Proofs for the Paper "Inference for Heteroskedastic PCA with Missing Data"

Yuling Yan* Yuxin Chen† Jianqing Fan‡

February 27, 2024

Contents

*Institute for Data, Systems, and Society, MIT, Cambridge, MA 02142, USA; Email: `yulingy@mit.edu`.

†Department of Statistics and Data Science, Wharton School, University of Pennsylvania, Philadelphia, PA, 19104, USA; Email: `yuxinc@wharton.upenn.edu`.

‡Department of Operations Research and Financial Engineering, Princeton University, Princeton, NJ 08544, USA; Email: `jqfan@princeton.edu`.

A Additional notation and organization of the appendix

For any matrix $\boldsymbol{U}$ with orthonormal columns, we denote by $\mathcal{P}_{\boldsymbol{U}}(\boldsymbol{M}) \coloneqq \boldsymbol{U}\boldsymbol{U}^\top \boldsymbol{M}$ the Euclidean projection of a matrix $\boldsymbol{M}$ onto the column space of $\boldsymbol{U}$, and let $\mathcal{P}_{\boldsymbol{U}^\perp}(\boldsymbol{M}) = \boldsymbol{M} - \mathcal{P}_{\boldsymbol{U}}(\boldsymbol{M})$ denote the Euclidean projection of $\boldsymbol{M}$ onto the orthogonal complement of the column space of $\boldsymbol{U}$. For any matrix $\boldsymbol{B} \in \mathbb{R}^{n_1 \times n_2}$ and some $l \in [n_1]$, define $\mathcal{P}_{-l,\cdot}(\boldsymbol{B})$ to be the orthogonal projection of the matrix $\boldsymbol{B}$ onto the subspace of matrix that vanishes outside the l-th row, namely,

$$[\mathcal{P}_{-l,\cdot}(\boldsymbol{B})]_{i,j} = \begin{cases} B_{i,j}, & \text{if } i \neq l, \\ 0, & \text{otherwise}, \end{cases} \qquad \forall (i,j) \in [n_1] \times [n_2]. \tag{A.1}$$

For any point $x \in \mathbb{R}^d$ and any non-empty convex set $\mathcal{C} \in \mathscr{C}^d$ satisfying $\mathcal{C} \neq \mathbb{R}^d$, let us define the signed distance function as follows

$$\delta_{\mathcal{C}}(x) \coloneqq \begin{cases} -\mathsf{dist}\left(x, \mathbb{R}^d \setminus \mathcal{C}\right), & \text{if } x \in \mathcal{C}; \\ \mathsf{dist}\left(x, \mathcal{C}\right), & \text{if } x \notin \mathcal{C}. \end{cases} \tag{A.2}$$

Here, $\mathsf{dist}(x, \mathcal{A})$ is the Euclidean distance between a point $x \in \mathbb{R}^d$ and a non-empty set $\mathcal{A} \subseteq \mathbb{R}^d$. Also, for any $\varepsilon \in \mathbb{R}$, define

$$\mathcal{C}^\varepsilon \coloneqq \left\{x \in \mathbb{R}^d : \delta_{\mathcal{C}}(x) \leq \varepsilon\right\} \tag{A.3}$$

for any non-empty convex set $\mathcal{C} \in \mathscr{C}^d$ satisfying $\mathcal{C} \neq \mathbb{R}^d$, and define $\varnothing^\varepsilon = \varnothing$ and $(\mathbb{R}^d)^\varepsilon = \mathbb{R}^d$.

Organization of the appendix. The rest of the appendix is organized as follows. Appendix C is devoted to proving Theorem 5. Appendices D and E are dedicated to establishing our theory for HeteroPCA (i.e., Theorems 11-14).

B A list of general theorems

For simplicity of presentation, the theorems presented in the main text (i.e., Section 3) concentrate on the scenario where $\kappa, \mu, r \asymp 1$. Note, however, that our theoretical framework can certainly allow these parameters to grow with the problem dimension. In this section, we provide a list of theorems accommodating more general scenarios; all subsequent proofs are dedicated to establishing these general theorems.

The following theorems generalize Theorem 1, Theorem 2, Theorem 3 and Theorem 4, respectively.

Theorem 11. *Suppose that $p < 1-\delta$ for some arbitrary constant $0<\delta<1$ or $p=1$. In addition, suppose that Assumption 1 holds and $n \gtrsim \kappa^8\mu^2 r^4 \kappa_\omega^2 \log^4(n+d)$, $d \gtrsim \kappa^7\mu^3 r^{7/2}\kappa_\omega^2\log^5(n+d)$,*

$$\frac{\omega_{\max}^2}{p\sigma_r^{\star 2}}\sqrt{\frac{d}{n}} \lesssim \frac{1}{\kappa^4\mu^{3/2}r^{11/4}\kappa_\omega \log^{7/2}(n+d)}, \qquad \frac{\omega_{\max}}{\sigma_r^\star}\sqrt{\frac{d}{np}} \lesssim \frac{1}{\kappa^{3/2}\mu r^{5/4}\kappa_\omega^{1/2}\log^3(n+d)},$$

$$ndp^2 \gtrsim \kappa^9\mu^4 r^{13/2}\kappa_\omega^2\log^9(n+d), \qquad np \gtrsim \kappa^9\mu^3 r^{11/2}\kappa_\omega^2\log^7(n+d).$$

Then the estimate $\boldsymbol{U}$ returned by Algorithm 2 with number of iterations satisfying (3.6) satisfies

$$\sup_{1\le l\le d}\ \sup_{\mathcal{C}\in\mathscr{C}^r}\left|\mathbb{P}\left(\left[\boldsymbol{U}\mathsf{sgn}\left(\boldsymbol{U}^\top\boldsymbol{U}^\star\right)-\boldsymbol{U}^\star\right]_{l,\cdot}\in\mathcal{C}\right)-\mathcal{N}\left(\boldsymbol{0},\boldsymbol{\Sigma}_{U,l}^\star\right)\{\mathcal{C}\}\right| = o(1).$$

where $\mathscr{C}^r$ represents the the set of all convex sets in $\mathbb{R}^r$, and for each $l\in[d]$,

$$\boldsymbol{\Sigma}_{U,l}^\star := \left(\frac{1-p}{np}\left\|\boldsymbol{U}_{l,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2^2+\frac{\omega_l^{\star 2}}{np}\right)(\boldsymbol{\Sigma}^\star)^{-2}+\frac{2(1-p)}{np}\boldsymbol{U}_{l,\cdot}^{\star\top}\boldsymbol{U}_{l,\cdot}^\star+(\boldsymbol{\Sigma}^\star)^{-2}\boldsymbol{U}^{\star\top}\mathsf{diag}\left\{\left[d_{l,i}^\star\right]_{1\le i\le d}\right\}\boldsymbol{U}^\star(\boldsymbol{\Sigma}^\star)^{-2}$$

where

$$d_{l,i}^\star := \frac{1}{np^2}\left[\omega_l^{\star 2}+(1-p)\left\|\boldsymbol{U}_{l,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2^2\right]\left[\omega_i^{\star 2}+(1-p)\left\|\boldsymbol{U}_{i,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2^2\right]+\frac{2(1-p)^2}{np^2}S_{l,i}^{\star 2}.$$

Theorem 12. *Suppose that the conditions of Theorem 11 hold. Further assume that $n \gtrsim \kappa^{12}\mu^3 r^{11/2}\kappa_\omega\log^5(n+d)$,*

$$\frac{\omega_{\max}^2}{p\sigma_r^{\star 2}}\sqrt{\frac{d}{n}} \lesssim \frac{1}{\kappa^{9/2}\mu^{3/2}r^{9/4}\kappa_\omega^{3/2}\log^{7/2}(n+d)}, \qquad \frac{\omega_{\max}}{\sigma_r^\star}\sqrt{\frac{d}{np}} \lesssim \frac{1}{\kappa^5\mu^{3/2}r^{9/4}\kappa_\omega^{3/2}\log^3(n+d)},$$

$$ndp^2 \gtrsim \kappa^{11}\mu^5 r^{13/2}\kappa_\omega^3\log^9(n+d), \qquad np \gtrsim \kappa^{11}\mu^4 r^{11/2}\kappa_\omega^3\log^7(n+d).$$

Then the confidence region $\mathsf{CR}_{U,l}^{1-\alpha}$ computed in Algorithm 3 obeys

$$\mathbb{P}\left(\boldsymbol{U}_{l,\cdot}^\star\mathsf{sgn}\left(\boldsymbol{U}^{\star\top}\boldsymbol{U}\right)\in\mathsf{CR}_{U,l}^{1-\alpha}\right) = 1-\alpha+o(1).$$

Theorem 13. *Suppose that $p < 1-\delta$ for some arbitrary constant $0<\delta<1$ or $p=1$. Consider any $1\le i,j\le d$. Assume that $\boldsymbol{U}^\star$ is μ-incoherent and satisfies the following condition*

$$\left\|\boldsymbol{U}_{i,\cdot}^\star\right\|_2+\left\|\boldsymbol{U}_{j,\cdot}^\star\right\|_2 \gtrsim \kappa r^{1/2}\log^{5/2}(n+d)\left(\frac{\omega_{\max}}{\sigma_r^\star}\sqrt{\frac{d}{np}}+\kappa_\omega\frac{\omega_{\max}^2}{p\sigma_r^{\star 2}}\sqrt{\frac{d}{n}}+\frac{\kappa\mu r\kappa_\omega\log(n+d)}{\sqrt{ndp^2}}\right)\sqrt{\frac{r}{d}}.$$

In addition, suppose that Assumption 1 holds and $n \gtrsim r\log^4(n+d)$, $d \gtrsim \kappa^8\mu^3 r^3\kappa_\omega^2\log^5(n+d)$,

$$\frac{\omega_{\max}^2}{p\sigma_r^{\star 2}}\sqrt{\frac{d}{n}} \lesssim \frac{1}{\kappa^{3/2}\mu r\kappa_\omega^{1/2}\log^{7/2}(n+d)}, \qquad \frac{\omega_{\max}}{\sigma_r^\star}\sqrt{\frac{d}{np}} \lesssim \frac{1}{\kappa^4\mu r\kappa_\omega^{1/2}\log^3(n+d)},$$

$$ndp^2 \gtrsim \kappa^{10}\mu^4 r^4 \kappa_\omega^2 \log^7(n+d), \qquad np \gtrsim \kappa^{10}\mu^3 r^3 \kappa_\omega^2 \log^7(n+d),$$

Then the matrix $\boldsymbol{S}$ computed by Algorithm 2 with number of iterations satisfying (3.6) obeys

$$\sup_{t\in\mathbb{R}}\left|\mathbb{P}\left(\frac{S_{i,j}-S_{i,j}^\star}{\sqrt{v_{i,j}^\star}}\le t\right)-\Phi(t)\right| = o(1).$$

Theorem 14. *Suppose that the conditions of Theorem 13 hold. Further assume that $n \gtrsim \kappa^9\mu^3 r^4 \kappa_\omega^3 \log^4(n+d)$,*

$$\frac{\omega_{\max}^2}{p\sigma_r^{\star 2}}\sqrt{\frac{d}{n}} \lesssim \frac{1}{\kappa^3\mu^{3/2}r^{5/2}\kappa_\omega^{3/2}\log^3(n+d)}, \qquad \frac{\omega_{\max}}{\sigma_r^\star}\sqrt{\frac{d}{np}} \lesssim \frac{1}{\kappa^{7/2}\mu^{3/2}r^{5/2}\kappa_\omega^{3/2}\log^{5/2}(n+d)},$$

$$ndp^2 \gtrsim \kappa^8\mu^5 r^7\kappa_\omega^3\log^8(n+d), \qquad np \gtrsim \kappa^8\mu^4 r^6\kappa_\omega^3\log^6(n+d),$$

and

$$\left\|\boldsymbol{U}_{i,\cdot}^\star\right\|_2+\left\|\boldsymbol{U}_{j,\cdot}^\star\right\|_2 \gtrsim \kappa^2\mu^{3/2}r^{5/2}\kappa_\omega^{3/2}\log^{5/2}(n+d)\left[\frac{\kappa\mu r\log(n+d)}{\sqrt{ndp}}+\frac{\omega_{\max}^2}{p\sigma_r^{\star 2}}\sqrt{\frac{d}{n}}+\frac{\omega_{\max}}{\sigma_r^\star}\sqrt{\frac{d}{np}}\right]\sqrt{\frac{r}{d}}.$$

Then the confidence interval computed in Algorithm 4 obeys

$$\mathbb{P}\left(S_{i,j}^\star\in\mathsf{CI}_{i,j}^{1-\alpha}\right)=1-\alpha+o(1).$$

Remark 1. While the above theorems allow the problem parameters $\kappa,\mu,r,\kappa_\omega$ to grow, our results remain suboptimal in terms of the dependency on these parameters. For instance, our theorems require stringent dependency on the condition number κ, which has been a common issue in analyzing spectral methods for low-rank matrix estimation using leave-one-out arguments (e.g., in order to obtain $\ell_{2,\infty}$ guarantees for matrix completion, the sample complexity requirement in the prior work by Chen et al. (2020) scales as κ^{10}). It is also worth noting that inference and uncertainty quantification might require stronger conditions compared to the estimation task, which further complicates matters (e.g., the multivariate Berry-Esseen theorem (cf. Theorem 15) employed in the current paper might already exhibit suboptimal scaling with r). Improving the dependency on all these parameters is a fundamentally important direction for future investigation.

C Analysis for HeteroPCA applied to subspace estimation (Theorem 5)

This section outlines the proof that establishes our statistical guarantees stated in Theorem 5. We shall begin by isolating several useful lemmas, and then combine these lemmas to complete the proof.

C.1 A few key lemmas

We now state below a couple of key lemmas that lead to improved statistical guarantees of Algorithm 5. From now on, we will denote

$$\boldsymbol{G}=\boldsymbol{G}^{t_0}, \qquad \boldsymbol{U}=\boldsymbol{U}^{t_0} \qquad \text{and} \qquad \boldsymbol{\Sigma}=(\boldsymbol{\Lambda}^{t_0})^{1/2}$$

and let

$$\boldsymbol{H} \coloneqq \boldsymbol{U}^\top\boldsymbol{U}^\natural.$$

To begin with, the first lemma controls the discrepancy between the sample gram matrix $\boldsymbol{G}$ and the ground truth $\boldsymbol{G}^\natural$, which improves upon prior theory developed in Zhang et al. (2022).

Lemma 1. *Suppose that the assumptions of Theorem 5 hold. Suppose that the number of iterations exceeds $t_0 \geq \log\left(\frac{\sigma_1^{\star 2}}{\zeta_{\mathsf{op}}}\right)$. Then with probability exceeding $1 - O(n^{-10})$, the iterate $\boldsymbol{G} \coloneqq \boldsymbol{G}^{t_0}$ computed in Algorithm 5 satisfies*

$$\left\|\mathcal{P}_{\mathsf{diag}}\left(\boldsymbol{G} - \boldsymbol{G}^{\natural}\right)\right\| \lesssim \kappa^{\natural 2}\sqrt{\frac{\mu^{\natural} r}{n_1}}\zeta_{\mathsf{op}}, \tag{C.1}$$

$$\left\|\boldsymbol{G} - \boldsymbol{G}^{\natural}\right\| \lesssim \zeta_{\mathsf{op}}, \tag{C.2}$$

where ζ_{op} is defined in (6.10). In addition, for each $m \in [n_1]$,

$$\left|G_{m,m} - G_{m,m}^{\natural}\right| \lesssim \kappa^{\natural 2}\zeta_{\mathsf{op}}\left(\left\|\boldsymbol{U}_{m,\cdot}^{\natural}\right\|_2 + \left\|\boldsymbol{U}_{m,\cdot}\right\|_2\right).$$

Proof. See Appendix C.3.2. □

Lemma 1 makes clear that $\boldsymbol{G}$ converges to $\boldsymbol{G}^{\natural}$ as the signal-to-noise ratio increases. We pause to compare this lemma with its counterpart in Zhang et al. (2022). Specifically, Zhang et al. (2022, Theorem 7) and its proof demonstrated that $\|\mathcal{P}_{\mathsf{diag}}(\boldsymbol{G} - \boldsymbol{G}^{\natural})\| \lesssim \zeta_{\mathsf{op}}$. In comparison our result (C.1) strengthens the prior estimation bound by a factor of $1/\sqrt{n_1}$ for the scenario with $\kappa^{\natural}, \mu^{\natural}, r = O(1)$. This improvement serves as one of the key analysis ingredients that allows us to sharpen the statistical guarantees for HeteroPCA.

The above bound on the difference between $\boldsymbol{G}$ and $\boldsymbol{G}^{\natural}$ in turn allows one to develop perturbation bounds for the eigenspace measured by the spectral norm, as stated in the following lemma. In the meantime, this lemma also contains some basic facts regarding $\boldsymbol{H}$ and $\boldsymbol{R}_{\boldsymbol{U}}$.

Lemma 2. *Suppose that the assumptions of Theorem 5 hold, and recall the definition of ζ_{op} in (6.10). Then with probability exceeding $1 - O(n^{-10})$, we have*

$$\left\|\boldsymbol{U}\boldsymbol{H} - \boldsymbol{U}^{\star}\right\| \lesssim \frac{\zeta_{\mathsf{op}}}{\sigma_r^{\natural 2}} \qquad \text{and} \qquad \left\|\boldsymbol{U}\boldsymbol{R}_{\boldsymbol{U}} - \boldsymbol{U}^{\star}\right\| \lesssim \frac{\zeta_{\mathsf{op}}}{\sigma_r^{\natural 2}}, \tag{C.3}$$

$$\left\|\boldsymbol{H} - \boldsymbol{R}_{\boldsymbol{U}}\right\| \lesssim \frac{\zeta_{\mathsf{op}}^2}{\sigma_r^{\natural 4}} \qquad \text{and} \qquad \left\|\boldsymbol{H}^{\top}\boldsymbol{H} - \boldsymbol{I}_r\right\| \lesssim \frac{\zeta_{\mathsf{op}}^2}{\sigma_r^{\natural 4}},$$

and

$$\frac{1}{2} \leq \sigma_i\left(\boldsymbol{H}\right) \leq 2, \qquad \forall 1 \leq i \leq r.$$

Proof. See Appendix C.3.3. □

While the above two lemmas focus on spectral norm metrics, the following lemma takes one substantial step further by characterizing the difference between $\boldsymbol{G}$ and $\boldsymbol{G}^{\natural}$ in each row, when projected onto the subspace spanned by $\boldsymbol{U}^{\natural}$.

Lemma 3. *Suppose that $n_1 \gtrsim \kappa^{\natural 4}\mu^{\natural} r$. Then with probability exceeding $1 - O(n^{-10})$, we have*

$$\begin{aligned}\left\|\left(\boldsymbol{G} - \boldsymbol{G}^{\natural}\right)_{m,\cdot}\boldsymbol{U}^{\natural}\right\|_2 \lesssim \zeta_{\mathsf{op},m}\sqrt{\frac{\mu^{\natural} r}{n_1}} + \sigma\sigma_1^{\natural}\sqrt{\mu^{\natural} r \log n}\left\|\boldsymbol{U}_{m,\cdot}^{\natural}\right\|_2 + \kappa^{\natural 2}\zeta_{\mathsf{op}}\left\|\boldsymbol{U}_{m,\cdot}^{\natural}\right\|_2^2 \\ + \kappa^{\natural 2}\zeta_{\mathsf{op}}\left\|\boldsymbol{U}_{m,\cdot}\boldsymbol{H} - \boldsymbol{U}_{m,\cdot}^{\natural}\right\|_2\left\|\boldsymbol{U}_{m,\cdot}^{\natural}\right\|_2\end{aligned}$$

simultaneously for each $m \in [n_1]$, with ζ_{op} and $\zeta_{\mathsf{op},m}$ defined in (6.10).

Proof. See Appendix C.3.4. □

Next, we present a crucial technical lemma that uncovers the intertwined relation between $\boldsymbol{U}\boldsymbol{\Sigma}^2\boldsymbol{H}$ and $\boldsymbol{G}\boldsymbol{U}^{\natural}$ in an $\ell_{2,\infty}$ sense. To establish the $\ell_{2,\infty}$ bound in this lemma, we invoke the powerful leave-one-out analysis framework to decouple complicated statistical dependency.

Lemma 4. *Suppose that $n_1 \gtrsim \kappa^{\natural 4}\mu^{\natural} r$ and $\zeta_{\mathsf{op}} \ll \sigma_r^{\natural 2}/\kappa^{\natural 2}$. Then with probability exceeding $1 - O(n^{-10})$,*

$$\begin{aligned}\left\|\left(\boldsymbol{U}\boldsymbol{\Sigma}^2\boldsymbol{H} - \boldsymbol{G}\boldsymbol{U}^{\natural}\right)_{m,\cdot}\right\|_2 &= \left\|\boldsymbol{G}_{m,\cdot}\left(\boldsymbol{U}\boldsymbol{H} - \boldsymbol{U}^{\natural}\right)\right\|_2 \\ &\lesssim \zeta_{\mathsf{op},m}\left(\kappa^{\natural 2}\frac{\zeta_{\mathsf{op}}}{\sigma_r^{\natural 2}}\sqrt{\frac{\mu^{\natural} r}{n_1}} + \left\|\boldsymbol{U}\boldsymbol{H} - \boldsymbol{U}^{\natural}\right\|_{2,\infty}\right) + \kappa^{\natural 2}\frac{\zeta_{\mathsf{op}}^2}{\sigma_r^{\natural 2}}\left\|\boldsymbol{U}_{m,\cdot}^{\natural}\right\|_2 \\ &\quad + \kappa^{\natural 2}\zeta_{\mathsf{op}}\left\|\boldsymbol{U}_{m,\cdot}^{\natural}\right\|_2\left\|\boldsymbol{U}_{m,\cdot}\boldsymbol{H} - \boldsymbol{U}_{m,\cdot}^{\natural}\right\|_2 + \kappa^{\natural 2}\zeta_{\mathsf{op}}\left\|\boldsymbol{U}_{m,\cdot}\boldsymbol{H} - \boldsymbol{U}_{m,\cdot}^{\natural}\right\|_2^2.\end{aligned}$$

holds simultaneously for each $m \in [n_1]$. Here, ζ_{op} and $\zeta_{\mathsf{op},m}$ are quantities defined in (6.10).

