widehat{Z}_j - Z_j^* O b_2\right\|_{\mathrm{F}}^2\\ &\leq 4d\left(\mathbb{I}\left\{8\sigma\left\|\Xi_j\right\|\geq\gamma np\right\}+\mathbb{I}\left\{8\sqrt{n}\sigma\left\|[(A\otimes J_d)\circ\mathcal{W}]_{j\cdot}\Delta^*\right\|\geq\rho np\right\}+\mathbb{I}\left\{4\sqrt{n}\left\|\Delta_j\right\|\geq\rho\right\}\right)\\ &\leq 4d\left(\mathbb{I}\left\{8\sigma\left\|\Xi_j\right\|\geq\gamma np\right\}+\frac{64\sigma^2 n}{(\rho np)^2}\left\|[(A\otimes J_d)\circ\mathcal{W}]_{j\cdot}\Delta^*\right\|_{\mathrm{F}}^2+16n\rho^{-2}\left\|\Delta_j\right\|_{\mathrm{F}}^2\right).\end{aligned}$$

Combining these two cases together, we have

$$\begin{aligned}&\left\|\widehat{Z}_j - Z_j^* O b_2\right\|_{\mathrm{F}}^2\\ &\leq \frac{1+\eta}{\left(np-p-c_2\sqrt{np}\right)^2\left(1-c_2\left(\sqrt{\frac{\log n}{np}}+\frac{1}{\log(np)}\right)\right)^2\left(\frac{1}{np+t}-\frac{\gamma+2\rho}{np}\right)^2}\frac{\sigma^2}{4(np)^2}\left\|\Xi_j-\Xi_j^{\mathrm{T}}\right\|_{\mathrm{F}}^2\\ &\quad+3(1+\eta^{-1})\frac{16\sigma^2t^2}{(np)^4}\left\|\Xi_j\right\|_{\mathrm{F}}^2+3(1+\eta^{-1})\frac{256\sigma^2 n}{(np)^2}\left\|[(A\otimes J_d)\circ\mathcal{W}]_{j\cdot}\Delta^*\right\|_{\mathrm{F}}^2\end{aligned}$$

$$
\begin{aligned}
&+ 3(1+\eta^{-1})64n \|\Delta_j\|_{\mathrm{F}}^2 \\
&+ 4d\left(\mathbb{I}\{8\sigma\|\Xi_j\| \geq \gamma np\} + \frac{64\sigma^2 n}{(\rho np)^2}\|[(A\otimes J_d)\circ\mathcal{W}]_{j\cdot}\Delta^*\|_{\mathrm{F}}^2 + 16n\rho^{-2}\|\Delta_j\|_{\mathrm{F}}^2\right) \\
= &\frac{1+\eta}{\left(np-p-c_2\sqrt{np}\right)^2\left(1-c_2\left(\sqrt{\frac{\log n}{np}}+\frac{1}{\log(np)}\right)\right)^2\left(\frac{1}{np+t}-\frac{\gamma+2\rho}{np}\right)^2}\frac{\sigma^2}{4(np)^2}\left\|\Xi_j-\Xi_j^{\mathrm{T}}\right\|_{\mathrm{F}}^2 \\
&+ 3(1+\eta^{-1})\frac{16\sigma^2t^2}{(np)^4}\|\Xi_j\|_{\mathrm{F}}^2 + 4d\mathbb{I}\{8\sigma\|\Xi_j\|\geq\gamma np\} \\
&+ \frac{256\sigma^2 n}{(np)^2}\left(3(1+\eta^{-1})+d\rho^{-2}\right)\|[(A\otimes J_d)\circ\mathcal{W}]_{j\cdot}\Delta^*\|_{\mathrm{F}}^2 \\
&+ 64n\left(3(1+\eta^{-1})+d\rho^{-2}\right)\|\Delta_j\|_{\mathrm{F}}^2.
\end{aligned}
$$