Proof. See Appendix C.3.5. □

Furthermore, the lemma below reveals that $\boldsymbol{\Sigma}^2$ and $\boldsymbol{\Sigma}^{\natural 2}$ remain close even after $\boldsymbol{\Sigma}^2$ is rotated by the rotation matrix $\boldsymbol{R}_{\boldsymbol{U}}$.

Lemma 5. *Suppose that $n_1 \gtrsim \mu^{\natural 2} r \log^2 n$ and $n_2 \gtrsim r \log^4 n$. Then with probability exceeding $1 - O(n^{-10})$ we have*

$$\left\|\boldsymbol{R}_{\boldsymbol{U}}^{\top}\boldsymbol{\Sigma}^2\boldsymbol{R}_{\boldsymbol{U}} - \boldsymbol{\Sigma}^{\natural 2}\right\| \lesssim \kappa^{\natural 2}\sqrt{\frac{\mu^{\natural} r}{n_1}}\zeta_{\mathsf{op}} + \kappa^{\natural 2}\frac{\zeta_{\mathsf{op}}^2}{\sigma_r^{\natural 2}},$$

where ζ_{op} is defined in (6.10).

Proof. See Appendix C.3.6. □

With the above auxiliary results in place, we can put them together to yield the following $\ell_{2,\infty}$ statistical guarantees.

Lemma 6. *Suppose that $\zeta_{\mathsf{op}} \ll \sigma_r^{\natural 2}/\kappa^{\natural 2}$ and that*

$$n_1 \gtrsim \kappa^{\natural 4}\mu^{\natural} r + \mu^{\natural 2} r \log^2 n, \qquad n_2 \gtrsim r \log^4 n.$$

Then with probability at least $1 - O(n^{-10})$,

$$\left\|\boldsymbol{U}_{m,\cdot}\boldsymbol{R}_{\boldsymbol{U}} - \boldsymbol{U}_{m,\cdot}^{\natural}\right\|_2 \lesssim \frac{\zeta_{\mathsf{op},m}}{\sigma_r^{\natural 2}}\sqrt{\frac{\mu^{\natural} r}{n_1}} + \left\|\boldsymbol{U}_{m,\cdot}^{\natural}\right\|_2\left(\kappa^{\natural 2}\sqrt{\frac{\mu^{\natural} r}{n_1}}\frac{\zeta_{\mathsf{op}}}{\sigma_r^{\natural 2}} + \kappa^{\natural 2}\frac{\zeta_{\mathsf{op}}^2}{\sigma_r^{\natural 4}}\right)$$

holds simultaneously for each $m \in [n_1]$, and

$$\left\|\boldsymbol{U}\boldsymbol{R}_{\boldsymbol{U}} - \boldsymbol{U}^{\natural}\right\|_{2,\infty} \lesssim \frac{\zeta_{\mathsf{op}}}{\sigma_r^{\natural 2}}\sqrt{\frac{\mu^{\natural} r}{n_1}},$$

where ζ_{op} and $\zeta_{\mathsf{op},m}$ are defined in (6.10).

Proof. See Appendix C.3.7. □

C.2 Proof of Theorem 5

Armed with the above lemmas, we are in a position to establish Theorem 5. It is worth noting that, while Lemma 6 delivers $\ell_{2,\infty}$ perturbation bounds for the eigenspace, it falls short of revealing the relation between the estimation error and the desired approximation

$$\boldsymbol{Z} = \boldsymbol{E}\boldsymbol{V}^{\natural}(\boldsymbol{\Sigma}^{\natural})^{-1} + \mathcal{P}_{\mathsf{off\text{-}diag}}(\boldsymbol{E}\boldsymbol{E}^{\top})\boldsymbol{U}^{\natural}(\boldsymbol{\Sigma}^{\natural})^{-2}. \tag{C.4}$$

In order to justify the tightness of this approximation $\boldsymbol{Z}$, we intend to establish each of the following steps:

$$\boldsymbol{U}\boldsymbol{R}_{\boldsymbol{U}}\boldsymbol{\Sigma}^{\natural 2} \overset{\text{Step 1}}{\approx} \boldsymbol{U}\boldsymbol{\Sigma}^2\boldsymbol{R}_{\boldsymbol{U}} \overset{\text{Step 2}}{\approx} \boldsymbol{U}\boldsymbol{\Sigma}^2\boldsymbol{H} \overset{\text{Step 3}}{\approx} \boldsymbol{G}\boldsymbol{U}^{\natural} \overset{\text{Step 4}}{\approx} \boldsymbol{U}^{\natural}\boldsymbol{\Sigma}^{\natural 2} + \underbrace{\boldsymbol{E}\boldsymbol{V}^{\natural}\boldsymbol{\Sigma}^{\natural} + \mathcal{P}_{\mathsf{off\text{-}diag}}\left(\boldsymbol{E}\boldsymbol{E}^{\top}\right)\boldsymbol{U}^{\natural}}_{=\boldsymbol{Z}(\boldsymbol{\Sigma}^{\natural})^2}, \tag{C.5}$$

which would in turn ensure that

$$\boldsymbol{U}\boldsymbol{R}_{\boldsymbol{U}} \approx \boldsymbol{U}^{\natural} + \boldsymbol{Z}$$

as advertised.

C.2.1 Step 1: establishing the proximity of $\boldsymbol{U}\boldsymbol{R}_{\boldsymbol{U}}\boldsymbol{\Sigma}^{\natural 2}$ and $\boldsymbol{U}\boldsymbol{\Sigma}^{2}\boldsymbol{R}_{\boldsymbol{U}}$.

For each $m \in [n_1]$, Lemma 5 tells us that

$$\begin{aligned}
\left\|\boldsymbol{U}_{m,\cdot}\boldsymbol{R}_{\boldsymbol{U}}\boldsymbol{\Sigma}^{\natural 2}-\boldsymbol{U}_{m,\cdot}\boldsymbol{\Sigma}^{2}\boldsymbol{R}_{\boldsymbol{U}}\right\|_2 &= \left\|\boldsymbol{U}_{m,\cdot}\boldsymbol{R}_{\boldsymbol{U}}\left(\boldsymbol{\Sigma}^{\natural 2}-\boldsymbol{R}_{\boldsymbol{U}}^{\top}\boldsymbol{\Sigma}^{2}\boldsymbol{R}_{\boldsymbol{U}}\right)\right\|_2 \le \left\|\boldsymbol{U}_{m,\cdot}\right\|_2\left\|\boldsymbol{\Sigma}^{\natural 2}-\boldsymbol{R}_{\boldsymbol{U}}^{\top}\boldsymbol{\Sigma}^{2}\boldsymbol{R}_{\boldsymbol{U}}\right\| \\
&\lesssim \left(\left\|\boldsymbol{U}_{m,\cdot}^{\natural}\right\|_2+\frac{\zeta_{\mathsf{op},m}}{\sigma_r^{\natural 2}}\sqrt{\frac{\mu^{\natural} r}{n_1}}\right)\left\|\boldsymbol{\Sigma}^{\natural 2}-\boldsymbol{R}_{\boldsymbol{U}}^{\top}\boldsymbol{\Sigma}^{2}\boldsymbol{R}_{\boldsymbol{U}}\right\| \\
&\lesssim \left(\left\|\boldsymbol{U}_{m,\cdot}^{\natural}\right\|_2+\frac{\zeta_{\mathsf{op},m}}{\sigma_r^{\natural 2}}\sqrt{\frac{\mu^{\natural} r}{n_1}}\right)\left(\kappa^{\natural 2}\sqrt{\frac{\mu^{\natural} r}{n_1}}\zeta_{\mathsf{op}}+\kappa^{\natural 2}\frac{\zeta_{\mathsf{op}}^{2}}{\sigma_r^{\natural 2}}\right). \qquad \text{(C.6)}
\end{aligned}$$

Here, the penultimate inequality follows since — according to Lemma 6 — we have

$$\begin{aligned}
\left\|\boldsymbol{U}_{m,\cdot}\right\|_2 = \left\|\boldsymbol{U}_{m,\cdot}\boldsymbol{R}_{\boldsymbol{U}}\right\|_2 &\le \left\|\boldsymbol{U}_{m,\cdot}^{\natural}\right\|_2+\left\|\boldsymbol{U}_{m,\cdot}\boldsymbol{R}_{\boldsymbol{U}}-\boldsymbol{U}_{m,\cdot}^{\natural}\right\|_2 \\
&\lesssim \left\|\boldsymbol{U}_{m,\cdot}^{\natural}\right\|_2+\frac{\zeta_{\mathsf{op},m}}{\sigma_r^{\natural 2}}\sqrt{\frac{\mu^{\natural} r}{n_1}}+\left\|\boldsymbol{U}_{m,\cdot}^{\natural}\right\|_2\left(\kappa^{\natural 2}\sqrt{\frac{\mu^{\natural} r}{n_1}}\frac{\zeta_{\mathsf{op}}}{\sigma_r^{\natural 2}}+\kappa^{\natural 2}\frac{\zeta_{\mathsf{op}}^{2}}{\sigma_r^{\natural 4}}\right) \\
&\lesssim \left\|\boldsymbol{U}_{m,\cdot}^{\natural}\right\|_2+\frac{\zeta_{\mathsf{op},m}}{\sigma_r^{\natural 2}}\sqrt{\frac{\mu^{\natural} r}{n_1}} \qquad \text{(C.7)}
\end{aligned}$$

under the condition $\zeta_{\mathsf{op}}/\sigma_r^{\natural 2} \lesssim 1/\kappa^{\natural 2}$ and $n_1 \gtrsim \mu^{\natural} r$, while the last relation arises from Lemma 5.

C.2.2 Step 2: replacing $\boldsymbol{U}\boldsymbol{\Sigma}^{2}\boldsymbol{R}_{\boldsymbol{U}}$ with $\boldsymbol{U}\boldsymbol{\Sigma}^{2}\boldsymbol{H}$.

Given that $\boldsymbol{R}_{\boldsymbol{U}}$ and $\boldsymbol{H}$ are fairly close, one can expect that replacing $\boldsymbol{R}_{\boldsymbol{U}}$ with $\boldsymbol{H}$ in $\boldsymbol{U}\boldsymbol{\Sigma}^{2}\boldsymbol{R}_{\boldsymbol{U}}$ does not change the matrix by much. To formalize this, we invoke Lemma 2 to reach

$$\begin{aligned}
\left\|\boldsymbol{U}_{m,\cdot}\boldsymbol{\Sigma}^{2}\boldsymbol{H}-\boldsymbol{U}_{m,\cdot}\boldsymbol{\Sigma}^{2}\boldsymbol{R}_{\boldsymbol{U}}\right\|_2 &\le \left\|\boldsymbol{U}_{m,\cdot}\right\|_2\left\|\boldsymbol{\Sigma}^{2}\right\|\left\|\boldsymbol{H}-\boldsymbol{R}_{\boldsymbol{U}}\right\| \\
&\lesssim \left(\left\|\boldsymbol{U}_{m,\cdot}^{\natural}\right\|_2+\frac{\zeta_{\mathsf{op},m}}{\sigma_r^{\natural 2}}\sqrt{\frac{\mu^{\natural} r}{n_1}}\right)\sigma_1^{\natural 2}\frac{\zeta_{\mathsf{op}}^{2}}{\sigma_r^{\natural 4}} \\
&\asymp \kappa^{\natural 2}\frac{\zeta_{\mathsf{op}}^{2}}{\sigma_r^{\natural 2}}\left(\left\|\boldsymbol{U}_{m,\cdot}^{\natural}\right\|_2+\frac{\zeta_{\mathsf{op},m}}{\sigma_r^{\natural 2}}\sqrt{\frac{\mu^{\natural} r}{n_1}}\right). \qquad \text{(C.8)}
\end{aligned}$$

Here, the penultimate relation follows from Lemma 2, (C.7), as well as a direct consequence of Lemma 5:

$$\left\|\boldsymbol{\Sigma}^{2}\right\| = \left\|\boldsymbol{R}_{\boldsymbol{U}}^{\top}\boldsymbol{\Sigma}^{2}\boldsymbol{R}_{\boldsymbol{U}}\right\| \le \left\|\boldsymbol{\Sigma}^{\natural 2}\right\|+\left\|\boldsymbol{R}_{\boldsymbol{U}}^{\top}\boldsymbol{\Sigma}^{2}\boldsymbol{R}_{\boldsymbol{U}}-\boldsymbol{\Sigma}^{\natural 2}\right\| \le \sigma_1^{\natural 2}+\kappa^{\natural 2}\sqrt{\frac{\mu^{\natural} r}{n_1}}\zeta_{\mathsf{op}}+\kappa^{\natural 2}\frac{\zeta_{\mathsf{op}}^{2}}{\sigma_r^{\natural 2}} \asymp \sigma_1^{\natural 2}, \qquad \text{(C.9)}$$

provided that $\zeta_{\mathsf{op}}/\sigma_r^{\natural 2} \lesssim 1$ and $n_1 \gtrsim \mu^{\natural} r$.

C.2.3 Step 3: establishing the proximity of $\boldsymbol{U}\boldsymbol{\Sigma}^{2}\boldsymbol{H}$ and $\boldsymbol{G}\boldsymbol{U}^{\natural}$.

It is readily seen from Lemma 4 that

$$\begin{aligned}
\left\|\left(\boldsymbol{U}\boldsymbol{\Sigma}^{2}\boldsymbol{H}-\boldsymbol{G}\boldsymbol{U}^{\natural}\right)_{m,\cdot}\right\|_2 &\lesssim \zeta_{\mathsf{op},m}\left(\kappa^{\natural 2}\frac{\zeta_{\mathsf{op}}}{\sigma_r^{\natural 2}}\sqrt{\frac{\mu^{\natural} r}{n_1}}+\left\|\boldsymbol{U}\boldsymbol{H}-\boldsymbol{U}^{\natural}\right\|_{2,\infty}\right)+\kappa^{\natural 2}\frac{\zeta_{\mathsf{op}}^{2}}{\sigma_r^{\natural 2}}\left\|\boldsymbol{U}_{m,\cdot}^{\natural}\right\|_2 \\
&\quad+\kappa^{\natural 2}\zeta_{\mathsf{op}}\left\|\boldsymbol{U}_{m,\cdot}^{\natural}\right\|_2\left\|\boldsymbol{U}_{m,\cdot}\boldsymbol{H}-\boldsymbol{U}_{m,\cdot}^{\natural}\right\|_2+\kappa^{\natural 2}\zeta_{\mathsf{op}}\left\|\boldsymbol{U}_{m,\cdot}\boldsymbol{H}-\boldsymbol{U}_{m,\cdot}^{\natural}\right\|_2^{2} \\
&\lesssim \zeta_{\mathsf{op},m}\kappa^{\natural 2}\frac{\zeta_{\mathsf{op}}}{\sigma_r^{\natural 2}}\sqrt{\frac{\mu^{\natural} r}{n_1}}+\kappa^{\natural 2}\frac{\zeta_{\mathsf{op}}^{2}}{\sigma_r^{\natural 2}}\left\|\boldsymbol{U}_{m,\cdot}^{\natural}\right\|_2
\end{aligned}$$

$$
\begin{aligned}
&+\kappa^{\natural 2}\zeta_{\mathsf{op}}\left\|\boldsymbol{U}_{m,\cdot}^{\natural}\right\|_2\left(\frac{\zeta_{\mathsf{op},m}}{\sigma_r^{\natural 2}}\sqrt{\frac{\mu^{\natural}r}{n_1}}+\left\|\boldsymbol{U}_{m,\cdot}^{\natural}\right\|_2\kappa^{\natural 2}\sqrt{\frac{\mu^{\natural}r}{n_1}}\frac{\zeta_{\mathsf{op}}}{\sigma_r^{\natural 2}}+\left\|\boldsymbol{U}_{m,\cdot}^{\natural}\right\|_2\kappa^{\natural 2}\frac{\zeta_{\mathsf{op}}^2}{\sigma_r^{\natural 4}}\right)\\
&+\kappa^{\natural 2}\zeta_{\mathsf{op}}\left(\frac{\zeta_{\mathsf{op},m}}{\sigma_r^{\natural 2}}\sqrt{\frac{\mu^{\natural}r}{n_1}}+\left\|\boldsymbol{U}_{m,\cdot}^{\natural}\right\|_2\kappa^{\natural 2}\sqrt{\frac{\mu^{\natural}r}{n_1}}\frac{\zeta_{\mathsf{op}}}{\sigma_r^{\natural 2}}+\left\|\boldsymbol{U}_{m,\cdot}^{\natural}\right\|_2\kappa^{\natural 2}\frac{\zeta_{\mathsf{op}}^2}{\sigma_r^{\natural 4}}\right)^2\\
\lesssim\;&\zeta_{\mathsf{op},m}\kappa^{\natural 2}\frac{\zeta_{\mathsf{op}}}{\sigma_r^{\natural 2}}\sqrt{\frac{\mu^{\natural}r}{n_1}}+\kappa^{\natural 2}\frac{\zeta_{\mathsf{op}}^2}{\sigma_r^{\natural 2}}\left\|\boldsymbol{U}_{m,\cdot}^{\natural}\right\|_2, \qquad\qquad (\text{C.10})
\end{aligned}
$$

where the penultimate relation follows from Lemma 6, while the last relation holds provided that $n_1 \gtrsim \kappa^{\natural 4}\mu^{\natural}r$ and $\zeta_{\mathsf{op}} \lesssim \sigma_r^{\natural 2}$.

C.2.4 Step 4: investigating the statistical properties of $\boldsymbol{G}\boldsymbol{U}^{\natural}$.

It then remains to decompose $\boldsymbol{G}\boldsymbol{U}^{\natural}$ as claimed in (C.5). To begin with, we make the observation that

$$
\begin{aligned}
\boldsymbol{G}&=\mathcal{P}_{\mathsf{off\text{-}diag}}\left[\left(\boldsymbol{M}^{\natural}+\boldsymbol{E}\right)\left(\boldsymbol{M}^{\natural}+\boldsymbol{E}\right)^{\top}\right]+\mathcal{P}_{\mathsf{diag}}\left(\boldsymbol{G}^{\natural}\right)+\mathcal{P}_{\mathsf{diag}}\left(\boldsymbol{G}-\boldsymbol{G}^{\natural}\right)\\
&=\boldsymbol{G}^{\natural}+\mathcal{P}_{\mathsf{off\text{-}diag}}\left[\boldsymbol{E}\boldsymbol{M}^{\natural\top}+\boldsymbol{M}^{\natural}\boldsymbol{E}^{\top}+\boldsymbol{E}\boldsymbol{E}^{\top}\right]+\mathcal{P}_{\mathsf{diag}}\left(\boldsymbol{G}-\boldsymbol{G}^{\natural}\right),
\end{aligned}
$$

which together with the eigen-decomposition $\boldsymbol{G}^{\natural}=\boldsymbol{U}^{\natural}\boldsymbol{\Sigma}^{\natural 2}\boldsymbol{U}^{\natural\top}$ and the definition (C.4) of $\boldsymbol{Z}$ allows one to derive

$$
\begin{aligned}
\boldsymbol{G}\boldsymbol{U}^{\natural}-\left(\boldsymbol{U}^{\natural}+\boldsymbol{Z}\right)\boldsymbol{\Sigma}^{\natural 2}&=\boldsymbol{U}^{\natural}\boldsymbol{\Sigma}^{\natural 2}+\mathcal{P}_{\mathsf{off\text{-}diag}}\left[\boldsymbol{E}\boldsymbol{M}^{\natural\top}+\boldsymbol{M}^{\natural}\boldsymbol{E}^{\top}+\boldsymbol{E}\boldsymbol{E}^{\top}\right]\boldsymbol{U}^{\natural}+\mathcal{P}_{\mathsf{diag}}\left(\boldsymbol{G}-\boldsymbol{G}^{\natural}\right)\boldsymbol{U}^{\natural}\\
&\quad-\boldsymbol{U}^{\natural}\boldsymbol{\Sigma}^{\natural 2}-\left[\boldsymbol{E}\boldsymbol{M}^{\natural\top}+\mathcal{P}_{\mathsf{off\text{-}diag}}\left(\boldsymbol{E}\boldsymbol{E}^{\top}\right)\right]\boldsymbol{U}^{\natural}\\
&=\underbrace{\boldsymbol{M}^{\natural}\boldsymbol{E}^{\top}\boldsymbol{U}^{\natural}}_{=:\boldsymbol{R}_1}-\underbrace{\mathcal{P}_{\mathsf{diag}}\left(\boldsymbol{E}\boldsymbol{M}^{\natural\top}+\boldsymbol{M}^{\natural}\boldsymbol{E}^{\top}\right)\boldsymbol{U}^{\natural}}_{=:\boldsymbol{R}_2}+\underbrace{\mathcal{P}_{\mathsf{diag}}\left(\boldsymbol{G}-\boldsymbol{G}^{\natural}\right)\boldsymbol{U}^{\natural}}_{=:\boldsymbol{R}_3}.
\end{aligned}
$$

This motivates us to control the terms $\boldsymbol{R}_1$, $\boldsymbol{R}_2$ and $\boldsymbol{R}_3$ separately.

- Regarding $\boldsymbol{R}_1$, note that we have shown in (C.45) that

  $$
  \left\|\boldsymbol{U}^{\natural\top}\boldsymbol{E}\boldsymbol{V}^{\natural}\right\|\lesssim\sigma\sqrt{r\log n}+\frac{B\mu^{\natural}r\log n}{\sqrt{n_1n_2}}\lesssim\sigma\sqrt{r\log n}+\frac{\sigma\mu^{\natural}r}{\sqrt[4]{n_1n_2}}
  $$

  with probability exceeding $1-O(n^{-10})$, where the last inequality holds under Assumption 3. Therefore, we can derive

  $$
  \begin{aligned}
  \left\|\boldsymbol{e}_m^{\top}\boldsymbol{R}_1\right\|_2&=\left\|\boldsymbol{U}_{m,\cdot}^{\natural}\boldsymbol{\Sigma}^{\natural}\left(\boldsymbol{U}^{\natural\top}\boldsymbol{E}\boldsymbol{V}^{\natural}\right)^{\top}\right\|_2\le\left\|\boldsymbol{U}_{m,\cdot}^{\natural}\right\|_2\left\|\boldsymbol{\Sigma}^{\natural}\right\|\left\|\boldsymbol{U}^{\natural\top}\boldsymbol{E}\boldsymbol{V}^{\natural}\right\|\\
  &\lesssim\left\|\boldsymbol{U}_{m,\cdot}^{\natural}\right\|_2\sigma_1^{\natural}\left(\sigma\sqrt{r\log n}+\frac{\sigma\mu^{\natural}r}{\sqrt[4]{n_1n_2}}\right)\lesssim\left\|\boldsymbol{U}_{m,\cdot}^{\natural}\right\|_2\sqrt{\frac{\mu^{\natural}r}{n_1}}\zeta_{\mathsf{op}},
  \end{aligned}
  $$

  where the last relation holds provided that $n_1n_2 \gtrsim \mu^{\natural 2}r^2$.

- With regards to $\boldsymbol{R}_2$, we observe that

  $$
  \left\|\boldsymbol{e}_m^{\top}\boldsymbol{R}_2\right\|_2=\left|\sum_{j=1}^{n_2}E_{m,j}M_{m,j}^{\natural}\right|\left\|\boldsymbol{U}_{m,\cdot}^{\natural}\right\|_2
  $$

  Recalling from (C.47) that

  $$
  \max_{1\le i\le n_1}\left|\sum_{j=1}^{n_2}E_{i,j}M_{i,j}^{\natural}\right|\lesssim\frac{\sqrt{\mu^{\natural}r}}{n_1}\zeta_{\mathsf{op}}
  $$

holds with probability exceeding $1-O(n^{-10})$, we can upper bound

$$\left\|\boldsymbol{e}_m^\top \boldsymbol{R}_2\right\|_2 \lesssim \left\|\boldsymbol{U}^\natural_{m,\cdot}\right\|_2 \frac{\sqrt{\mu^\natural r}}{n_1}\zeta_{\mathsf{op}}.$$

- It remains to control $\|\boldsymbol{R}_3\|_{2,\infty}$, towards which we can apply Lemma 1 to reach

$$\left\|\boldsymbol{e}_m^\top \boldsymbol{R}_3\right\|_2 = \left|G_{m,m} - G^\natural_{m,m}\right| \left\|\boldsymbol{U}^\natural_{m,\cdot}\right\|_2 \lesssim \kappa^{\natural 2}\sqrt{\frac{\mu^\natural r}{n_1}}\zeta_{\mathsf{op}} \left\|\boldsymbol{U}^\natural_{m,\cdot}\right\|_2.$$

Taking the preceding bounds on $\|\boldsymbol{e}_m^\top \boldsymbol{R}_1\|_2$, $\|\boldsymbol{e}_m^\top \boldsymbol{R}_2\|_2$ and $\|\boldsymbol{e}_m^\top \boldsymbol{R}_3\|_2$ collectively, we arrive at

$$\left\|\boldsymbol{G}_{m,\cdot}\boldsymbol{U}^\natural - \left(\boldsymbol{U}^\natural + \boldsymbol{Z}\right)_{m,\cdot}\boldsymbol{\Sigma}^{\natural 2}\right\|_2 \leq \left\|\boldsymbol{e}_m^\top \boldsymbol{R}_1\right\|_2 + \left\|\boldsymbol{e}_m^\top \boldsymbol{R}_2\right\|_2 + \left\|\boldsymbol{e}_m^\top \boldsymbol{R}_3\right\|_2 \lesssim \kappa^{\natural 2}\sqrt{\frac{\mu^\natural r}{n_1}}\zeta_{\mathsf{op}}\left\|\boldsymbol{U}^\natural_{m,\cdot}\right\|_2. \tag{C.11}$$

C.2.5 Step 5: putting everything together.

To finish up, combine (C.6), (C.8), (C.10) and (C.11) to conclude that

$$\begin{aligned}
\left\|\left(\boldsymbol{U}\boldsymbol{R}_{\boldsymbol{U}} - \boldsymbol{U}^\star - \boldsymbol{Z}\right)_{m,\cdot}\right\|_2 &\leq \frac{1}{\sigma_r^{\natural 2}}\left\|\boldsymbol{U}_{m,\cdot}\boldsymbol{R}_{\boldsymbol{U}}\boldsymbol{\Sigma}^{\natural 2} - \left(\boldsymbol{U}^\natural + \boldsymbol{Z}\right)_{m,\cdot}\boldsymbol{\Sigma}^{\natural 2}\right\|_2 \\
&\leq \frac{1}{\sigma_r^{\natural 2}}\left\|\boldsymbol{U}_{m,\cdot}\boldsymbol{R}_{\boldsymbol{U}}\boldsymbol{\Sigma}^{\natural 2} - \boldsymbol{U}_{m,\cdot}\boldsymbol{\Sigma}^2\boldsymbol{R}_{\boldsymbol{U}}\right\|_2 + \frac{1}{\sigma_r^{\natural 2}}\left\|\boldsymbol{U}_{m,\cdot}\boldsymbol{\Sigma}^2\boldsymbol{H} - \boldsymbol{U}_{m,\cdot}\boldsymbol{\Sigma}^2\boldsymbol{R}_{\boldsymbol{U}}\right\|_2 \\
&\quad + \frac{1}{\sigma_r^{\natural 2}}\left\|\boldsymbol{U}_{m,\cdot}\boldsymbol{\Sigma}^2\boldsymbol{H} - \boldsymbol{G}_{m,\cdot}\boldsymbol{U}^\natural\right\|_2 + \frac{1}{\sigma_r^{\natural 2}}\left\|\boldsymbol{G}_{m,\cdot}\boldsymbol{U}^\natural - \left(\boldsymbol{U}^\natural + \boldsymbol{Z}\right)_{m,\cdot}\boldsymbol{\Sigma}^{\natural 2}\right\|_2 \\
&\lesssim \left\|\boldsymbol{U}^\natural_{m,\cdot}\right\|_2\left(\kappa^{\natural 2}\sqrt{\frac{\mu^\natural r}{n_1}}\frac{\zeta_{\mathsf{op}}}{\sigma_r^{\natural 2}} + \kappa^{\natural 2}\frac{\zeta_{\mathsf{op}}^2}{\sigma_r^{\natural 4}}\right) + \kappa^{\natural 2}\frac{\zeta_{\mathsf{op}}\zeta_{\mathsf{op},m}}{\sigma_r^{\natural 4}}\sqrt{\frac{\mu^\natural r}{n_1}}
\end{aligned}$$

as claimed, provided that $\zeta_{\mathsf{op}} \lesssim \sigma_r^{\natural 2}$

C.3 Proof of auxiliary lemmas

In this section, we establish the auxiliary lemmas needed in the proof of Theorem 5. At the core of our analysis lies a leave-one-out (and leave-two-out) analysis framework that has been previously adopted to analyze spectral methods (Abbe et al., 2020; Cai et al., 2021; Chen et al., 2021, 2019b), which we shall introduce below.

C.3.1 Preparation: leave-one-out and leave-two-out auxiliary estimates

To begin with, let us introduce the leave-one-out and leave-two-out auxiliary sequences and the associated estimates, which play a pivotal role in enabling fine-grained statistical analysis.

Leave-one-out auxiliary estimates. For each $m \in [n_1]$, define

$$\boldsymbol{M}^{(m)} := \mathcal{P}_{-m,\cdot}\left(\boldsymbol{M}\right) + \mathcal{P}_{m,\cdot}\left(\boldsymbol{M}^\natural\right) \qquad \text{and} \qquad \boldsymbol{G}^{(m)} := \mathcal{P}_{\mathsf{off\text{-}diag}}\left(\boldsymbol{M}^{(m)}\boldsymbol{M}^{(m)\top}\right) + \mathcal{P}_{\mathsf{diag}}\left(\boldsymbol{M}^\natural\boldsymbol{M}^{\natural\top}\right), \tag{C.12}$$

where we recall that $\mathcal{P}_{-m,\cdot}(\boldsymbol{M}) \in \mathbb{R}^{n_1 \times n_2}$ is obtained by setting to zero all entries in the m-th row of $\boldsymbol{M}$, and $\mathcal{P}_{m,\cdot}(\boldsymbol{M}) = \boldsymbol{M} - \mathcal{P}_{-m,\cdot}(\boldsymbol{M})$. Throughout this section, we let $\boldsymbol{U}^{(m)}\boldsymbol{\Lambda}^{(m)}\boldsymbol{U}^{(m)\top}$ be the top-r eigen-decomposition of $\boldsymbol{G}^{(m)}$, and we define

$$\boldsymbol{H}^{(m)} := \boldsymbol{U}^{(m)\top}\boldsymbol{U}^\natural. \tag{C.13}$$

Two features are particularly worth emphasizing:

- The matrices $\boldsymbol{U}^{(m)}$, $\boldsymbol{G}^{(m)}$ and $\boldsymbol{H}^{(m)}$ are all statistically independent of the m-th row of the noise matrix $\boldsymbol{E}$, given that $\boldsymbol{M}^{(m)}$ does not contain any randomness arising in the m-th row of $\boldsymbol{E}$.
- The estimate $\boldsymbol{U}^{(m)}$ (resp. $\boldsymbol{H}^{(m)}$ and $\boldsymbol{G}^{(m)}$) is expected to be extremely close to the original estimate $\boldsymbol{U}$ (resp. $\boldsymbol{H} = \boldsymbol{U}^\top \boldsymbol{U}^\natural$ and $\boldsymbol{G}$), given that we have only dropped a small fraction of data when generating the leave-one-out estimate.

Informally, the above two features taken together allow one to show the weak statistical dependency between $(\boldsymbol{U}, \boldsymbol{G}, \boldsymbol{H})$ and the m-th row of $\boldsymbol{E}$.

Leave-two-out auxiliary estimates. As it turns out, we are also in need of a collection of slightly more complicated leave-two-out estimates to assist in our analysis. Specifically, for each $m \in [n_1]$ and $l \in [n_2]$, we let

$$\boldsymbol{M}^{(m,l)} = \mathcal{P}_{-m,-l}\left(\boldsymbol{M}\right) + \mathcal{P}_{m,l}\left(\boldsymbol{M}^\natural\right) \qquad \text{and} \qquad \boldsymbol{G}^{(m,l)} = \mathcal{P}_{\mathsf{off\text{-}diag}}\left(\boldsymbol{M}^{(m,l)}\boldsymbol{M}^{(m,l)\top}\right) + \mathcal{P}_{\mathsf{diag}}\left(\boldsymbol{M}^\natural \boldsymbol{M}^{\natural\top}\right). \tag{C.14}$$

Here, $\mathcal{P}_{-m,-l}\left(\boldsymbol{M}\right) \in \mathbb{R}^{n_1 \times n_2}$ is obtained by zeroing out all entries in the m-th row and the l-th column of $\boldsymbol{M}$, and $\mathcal{P}_{m,l}\left(\boldsymbol{M}\right) = \boldsymbol{M} - \mathcal{P}_{-m,-l}\left(\boldsymbol{M}\right)$. We let $\boldsymbol{U}^{(m,l)}\boldsymbol{\Lambda}^{(m,l)}\boldsymbol{U}^{(m,l)\top}$ represent the top-r eigen-decomposition of $\boldsymbol{G}^{(m,l)}$, and define

$$\boldsymbol{H}^{(m,l)} \coloneqq \boldsymbol{U}^{(m,l)\top}\boldsymbol{U}^\natural. \tag{C.15}$$

Akin to the leave-one-out counterpart, the matrices $\boldsymbol{U}^{(m,l)}$, $\boldsymbol{G}^{(m,l)}$ and $\boldsymbol{H}^{(m,l)}$ we construct are all statistically independent of the m-th row and the l-th column of $\boldsymbol{E}$, and we shall also exploit the proximity of $\boldsymbol{U}^{(m,l)}$, $\boldsymbol{U}^{(m)}$ and $\boldsymbol{U}$ in the subsequent analysis.

Useful properties concerning these auxiliary estimates. In the sequel, we collect a couple of useful lemmas that are concerned with the leave-one-out and leave-two-out estimates, several of which are adapted from Cai et al. (2021). The first lemma controls the difference between the leave-one-out gram matrix $\boldsymbol{G}^{(m)}$ and the original gram matrix $\boldsymbol{G}$, as well as the difference between $\boldsymbol{G}^{(m)}$ and the ground-truth gram matrix.

Lemma 7. *Suppose that $n_1 \gtrsim \kappa^{\natural 4}\mu^\natural r$. Then with probability exceeding $1 - O(n^{-10})$,*

$$\left\|\boldsymbol{G}^{(m)} - \boldsymbol{G}\right\| \lesssim \sigma^2\sqrt{n_1 n_2}\log n + \kappa^{\natural 2}\sigma\sigma_1^\natural\sqrt{\mu^\natural r \log n}$$
$$\left\|\boldsymbol{G}^{(m)} - \boldsymbol{G}^\natural\right\| \lesssim \zeta_{\mathsf{op}} \asymp \sigma^2\sqrt{n_1 n_2}\log n + \sigma\sigma_1^\natural\sqrt{n_1 \log n}$$

hold simultaneously for all $1 \leq m \leq n_1$.

Proof. See Appendix C.3.8. ☐

The next lemma confirms that the leave-one-out estimate $\boldsymbol{U}^{(m)}$ and the leave-two-out estimate $\boldsymbol{U}^{(m,l)}$ are exceedingly close.

Lemma 8. *Suppose that the assumptions of Theorem 5 hold. Then with probability exceeding $1 - O(n^{-10})$, we have*

$$\begin{aligned}\left\|\boldsymbol{U}^{(m)}\boldsymbol{U}^{(m)\top} - \boldsymbol{U}^{(m,l)}\boldsymbol{U}^{(m,l)\top}\right\| &\lesssim \frac{1}{\sigma_r^{\natural 2}}\left(B\log n + \sigma\sqrt{n_1 \log n}\right)^2 \left\|\boldsymbol{U}^{(m)}\boldsymbol{H}^{(m)}\right\|_{2,\infty} + \frac{\sigma^2}{\sigma_r^{\natural 2}} \\ &\quad + \frac{1}{\sigma_r^{\natural 2}}\left(B\log n + \sigma\sqrt{n_1 \log n}\right)\left\|\boldsymbol{M}^{\natural\top}\right\|_{2,\infty}\end{aligned}$$

simultaneously for all $1 \leq m \leq n_1$ and $1 \leq l \leq n_2$.

Proof. See Appendix C.3.9. ☐

The following two lemmas study the size of $\boldsymbol{U}^{(m)}\boldsymbol{H}^{(m)} - \boldsymbol{U}^\natural$ and $\boldsymbol{U}^{(m)}\boldsymbol{H}^{(m)}$ when projected towards several important directions.

Lemma 9. *Suppose that the assumptions of Theorem 5 hold. Then with probability exceeding* $1 - O(n^{-10})$, *we have, for all* $l \in [n_2]$ *and all* $m \in [n_1]$,

$$\begin{aligned}\left\| \boldsymbol{e}_l^\top \left[\mathcal{P}_{-m,\cdot}\left(\boldsymbol{M}\right)\right]^\top \left(\boldsymbol{U}^{(m)}\boldsymbol{H}^{(m)} - \boldsymbol{U}^\natural\right)\right\|_2 &\lesssim \frac{\zeta_{\mathsf{op}}^2}{\sigma_r^{\natural 4}} \left\|\boldsymbol{M}^{\natural\top}\right\|_{2,\infty} + \left(B\log n + \sigma\sqrt{n_1 \log n}\right)\left\|\boldsymbol{U}^{(m)}\boldsymbol{H}^{(m)} - \boldsymbol{U}^\natural\right\|_{2,\infty} \\ &+ \left(\left\|\boldsymbol{M}^{\natural\top}\right\|_{2,\infty} + B\log n + \sigma\sqrt{n_1\log n}\right)\left\|\boldsymbol{U}^{(m)}\boldsymbol{U}^{(m)\top} - \boldsymbol{U}^{(m,l)}\boldsymbol{U}^{(m,l)\top}\right\|; \qquad \text{(C.16)}\end{aligned}$$

$$\left\|\left[\mathcal{P}_{-m,\cdot}\left(\boldsymbol{M}\right)\right]^\top \boldsymbol{U}^{(m)}\boldsymbol{H}^{(m)}\right\|_{2,\infty} \lesssim \left(B\log n + \sigma\sqrt{n_1\log n}\right)\left\|\boldsymbol{U}^{(m)}\boldsymbol{H}^{(m)}\right\|_{2,\infty} + \sigma_1^\natural \sqrt{\frac{\mu^\natural r}{n_2}}. \qquad \text{(C.17)}$$

Proof. See Appendix C.3.10. □

Lemma 10. *Suppose that the assumptions of Theorem 5 hold. Then with probability exceeding* $1 - O(n^{-10})$, *we have*

$$\begin{aligned}\left\|\boldsymbol{E}_{m,\cdot}\left[\mathcal{P}_{-m,\cdot}\left(\boldsymbol{M}\right)\right]^\top \left(\boldsymbol{U}^{(m)}\boldsymbol{H}^{(m)} - \boldsymbol{U}^\natural\right)\right\|_2 &\lesssim \frac{\zeta_{\mathsf{op}}}{\sigma_r^{\natural 2}}\sigma_m \sigma_1^\natural \sqrt{n_1\log n}\sqrt{\frac{\mu^\natural r}{n_1}} \\ &+ \zeta_{\mathsf{op},m}\left(\left\|\boldsymbol{U}^{(m)}\boldsymbol{H}^{(m)} - \boldsymbol{U}^\natural\right\|_{2,\infty} + \frac{\zeta_{\mathsf{op}}}{\sigma_r^{\natural 2}}\left\|\boldsymbol{U}^{(m)}\boldsymbol{H}^{(m)}\right\|_{2,\infty}\right); \qquad \text{(C.18)}\end{aligned}$$

$$\left\|\boldsymbol{E}_{m,\cdot}\left[\mathcal{P}_{-m,\cdot}\left(\boldsymbol{M}\right)\right]^\top \boldsymbol{U}^{(m)}\boldsymbol{H}^{(m)}\right\|_2 \lesssim \zeta_{\mathsf{op}}\left(\left\|\boldsymbol{U}^{(m)}\boldsymbol{H}^{(m)}\right\|_{2,\infty} + \sqrt{\frac{\mu^\natural r}{n_1}}\right) \qquad \text{(C.19)}$$

simultaneously for all $1 \le m \le n_1$.

Proof. See Appendix C.3.11. □

Finally, we justify the proximity of the leave-one-out estimate $\boldsymbol{U}^{(m)}$ and the original estimate $\boldsymbol{U}$, which turns out to be a consequence of the preceding results.

Lemma 11. *Suppose that the assumptions of Theorem 5 hold. Then with probability exceeding* $1 - O(n^{-10})$,

$$\left\|\boldsymbol{U}^{(m)}\boldsymbol{U}^{(m)\top} - \boldsymbol{U}\boldsymbol{U}^\top\right\| \lesssim \kappa^{\natural 2}\frac{\zeta_{\mathsf{op}}}{\sigma_r^{\natural 2}}\left(\left\|\boldsymbol{U}\boldsymbol{H}\right\|_{2,\infty} + \sqrt{\frac{\mu^\natural r}{n_1}}\right)$$

holds simultaneously for all $1 \le m \le n_1$.

Proof. See Appendix C.3.12. □

C.3.2 Proof of Lemma 1

Recall that the initialization $\boldsymbol{G}^0$ is obtained by dropping all diagonal entries of $\boldsymbol{M}\boldsymbol{M}^\top$, namely,

$$\boldsymbol{G}^0 = \mathcal{P}_{\mathsf{off\text{-}diag}}\left(\boldsymbol{M}\boldsymbol{M}^\top\right),$$

and $\boldsymbol{U}^0$ consists of the top-r eigenvectors of $\boldsymbol{G}^0$. These are precisely the subjects studied in Cai et al. (2021). In light of this, we state below the following two lemmas borrowed from Cai et al. (2021), which assist in proving Lemma 1.

Lemma 12. *Suppose that the assumptions of Theorem 5 hold. With probability exceeding* $1 - O(n^{-10})$,

$$\left\|\boldsymbol{G}^0 - \mathcal{P}_{\mathsf{off\text{-}diag}}\left(\boldsymbol{G}^\natural\right)\right\| = \left\|\mathcal{P}_{\mathsf{off\text{-}diag}}\left(\boldsymbol{G}^0 - \boldsymbol{G}^\natural\right)\right\| \le C^{\mathsf{dd}}\zeta_{\mathsf{op}}$$
$$\left\|\boldsymbol{G}^0 - \boldsymbol{G}^\natural\right\| \le C^{\mathsf{dd}}\zeta_{\mathsf{op}} + \left\|\boldsymbol{M}^\natural\right\|_{2,\infty}^2$$

hold for some universal constant $C^{\mathsf{dd}} > 0$.

Proof. See Cai et al. (2021, Lemma 1). $\square$

Lemma 13. *Suppose that the assumptions of Theorem 5 hold. With probability exceeding* $1 - O(n^{-10})$,

$$\left\|\boldsymbol{U}^0\boldsymbol{R}^0 - \boldsymbol{U}^\natural\right\| \le C_{\mathsf{op}}^{\mathsf{dd}} \frac{\zeta_{\mathsf{op}} + \left\|\boldsymbol{M}^\natural\right\|_{2,\infty}^2}{\sigma_r^{\natural 2}}$$

$$\left\|\boldsymbol{U}^0\boldsymbol{R}^0 - \boldsymbol{U}^\natural\right\|_{2,\infty} \le C_{\infty}^{\mathsf{dd}} \frac{\kappa^{\natural 2}\left(\zeta_{\mathsf{op}} + \left\|\boldsymbol{M}^\natural\right\|_{2,\infty}^2\right)}{\sigma_r^{\natural 2}}\sqrt{\frac{\mu^\natural r}{n_1}}$$

hold for some universal constants $C_{\mathsf{op}}^{\mathsf{dd}}, C_{\infty}^{\mathsf{dd}} > 0$*. Here,* $\boldsymbol{R}_0$ *is defined to be the following rotation matrix*

$$\boldsymbol{R}^0 := \arg\min_{\boldsymbol{O}\in\mathcal{O}^{r\times r}} \left\|\boldsymbol{U}^0\boldsymbol{O} - \boldsymbol{U}^\natural\right\|_{\mathrm{F}}^2. \tag{C.20}$$

Proof. See Cai et al. (2021, Theorem 1). $\square$

Armed with the above lemmas, we are ready to show by induction that: for $1 \le s \le t_0$,

$$\left\|\mathcal{P}_{\mathsf{diag}}\left(\boldsymbol{G}^s - \boldsymbol{G}^\natural\right)\right\| \le C_0\kappa^{\natural 2}\sqrt{\frac{\mu^\natural r}{n_1}}\left\|\boldsymbol{G}^{s-1} - \boldsymbol{G}^\natural\right\| \tag{C.21}$$

for some sufficiently large constant $C_0 > 0$.