As a result, we have

$$
\begin{aligned}
&\ell^{\mathrm{od}}(\widehat{Z},Z^*) \\
&\leq \frac{1}{n}\sum_{j\in[n]}\left\|\widehat{Z}_j - Z_j^* Ob_2\right\|_{\mathrm{F}}^2 \\
&\leq \frac{1+\eta}{\left(np-p-c_2\sqrt{np}\right)^2\left(1-c_2\left(\sqrt{\frac{\log n}{np}}+\frac{1}{\log(np)}\right)\right)^2\left(\frac{1}{np+t}-\frac{\gamma+2\rho}{np}\right)^2} \\
&\quad\times\frac{\sigma^2}{4(np)^2}\frac{1}{n}\sum_{j\in[n]}\|\Xi_j-\Xi_j^{\mathrm{T}}\|_{\mathrm{F}}^2 \\
&\quad+ 3(1+\eta^{-1})\frac{16\sigma^2t^2}{(np)^4}\frac{1}{n}\sum_{j\in[n]}\|\Xi_j\|_{\mathrm{F}}^2 + 4d\frac{1}{n}\sum_{j\in[n]}\mathbb{I}\{8\sigma\|\Xi_j\|\geq\gamma np\} \\
&\quad+ \frac{256\sigma^2}{(np)^2}\left(3(1+\eta^{-1})+d\rho^{-2}\right)\sum_{j\in[n]}\|[(A\otimes J_d)\circ\mathcal{W}]_{j\cdot}\Delta^*\|_{\mathrm{F}}^2 \\
&\quad+ 64\left(3(1+\eta^{-1})+d\rho^{-2}\right)\sum_{j\in[n]}\|\Delta_j\|_{\mathrm{F}}^2.
\end{aligned}
$$

In the rest of the proof, we are going to simplify the display above. Specifically, we are going to upper bound $\sum_{j\in[n]}\|\Xi_j-\Xi_j^{\mathrm{T}}\|_{\mathrm{F}}^2$, $\sum_{j\in[n]}\|\Xi_j\|_{\mathrm{F}}^2$, $\sum_{j\in[n]}\mathbb{I}\{8\sigma\|\Xi_j\|\geq\gamma np\}$, $\sum_{j\in[n]}\|[(A\otimes J_d)\circ\mathcal{W}]_{j\cdot}\Delta^*\|_{\mathrm{F}}^2$, and $\sum_{j\in[n]}\|\Delta_j\|_{\mathrm{F}}^2$.

For $\sum_{j\in[n]}\|\Xi_j-\Xi_j^{\mathrm{T}}\|_{\mathrm{F}}^2$ and $\sum_{j\in[n]}\|\Xi_j\|_{\mathrm{F}}^2$, note that they are the left-hand sides of (59) and (60), respectively. Hence, they can be upper bounded by the right-hand sides of (59) and (60), respectively. For $\sum_{j\in[n]}\mathbb{I}\{8\sigma\|\Xi_j\|\geq\gamma np\}$, according to Lemma B.3, if $\frac{\gamma^2 np}{d^2\sigma^2} > c_3$ for some $c_3>0$, we have

$$
\sum_{j\in[n]}\mathbb{I}\{8\sigma\|\Xi_j\|\geq\gamma np\} \leq \frac{16\sigma^2}{\gamma^2 p}\exp\left(-\sqrt{\frac{\gamma^2 np}{16\sigma^2}}\right)
$$

with probability at least $1-\exp\left(-\sqrt{\frac{\gamma^2 np}{16\sigma^2}}\right)$. When $c_3$ is sufficiently large, it follows that

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

by the same argument as in the proof of Theorem 2.1. For $\sum_{j\in[n]}\|[(A\otimes J_d)\circ\mathcal{W}]_{j\cdot}\Delta^*\|_{\mathrm{F}}^2$, we have

$$\begin{aligned}\sum_{j\in[n]}\|[(A\otimes J_d)\circ\mathcal{W}]_{j\cdot}\Delta^*\|_{\mathrm{F}}^2 &= \|(A\otimes J_d)\circ\mathcal{W}\Delta^*\|_{\mathrm{F}}^2\\ &\leq \|(A\otimes J_d)\circ\mathcal{W}\|^2\|\Delta^*\|_{\mathrm{F}}^2\\ &\leq d\|(A\otimes J_d)\circ\mathcal{W}\|^2\|\Delta^*\|^2\\ &\leq c_2 d\left(\sqrt{dnp}\frac{2c_2\sqrt{np}+2p}{np}\sqrt{d}\right)^2,\end{aligned}$$

where in the second to last inequality we use the fact that $\Delta^*$ is rank-$d$ and in the last inequality we use (67). For $\sum_{j\in[n]}\|\Delta_j\|_{\mathrm{F}}^2$, we have $\sum_{j\in[n]}\|\Delta_j\|_{\mathrm{F}}^2=\|\Delta\|_{\mathrm{F}}^2\leq d\|\Delta\|^2\leq d\left(c_2\frac{\sigma^2 d+\sigma\sqrt{d}}{np}\right)^2$ where the last inequality is due to (55).