Base case: $s = 1$. Let us start with the base case by making the observation that

$$\begin{aligned}\left\|\mathcal{P}_{\mathsf{diag}}\left(\boldsymbol{G}^1 - \boldsymbol{G}^\natural\right)\right\| &= \left\|\mathcal{P}_{\mathsf{diag}}\left(\boldsymbol{U}^0\boldsymbol{\Lambda}^0\boldsymbol{U}^{0\top} - \boldsymbol{G}^\natural\right)\right\| = \left\|\mathcal{P}_{\mathsf{diag}}\left(\mathcal{P}_{\boldsymbol{U}^0}(\boldsymbol{G}^0) - \boldsymbol{G}^\natural\right)\right\| \\ &\le \underbrace{\left\|\mathcal{P}_{\mathsf{diag}}\left[\mathcal{P}_{\boldsymbol{U}^0}\left(\boldsymbol{G}^0 - \boldsymbol{G}^\natural\right)\right]\right\|}_{=:\alpha_1} + \underbrace{\left\|\mathcal{P}_{\mathsf{diag}}\left(\mathcal{P}_{\boldsymbol{U}_\perp^0}\boldsymbol{G}^\natural\right)\right\|}_{=:\alpha_2},\end{aligned}$$

where $\boldsymbol{U}_\perp^0 \in \mathbb{R}^{n_1\times(n_1-r)}$ is a matrix whose orthonormal columns span the orthogonal complement of the column space of $\boldsymbol{U}^0$.

- Regarding the first term α_1, we have seen from Lemma 13 that

$$\begin{aligned}\left\|\boldsymbol{U}^0\right\|_{2,\infty} &= \left\|\boldsymbol{U}^0\boldsymbol{R}^0\right\|_{2,\infty} \le \left\|\boldsymbol{U}^0\boldsymbol{R}^0 - \boldsymbol{U}^\natural\right\|_{2,\infty} + \left\|\boldsymbol{U}^\natural\right\|_{2,\infty} \le C_\infty^{\mathsf{dd}}\frac{\kappa^{\natural 2}\left(\zeta_{\mathsf{op}} + \left\|\boldsymbol{M}^\natural\right\|_{2,\infty}^2\right)}{\sigma_r^{\natural 2}}\sqrt{\frac{\mu^\natural r}{n_1}} + \sqrt{\frac{\mu^\natural r}{n_1}} \\ &\overset{\text{(i)}}{\le} C_\infty^{\mathsf{dd}}\kappa^{\natural 2}\left(\frac{\zeta_{\mathsf{op}}}{\sigma_r^{\natural 2}} + \frac{\mu^\natural r}{d}\right)\sqrt{\frac{\mu^\natural r}{n_1}} + \sqrt{\frac{\mu^\natural r}{n_1}} \overset{\text{(ii)}}{\le} \sqrt{\frac{4\mu^\natural r}{n_1}},\end{aligned} \tag{C.22}$$

  where (i) holds since

$$\left\|\boldsymbol{M}^\natural\right\|_{2,\infty} = \left\|\boldsymbol{U}^\natural\boldsymbol{\Sigma}^\natural\boldsymbol{V}^{\natural\top}\right\|_{2,\infty} \le \left\|\boldsymbol{U}^\natural\right\|_{2,\infty}\left\|\boldsymbol{\Sigma}^\natural\right\|\left\|\boldsymbol{V}^\natural\right\| \le \sqrt{\frac{\mu^\natural r}{n_1}}\sigma_1^\natural, \tag{C.23}$$

  and (ii) holds provided that $\zeta_{\mathsf{op}}/\sigma_r^{\natural 2} \ll 1/\kappa^{\natural 2}$ and $d \gg \kappa^{\natural 2}\mu^\natural r$. This in turn allows us to use Zhang et al. (2022, Lemma 1) to reach

$$\alpha_1 \le 2\sqrt{\frac{\mu^\natural r}{n_1}}\left\|\boldsymbol{G}^0 - \boldsymbol{G}^\natural\right\|. \tag{C.24}$$

- Regarding the second term α_2, invoke Zhang et al. (2022, Lemma 1) again to arrive at

$$\alpha_2 = \left\|\mathcal{P}_{\mathsf{diag}}\left(\mathcal{P}_{\boldsymbol{U}_\perp^0}\boldsymbol{G}^\natural\mathcal{P}_{\boldsymbol{U}^\natural}\right)\right\| \le \sqrt{\frac{\mu^\natural r}{n_1}}\left\|\mathcal{P}_{\boldsymbol{U}_\perp^0}\boldsymbol{G}^\natural\right\|, \tag{C.25}$$

which has made use of the fact that $\boldsymbol{G}^{\natural}\mathcal{P}_{\boldsymbol{U}^{\natural}} = \boldsymbol{G}^{\natural}$. Moreover, it is seen that

$$\left\|\mathcal{P}_{\boldsymbol{U}_{\perp}^{0}}\boldsymbol{G}^{\natural}\right\| = \left\|\boldsymbol{U}_{\perp}^{0}\boldsymbol{U}_{\perp}^{0\top}\boldsymbol{U}^{\natural}\boldsymbol{\Sigma}^{\natural 2}\boldsymbol{U}^{\natural\top}\right\| \leq \sigma_1^{\natural 2}\left\|\boldsymbol{U}_{\perp}^{0\top}\boldsymbol{U}^{\natural}\right\| = \sigma_1^{\natural 2}\left\|\sin\boldsymbol{\Theta}\left(\boldsymbol{U}^{0},\boldsymbol{U}^{\natural}\right)\right\|, \tag{C.26}$$

where $\boldsymbol{\Theta}\left(\boldsymbol{U}^{0},\boldsymbol{U}^{\natural}\right)$ is a diagonal matrix whose diagonal entries correspond to the principal angles between $\boldsymbol{U}^{0}$ and $\boldsymbol{U}^{\natural}$, and the last identity follows from Chen et al. (2021, Lemma 2.1.2). In view of the Davis-Kahan $\sin\boldsymbol{\Theta}$ Theorem (Chen et al., 2021, Theorem 2.2.1), we can demonstrate that

$$\left\|\sin\boldsymbol{\Theta}\left(\boldsymbol{U}^{0},\boldsymbol{U}^{\natural}\right)\right\| \leq \frac{\sqrt{2}\left\|\boldsymbol{G}^{0}-\boldsymbol{G}^{\natural}\right\|}{\lambda_r\left(\boldsymbol{G}^{0}\right)-\lambda_{r+1}\left(\boldsymbol{G}^{\natural}\right)} = \frac{\sqrt{2}\left\|\boldsymbol{G}^{0}-\boldsymbol{G}^{\natural}\right\|}{\lambda_r\left(\boldsymbol{G}^{0}\right)} \leq \frac{2\sqrt{2}\left\|\boldsymbol{G}^{0}-\boldsymbol{G}^{\natural}\right\|}{\sigma_r^{\natural 2}}. \tag{C.27}$$

Here, the identity comes from the fact $\lambda_{r+1}(\boldsymbol{G}^{\natural}) = 0$; the last inequality follows from a direct application of Weyl's inequality:

$$\lambda_r\left(\boldsymbol{G}^{0}\right) \geq \lambda_r\left(\boldsymbol{G}^{\natural}\right) - \left\|\boldsymbol{G}^{0}-\boldsymbol{G}^{\natural}\right\| \overset{\text{(i)}}{\geq} \sigma_r^{\natural 2} - \zeta_{\mathsf{op}} - \left\|\boldsymbol{M}^{\natural}\right\|_{2,\infty}^{2} \overset{\text{(ii)}}{\geq} \sigma_r^{\natural 2} - \zeta_{\mathsf{op}} - \frac{\mu^{\natural}r}{n_1}\sigma_1^{\natural 2} \overset{\text{(iii)}}{\geq} \frac{1}{2}\sigma_r^{\natural 2},$$

where (i) is a consequence of Lemma 12, (ii) follows from (C.23), and (iii) holds as long as $\zeta_{\mathsf{op}} \ll \sigma_r^{\natural 2}$ and $n_1 \gg \kappa^{\natural 2}\mu^{\natural}r$. Combine (C.25), (C.26) and (C.27) to reach

$$\alpha_2 \leq 2\sqrt{2}\kappa^{\natural 2}\sqrt{\frac{\mu^{\natural}r}{n_1}}\left\|\boldsymbol{G}^{0}-\boldsymbol{G}^{\natural}\right\|.$$

Combine the above bounds on α_1 and α_2 to yield

$$\left\|\mathcal{P}_{\mathsf{diag}}\left(\boldsymbol{G}^{1}-\boldsymbol{G}^{\natural}\right)\right\| \leq \alpha_1+\alpha_2 \leq \left(2+2\sqrt{2}\kappa^{\natural 2}\right)\sqrt{\frac{\mu^{\natural}r}{n_1}}\left\|\boldsymbol{G}^{0}-\boldsymbol{G}^{\natural}\right\| \leq C_0\kappa^{\natural 2}\sqrt{\frac{\mu^{\natural}r}{n_1}}\left\|\boldsymbol{G}^{0}-\boldsymbol{G}^{\natural}\right\|,$$

with the proviso that the constant C_0 is sufficiently large.

Induction step. For any given $t > 1$, Suppose that (C.21) holds for all $s = 1, 2\ldots, t$, and we'd like to show that it continues to hold for $s = t+1$. From the induction hypothesis, we know that for any $1 \leq \tau \leq t$, one has

$$\begin{aligned}\left\|\mathcal{P}_{\mathsf{diag}}\left(\boldsymbol{G}^{\tau}-\boldsymbol{G}^{\natural}\right)\right\| &\leq C_0\kappa^{\natural 2}\sqrt{\frac{\mu^{\natural}r}{n_1}}\left\|\boldsymbol{G}^{\tau-1}-\boldsymbol{G}^{\natural}\right\| \leq C_0\kappa^{\natural 2}\sqrt{\frac{\mu^{\natural}r}{n_1}}\left(\left\|\mathcal{P}_{\mathsf{diag}}\left(\boldsymbol{G}^{\tau-1}-\boldsymbol{G}^{\natural}\right)\right\|+\left\|\mathcal{P}_{\mathsf{off\text{-}diag}}\left(\boldsymbol{G}^{\tau-1}-\boldsymbol{G}^{\natural}\right)\right\|\right)\\ &= C_0\kappa^{\natural 2}\sqrt{\frac{\mu^{\natural}r}{n_1}}\left(\left\|\mathcal{P}_{\mathsf{diag}}\left(\boldsymbol{G}^{\tau-1}-\boldsymbol{G}^{\natural}\right)\right\|+\left\|\mathcal{P}_{\mathsf{off\text{-}diag}}\left(\boldsymbol{G}^{0}-\boldsymbol{G}^{\natural}\right)\right\|\right),\end{aligned} \tag{C.28}$$

where the last line holds since, by construction, $\mathcal{P}_{\mathsf{off\text{-}diag}}\left(\boldsymbol{G}^{\tau}\right) = \mathcal{P}_{\mathsf{off\text{-}diag}}\left(\boldsymbol{G}^{0}\right)$ for all τ. Applying the above inequality recursively gives

$$\begin{aligned}\left\|\mathcal{P}_{\mathsf{diag}}\left(\boldsymbol{G}^{t}-\boldsymbol{G}^{\natural}\right)\right\| &\leq C_0\kappa^{\natural 2}\sqrt{\frac{\mu^{\natural}r}{n_1}}\left(\left\|\mathcal{P}_{\mathsf{diag}}\left(\boldsymbol{G}^{t-1}-\boldsymbol{G}^{\natural}\right)\right\|+\left\|\mathcal{P}_{\mathsf{off\text{-}diag}}\left(\boldsymbol{G}^{0}-\boldsymbol{G}^{\natural}\right)\right\|\right) \leq \cdots\\ &\leq \left(C_0\kappa^{\natural 2}\sqrt{\frac{\mu^{\natural}r}{n_1}}\right)^{t}\left\|\mathcal{P}_{\mathsf{diag}}\left(\boldsymbol{G}^{0}-\boldsymbol{G}^{\natural}\right)\right\| + \sum_{i=1}^{t}\left(C_0\kappa^{\natural 2}\sqrt{\frac{\mu^{\natural}r}{n_1}}\right)^{i}\left\|\mathcal{P}_{\mathsf{off\text{-}diag}}\left(\boldsymbol{G}^{0}-\boldsymbol{G}^{\natural}\right)\right\|\\ &\leq \left(C_0\kappa^{\natural 2}\sqrt{\frac{\mu^{\natural}r}{n_1}}\right)^{t}\left\|\boldsymbol{M}^{\natural}\right\|_{2,\infty}^{2} + 2C_0\kappa^{\natural 2}\sqrt{\frac{\mu^{\natural}r}{n_1}}\left\|\mathcal{P}_{\mathsf{off\text{-}diag}}\left(\boldsymbol{G}^{0}-\boldsymbol{G}^{\natural}\right)\right\|\\ &\leq \left(C_0\kappa^{\natural 2}\sqrt{\frac{\mu^{\natural}r}{n_1}}\right)^{t}\frac{\mu^{\natural}r}{n_1}\sigma_1^{\natural 2} + 2C_0C^{\mathsf{dd}}\kappa^{\natural 2}\sqrt{\frac{\mu^{\natural}r}{n_1}}\zeta_{\mathsf{op}}.\end{aligned} \tag{C.29}$$

Here, the penultimate line holds as long as $n_1 \gg \kappa^{\natural 4}\mu^\natural r$, and the last line follows from (C.23) and Lemma 12. An immediate consequence is that

$$\begin{aligned}\left\|\mathcal{P}_{\mathsf{diag}}\left(\boldsymbol{G}^t\right)\right\| &\le \left\|\mathcal{P}_{\mathsf{diag}}\left(\boldsymbol{G}^\natural\right)\right\| + \left\|\mathcal{P}_{\mathsf{diag}}\left(\boldsymbol{G}^t-\boldsymbol{G}^\natural\right)\right\| = \left\|\boldsymbol{M}^\natural\right\|_{2,\infty}^2 + \left\|\mathcal{P}_{\mathsf{diag}}\left(\boldsymbol{G}^t-\boldsymbol{G}^\natural\right)\right\| \\ &\le \frac{\mu^\natural r}{n_1}\sigma_1^{\natural 2} + \left(C_0\kappa^{\natural 2}\sqrt{\frac{\mu^\natural r}{n_1}}\right)^t \frac{\mu^\natural r}{n_1}\sigma_1^{\natural 2} + 2C_0C^{\mathsf{dd}}\kappa^{\natural 2}\sqrt{\frac{\mu^\natural r}{n_1}}\zeta_{\mathsf{op}} \\ &\le \frac{2\mu^\natural r}{n_1}\sigma_1^{\natural 2} + 2C_0C^{\mathsf{dd}}\kappa^{\natural 2}\sqrt{\frac{\mu^\natural r}{n_1}}\zeta_{\mathsf{op}},\end{aligned}$$

where the second line follows from (C.23) and (C.29), and the last relation is valid as long as $n_1 \gg \kappa^{\natural 4}\mu^\natural r$. This together with the fact $\mathcal{P}_{\mathsf{off\text{-}diag}}\left(\boldsymbol{G}^t\right) = \mathcal{P}_{\mathsf{off\text{-}diag}}\left(\boldsymbol{G}^0\right)$ (by construction) allows one to obtain

$$\left\|\boldsymbol{G}^t-\boldsymbol{G}^0\right\| = \left\|\mathcal{P}_{\mathsf{diag}}\left(\boldsymbol{G}^t\right)\right\| \le \frac{2\mu^\natural r}{n_1}\sigma_1^{\natural 2} + 2C_0C^{\mathsf{dd}}\kappa^{\natural 2}\sqrt{\frac{\mu^\natural r}{n_1}}\zeta_{\mathsf{op}}. \tag{C.30}$$

In view of Weyl's inequality, we have

$$\begin{aligned}\lambda_r\left(\boldsymbol{G}^t\right) &\ge \lambda_r\left(\boldsymbol{G}^\natural\right) - \left\|\boldsymbol{G}^t-\boldsymbol{G}^\natural\right\| \ge \lambda_r\left(\boldsymbol{G}^\natural\right) - \left\|\boldsymbol{G}^t-\boldsymbol{G}^0\right\| - \left\|\boldsymbol{G}^\natural-\boldsymbol{G}^0\right\| \\ &\overset{\text{(i)}}{\ge} \sigma_r^{\natural 2} - \frac{2\mu^\natural r}{n_1}\sigma_1^{\natural 2} - 2C_0C^{\mathsf{dd}}\kappa^{\natural 2}\sqrt{\frac{\mu^\natural r}{n_1}}\zeta_{\mathsf{op}} - C^{\mathsf{dd}}\zeta_{\mathsf{op}} - \left\|\boldsymbol{M}^\natural\right\|_{2,\infty}^2 \\ &\overset{\text{(ii)}}{\ge} \sigma_r^{\natural 2} - \frac{3\mu^\natural r}{n_1}\sigma_1^{\natural 2} - 2C_0C^{\mathsf{dd}}\kappa^{\natural 2}\sqrt{\frac{\mu^\natural r}{n_1}}\zeta_{\mathsf{op}} - C^{\mathsf{dd}}\zeta_{\mathsf{op}} \overset{\text{(iii)}}{\ge} \frac{\sigma_r^{\natural 2}}{2}.\end{aligned} \tag{C.31}$$

Here, (i) follows from (C.30) and Lemma 12, (ii) comes from (C.23), and (iii) is guaranteed to hold as long as $n_1 \gg \kappa^{\natural 4}\mu^\natural r$ and $\zeta_{\mathsf{op}} \ll \sigma_r^{\natural 2}$. Then by virtue of Davis Kahan's $\sin\boldsymbol{\Theta}$ Theorem (Chen et al., 2021, Theorem 2.2.1),

$$\left\|\boldsymbol{U}^t\boldsymbol{R}^t-\boldsymbol{U}^0\right\| \le \frac{\left\|\boldsymbol{G}^t-\boldsymbol{G}^0\right\|}{\lambda_r\left(\boldsymbol{G}^t\right)-\lambda_{r+1}\left(\boldsymbol{G}^\natural\right)} \le \frac{4\kappa^{\natural 2}\mu^\natural r}{n_1} + 4C_0C^{\mathsf{dd}}\kappa^{\natural 2}\sqrt{\frac{\mu^\natural r}{n_1}}\frac{\zeta_{\mathsf{op}}}{\sigma_r^{\natural 2}},$$

where the last relation arises from $\lambda_{r+1}(\boldsymbol{G}^\natural) = 0$, (C.30) and (C.31). This indicates that

$$\begin{aligned}\left\|\boldsymbol{U}^t\right\|_{2,\infty} &\le \left\|\boldsymbol{U}^0\right\|_{2,\infty} + \left\|\boldsymbol{U}^t\boldsymbol{R}^t-\boldsymbol{U}^0\right\|_{2,\infty} \\ &\le \sqrt{\frac{4\mu^\natural r}{n_1}} + \frac{4\kappa^{\natural 2}\mu^\natural r}{n_1} + 4C_0C^{\mathsf{dd}}\kappa^{\natural 2}\sqrt{\frac{\mu^\natural r}{n_1}}\frac{\zeta_{\mathsf{op}}}{\sigma_r^{\natural 2}} \\ &\le 3\sqrt{\frac{\mu^\natural r}{n_1}},\end{aligned} \tag{C.32}$$

where the penultimate relation follows from (C.22), and the last relation holds provided that $n_1 \gg \kappa^{\natural 4}\mu^\natural r$ and $\zeta_{\mathsf{op}}/\sigma_r^{\natural 2} \ll 1/\kappa^{\natural 2}$.

To proceed, we recall that the diagonal entries of $\boldsymbol{G}^{t+1}$ are set to be the diagonal entries of $\boldsymbol{U}^t\boldsymbol{\Lambda}^t\boldsymbol{U}^{t\top}$, thus revealing that

$$\begin{aligned}\left\|\mathcal{P}_{\mathsf{diag}}\left(\boldsymbol{G}^{t+1}-\boldsymbol{G}^\natural\right)\right\| &= \left\|\mathcal{P}_{\mathsf{diag}}\left(\boldsymbol{U}^t\boldsymbol{\Lambda}^t\boldsymbol{U}^{t\top}-\boldsymbol{G}^\natural\right)\right\| = \left\|\mathcal{P}_{\mathsf{diag}}\left(\mathcal{P}_{\boldsymbol{U}^t}\left(\boldsymbol{G}^t\right)-\boldsymbol{G}^\natural\right)\right\| \\ &\le \underbrace{\left\|\mathcal{P}_{\mathsf{diag}}\left[\mathcal{P}_{\boldsymbol{U}^t}\left(\boldsymbol{G}^t-\boldsymbol{G}^\natural\right)\right]\right\|}_{=:\beta_1} + \underbrace{\left\|\mathcal{P}_{\mathsf{diag}}\left(\mathcal{P}_{\boldsymbol{U}_\perp^t}\boldsymbol{G}^\natural\right)\right\|}_{=:\beta_2}.\end{aligned}$$

Here, $\boldsymbol{U}_{\perp}^{t}$ represents the orthogonal complement of the subspace $\boldsymbol{U}^{t}$. Similar to how we bound α_1 and α_2 for the base case, we can invoke Zhang et al. (2022, Lemma 1) and (C.32) to reach

$$\beta_1 \leq 3\sqrt{\frac{\mu^{\natural} r}{n_1}} \left\|\boldsymbol{G}^{t} - \boldsymbol{G}^{\natural}\right\|, \tag{C.33}$$

$$\beta_2 \leq \sqrt{\frac{\mu^{\natural} r}{n_1}} \left\|\mathcal{P}_{\boldsymbol{U}_{\perp}^{t}} \boldsymbol{G}^{\natural}\right\| \leq \sigma_1^{\natural 2} \left\|\left(\boldsymbol{U}_{\perp}^{t}\right)^{\top} \boldsymbol{U}^{\natural}\right\| = \sigma_1^{\natural 2} \left\|\sin \boldsymbol{\Theta}\left(\boldsymbol{U}^{t}, \boldsymbol{U}^{\natural}\right)\right\|, \tag{C.34}$$

where $\boldsymbol{\Theta}\left(\boldsymbol{U}^{t}, \boldsymbol{U}^{\natural}\right)$ is a diagonal matrix whose diagonal entries are the principal angles between $\boldsymbol{U}^{t}$ and $\boldsymbol{U}^{\natural}$. Apply the Davis Kahan $\sin \boldsymbol{\Theta}$ Theorem (Chen et al., 2021, Theorem 2.2.1) to obtain

$$\left\|\sin \boldsymbol{\Theta}\left(\boldsymbol{U}^{t}, \boldsymbol{U}^{\natural}\right)\right\| \leq \frac{\left\|\boldsymbol{G}^{t} - \boldsymbol{G}^{\natural}\right\|}{\lambda_r\left(\boldsymbol{G}^{t}\right) - \lambda_{r+1}\left(\boldsymbol{G}^{\natural}\right)} \leq \frac{2\left\|\boldsymbol{G}^{t} - \boldsymbol{G}^{\natural}\right\|}{\sigma_r^{\natural 2}}. \tag{C.35}$$

Here, the last inequality comes from $\lambda_{r+1}(\boldsymbol{G}^{\natural}) = 0$ and (C.31). Combine (C.34) and (C.35) to reach

$$\beta_2 \leq \sqrt{\frac{\mu^{\natural} r}{n_1}} \sigma_1^{\natural 2} \frac{2\left\|\boldsymbol{G}^{t} - \boldsymbol{G}^{\natural}\right\|}{\sigma_r^{\natural 2}} \leq 2\kappa^{\natural 2} \sqrt{\frac{\mu^{\natural} r}{n_1}} \left\|\boldsymbol{G}^{t} - \boldsymbol{G}^{\natural}\right\|.$$

Taking together the preceding bounds on β_1 and β_2, we arrive at

$$\left\|\mathcal{P}_{\mathsf{diag}}\left(\boldsymbol{G}^{t+1} - \boldsymbol{G}^{\natural}\right)\right\| \leq \beta_1 + \beta_2 \leq \left(3 + 2\kappa^{\natural 2}\right) \sqrt{\frac{\mu^{\natural} r}{n_1}} \left\|\boldsymbol{G}^{t} - \boldsymbol{G}^{\natural}\right\| \leq C_0 \kappa^{\natural 2} \sqrt{\frac{\mu^{\natural} r}{n_1}} \left\|\boldsymbol{G}^{t} - \boldsymbol{G}^{\natural}\right\|,$$

provided that the constant C_0 is large enough.

Invoking the inequality (C.21) to establish the lemma. The above induction steps taken together establish the hypothesis (C.21), namely, for all $t \geq 1$,

$$\left\|\mathcal{P}_{\mathsf{diag}}\left(\boldsymbol{G}^{t} - \boldsymbol{G}^{\natural}\right)\right\| \leq C_0 \kappa^{\natural 2} \sqrt{\frac{\mu^{\natural} r}{n_1}} \left\|\boldsymbol{G}^{t-1} - \boldsymbol{G}^{\natural}\right\|.$$

Follow the same procedure used to derive (C.29), we can show that

$$\left\|\mathcal{P}_{\mathsf{diag}}\left(\boldsymbol{G}^{t} - \boldsymbol{G}^{\natural}\right)\right\| \leq \left(C_0 \kappa^{\natural 2} \sqrt{\frac{\mu^{\natural} r}{n_1}}\right)^{t} \frac{\mu^{\natural} r}{n_1} \sigma_1^{\natural 2} + 2 C_0 C^{\mathsf{dd}} \kappa^{\natural 2} \sqrt{\frac{\mu^{\natural} r}{n_1}} \zeta_{\mathsf{op}}.$$

If the number of iterations t_0 satisfies

$$t_0 \geq \frac{\log\left(\left(C^{\mathsf{dd}}\right)^{-1} (\kappa^{\natural})^{-2} \sqrt{\mu^{\natural} r / n_1} \sigma_1^{\natural 2} / \zeta_{\mathsf{op}}\right)}{\log\left((\kappa^{\natural})^{-2} \sqrt{n_1 / \left(\mu^{\natural} r\right)}\right)},$$

we can guarantee that

$$\left\|\mathcal{P}_{\mathsf{diag}}\left(\boldsymbol{G}^{t_0} - \boldsymbol{G}^{\natural}\right)\right\| \leq 3 C_0 C^{\mathsf{dd}} \kappa^{\natural 2} \sqrt{\frac{\mu^{\natural} r}{n_1}} \zeta_{\mathsf{op}}$$

as long as $n_1 \gg \kappa^{\natural 4} \mu^{\natural} r$. In addition, note that when $n_1 \gg \kappa^{\natural 4} \mu^{\natural} r$, one has

$$\frac{\log\left(\left(C^{\mathsf{dd}}\right)^{-1} (\kappa^{\natural})^{-2} \sqrt{\mu^{\natural} r / n_1} \sigma_1^{\natural 2} / \zeta_{\mathsf{op}}\right)}{\log\left((\kappa^{\natural})^{-2} \sqrt{n_1 / \left(\mu^{\natural} r\right)}\right)} \leq \log\left(\frac{\sigma_1^{\natural 2}}{\zeta_{\mathsf{op}}}\right).$$

Therefore, it suffices to take $t_0 \geq \log\left(\frac{\sigma_1^{\natural 2}}{\zeta_{\mathsf{op}}}\right)$ as claimed.

To finish up, we observe that

$$\begin{aligned}\left\|\boldsymbol{G}^{t_0}-\boldsymbol{G}^{\natural}\right\| &\leq \left\|\mathcal{P}_{\mathsf{diag}}\left(\boldsymbol{G}^{t_0}-\boldsymbol{G}^{\natural}\right)\right\|+\left\|\mathcal{P}_{\mathsf{off\text{-}diag}}\left(\boldsymbol{G}^{t_0}-\boldsymbol{G}^{\natural}\right)\right\| \\ &\overset{\text{(i)}}{=} \left\|\mathcal{P}_{\mathsf{diag}}\left(\boldsymbol{G}^{t_0}-\boldsymbol{G}^{\natural}\right)\right\|+\left\|\boldsymbol{G}^{0}-\mathcal{P}_{\mathsf{off\text{-}diag}}\left(\boldsymbol{G}^{\natural}\right)\right\| \\ &\overset{\text{(ii)}}{\lesssim} \kappa^{\natural 2}\sqrt{\frac{\mu^{\natural} r}{n_1}}\zeta_{\mathsf{op}}+\zeta_{\mathsf{op}} \overset{\text{(iii)}}{\lesssim} \zeta_{\mathsf{op}}.\end{aligned}$$

Here, (i) follows from the construction $\boldsymbol{G}^0=\mathcal{P}_{\mathsf{off\text{-}diag}}\left(\boldsymbol{G}^0\right)=\mathcal{P}_{\mathsf{off\text{-}diag}}\left(\boldsymbol{G}^{t_0}\right)$; (ii) follows from Lemma 12; and (iii) holds provided that $n_1 \gtrsim \kappa^{\natural 4}\mu^{\natural} r$.

Last but not least, for each $m\in[n_1]$ and any given $t>1$, we have

$$\begin{aligned}\left|G_{m,m}^{t}-G_{m,m}^{\natural}\right| &= \left|\boldsymbol{e}_m^{\top}\left(\boldsymbol{G}^{t}-\boldsymbol{G}^{\natural}\right)\boldsymbol{e}_m\right| \leq \left|\boldsymbol{e}_m^{\top}\mathcal{P}_{\boldsymbol{U}^{t_0}}\left(\boldsymbol{G}^{t}-\boldsymbol{G}^{\natural}\right)\boldsymbol{e}_m\right|+\left|\boldsymbol{e}_m^{\top}\mathcal{P}_{\boldsymbol{U}_{\perp}^{t}}\left(\boldsymbol{G}^{t}-\boldsymbol{G}^{\natural}\right)\boldsymbol{e}_m\right| \\ &= \underbrace{\left|\boldsymbol{e}_m^{\top}\mathcal{P}_{\boldsymbol{U}^{t}}\left(\boldsymbol{G}^{t}-\boldsymbol{G}^{\natural}\right)\boldsymbol{e}_m\right|}_{=:\gamma_1}+\underbrace{\left|\boldsymbol{e}_m^{\top}\mathcal{P}_{\boldsymbol{U}_{\perp}^{t}}\boldsymbol{G}^{\natural}\boldsymbol{e}_m\right|}_{=:\gamma_2}.\end{aligned}$$

The first term γ_1 can be upper bounded by

$$\begin{aligned}\gamma_1 &= \left|\boldsymbol{e}_m^{\top}\boldsymbol{U}^{t}\boldsymbol{U}^{t\top}\left(\boldsymbol{G}^{t}-\boldsymbol{G}^{\natural}\right)\boldsymbol{e}_m\right| = \left|\boldsymbol{U}_{m,\cdot}^{t}\boldsymbol{U}^{t\top}\left(\boldsymbol{G}^{t}-\boldsymbol{G}^{\natural}\right)_{\cdot,m}\right| \leq \left\|\boldsymbol{U}_{m,\cdot}^{t}\right\|_2\left\|\left(\boldsymbol{G}^{t}-\boldsymbol{G}^{\natural}\right)_{\cdot,m}\right\|_2 \\ &\leq \left\|\boldsymbol{U}_{m,\cdot}^{t}\right\|_2\left\|\left(\boldsymbol{G}^{t}-\boldsymbol{G}^{\natural}\right)\right\|_2.\end{aligned}$$

The second term γ_2 admits the following upper bound

$$\begin{aligned}\gamma_2 &= \left|\boldsymbol{e}_m^{\top}\boldsymbol{U}_{\perp}^{t}\boldsymbol{U}_{\perp}^{t\top}\boldsymbol{U}^{\natural}\boldsymbol{\Sigma}^{\natural 2}\boldsymbol{U}_{m,\cdot}^{\natural\top}\right| \leq \sigma_1^{\natural 2}\left\|\boldsymbol{U}_{\perp}^{t\top}\boldsymbol{U}^{\natural}\right\|\left\|\boldsymbol{U}_{m,\cdot}^{\natural}\right\|_2 \leq \sigma_1^{\natural 2}\left\|\sin\boldsymbol{\Theta}\left(\boldsymbol{U}^{t},\boldsymbol{U}^{\natural}\right)\right\|\left\|\boldsymbol{U}_{m,\cdot}^{\natural}\right\|_2 \\ &\leq 2\kappa^{\natural 2}\left\|\boldsymbol{G}^{t}-\boldsymbol{G}^{\natural}\right\|\left\|\boldsymbol{U}_{m,\cdot}^{\natural}\right\|_2,\end{aligned}$$

where the last relation follows from (C.35). Taking the bounds on γ_1 and γ_2 collectively and let $t=t_0$ gives

$$\left|G_{m,m}^{t_0}-G_{m,m}^{\natural}\right| \lesssim \kappa^{\natural 2}\left\|\boldsymbol{G}^{t}-\boldsymbol{G}^{\natural}\right\|\left(\left\|\boldsymbol{U}_{m,\cdot}^{\natural}\right\|_2+\left\|\boldsymbol{U}_{m,\cdot}^{t}\right\|_2\right) \lesssim \kappa^{\natural 2}\zeta_{\mathsf{op}}\left(\left\|\boldsymbol{U}_{m,\cdot}^{\natural}\right\|_2+\left\|\boldsymbol{U}_{m,\cdot}^{t}\right\|_2\right).$$

C.3.3 Proof of Lemma 2

We first apply Weyl's inequality to demonstrate that

$$\lambda_r\left(\boldsymbol{G}\right) \geq \sigma_r^{\natural 2}-\left\|\boldsymbol{G}-\boldsymbol{G}^{\natural}\right\| \geq \sigma_r^{\natural 2}-\widetilde{C}\zeta_{\mathsf{op}} \geq \frac{1}{2}\sigma_r^{\natural 2}$$

for some constant $\widetilde{C}>0$, where the penultimate step comes from Lemma 1, and the last step holds true provided that $\zeta_{\mathsf{op}} \ll \sigma_r^{\natural 2}$. In view of the Davis-Kahan $\sin\boldsymbol{\Theta}$ Theorem (Chen et al., 2021, Theorem 2.2.1), we obtain

$$\left\|\boldsymbol{U}\boldsymbol{U}^{\top}-\boldsymbol{U}^{\natural}\boldsymbol{U}^{\natural\top}\right\| \leq \frac{\sqrt{2}\left\|\boldsymbol{G}-\boldsymbol{G}^{\natural}\right\|}{\lambda_r\left(\boldsymbol{G}\right)-\lambda_{r+1}\left(\boldsymbol{G}^{\natural}\right)} \leq \frac{2\sqrt{2}\left\|\boldsymbol{G}-\boldsymbol{G}^{\natural}\right\|}{\sigma_r^{\natural 2}} \lesssim \frac{\zeta_{\mathsf{op}}}{\sigma_r^{\natural 2}}, \tag{C.36}$$

which immediately leads to the advertised bound on $\left\|\boldsymbol{U}\boldsymbol{H}-\boldsymbol{U}^{\natural}\right\|$ as follows

$$\left\|\boldsymbol{U}\boldsymbol{H}-\boldsymbol{U}^{\natural}\right\| = \left\|\boldsymbol{U}\boldsymbol{U}^{\top}\boldsymbol{U}^{\natural}-\boldsymbol{U}^{\natural}\boldsymbol{U}^{\natural\top}\boldsymbol{U}^{\natural}\right\| \leq \left\|\boldsymbol{U}\boldsymbol{U}^{\top}-\boldsymbol{U}^{\natural}\boldsymbol{U}^{\natural\top}\right\| \lesssim \frac{\zeta_{\mathsf{op}}}{\sigma_r^{\natural 2}}.$$

Next, we turn attention to $\left\|\boldsymbol{H}-\boldsymbol{R}_{\boldsymbol{U}}\right\|$. Given that both $\boldsymbol{U}$ and $\boldsymbol{U}^{\natural}$ have orthonormal columns, the SVD of $\boldsymbol{H}=\boldsymbol{U}^{\top}\boldsymbol{U}^{\natural}$ can be written as

$$\boldsymbol{H}=\boldsymbol{X}(\cos\boldsymbol{\Theta})\boldsymbol{Y}^{\top},$$

where $\boldsymbol{X}, \boldsymbol{Y} \in \mathbb{R}^{r\times r}$ are orthonormal matrices and $\boldsymbol{\Theta}$ is a diagonal matrix composed of the principal angles between $\boldsymbol{U}$ and $\boldsymbol{U}^{\natural}$ (see Chen et al. (2021, Section 2.1)). It is well known that one can write $\boldsymbol{R}_{\boldsymbol{U}} = \mathsf{sgn}(\boldsymbol{H}) = \boldsymbol{X}\boldsymbol{Y}^{\top}$, and therefore,

$$\begin{aligned}\|\boldsymbol{H}-\boldsymbol{R}_{\boldsymbol{U}}\| &= \left\|\boldsymbol{X}\left(\cos\boldsymbol{\Theta}-\boldsymbol{I}_r\right)\boldsymbol{Y}^{\top}\right\| = \|\boldsymbol{I}_r-\cos\boldsymbol{\Theta}\| \\ &= \left\|2\sin^2\left(\boldsymbol{\Theta}/2\right)\right\| \lesssim \|\sin\boldsymbol{\Theta}\|^2 \asymp \left\|\boldsymbol{U}\boldsymbol{U}^{\top}-\boldsymbol{U}^{\natural}\boldsymbol{U}^{\natural\top}\right\|^2 \lesssim \frac{\zeta_{\mathsf{op}}^2}{\sigma_r^{\natural 4}}.\end{aligned}$$

Here the penultimate relation comes from Chen et al. (2021, Lemma 2.1.2) and the last relation invokes (C.36). Given that $\boldsymbol{R}_{\boldsymbol{U}}$ is a square orthonormal matrix, we immediately have

$$\begin{aligned}\sigma_{\max}(\boldsymbol{H}_{\boldsymbol{U}}) &\leq \sigma_{\max}(\boldsymbol{R}_{\boldsymbol{U}}) + \|\boldsymbol{H}_{\boldsymbol{U}}-\boldsymbol{R}_{\boldsymbol{U}}\| \leq 1 + O\Big(\frac{\sigma^2 n}{\sigma_r^{\natural 2}}\Big) \leq 2, \\ \sigma_r(\boldsymbol{H}_{\boldsymbol{U}}) &\geq \sigma_{\max}(\boldsymbol{R}_{\boldsymbol{U}}) - \|\boldsymbol{H}_{\boldsymbol{U}}-\boldsymbol{R}_{\boldsymbol{U}}\| \geq 1 - O\Big(\frac{\sigma^2 n}{\sigma_r^{\natural 2}}\Big) \geq 1/2,\end{aligned}$$

provided that $\zeta_{\mathsf{op}} \ll \sigma_r^{\natural 2}$. In addition, we can similarly derive

$$\left\|\boldsymbol{H}^{\top}\boldsymbol{H}-\boldsymbol{I}_r\right\| = \left\|\boldsymbol{Y}\left[\cos^2\boldsymbol{\Theta}-\boldsymbol{I}_r\right]\boldsymbol{Y}^{\top}\right\| = \left\|\cos^2\boldsymbol{\Theta}-\boldsymbol{I}_r\right\| = \left\|\sin^2\boldsymbol{\Theta}\right\| = \|\sin\boldsymbol{\Theta}\|^2 \lesssim \frac{\zeta_{\mathsf{op}}^2}{\sigma_r^{\natural 4}}.$$

An immediate consequence of the proximity between $\boldsymbol{H}$ and $\boldsymbol{R}_{\boldsymbol{U}}$ is that

$$\begin{aligned}\left\|\boldsymbol{U}\boldsymbol{H}-\boldsymbol{U}^{\natural}\right\| &\leq \left\|\boldsymbol{U}\boldsymbol{R}_{\boldsymbol{U}}-\boldsymbol{U}^{\natural}\right\| + \left\|\boldsymbol{U}\left(\boldsymbol{H}-\boldsymbol{R}_{\boldsymbol{U}}\right)\right\| \leq \left\|\boldsymbol{U}\boldsymbol{R}_{\boldsymbol{U}}-\boldsymbol{U}^{\natural}\right\| + \|\boldsymbol{H}-\boldsymbol{R}_{\boldsymbol{U}}\| \\ &\lesssim \frac{\zeta_{\mathsf{op}}}{\sigma_r^{\natural 2}} + \frac{\zeta_{\mathsf{op}}^2}{\sigma_r^{\natural 4}} \lesssim \frac{\zeta_{\mathsf{op}}}{\sigma_r^{\natural 2}},\end{aligned}$$

provided that $\zeta_{\mathsf{op}} \ll \sigma_r^{\natural 2}$.

C.3.4 Proof of Lemma 3

We begin with the triangle inequality:

$$\begin{aligned}\left\|\left(\boldsymbol{G}-\boldsymbol{G}^{\natural}\right)_{m,\cdot}\boldsymbol{U}^{\natural}\right\|_2 &\leq \left\|\left[\mathcal{P}_{\mathsf{off\text{-}diag}}\left(\boldsymbol{G}-\boldsymbol{G}^{\natural}\right)\right]_{m,\cdot}\boldsymbol{U}^{\natural}\right\|_2 + \left\|\left[\mathcal{P}_{\mathsf{diag}}\left(\boldsymbol{G}-\boldsymbol{G}^{\natural}\right)\right]_{m,\cdot}\boldsymbol{U}^{\natural}\right\|_2 \\ &= \underbrace{\left\|\left[\mathcal{P}_{\mathsf{off\text{-}diag}}\left(\boldsymbol{G}^{0}-\boldsymbol{G}^{\natural}\right)\right]_{m,\cdot}\boldsymbol{U}^{\natural}\right\|_2}_{=:\alpha_1} + \underbrace{\left\|\left[\mathcal{P}_{\mathsf{diag}}\left(\boldsymbol{G}-\boldsymbol{G}^{\natural}\right)\right]_{m,\cdot}\boldsymbol{U}^{\natural}\right\|_2}_{=:\alpha_2},\end{aligned}$$

where the last inequality makes use of our construction $\mathcal{P}_{\mathsf{off\text{-}diag}}\left(\boldsymbol{G}\right) = \mathcal{P}_{\mathsf{off\text{-}diag}}\left(\boldsymbol{G}^{0}\right)$. Given that the properties of $\boldsymbol{G}^{0}$ (which is the diagonal-deleted version of the sample gram matrix) have been studied in Cai et al. (2021), we can readily borrow Cai et al. (2021, Lemma 2) to bound

$$\alpha_1 \lesssim \left(\sigma\sigma_m\sqrt{n_1 n_2\log n} + \sigma\sqrt{n_1\log n}\left\|\boldsymbol{M}_{m,\cdot}^{\natural}\right\|_2 + B\log n\left\|\boldsymbol{M}_{m,\cdot}^{\natural}\right\|_{\infty}\right)\left\|\boldsymbol{U}^{\natural}\right\|_{2,\infty} + \sigma_m\sqrt{n_2\log n}\left\|\boldsymbol{M}^{\natural\top}\right\|_{2,\infty}$$

with probability exceeding $1-O(n^{-10})$. The careful reader might remark that the above bound is slightly different from Cai et al. (2021, Lemma 2) in the sense that the bound therein contains an additional term $\|\boldsymbol{M}^{\natural}\|_{2,\infty}^2\sqrt{\mu^{\natural}r/n_1}$; note, however, that this extra term is caused by the effect of the diagonal part $\mathcal{P}_{\mathsf{diag}}(\boldsymbol{G}^{\natural})$ in their analysis, which has been removed in the above term α_1. The interested reader is referred to Cai et al. (2021, Appendix B.3) for details. In view of the bounds

$$\left\|\boldsymbol{M}_{m,\cdot}^{\natural}\right\|_2 = \left\|\boldsymbol{U}_{m,\cdot}^{\natural}\boldsymbol{\Sigma}^{\natural}\boldsymbol{V}^{\natural\top}\right\|_2 \leq \left\|\boldsymbol{U}_{m,\cdot}^{\natural}\right\|_2\left\|\boldsymbol{\Sigma}^{\natural}\right\|\left\|\boldsymbol{V}^{\natural}\right\| \leq \sigma_1^{\natural}\left\|\boldsymbol{U}_{m,\cdot}^{\natural}\right\|_2, \tag{C.37}$$

$$\left\|\boldsymbol{M}_{m,\cdot}^{\natural}\right\|_{\infty} = \left\|\boldsymbol{U}_{m,\cdot}^{\natural}\boldsymbol{\Sigma}^{\natural}\boldsymbol{V}^{\natural\top}\right\|_{\infty} \leq \left\|\boldsymbol{U}_{m,\cdot}^{\natural}\right\|_2\left\|\boldsymbol{\Sigma}^{\natural}\right\|\left\|\boldsymbol{V}^{\natural}\right\|_{2,\infty} \leq \sigma_1^{\natural}\left\|\boldsymbol{U}_{m,\cdot}^{\natural}\right\|_2\sqrt{\frac{\mu^{\natural}r}{n_2}}, \tag{C.38}$$