Using the above results, we have

$$\begin{aligned}&\ell^{\mathrm{od}}(\widehat{Z},Z^*)\\ &\leq\frac{1+\eta}{\left(np-p-c_2\sqrt{np}\right)^2\left(1-c_2\left(\sqrt{\frac{\log n}{np}}+\frac{1}{\log(np)}\right)\right)^2\left(\frac{1}{np+t}-\frac{\gamma+2\rho}{np}\right)^2}\\ &\quad\times\frac{\sigma^2}{4(np)^2}2d(d-1)np\left(1+c_2'\sqrt{\frac{\log n}{n}}\right)\\ &\quad+3(1+\eta^{-1})\frac{16\sigma^2 t^2}{(np)^4}d^2 np\left(1+c_2'\sqrt{\frac{\log n}{n}}\right)+4d\left(\frac{\sigma^2}{\gamma^2 np}\right)^3\\ &\quad+\frac{256\sigma^2}{(np)^2}\left(3(1+\eta^{-1})+d\rho^{-2}\right)c_2 d\left(\sqrt{dnp}\frac{2c_2\sqrt{np}+2p}{np}\sqrt{d}\right)^2\\ &\quad+64\left(3(1+\eta^{-1})+d\rho^{-2}\right)d\left(c_2\frac{\sigma^2 d+\sigma\sqrt{d}}{np}\right)^2.\end{aligned}$$

Note that $\frac{1}{(1-x)^2}\leq 1+16x$ for any $0\leq x\leq\frac{1}{2}$. When $\frac{np}{\log n}$ is greater than some sufficiently large constant, we have $\left(1-c_2\left(\sqrt{\frac{\log n}{np}}+\frac{1}{\log(np)}\right)\right)^{-2}\leq 16c_2\left(\sqrt{\frac{\log n}{np}}+\frac{1}{\log(np)}\right)$ and $\left(1-c_2\frac{1}{\sqrt{np}}-\frac{1}{n}\right)^{-2}\leq 16\left(c_2\frac{1}{\sqrt{np}}+\frac{1}{n}\right)$. When $\frac{np}{d\sigma^2}$ is also greater than some sufficiently large constant, we have $\left(\frac{np}{np+t}-\gamma-2\rho\right)^{-2}\leq 16\left(\frac{t}{np+t}+\gamma+2\rho\right)\leq 16\left(\frac{t}{np}+\gamma+2\rho\right)\leq 16\left(\frac{p+c_2\sqrt{np}+c_2\sigma\sqrt{dnp}}{np}+\gamma+2\rho\right)$, using the definition of $t$ in (62). We then have

$\ell^{\mathrm{od}}(\widehat{Z},Z^*)$

$$\leq 16^3 c_2(1+\eta)\left(c_2\frac{1}{\sqrt{np}}+\frac{1}{n}\right)\left(\sqrt{\frac{\log n}{np}}+\frac{1}{\log(np)}\right)\left(\frac{p+c_2\sqrt{np}+c_2\sigma\sqrt{dnp}}{np}+\gamma+2\rho\right)$$
$$\times\left(1+c_2'\sqrt{\frac{\log n}{n}}\right)\frac{d(d-1)\sigma^2}{2np}$$
$$+3(1+\eta^{-1})\left(\frac{p+c_2\sqrt{np}+c_2\sigma\sqrt{dnp}}{np}\right)^2\left(1+c_2'\sqrt{\frac{\log n}{n}}\right)\frac{16}{np}\frac{d^2\sigma^2}{np}$$
$$+4\gamma^{-6}\left(\frac{\sigma^2}{np}\right)^2\frac{d\sigma^2}{np}+256c_2\left(3(1+\eta^{-1})+d\rho^{-2}\right)\left(\frac{2c_2}{\sqrt{np}}+\frac{2}{n\sqrt{np}}\right)^2\frac{d^2\sigma^2}{np}$$
$$+64\left(3(1+\eta^{-1})+d\rho^{-2}\right)\left(c_2\frac{\sigma\sqrt{d}+1}{\sqrt{np}}\right)^2\frac{d^2\sigma^2}{np}.$$