and

$$\left\|\boldsymbol{M}^{\natural\top}\right\|_{2,\infty}=\left\|\boldsymbol{V}^{\natural}\boldsymbol{\Sigma}^{\natural}\boldsymbol{U}^{\natural\top}\right\|_{2,\infty}\le\left\|\boldsymbol{V}^{\natural}\right\|_{2,\infty}\left\|\boldsymbol{\Sigma}^{\natural}\right\|\left\|\boldsymbol{U}^{\natural}\right\|\le\sigma_1^{\natural}\sqrt{\frac{\mu^{\natural}r}{n_2}},\tag{C.39}$$

we further have

$$\begin{aligned}\alpha_1&\lesssim\left(\sigma\sigma_m\sqrt{n_1n_2\log n}+\sigma\sigma_1^{\natural}\sqrt{n_1\log n}\left\|\boldsymbol{U}_{m,\cdot}^{\natural}\right\|_2+\sigma\sigma_1^{\natural}\sqrt[4]{n_1\mu^{\natural}r\log n}\left\|\boldsymbol{U}_{m,\cdot}^{\natural}\right\|_2\right)\left\|\boldsymbol{U}^{\natural}\right\|_{2,\infty}\\&\quad+\sigma_m\sigma_1^{\natural}\sqrt{\mu^{\natural}r\log n}\\&\lesssim\zeta_{\mathsf{op},m}\sqrt{\frac{\mu^{\natural}r}{n_1}}+\sigma\sigma_1^{\natural}\sqrt{\mu^{\natural}r\log n}\left\|\boldsymbol{U}_{m,\cdot}^{\natural}\right\|_2,\end{aligned}$$

provided that $n_1\gtrsim\mu^{\natural}r$. In addition, it is seen that

$$\begin{aligned}\alpha_2=\left|G_{m,m}-G_{m,m}^{\natural}\right|\left\|\boldsymbol{U}_{m,\cdot}^{\natural}\right\|_2&\lesssim\kappa^{\natural2}\zeta_{\mathsf{op}}\left(\left\|\boldsymbol{U}_{m,\cdot}^{\natural}\right\|_2+\left\|\boldsymbol{U}_{m,\cdot}\right\|_2\right)\left\|\boldsymbol{U}_{m,\cdot}^{\natural}\right\|_2\\&\lesssim\kappa^{\natural2}\zeta_{\mathsf{op}}\left\|\boldsymbol{U}_{m,\cdot}^{\natural}\right\|_2^2+\kappa^{\natural2}\zeta_{\mathsf{op}}\left\|\boldsymbol{U}_{m,\cdot}\boldsymbol{H}-\boldsymbol{U}_{m,\cdot}^{\natural}\right\|_2\left\|\boldsymbol{U}_{m,\cdot}^{\natural}\right\|_2,\end{aligned}$$

where the penultimate relation invokes Lemma 1. We can thus conclude that

$$\begin{aligned}\left\|\left(\boldsymbol{G}-\boldsymbol{G}^{\natural}\right)_{m,\cdot}\boldsymbol{U}^{\natural}\right\|_2&\le\alpha_1+\alpha_2\\&\lesssim\zeta_{\mathsf{op},m}\sqrt{\frac{\mu^{\natural}r}{n_1}}+\sigma\sigma_1^{\natural}\sqrt{\mu^{\natural}r\log n}\left\|\boldsymbol{U}_{m,\cdot}^{\natural}\right\|_2+\kappa^{\natural2}\zeta_{\mathsf{op}}\left\|\boldsymbol{U}_{m,\cdot}^{\natural}\right\|_2^2\\&\quad+\kappa^{\natural2}\zeta_{\mathsf{op}}\left\|\boldsymbol{U}_{m,\cdot}\boldsymbol{H}-\boldsymbol{U}_{m,\cdot}^{\natural}\right\|_2\left\|\boldsymbol{U}_{m,\cdot}^{\natural}\right\|_2.\end{aligned}$$

C.3.5 Proof of Lemma 4

Recall that for each $1\le m\le n_1$, we employ the notation $\mathcal{P}_{-m,\cdot}(\boldsymbol{M})$ to represent an $n_1\times n_2$ matrix such that

$$\left[\mathcal{P}_{-m,\cdot}(\boldsymbol{M})\right]_{i,\cdot}=\begin{cases}\boldsymbol{0}, & \text{if } i=m,\\ \boldsymbol{M}_{i,\cdot}, & \text{if } i\neq m.\end{cases}$$

In other words, it is obtained by zeroing out the m-th row of $\boldsymbol{M}$. Simple algebra then reveals that the m-th row of $\boldsymbol{G}$ can be decomposed into

$$\begin{aligned}\boldsymbol{G}_{m,\cdot}&=\boldsymbol{M}_{m,\cdot}\left[\mathcal{P}_{-m,\cdot}(\boldsymbol{M})\right]^{\top}+G_{m,m}\boldsymbol{e}_m^{\top}\\&=\boldsymbol{M}_{m,\cdot}^{\natural}\left[\mathcal{P}_{-m,\cdot}\left(\boldsymbol{M}^{\natural}\right)\right]^{\top}+\boldsymbol{M}_{m,\cdot}^{\natural}\left[\mathcal{P}_{-m,\cdot}(\boldsymbol{E})\right]^{\top}+\boldsymbol{E}_{m,\cdot}\left[\mathcal{P}_{-m,\cdot}(\boldsymbol{M})\right]^{\top}+G_{m,m}\boldsymbol{e}_m^{\top}\\&=\boldsymbol{M}_{m,\cdot}^{\natural}\boldsymbol{M}^{\natural\top}+\boldsymbol{M}_{m,\cdot}^{\natural}\left[\mathcal{P}_{-m,\cdot}(\boldsymbol{E})\right]^{\top}+\boldsymbol{E}_{m,\cdot}\left[\mathcal{P}_{-m,\cdot}(\boldsymbol{M})\right]^{\top}+\left(G_{m,m}-G_{m,m}^{\natural}\right)\boldsymbol{e}_m^{\top},\end{aligned}$$

where we have used $\boldsymbol{G}^{\natural}=\boldsymbol{M}^{\natural}\boldsymbol{M}^{\natural\top}$. Apply the triangle inequality once again to yield

$$\begin{aligned}\left\|\boldsymbol{G}_{m,\cdot}\left(\boldsymbol{U}\boldsymbol{H}-\boldsymbol{U}^{\natural}\right)\right\|_2&\le\underbrace{\left\|\boldsymbol{M}_{m,\cdot}^{\natural}\boldsymbol{M}^{\natural\top}\left(\boldsymbol{U}\boldsymbol{H}-\boldsymbol{U}^{\natural}\right)\right\|_2}_{=:\alpha_1}+\underbrace{\left\|\boldsymbol{M}_{m,\cdot}^{\natural}\left[\mathcal{P}_{-m,\cdot}(\boldsymbol{E})\right]^{\top}\left(\boldsymbol{U}\boldsymbol{H}-\boldsymbol{U}^{\natural}\right)\right\|_2}_{=:\alpha_2}\\&\quad+\underbrace{\left\|\boldsymbol{E}_{m,\cdot}\left[\mathcal{P}_{-m,\cdot}(\boldsymbol{M})\right]^{\top}\left(\boldsymbol{U}\boldsymbol{H}-\boldsymbol{U}^{\natural}\right)\right\|_2}_{=:\alpha_3}+\underbrace{\left\|\left(G_{m,m}-G_{m,m}^{\natural}\right)\boldsymbol{e}_m^{\top}\left(\boldsymbol{U}\boldsymbol{H}-\boldsymbol{U}^{\natural}\right)\right\|_2}_{=:\alpha_4}\end{aligned}$$

for each $1\le m\le n_1$. In what follows, we shall bound the terms $\alpha_1,\cdots,\alpha_4$ separately.

- Let us begin with α_1. Write

$$\left\|\boldsymbol{U}^{\natural\top}\left(\boldsymbol{U}\boldsymbol{H}-\boldsymbol{U}^{\natural}\right)\right\|=\left\|\boldsymbol{U}^{\natural\top}\boldsymbol{U}\boldsymbol{U}^{\top}\boldsymbol{U}^{\natural}-\boldsymbol{I}_r\right\|=\left\|\boldsymbol{H}^{\top}\boldsymbol{H}-\boldsymbol{I}_r\right\|\lesssim\frac{\zeta_{\mathsf{op}}^2}{\sigma_r^{\natural4}},\tag{C.40}$$

where the last relation follows from Lemma 2. This allows us to upper bound α_1 as follows:

$$\begin{aligned}\alpha_1 &= \left\|\boldsymbol{e}_m^\top \boldsymbol{M}^\natural \boldsymbol{M}^{\natural\top}\left(\boldsymbol{U}\boldsymbol{H}-\boldsymbol{U}^\natural\right)\right\|_2 \le \left\|\boldsymbol{e}_m^\top \boldsymbol{U}^\natural \boldsymbol{\Sigma}^{\natural 2}\boldsymbol{U}^{\natural\top}\left(\boldsymbol{U}\boldsymbol{H}-\boldsymbol{U}^\natural\right)\right\|_2 \\ &\le \left\|\boldsymbol{U}_{m,\cdot}^\natural\right\|_2 \left\|\boldsymbol{\Sigma}^\natural\right\|^2 \left\|\boldsymbol{U}^{\natural\top}\left(\boldsymbol{U}\boldsymbol{H}-\boldsymbol{U}^\natural\right)\right\| \\ &\lesssim \sigma_1^{\natural 2}\frac{\zeta_{\mathsf{op}}^2}{\sigma_r^{\natural 4}}\left\|\boldsymbol{U}_{m,\cdot}^\natural\right\|_2 \lesssim \kappa^{\natural 2}\frac{\zeta_{\mathsf{op}}^2}{\sigma_r^{\natural 2}}\left\|\boldsymbol{U}_{m,\cdot}^\natural\right\|_2.\end{aligned}$$

- Regarding α_2, it is readily seen that

$$\alpha_2 \le \left\|\boldsymbol{M}_{m,\cdot}^\natural\left[\mathcal{P}_{-m,\cdot}\left(\boldsymbol{E}\right)\right]^\top\right\|_2 \left\|\boldsymbol{U}\boldsymbol{H}-\boldsymbol{U}^\natural\right\|.$$

Regarding the first term $\|\boldsymbol{M}_{m,\cdot}^\natural[\mathcal{P}_{-m,\cdot}(\boldsymbol{E})]^\top\|_2$, we notice that

$$\begin{aligned}\left\|\boldsymbol{M}_{m,\cdot}^\natural\left[\mathcal{P}_{-m,\cdot}\left(\boldsymbol{E}\right)\right]^\top\right\|_2^2 &\le \left\|\boldsymbol{M}_{m,\cdot}^\natural \boldsymbol{E}^\top\right\|_2^2 = \sum_{i=1}^{n_1}\left(\sum_{j=1}^{n_2} M_{m,j}^\natural E_{i,j}\right)^2 \\ &\lesssim \left\|\boldsymbol{M}_{m,\cdot}^\natural\right\|_2^2\left(\sigma^2 n_1 + \sigma^2\log^2 n\right) + \left\|\boldsymbol{M}_{m,\cdot}^\natural\right\|_\infty^2 B^2\log^3 n \\ &\lesssim \left\|\boldsymbol{U}_{m,\cdot}^\natural\right\|_2^2 \sigma_1^{\natural 2}\left(\sigma^2 n_1 + \sigma^2\log^2 n\right) + \frac{\mu^\natural r}{n_2}\left\|\boldsymbol{U}_{m,\cdot}^\natural\right\|_2^2 \sigma_1^{\natural 2} B^2\log^3 n \\ &\lesssim \sigma^2\sigma_1^{\natural 2} n_1 \left\|\boldsymbol{U}_{m,\cdot}^\natural\right\|_2^2,\end{aligned}$$

where the second line follows from Cai et al. (2021, Lemma 14), the third line comes from (C.37) and (C.38), and the last line holds provided that $n_1 \gtrsim \mu^{\natural 2} r\log^2 n$, $n_2 \gtrsim r\log^2 n$ and $B \lesssim \sigma\sqrt[4]{n_1 n_2}/\sqrt{\log n}$. Therefore, we can demonstrate that

$$\alpha_2 \lesssim \left\|\boldsymbol{M}_{m,\cdot}^\natural\left[\mathcal{P}_{-m,\cdot}\left(\boldsymbol{E}\right)\right]^\top\right\|_2 \left\|\boldsymbol{U}\boldsymbol{H}-\boldsymbol{U}^\natural\right\| \overset{\text{(i)}}{\lesssim} \sigma\sigma_1^\natural\sqrt{n_1}\left\|\boldsymbol{U}_{m,\cdot}^\natural\right\|_2 \frac{\zeta_{\mathsf{op}}}{\sigma_r^{\natural 2}} \overset{\text{(ii)}}{\lesssim} \left\|\boldsymbol{U}_{m,\cdot}^\natural\right\|_2 \frac{\zeta_{\mathsf{op}}^2}{\sigma_r^{\natural 2}}.$$

Here, (i) utilizes Lemma 2 and (ii) comes from the definition of $\zeta_{\mathsf{op}} \ge \sigma\sigma_1^\natural\sqrt{n_1\log n}$.

- When it comes to α_3, our starting point is

$$\alpha_3 \le \underbrace{\left\|\boldsymbol{E}_{m,\cdot}\left[\mathcal{P}_{-m,\cdot}\left(\boldsymbol{M}\right)\right]^\top\left(\boldsymbol{U}^{(m)}\boldsymbol{H}^{(m)}-\boldsymbol{U}^\natural\right)\right\|_2}_{=:\beta_1} + \underbrace{\left\|\boldsymbol{E}_{m,\cdot}\left[\mathcal{P}_{-m,\cdot}\left(\boldsymbol{M}\right)\right]^\top\left(\boldsymbol{U}\boldsymbol{H}-\boldsymbol{U}^{(m)}\boldsymbol{H}^{(m)}\right)\right\|_2}_{=:\beta_2}.$$

  - The first term β_1 can be controlled using Lemma 10 as follows

$$\begin{aligned}\beta_1 &\lesssim \zeta_{\mathsf{op},m}\left(\left\|\boldsymbol{U}^{(m)}\boldsymbol{H}^{(m)}-\boldsymbol{U}^\natural\right\|_{2,\infty} + \frac{\zeta_{\mathsf{op}}}{\sigma_r^{\natural 2}}\left\|\boldsymbol{U}^{(m)}\boldsymbol{H}^{(m)}\right\|_{2,\infty}\right) \\ &\quad + \frac{\zeta_{\mathsf{op}}}{\sigma_r^{\natural 2}}\sigma_m\sigma_1^\natural\sqrt{n_1\log n}\sqrt{\frac{\mu^\natural r}{n_1}}. \qquad \text{(C.41)}\end{aligned}$$

  Additionally, it follows from Lemma 11 that

$$\begin{aligned}\left\|\boldsymbol{U}^{(m)}\boldsymbol{H}^{(m)}-\boldsymbol{U}\boldsymbol{H}\right\| &\le \left\|\boldsymbol{U}^{(m)}\boldsymbol{U}^{(m)\top}-\boldsymbol{U}\boldsymbol{U}^\top\right\| \le \kappa^{\natural 2}\frac{\zeta_{\mathsf{op}}}{\sigma_r^{\natural 2}}\left(\left\|\boldsymbol{U}\boldsymbol{H}\right\|_{2,\infty} + \sqrt{\frac{\mu^\natural r}{n_1}}\right) \\ &\le \kappa^{\natural 2}\frac{\zeta_{\mathsf{op}}}{\sigma_r^{\natural 2}}\left(\left\|\boldsymbol{U}\boldsymbol{H}-\boldsymbol{U}^\natural\right\|_{2,\infty} + \left\|\boldsymbol{U}^\natural\right\|_{2,\infty} + \sqrt{\frac{\mu^\natural r}{n_1}}\right) \\ &\asymp \kappa^{\natural 2}\frac{\zeta_{\mathsf{op}}}{\sigma_r^{\natural 2}}\left(\left\|\boldsymbol{U}\boldsymbol{H}-\boldsymbol{U}^\natural\right\|_{2,\infty} + \sqrt{\frac{\mu^\natural r}{n_1}}\right),\end{aligned}$$

which together with the triangle inequality gives

$$
\begin{aligned}
\left\|\boldsymbol{U}^{(m)}\boldsymbol{H}^{(m)}-\boldsymbol{U}^{\natural}\right\|_{2,\infty} &\leq \left\|\boldsymbol{U}^{(m)}\boldsymbol{H}^{(m)}-\boldsymbol{U}\boldsymbol{H}\right\|+\left\|\boldsymbol{U}\boldsymbol{H}-\boldsymbol{U}^{\natural}\right\|_{2,\infty} \\
&\lesssim \kappa^{\natural 2}\frac{\zeta_{\mathsf{op}}}{\sigma_r^{\natural 2}}\left(\left\|\boldsymbol{U}\boldsymbol{H}-\boldsymbol{U}^{\natural}\right\|_{2,\infty}+\sqrt{\frac{\mu^{\natural}r}{n_1}}\right)+\left\|\boldsymbol{U}\boldsymbol{H}-\boldsymbol{U}^{\natural}\right\|_{2,\infty} \\
&\asymp \kappa^{\natural 2}\frac{\zeta_{\mathsf{op}}}{\sigma_r^{\natural 2}}\sqrt{\frac{\mu^{\natural}r}{n_1}}+\left\|\boldsymbol{U}\boldsymbol{H}-\boldsymbol{U}^{\natural}\right\|_{2,\infty} \qquad \text{(C.42)}
\end{aligned}
$$

as long as $\zeta_{\mathsf{op}} \ll \sigma_r^{\natural 2}/\kappa^{\natural 2}$. An immediate consequence is that

$$
\begin{aligned}
\left\|\boldsymbol{U}^{(m)}\boldsymbol{H}^{(m)}\right\|_{2,\infty} &\leq \left\|\boldsymbol{U}^{\natural}\right\|_{2,\infty}+\left\|\boldsymbol{U}^{(m)}\boldsymbol{H}^{(m)}-\boldsymbol{U}^{\natural}\right\|_{2,\infty} \\
&\lesssim \sqrt{\frac{\mu^{\natural}r}{n_1}}+\kappa^{\natural 2}\frac{\zeta_{\mathsf{op}}}{\sigma_r^{\natural 2}}\sqrt{\frac{\mu^{\natural}r}{n_1}}+\left\|\boldsymbol{U}\boldsymbol{H}-\boldsymbol{U}^{\natural}\right\|_{2,\infty} \\
&\asymp \sqrt{\frac{\mu^{\natural}r}{n_1}}+\left\|\boldsymbol{U}\boldsymbol{H}-\boldsymbol{U}^{\natural}\right\|_{2,\infty}, \qquad \text{(C.43)}
\end{aligned}
$$

with the proviso that $\zeta_{\mathsf{op}} \ll \sigma_r^{\natural 2}/\kappa^{\natural 2}$. Therefore we can invoke (C.42) and (C.43) to refine (C.41) as

$$
\beta_1 \lesssim \zeta_{\mathsf{op},m}\left(\kappa^{\natural 2}\frac{\zeta_{\mathsf{op}}}{\sigma_r^{\natural 2}}\sqrt{\frac{\mu^{\natural}r}{n_1}}+\left\|\boldsymbol{U}\boldsymbol{H}-\boldsymbol{U}^{\natural}\right\|_{2,\infty}\right)
$$

provided that $\zeta_{\mathsf{op}} \ll \sigma_r^{\natural 2}$.

– With regards to the second term β_2, we see that

$$
\begin{aligned}
\beta_2 &= \left\|\boldsymbol{E}_{m,\cdot}\left[\mathcal{P}_{-m,\cdot}(\boldsymbol{M})\right]^{\top}\left(\boldsymbol{U}\boldsymbol{U}^{\top}-\boldsymbol{U}^{(m)}\boldsymbol{U}^{(m)\top}\right)\boldsymbol{U}^{\natural}\right\|_2 \\
&\leq \left\|\boldsymbol{E}_{m,\cdot}\left[\mathcal{P}_{-m,\cdot}(\boldsymbol{M})\right]^{\top}\right\|_2\left\|\boldsymbol{U}\boldsymbol{U}^{\top}-\boldsymbol{U}^{(m)}\boldsymbol{U}^{(m)\top}\right\|.
\end{aligned}
$$

It has already been proved in Cai et al. (2021, Appendix C.2) that with probability exceeding $1-O(n^{-10})$,

$$
\begin{aligned}
\left\|\boldsymbol{E}_{m,\cdot}\left[\mathcal{P}_{-m,\cdot}(\boldsymbol{M})\right]^{\top}\right\|_2 &\lesssim \sigma\sigma_m\sqrt{n_1 n_2\log n}+\sigma_m\sqrt{n_2\log n}\left\|\boldsymbol{M}^{\natural\top}\right\|_{2,\infty} \\
&\lesssim \sigma\sigma_m\sqrt{n_1 n_2\log n}+\sigma_m\sqrt{\mu^{\natural}r\log n}\,\sigma_1^{\natural}
\end{aligned}
$$

where the second relation follows from (C.39). This together with Lemma 11 provides an upper bound on β_2 as follows:

$$
\begin{aligned}
\beta_2 &\leq \left\|\boldsymbol{E}_{m,\cdot}\left[\mathcal{P}_{-m,\cdot}(\boldsymbol{M})\right]^{\top}\right\|_2\left\|\boldsymbol{U}\boldsymbol{H}-\boldsymbol{U}^{(m)}\boldsymbol{H}^{(m)}\right\| \\
&\lesssim \left(\sigma\sigma_m\sqrt{n_1 n_2\log n}+\sigma_m\sqrt{\mu^{\natural}r\log n}\,\sigma_1^{\natural}\right)\kappa^{\natural 2}\frac{\zeta_{\mathsf{op}}}{\sigma_r^{\natural 2}}\left(\left\|\boldsymbol{U}\boldsymbol{H}\right\|_{2,\infty}+\sqrt{\frac{\mu^{\natural}r}{n_1}}\right) \\
&\lesssim \left(\sigma\sigma_m\sqrt{n_1 n_2\log n}+\sigma_m\sqrt{\mu^{\natural}r\log n}\,\sigma_1^{\natural}\right)\kappa^{\natural 2}\frac{\zeta_{\mathsf{op}}}{\sigma_r^{\natural 2}}\left(\left\|\boldsymbol{U}\boldsymbol{H}-\boldsymbol{U}^{\natural}\right\|_{2,\infty}+\sqrt{\frac{\mu^{\natural}r}{n_1}}\right).
\end{aligned}
$$

– Combine the preceding bounds on β_1 and β_2 to arrive at

$$
\alpha_3 \leq \beta_1+\beta_2 \lesssim \zeta_{\mathsf{op},m}\left(\kappa^{\natural 2}\frac{\zeta_{\mathsf{op}}}{\sigma_r^{\natural 2}}\sqrt{\frac{\mu^{\natural}r}{n_1}}+\left\|\boldsymbol{U}\boldsymbol{H}-\boldsymbol{U}^{\natural}\right\|_{2,\infty}\right),
$$

provided that $n_1 \gtrsim \mu^{\natural}r$ and $\zeta_{\mathsf{op}}/\sigma_r^{\natural 2} \lesssim 1/\kappa^{\natural 2}$.

- For α_4, Lemma 1 tells us that

$$\begin{aligned}\alpha_4 &\leq \left|G_{m,m} - G^{\natural}_{m,m}\right| \left\|\boldsymbol{U}_{m,\cdot}\boldsymbol{H} - \boldsymbol{U}^{\natural}_{m,\cdot}\right\|_2 \lesssim \kappa^{\natural 2}\zeta_{\mathsf{op}}\left(\left\|\boldsymbol{U}^{\natural}_{m,\cdot}\right\|_2 + \left\|\boldsymbol{U}_{m,\cdot}\right\|_2\right)\left\|\boldsymbol{U}_{m,\cdot}\boldsymbol{H} - \boldsymbol{U}^{\natural}_{m,\cdot}\right\|_2 \\ &\lesssim \kappa^{\natural 2}\zeta_{\mathsf{op}}\left\|\boldsymbol{U}^{\natural}_{m,\cdot}\right\|_2\left\|\boldsymbol{U}_{m,\cdot}\boldsymbol{H} - \boldsymbol{U}^{\natural}_{m,\cdot}\right\|_2 + \kappa^{\natural 2}\zeta_{\mathsf{op}}\left\|\boldsymbol{U}_{m,\cdot}\boldsymbol{H} - \boldsymbol{U}^{\natural}_{m,\cdot}\right\|_2^2\end{aligned}$$

Thus far, we have developed upper bounds on $\alpha_1, \cdots, \alpha_4$, which taken collectively lead to

$$\begin{aligned}\left\|\boldsymbol{G}_{m,\cdot}\left(\boldsymbol{U}\boldsymbol{H} - \boldsymbol{U}^{\natural}\right)\right\|_2 &\leq \alpha_1 + \alpha_2 + \alpha_3 + \alpha_4 \\ &\lesssim \zeta_{\mathsf{op},m}\left(\kappa^{\natural 2}\frac{\zeta_{\mathsf{op}}}{\sigma_r^{\natural 2}}\sqrt{\frac{\mu^{\natural} r}{n_1}} + \left\|\boldsymbol{U}\boldsymbol{H} - \boldsymbol{U}^{\natural}\right\|_{2,\infty}\right) + \kappa^{\natural 2}\frac{\zeta_{\mathsf{op}}^2}{\sigma_r^{\natural 2}}\left\|\boldsymbol{U}^{\natural}_{m,\cdot}\right\|_2 \\ &\quad + \kappa^{\natural 2}\zeta_{\mathsf{op}}\left\|\boldsymbol{U}^{\natural}_{m,\cdot}\right\|_2\left\|\boldsymbol{U}_{m,\cdot}\boldsymbol{H} - \boldsymbol{U}^{\natural}_{m,\cdot}\right\|_2 + \kappa^{\natural 2}\zeta_{\mathsf{op}}\left\|\boldsymbol{U}_{m,\cdot}\boldsymbol{H} - \boldsymbol{U}^{\natural}_{m,\cdot}\right\|_2^2.\end{aligned}$$

C.3.6 Proof of Lemma 5

To control the target quantity, we make note of the following decomposition

$$\left\|\boldsymbol{R}^{\top}\boldsymbol{\Sigma}^2\boldsymbol{R} - \boldsymbol{\Sigma}^{\natural 2}\right\| \leq \underbrace{\left\|\boldsymbol{R}^{\top}\boldsymbol{\Sigma}^2\boldsymbol{R} - \boldsymbol{H}^{\top}\boldsymbol{\Sigma}^2\boldsymbol{H}\right\|}_{=:\,\alpha_1} + \underbrace{\left\|\boldsymbol{H}^{\top}\boldsymbol{\Sigma}^2\boldsymbol{H} - \boldsymbol{U}^{\natural\top}\boldsymbol{G}\boldsymbol{U}^{\natural}\right\|}_{=:\,\alpha_2} + \underbrace{\left\|\boldsymbol{U}^{\natural\top}\boldsymbol{G}\boldsymbol{U}^{\natural} - \boldsymbol{\Sigma}^{\natural 2}\right\|}_{=:\,\alpha_3}.$$

In the sequel, we shall upper bound each of these terms separately.

Step 1: bounding α_1. Lemma 2 tells us that

$$\begin{aligned}\alpha_1 &\leq \left\|\left(\boldsymbol{H} - \boldsymbol{R}\right)^{\top}\boldsymbol{\Sigma}^2\boldsymbol{H}\right\| + \left\|\boldsymbol{R}^{\top}\boldsymbol{\Sigma}^2\left(\boldsymbol{H} - \boldsymbol{R}\right)\right\| \\ &\leq \left\|\left(\boldsymbol{H} - \boldsymbol{R}\right)\right\|\left\|\boldsymbol{\Lambda}\right\|\left(\left\|\boldsymbol{H}\right\| + \left\|\boldsymbol{R}\right\|\right) \\ &\lesssim \frac{\zeta_{\mathsf{op}}^2}{\sigma_r^{\natural 4}}\sigma_1^{\natural 2} \asymp \kappa^{\natural 2}\frac{\zeta_{\mathsf{op}}^2}{\sigma_r^{\natural 2}}.\end{aligned}$$

Here, the penultimate inequality results from an application of Weyl's inequality:

$$\left\|\boldsymbol{\Lambda}\right\| \leq \left\|\boldsymbol{\Lambda}^{\natural}\right\| + \left\|\boldsymbol{G} - \boldsymbol{G}^{\natural}\right\| \lesssim \sigma_1^{\natural 2} + \zeta_{\mathsf{op}} \asymp \sigma_1^{\natural 2},$$

where we have used Lemma 1 and the assumption $\zeta_{\mathsf{op}} \lesssim \sigma_1^{\natural 2}$.

Step 2: bounding α_2. Regarding α_2, it is easily seen that

$$\begin{aligned}\boldsymbol{H}^{\top}\boldsymbol{\Sigma}^2\boldsymbol{H} - \boldsymbol{U}^{\natural\top}\boldsymbol{G}\boldsymbol{U}^{\natural} &= \boldsymbol{U}^{\natural\top}\boldsymbol{U}\boldsymbol{\Sigma}^2\boldsymbol{U}^{\top}\boldsymbol{U}^{\natural} - \boldsymbol{U}^{\natural\top}\boldsymbol{G}\boldsymbol{U}^{\natural} = \boldsymbol{U}^{\natural\top}\left(\boldsymbol{U}\boldsymbol{\Sigma}^2\boldsymbol{U}^{\top} - \boldsymbol{G}\right)\boldsymbol{U}^{\natural} \\ &= -\boldsymbol{U}^{\natural\top}\boldsymbol{U}_{\perp}\boldsymbol{\Lambda}_{\perp}\boldsymbol{U}_{\perp}^{\top}\boldsymbol{U}^{\natural},\end{aligned}$$

where we denote the full SVD of $\boldsymbol{G}$ as

$$\boldsymbol{G} = \left[\begin{array}{cc}\boldsymbol{U} & \boldsymbol{U}_{\perp}\end{array}\right]\left[\begin{array}{cc}\boldsymbol{\Lambda} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{\Lambda}_{\perp}\end{array}\right]\left[\begin{array}{c}\boldsymbol{V}^{\top} \\ \boldsymbol{V}_{\perp}^{\top}\end{array}\right] = \boldsymbol{U}\boldsymbol{\Sigma}^2\boldsymbol{V}^{\top} + \boldsymbol{U}_{\perp}\boldsymbol{\Sigma}_{\perp}^2\boldsymbol{V}_{\perp}^{\top}.$$

From Chen et al. (2021, Lemma 2.1.2) and (C.36), we have learned that

$$\left\|\boldsymbol{U}^{\natural\top}\boldsymbol{U}_{\perp}\right\| = \left\|\boldsymbol{U}\boldsymbol{U}^{\top} - \boldsymbol{U}^{\natural}\boldsymbol{U}^{\natural\top}\right\| \lesssim \frac{\zeta_{\mathsf{op}}}{\sigma_r^{\natural 2}}.$$

In view of Weyl's inequality and Lemma 1, we obtain

$$\left\|\boldsymbol{\Lambda}_{\perp}\right\| \leq \lambda_{r+1}\left(\boldsymbol{G}^{\natural}\right) + \left\|\boldsymbol{G} - \boldsymbol{G}^{\natural}\right\| \lesssim \zeta_{\mathsf{op}},$$

thus indicating that

$$\alpha_2 = \left\|\boldsymbol{U}^{\natural\top}\boldsymbol{U}_{\perp}\boldsymbol{\Lambda}_{\perp}\boldsymbol{U}_{\perp}^{\top}\boldsymbol{U}^{\natural}\right\| \leq \left\|\boldsymbol{U}^{\natural\top}\boldsymbol{U}_{\perp}\right\|^2\left\|\boldsymbol{\Lambda}_{\perp}\right\| \lesssim \frac{\zeta_{\mathsf{op}}^3}{\sigma_r^{\natural 4}}.$$

Step 3: bounding α_3. Let us decompose

$$\begin{aligned}
\boldsymbol{G} &= \mathcal{P}_{\mathsf{off\text{-}diag}}\left[\left(\boldsymbol{M}^{\natural}+\boldsymbol{E}\right)\left(\boldsymbol{M}^{\natural}+\boldsymbol{E}\right)^{\top}\right]+\mathcal{P}_{\mathsf{diag}}\left(\boldsymbol{G}^{\natural}\right)+\mathcal{P}_{\mathsf{diag}}\left(\boldsymbol{G}-\boldsymbol{G}^{\natural}\right) \\
&= \boldsymbol{G}^{\natural}+\mathcal{P}_{\mathsf{off\text{-}diag}}\left(\boldsymbol{E}\boldsymbol{M}^{\natural\top}+\boldsymbol{M}^{\natural}\boldsymbol{E}^{\top}+\boldsymbol{E}\boldsymbol{E}^{\top}\right)+\mathcal{P}_{\mathsf{diag}}\left(\boldsymbol{G}-\boldsymbol{G}^{\natural}\right) \\
&= \boldsymbol{G}^{\natural}+\boldsymbol{E}\boldsymbol{M}^{\natural\top}+\boldsymbol{M}^{\natural}\boldsymbol{E}^{\top}+\mathcal{P}_{\mathsf{off\text{-}diag}}\left(\boldsymbol{E}\boldsymbol{E}^{\top}\right)+\mathcal{P}_{\mathsf{diag}}\left(\boldsymbol{G}-\boldsymbol{G}^{\natural}\right)-\mathcal{P}_{\mathsf{diag}}\left(\boldsymbol{E}\boldsymbol{M}^{\natural\top}+\boldsymbol{M}^{\natural}\boldsymbol{E}^{\top}\right),
\end{aligned}$$

which in turn implies that

$$\begin{aligned}
\boldsymbol{U}^{\natural\top}\boldsymbol{G}\boldsymbol{U}^{\natural}-\boldsymbol{\Sigma}^{\natural 2} &= \underbrace{\boldsymbol{U}^{\natural\top}\left(\boldsymbol{E}\boldsymbol{M}^{\natural\top}+\boldsymbol{M}^{\natural}\boldsymbol{E}^{\top}\right)\boldsymbol{U}^{\natural}}_{=:\boldsymbol{J}_1}+\underbrace{\boldsymbol{U}^{\natural\top}\mathcal{P}_{\mathsf{off\text{-}diag}}\left(\boldsymbol{E}\boldsymbol{E}^{\top}\right)\boldsymbol{U}^{\natural}}_{=:\boldsymbol{J}_2} \\
&\quad+\underbrace{\boldsymbol{U}^{\natural\top}\mathcal{P}_{\mathsf{diag}}\left(\boldsymbol{G}-\boldsymbol{G}^{\natural}\right)\boldsymbol{U}^{\natural}}_{=:\boldsymbol{J}_3}-\underbrace{\boldsymbol{U}^{\natural\top}\mathcal{P}_{\mathsf{diag}}\left(\boldsymbol{E}\boldsymbol{M}^{\natural\top}+\boldsymbol{M}^{\natural}\boldsymbol{E}^{\top}\right)\boldsymbol{U}^{\natural}}_{=:\boldsymbol{J}_4}.
\end{aligned}$$

We shall then bound the spectral norm of $\boldsymbol{J}_1$, $\boldsymbol{J}_2$, $\boldsymbol{J}_3$ and $\boldsymbol{J}_4$ separately.

Step 3.1: bounding $\|\boldsymbol{J}_1\|$. Note that

$$\left\|\boldsymbol{U}^{\natural\top}\boldsymbol{E}\boldsymbol{M}^{\natural\top}\boldsymbol{U}^{\natural}\right\|=\left\|\boldsymbol{U}^{\natural\top}\boldsymbol{E}\boldsymbol{V}^{\natural}\boldsymbol{\Sigma}^{\natural}\right\|\leq\sigma_1^{\natural}\left\|\boldsymbol{U}^{\natural\top}\boldsymbol{E}\boldsymbol{V}^{\natural}\right\|.$$

It boils down to controlling the spectral norm of $\boldsymbol{U}^{\natural\top}\boldsymbol{E}\boldsymbol{V}^{\natural}$. This matrix admits the following decomposition

$$\boldsymbol{U}^{\natural\top}\boldsymbol{E}\boldsymbol{V}^{\natural}=\sum_{i=1}^{n_1}\sum_{j=1}^{n_2}E_{i,j}\boldsymbol{U}_{i,\cdot}^{\natural\top}\boldsymbol{V}_{j,\cdot}^{\natural}, \tag{C.44}$$

which can be controlled through the matrix Bernstein inequality. To do so, we are in need of calculating the following quantities:

$$\begin{aligned}
L &:= \max_{(i,j)\in[n_1]\times[n_2]}\left\|E_{i,j}\boldsymbol{U}_{i,\cdot}^{\natural\top}\boldsymbol{V}_{j,\cdot}^{\natural}\right\|\leq B\sqrt{\frac{\mu^{\natural}r}{n_1}}\sqrt{\frac{\mu^{\natural}r}{n_2}}=\frac{B\mu^{\natural}r}{\sqrt{n_1n_2}}, \\
V &:= \max\left\{\left\|\sum_{i=1}^{n_1}\sum_{j=1}^{n_2}\mathbb{E}\left[\left(E_{i,j}(\boldsymbol{U}_{i,\cdot}^{\natural})^{\top}\boldsymbol{V}_{j,\cdot}^{\natural}\right)\left(E_{i,j}(\boldsymbol{U}_{i,\cdot}^{\natural})^{\top}\boldsymbol{V}_{j,\cdot}^{\natural}\right)^{\top}\right]\right\|,\right. \\
&\qquad\qquad\left.\left\|\sum_{i=1}^{n_1}\sum_{j=1}^{n_2}\mathbb{E}\left[\left(E_{i,j}(\boldsymbol{U}_{i,\cdot}^{\natural})^{\top}\boldsymbol{V}_{j,\cdot}^{\natural}\right)^{\top}\left(E_{i,j}(\boldsymbol{U}_{i,\cdot}^{\natural})^{\top}\boldsymbol{V}_{j,\cdot}^{\natural}\right)\right]\right\|\right\} \\
&\leq \max\left\{\left\|\sum_{i=1}^{n_1}\sum_{j=1}^{n_2}\sigma^2\left\|\boldsymbol{V}_{j,\cdot}^{\natural}\right\|_2^2(\boldsymbol{U}_{i,\cdot}^{\natural})^{\top}\boldsymbol{U}_{i,\cdot}^{\natural}\right\|,\left\|\sum_{i=1}^{n_1}\sum_{j=1}^{n_2}\sigma^2\left\|\boldsymbol{U}_{j,\cdot}^{\natural}\right\|_2^2(\boldsymbol{V}_{i,\cdot}^{\natural})^{\top}\boldsymbol{V}_{i,\cdot}^{\natural}\right\|\right\} \\
&= \sigma^2\max\left\{\left\|\boldsymbol{V}^{\natural}\right\|_{\mathrm{F}}^2\left\|\boldsymbol{U}^{\natural\top}\boldsymbol{U}^{\natural}\right\|,\left\|\boldsymbol{U}^{\natural}\right\|_{\mathrm{F}}^2\left\|\boldsymbol{V}^{\natural\top}\boldsymbol{V}^{\natural}\right\|\right\}=\sigma^2r,
\end{aligned}$$

where we have made use of Assumptions 2-3. Applying matrix Bernstein inequality (Tropp, 2015, Theorem 6.1.1) to the decomposition (C.44) then tells us that, with probability exceeding $1-O(n^{-10})$,

$$\left\|\boldsymbol{U}^{\natural\top}\boldsymbol{E}\boldsymbol{V}^{\natural}\right\|\lesssim\sqrt{V\log n}+L\log n\lesssim\sigma\sqrt{r\log n}+\frac{B\mu^{\natural}r\log n}{\sqrt{n_1n_2}}\lesssim\sigma\sqrt{r\log n}+\frac{\sigma\mu^{\natural}r\sqrt{\log n}}{\sqrt[4]{n_1n_2}}, \tag{C.45}$$

provided that $B\lesssim\sigma\sqrt[4]{n_1n_2}/\sqrt{\log n}$. Therefore, we obtain

$$\begin{aligned}
\|\boldsymbol{J}_1\| &\leq 2\left\|\boldsymbol{U}^{\natural\top}\boldsymbol{E}\boldsymbol{M}^{\natural\top}\boldsymbol{U}^{\natural}\right\|=2\left\|\boldsymbol{U}^{\natural\top}\boldsymbol{E}\boldsymbol{V}^{\natural\top}\boldsymbol{\Sigma}^{\natural}\right\|\leq 2\sigma_1^{\natural}\left\|\boldsymbol{U}^{\natural\top}\boldsymbol{E}\boldsymbol{V}^{\natural}\right\| \\
&\lesssim \sigma\sigma_1^{\natural}\sqrt{r\log n}+\frac{\sigma\sigma_1^{\natural}\mu^{\natural}r\sqrt{\log n}}{\sqrt[4]{n_1n_2}}\lesssim\sqrt{\frac{\mu^{\natural}r}{n_1}}\zeta_{\mathsf{op}}
\end{aligned}$$

as long as $n_1 n_2 \gtrsim \mu^{\natural 2} r^2$.