After rearrangement, there exists some constant $c_5>0$ such that

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

We can take $\gamma^2=\sqrt{d^2\sigma^2/np}$ (then $\frac{\gamma^2 np}{d^2\sigma^2}>c_3$ is guaranteed as long as $\frac{np}{d^2\sigma^2}>c_3^2$). We also take $\rho^2=\sqrt{(d+d\sigma^2)/np}$ and let $\eta=\rho^2$. They are guaranteed to be smaller than $1/8$ when $\frac{np}{d}$ and $\frac{np}{d^2\sigma^2}$ are greater than some large constant. Then, there exists some constant $c_6>0$ such that

$$\ell^{\mathrm{od}}(\widehat{Z},Z^*)\leq\left(1+c_5\left(\left(\frac{d+d\sigma^2}{np}\right)^{\frac{1}{2}}+\left(\frac{d^2\sigma^2}{np}\right)^{\frac{1}{4}}+\left(\frac{d+d\sigma^2}{np}\right)^{\frac{1}{4}}+\sqrt{\frac{\log n}{np}}+\frac{1}{\log(np)}\right.\right.$$
$$\left.\left.+d^{-3}\left(\frac{\sigma^2}{np}\right)^{\frac{1}{2}}+\sqrt{\frac{d\sigma^2}{np}}+(1+d)\sqrt{\frac{np}{d+d\sigma^2}}\left(\frac{1+d\sigma^2}{np}\right)\right)\right)\frac{d(d-1)\sigma^2}{2np}$$
$$\leq\left(1+c_6\left(\left(\frac{d+d^2\sigma^2}{np}\right)^{\frac{1}{4}}+\sqrt{\frac{\log n}{np}}+\frac{1}{\log(np)}\right)\right)\frac{d(d-1)\sigma^2}{2np}.$$

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

□

## APPENDIX C: CALCULATION FOR (18)

Recall the definitions of $Y^*$ and $Y$ in (17). First, we are going to show $v$, the leading eigenvector of $Y$, must be a linear combination of $e_1$ and $e_2$. Note that for any unit vector $x=(x_1,\ldots,x_n)^{\mathrm{T}}\in\mathbb{R}^n$, we have

$$x^{\mathrm{T}}Yx=x^{\mathrm{T}}Y^*x+x^{\mathrm{T}}(Y-Y^*)x=\left(-\sum_{2\leq j\leq n}x_j^2\right)+\frac{\delta}{2}(x_1+x_2)^2=-1+x_1^2+\frac{\delta}{2}(x_1+x_2)^2.$$

If $x$ maximizes the right-hand side over the unit sphere, it is obvious that neither $x_1$ nor $x_2$ can be 0. In addition, $x_1 x_2 \geq 0$ and $x_1^2 + x_2^2 = 1$ must be satisfied; otherwise the right-hand side can be made strictly larger. Then we can write $v = \alpha e_1 + \sqrt{1-\alpha^2} e_2$ where $\alpha \in [0,1]$. Since $Yv = \frac{\delta}{2}(\alpha + \sqrt{1-\alpha^2})e_1 + \left(\frac{\delta}{2}(\alpha + \sqrt{1-\alpha^2}) - \sqrt{1-\alpha^2}\right) e_2$, we have

$$\frac{\alpha}{\frac{\delta}{2}(\alpha + \sqrt{1-\alpha^2})} = \frac{\sqrt{1-\alpha^2}}{\left(\frac{\delta}{2}(\alpha + \sqrt{1-\alpha^2}) - \sqrt{1-\alpha^2}\right)}.$$

After rearrangement, this gives $\delta(2\alpha^2 - 1) = 2\alpha\sqrt{1-\alpha^2}$ which means $\alpha^2 > \frac{1}{2}$. Squaring it yields the equation $4(1+\delta^2)\alpha^4 - 4(1+\delta^2)\alpha^2 + \delta^2 = 0$ whose solution is $\alpha^2 = \frac{1}{2}\left(1 \pm \frac{1}{\sqrt{1+\delta^2}}\right)$. Since $\alpha^2 > \frac{1}{2}$, we have $\alpha^2 = \frac{1}{2}\left(1 + \frac{1}{\sqrt{1+\delta^2}}\right)$. Hence,

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

We can verify it is the eigenvector of $Y$ corresponding to the eigenvalue $\frac{1}{2}(\delta + \sqrt{1+\delta^2} - 1)$.