Step 3.2: bounding $\|\boldsymbol{J}_2\|$. This matrix $\boldsymbol{J}_2$ can be expressed as a sum of independent and zero-mean random matrices:

$$\boldsymbol{J}_2 = \sum_{l=1}^{n_2} \underbrace{\boldsymbol{U}^{\natural\top}\left(\boldsymbol{E}_{\cdot,l}\boldsymbol{E}_{\cdot,l}^{\top} - \boldsymbol{D}_l\right)\boldsymbol{U}^{\natural}}_{=:\boldsymbol{X}_l};$$

here and below, we denote

$$\boldsymbol{D}_l \coloneqq \mathsf{diag}\left\{E_{1,l}^2, \dots, E_{n_1,l}^2\right\}. \tag{C.46}$$

We will invoke the truncated matrix Bernstein inequality (Chen et al., 2021, Theorem 3.1.1) to bound its spectral norm. To do so, we need to calculate the following quantities:

- We first study

$$\begin{aligned} v &\coloneqq \left\|\sum_{l=1}^{n_2}\mathbb{E}\left[\boldsymbol{X}_l\boldsymbol{X}_l^{\top}\right]\right\| \le \sup_{\|\boldsymbol{v}\|_2=1}\sum_{l=1}^{n_2}\mathbb{E}\left[\boldsymbol{v}^{\top}\boldsymbol{U}^{\natural\top}\left(\boldsymbol{E}_{\cdot,l}\boldsymbol{E}_{\cdot,l}^{\top} - \boldsymbol{D}_l\right)\boldsymbol{U}^{\natural}\boldsymbol{U}^{\natural\top}\left(\boldsymbol{E}_{\cdot,l}\boldsymbol{E}_{\cdot,l}^{\top} - \boldsymbol{D}_l\right)\boldsymbol{U}^{\natural}\boldsymbol{v}\right] \\ &= \sup_{\|\boldsymbol{v}\|_2=1}\sum_{l=1}^{n_2}\sum_{k=1}^{r}\mathbb{E}\left[\boldsymbol{v}^{\top}\boldsymbol{U}^{\natural\top}\left(\boldsymbol{E}_{\cdot,l}\boldsymbol{E}_{\cdot,l}^{\top} - \boldsymbol{D}_l\right)\boldsymbol{U}^{\natural}_{\cdot,k}\boldsymbol{U}^{\natural\top}_{\cdot,k}\left(\boldsymbol{E}_{\cdot,l}\boldsymbol{E}_{\cdot,l}^{\top} - \boldsymbol{D}_l\right)\boldsymbol{U}^{\natural}\boldsymbol{v}\right] \\ &= \sup_{\|\boldsymbol{v}\|_2=1}\sum_{l=1}^{n_2}\sum_{k=1}^{r}\mathbb{E}\left[\left(\boldsymbol{v}^{\top}\boldsymbol{U}^{\natural\top}\left(\boldsymbol{E}_{\cdot,l}\boldsymbol{E}_{\cdot,l}^{\top} - \boldsymbol{D}_l\right)\boldsymbol{U}^{\natural}_{\cdot,k}\right)^2\right]. \end{aligned}$$

  For any $\boldsymbol{v} \in \mathbb{R}^r$ with unit norm, let $\boldsymbol{x} = \boldsymbol{U}^{\natural}\boldsymbol{v}$ and derive

$$\begin{aligned} &\sum_{l=1}^{n_2}\sum_{k=1}^{r}\mathbb{E}\left[\left(\boldsymbol{v}^{\top}\boldsymbol{U}^{\natural\top}\left(\boldsymbol{E}_{\cdot,l}\boldsymbol{E}_{\cdot,l}^{\top} - \boldsymbol{D}_l\right)\boldsymbol{U}^{\natural}_{\cdot,k}\right)^2\right] = \sum_{l=1}^{n_2}\sum_{k=1}^{r}\mathbb{E}\left[\left(\boldsymbol{x}^{\top}\left(\boldsymbol{E}_{\cdot,l}\boldsymbol{E}_{\cdot,l}^{\top} - \boldsymbol{D}_l\right)\boldsymbol{U}^{\natural}_{\cdot,k}\right)^2\right] \\ &\quad = \sum_{l=1}^{n_2}\sum_{k=1}^{r}\mathsf{var}\left(\boldsymbol{x}^{\top}\boldsymbol{E}_{\cdot,l}\boldsymbol{E}_{\cdot,l}^{\top}\boldsymbol{U}^{\natural}_{\cdot,k}\right) = \sum_{l=1}^{n_2}\sum_{k=1}^{r}\mathsf{var}\left(\sum_{i=1}^{n_1}\sum_{j=1}^{n_1}x_i U^{\natural}_{j,k}E_{i,l}E_{j,l}\right) \\ &\quad = \sum_{l=1}^{n_2}\sum_{k=1}^{r}\sum_{i=1}^{n_1}\sum_{j=1}^{n_1}x_i^2 U^{\natural 2}_{j,k}\sigma_{i,l}^2\sigma_{j,l}^2 \le \sigma^4 n_2\|\boldsymbol{x}\|_2^2\left\|\boldsymbol{U}^{\natural}\right\|_{\mathrm{F}}^2 = \sigma^4 n_2 r. \end{aligned}$$

  To sum up, we have obtained

$$v \le \sigma^4 n_2 r.$$

- Note that for each $l \in [n_2]$, one can decompose

$$\|\boldsymbol{X}_l\| \le \left\|\boldsymbol{U}^{\natural\top}\boldsymbol{E}_{\cdot,l}\right\|_2^2 + \left\|\boldsymbol{U}^{\natural\top}\boldsymbol{D}_l\boldsymbol{U}^{\natural}\right\| = \underbrace{\left\|\sum_{i=1}^{n_1}E_{i,l}\boldsymbol{U}^{\natural}_{i,\cdot}\right\|_2^2}_{=:\beta_1} + \underbrace{\left\|\sum_{i=1}^{n_1}E_{i,l}^2\boldsymbol{U}^{\natural\top}_{i,\cdot}\boldsymbol{U}^{\natural}_{i,\cdot}\right\|}_{=:\beta_2}.$$

  – For β_1, it is straightforward to calculate

$$\begin{aligned} L_{\beta_1} &\coloneqq \max_{i\in[n_1]}\left\|E_{i,l}\boldsymbol{U}^{\natural}_{i,\cdot}\right\|_2 \le B\left\|\boldsymbol{U}^{\natural}\right\|_{2,\infty} \le B\sqrt{\frac{\mu^{\natural}r}{n_1}}, \\ V_{\beta_1} &\coloneqq \sum_{i=1}^{n_1}\mathbb{E}\left[E_{i,l}^2\right]\left\|\boldsymbol{U}^{\natural}_{i,\cdot}\right\|_2^2 \le \sigma^2\left\|\boldsymbol{U}^{\natural}\right\|_{\mathrm{F}}^2 = \sigma^2 r. \end{aligned}$$

    The matrix Bernstein inequality (Tropp, 2015, Theorem 6.1.1) tells us that

$$\mathbb{P}\left(\beta_1 \ge t\right) \le r\exp\left(\frac{-t/2}{V_{\beta_1} + L_{\beta_1}\sqrt{t}/3}\right).$$

– Regarding β_2, we can also calculate

$$L_{\beta_2} \coloneqq \max_{i\in[n_1]} \left\| E_{i,l}^2 \boldsymbol{U}_{i,\cdot}^{\natural\top} \boldsymbol{U}_{i,\cdot}^{\natural} \right\|_2 \le B^2 \left\| \boldsymbol{U}^{\natural} \right\|_{2,\infty}^2 \le B^2 \frac{\mu^{\natural} r}{n_1},$$

$$V_{\beta_2} \coloneqq \left\| \sum_{i=1}^{n_1} \mathbb{E}\left[E_{i,l}^4\right] \left\| \boldsymbol{U}_{i,\cdot}^{\natural} \right\|_2^2 \boldsymbol{U}_{i,\cdot}^{\natural\top} \boldsymbol{U}_{i,\cdot}^{\natural} \right\| \le \sigma^2 B^2 \left\| \boldsymbol{U}^{\natural} \right\|_{2,\infty}^2 \le \sigma^2 B^2 \frac{\mu^{\natural} r}{n_1}.$$

By virtue of the matrix Bernstein inequality (Tropp, 2015, Theorem 6.1.1), we can obtain

$$\mathbb{P}\left(\beta_2 \ge t\right) \le r \exp\left(\frac{-t^2/2}{V_{\beta_2} + L_{\beta_2} t/3} \right).$$

– Combine these tail probability bounds for β_1 and β_2 to achieve

$$\begin{aligned}
\mathbb{P}\left(\|\boldsymbol{X}_l\| \ge t\right) &\le \mathbb{P}\left(\beta_1 \ge \frac{t}{2}\right) + \mathbb{P}\left(\beta_2 \ge \frac{t}{2}\right) \\
&\le r \exp\left(\frac{-t/4}{V_{\beta_1} + L_{\beta_1}\sqrt{t}/5} \right) + r \exp\left(\frac{-t^2/8}{V_{\beta_2} + L_{\beta_2} t/6} \right) \\
&\le r \exp\left(-\min\left\{ \frac{t}{8V_{\beta_1}}, \frac{5\sqrt{t}}{8L_{\beta_1}} \right\} \right) + r \exp\left(-\min\left\{ \frac{t^2}{16V_{\beta_2}}, \frac{3t}{8L_{\beta_2}} \right\} \right).
\end{aligned}$$

Therefore, we can set

$$L \coloneqq \widetilde{C}\left(V_{\beta_1} \log^2 n + L_{\beta_1}^2 \log^2 n + \sqrt{V_{\beta_2}} \log n + L_{\beta_2} \log n \right) \asymp \sigma^2 r \log^2 n + \frac{B^2 \mu^{\natural} r}{n_1} \log^2 n$$

for some sufficiently large constant $\widetilde{C} > 0$, so as to guarantee that

$$\mathbb{P}\left(\|\boldsymbol{X}_l\| \ge L\right) \le q_0 \coloneqq \frac{1}{n^{10}}, \qquad \text{for all } l \in [n_1].$$

- Based on the above choice of L, we can see that

$$\begin{aligned}
q_1 \coloneqq \left\| \mathbb{E}\left[\boldsymbol{X}_l \mathbb{1}\left\{ \|\boldsymbol{X}_l\| \ge L \right\} \right] \right\| &\le \mathbb{E}\left[\|\boldsymbol{X}_l\| \mathbb{1}\left\{ \|\boldsymbol{X}_l\| \ge L \right\} \right] \\
&= \int_0^\infty \mathbb{P}\left(\|\boldsymbol{X}_l\| \mathbb{1}\left\{ \|\boldsymbol{X}_l\| \ge L \right\} > t \right) \mathrm{d}t \\
&= \int_0^L \mathbb{P}\left(\|\boldsymbol{X}_l\| \ge L \right) \mathrm{d}t + \int_L^\infty \mathbb{P}\left(\|\boldsymbol{X}_l\| > t \right) \mathrm{d}t \\
&\le \frac{L}{n^{10}} + \int_L^\infty \mathbb{P}\left(\|\boldsymbol{X}_l\| > t \right) \mathrm{d}t,
\end{aligned}$$

where the second inequality uses Jensen's inequality. Notice that for $t \ge L$, we have $t \gg V_{\beta_1}$ and $t \gg \sqrt{V_{\beta_2}}$ as long as the constant $\widetilde{C}$ is sufficiently large. As a consequence,

$$\begin{aligned}
\mathbb{P}\left(\|\boldsymbol{X}_l\| \ge t\right) &\le r \exp\left(-\min\left\{ \frac{t}{8V_{\beta_1}}, \frac{5\sqrt{t}}{8L_{\beta_1}} \right\} \right) + r \exp\left(-\min\left\{ \frac{t^2}{16V_{\beta_2}}, \frac{3t}{8L_{\beta_2}} \right\} \right) \\
&\le r \exp\left(-\min\left\{ \frac{\sqrt{t}}{\sqrt{V_{\beta_1}}}, \frac{5\sqrt{t}}{8L_{\beta_1}} \right\} \right) + r \exp\left(-\min\left\{ \frac{t}{\sqrt{V_{\beta_2}}}, \frac{3t}{8L_{\beta_2}} \right\} \right) \\
&\le r \exp\left(-\frac{\sqrt{t}}{\max\left\{ \sqrt{V_{\beta_1}}, 2L_{\beta_1} \right\}} \right) + r \exp\left(-\frac{t}{\max\left\{ \sqrt{V_{\beta_2}}, 3L_{\beta_2} \right\}} \right),
\end{aligned}$$

provided that $\widetilde{C}$ is sufficiently large. Consequently, we can deduce that

$$\int_L^\infty \mathbb{P}\left(\|\boldsymbol{X}_l\| \ge t\right) \mathrm{d}t \le r \underbrace{\int_L^\infty \exp\left(-\frac{\sqrt{t}}{\max\left\{ \sqrt{V_{\beta_1}}, 2L_{\beta_1} \right\}} \right) \mathrm{d}t}_{=:I_1} + r \underbrace{\int_L^\infty \exp\left(-\frac{t}{\max\left\{ \sqrt{V_{\beta_2}}, 3L_{\beta_2} \right\}} \right) \mathrm{d}t}_{=:I_2}.$$

– The first integral obeys

$$
\begin{aligned}
I_1 &\overset{\text{(i)}}{=} 2\max\left\{\sqrt{V_{\beta_1}}, 2L_{\beta_1}\right\}\sqrt{L}\exp\left(-\frac{\sqrt{L}}{\max\left\{\sqrt{V_{\beta_1}}, 2L_{\beta_1}\right\}}\right) \\
&\quad + 2\max\left\{\sqrt{V_{\beta_1}}, 2L_{\beta_1}\right\}^2\exp\left(-\frac{\sqrt{L}}{\max\left\{\sqrt{V_{\beta_1}}, 2L_{\beta_1}\right\}}\right) \\
&\le \left(\frac{4}{\sqrt{\widetilde{C}}} + \frac{8}{\widetilde{C}}\right) L\exp\left(-\frac{\sqrt{\widetilde{C}}}{2}\log n\right) \overset{\text{(iii)}}{\le} \frac{L}{n^{20}}.
\end{aligned}
$$

Here, (i) follows from the following formula that holds for any constants $\alpha, \beta > 0$:

$$
\begin{aligned}
\int_\beta^\infty \exp\left(-\alpha\sqrt{x}\right)\mathrm{d}x &\overset{y=\sqrt{x}}{=} 2\int_{\sqrt{\beta}}^\infty y\exp\left(-\alpha y\right)\mathrm{d}y = \left[-\frac{2}{\alpha}y\exp\left(-\alpha y\right)\right]\Bigg|_{\sqrt{\beta}}^\infty + \frac{2}{\alpha}\int_{\sqrt{\beta}}^\infty \exp\left(-\alpha y\right)\mathrm{d}y \\
&= \frac{2\sqrt{\beta}}{\alpha}\exp\left(-\alpha\sqrt{\beta}\right) + \left[-\frac{2}{\alpha^2}\exp\left(-\alpha y\right)\right]\Bigg|_{\sqrt{\beta}}^\infty \\
&= \frac{2\sqrt{\beta}}{\alpha}\exp\left(-\alpha\sqrt{\beta}\right) + \frac{2}{\alpha^2}\exp\left(-\alpha\sqrt{\beta}\right);
\end{aligned}
$$

(ii) comes from the definition of L:

$$
\sqrt{L} \ge \sqrt{\widetilde{C}\left(V_{\beta_1}\log^2 n + L_{\beta_1}^2\log^2 n\right)} \ge \sqrt{\widetilde{C}}\max\left\{\sqrt{V_{\beta_1}}, L_{\beta_1}\right\}\log n;
$$

and (iii) holds provided that $\widetilde{C}$ is sufficiently large.

– The second integral satisfies

$$
I_2 = \max\left\{\sqrt{V_{\beta_2}}, 3L_{\beta_2}\right\}\exp\left(-\frac{L}{\max\left\{\sqrt{V_{\beta_2}}, 3L_{\beta_2}\right\}}\right) \le \frac{3L}{\widetilde{C}}\exp\left(-\frac{\widetilde{C}}{3}\log n\right) \le \frac{L}{n^{20}}.
$$

Here the penultimate inequality follows from

$$
L \ge \widetilde{C}\left(\sqrt{V_{\beta_2}}\log n + L_{\beta_2}\log n\right) \ge \widetilde{C}\max\left\{\sqrt{V_{\beta_2}}, L_{\beta_2}\right\}\log n,
$$

and the last inequality holds provided that $\widetilde{C}$ is sufficiently large.

Therefore we conclude that

$$
q_1 \le \frac{L}{n^{10}} + rI_1 + rI_2 \le \frac{L}{n^{10}} + 2r\frac{L}{n^{20}} \le \frac{L}{n^9}.
$$

- With the above quantities in mind, we are ready to use the truncated matrix Bernstein inequality (Chen et al., 2021, Theorem 3.1.1) to show that with probability exceeding $1 - O(n^{-10})$,

$$
\begin{aligned}
\|\boldsymbol{J}_2\| &\lesssim \sqrt{v\log n} + L\log n + nq_1 \lesssim \sigma^2\sqrt{n_2 r\log n} + \sigma^2 r\log^3 n + \frac{B^2\mu^\natural r}{n_1}\log^3 n \\
&\overset{\text{(i)}}{\lesssim} \sigma^2\sqrt{n_2 r\log n} + \sigma^2 r\log^3 n + \frac{\sigma^2\sqrt{n_1 n_2}\mu^\natural r}{n_1}\log^2 n \\
&\asymp \left(\sqrt{\frac{r}{n_1}} + \frac{r\log^2 n}{\sqrt{n_1 n_2}} + \frac{\mu^\natural r\log n}{n_1}\right)\sigma^2\sqrt{n_1 n_2}\log n \overset{\text{(ii)}}{\lesssim} \sqrt{\frac{r}{n_1}}\zeta_{\mathsf{op}}.
\end{aligned}
$$

Here, (i) makes use of the noise condition $B \lesssim \sigma\sqrt[4]{n_1 n_2}/\sqrt{\log n}$, whereas (ii) holds provided that $n_1 \gtrsim \mu^{\natural 2} r\log^2 dn$ and $n_2 \gtrsim r\log^4 n$.

Step 3.3: bounding $\|\boldsymbol{J}_3\|$. Regarding $\boldsymbol{J}_3$, it is easy to show that

$$\|\boldsymbol{J}_3\| \leq \left\|\mathcal{P}_{\mathsf{diag}}\left(\boldsymbol{G}-\boldsymbol{G}^{\natural}\right)\right\| \lesssim \kappa^{\natural 2}\sqrt{\frac{\mu^{\natural}r}{n_1}}\zeta_{\mathsf{op}},$$

which results from Lemma 1.

Step 3.4: bounding $\|\boldsymbol{J}_4\|$. We are now left with bounding $\boldsymbol{J}_4$. Note that

$$\|\boldsymbol{J}_4\| \leq \left\|\mathcal{P}_{\mathsf{diag}}\left(\boldsymbol{E}\boldsymbol{M}^{\natural\top}+\boldsymbol{M}^{\natural}\boldsymbol{E}^{\top}\right)\right\| \leq 2\max_{i\in[n_1]}\left|\sum_{j=1}^{n_2}E_{i,j}M_{i,j}^{\natural}\right|.$$

For any $i\in[n_1]$, it is straightforward to calculate that

$$\begin{aligned}
L_i &\coloneqq \max_{j\in[n_2]}\left|E_{i,j}M_{i,j}^{\natural}\right| \leq B\left\|\boldsymbol{M}^{\natural}\right\|_{\infty},\\
V_i &\coloneqq \mathsf{var}\left(\sum_{j=1}^{n_2}E_{i,j}M_{i,j}^{\natural}\right) = \sum_{j=1}^{n_2}\sigma_{i,j}^2\left|M_{i,j}^{\natural}\right|^2 \leq \sigma^2\left\|\boldsymbol{M}^{\natural}\right\|_{2,\infty}^2.
\end{aligned}$$

In view of the Bernstein inequality (Vershynin, 2018, Theorem 2.8.4),

$$\begin{aligned}
\left|\sum_{j=1}^{n_2}E_{i,j}M_{i,j}^{\natural}\right| &\lesssim \sqrt{V_i\log n}+L_i\log n \lesssim \sigma\left\|\boldsymbol{M}^{\natural}\right\|_{2,\infty}\sqrt{\log n}+B\left\|\boldsymbol{M}^{\natural}\right\|_{\infty}\log n\\
&\lesssim \sigma\sigma_1^{\natural}\left\|\boldsymbol{U}^{\natural}\right\|_{2,\infty}\sqrt{\log n}+\sigma\sqrt{n_2}\sqrt{\frac{\mu^{\natural}r}{n_1n_2}}\sigma_1^{\natural}\\
&\lesssim \sigma\sigma_1^{\natural}\sqrt{\frac{\mu^{\natural}r\log n}{n_1}} \lesssim \frac{\sqrt{\mu^{\natural}r}}{n_1}\zeta_{\mathsf{op}} \qquad \text{(C.47)}
\end{aligned}$$

with probability exceeding $1-O(n^{-11})$. Combining the above bounds and applying the union bound show that with probability exceeding $1-O(n^{-10})$,

$$\|\boldsymbol{J}_4\| \lesssim \frac{\sqrt{\mu^{\natural}r}}{n_1}\zeta_{\mathsf{op}}.$$

Step 3.5: putting all this together. Taking the previous bounds on the spectral norm of $\boldsymbol{J}_1$, $\boldsymbol{J}_2$, $\boldsymbol{J}_3$, $\boldsymbol{J}_4$ collectively yields

$$\alpha_3 \leq \|\boldsymbol{J}_1\|+\|\boldsymbol{J}_2\|+\|\boldsymbol{J}_3\|+\|\boldsymbol{J}_4\| \lesssim \kappa^{\natural 2}\sqrt{\frac{\mu^{\natural}r}{n_1}}\zeta_{\mathsf{op}}+\frac{\mu^{\natural}r}{n_1}\zeta_{\mathsf{op}} \asymp \kappa^{\natural 2}\sqrt{\frac{\mu^{\natural}r}{n_1}}\zeta_{\mathsf{op}},$$

where the last relation holds as long as $n_1 \gtrsim \mu^{\natural}r$.

Step 4: combining the bounds on α_1, α_2 and α_3. Taking the bounds on α_1, α_2, α_3 together leads to

$$\left\|\boldsymbol{R}^{\top}\boldsymbol{\Sigma}^2\boldsymbol{R}-\boldsymbol{\Sigma}^{\natural 2}\right\| \leq \alpha_1+\alpha_2+\alpha_3 \lesssim \kappa^{\natural 2}\sqrt{\frac{\mu^{\natural}r}{n_1}}\zeta_{\mathsf{op}}+\kappa^{\natural 2}\frac{\zeta_{\mathsf{op}}^2}{\sigma_r^{\natural 2}}+\frac{\zeta_{\mathsf{op}}^3}{\sigma_r^{\natural 4}} \asymp \kappa^{\natural 2}\sqrt{\frac{\mu^{\natural}r}{n_1}}\zeta_{\mathsf{op}}+\kappa^{\natural 2}\frac{\zeta_{\mathsf{op}}^2}{\sigma_r^{\natural 2}}$$

with probability exceeding $1-O(n^{-10})$, where the last relation is valid under the condition that $\zeta_{\mathsf{op}} \ll \sigma_r^{\natural 2}$.

C.3.7 Proof of Lemma 6

To begin with, observe that $\boldsymbol{G}^\natural = \boldsymbol{U}^\natural \boldsymbol{\Sigma}^{\natural 2} \boldsymbol{U}^{\natural\top}$ (and hence $\boldsymbol{G}^\natural \boldsymbol{U}^\natural (\boldsymbol{\Sigma}^\natural)^{-2} = \boldsymbol{U}^\natural$), which allows us to decompose

$$\begin{aligned}
\left\|\boldsymbol{U}_{m,\cdot}\boldsymbol{H} - \boldsymbol{U}^\natural_{m,\cdot}\right\|_2 &= \left\|\left(\boldsymbol{U}\boldsymbol{H} - \boldsymbol{G}\boldsymbol{U}^\natural\left(\boldsymbol{\Sigma}^\natural\right)^{-2} + \boldsymbol{G}\boldsymbol{U}^\natural\left(\boldsymbol{\Sigma}^\natural\right)^{-2} - \boldsymbol{G}^\natural\boldsymbol{U}^\natural\left(\boldsymbol{\Sigma}^\natural\right)^{-2}\right)_{m,\cdot}\right\|_2 \\
&\leq \underbrace{\left\|\left(\boldsymbol{U}\boldsymbol{H} - \boldsymbol{G}\boldsymbol{U}^\natural\left(\boldsymbol{\Sigma}^\natural\right)^{-2}\right)_{m,\cdot}\right\|_2}_{=:\,\alpha_1} + \underbrace{\left\|\left(\boldsymbol{G} - \boldsymbol{G}^\natural\right)_{m,\cdot}\boldsymbol{U}^\natural\left(\boldsymbol{\Sigma}^\natural\right)^{-2}\right\|_2}_{=:\,\alpha_2}.
\end{aligned}$$

We then proceed to bound the terms α_1 the α_2.

- Regarding α_1, we have the following decomposition

$$\begin{aligned}
\alpha_1 &\leq \left\|\left(\boldsymbol{U}\boldsymbol{H}\boldsymbol{\Sigma}^{\natural 2} - \boldsymbol{G}\boldsymbol{U}^\natural\right)_{m,\cdot}\left(\boldsymbol{\Sigma}^\natural\right)^{-2}\right\|_2 \leq \frac{1}{\sigma_r^{\natural 2}}\left\|\left(\boldsymbol{U}\boldsymbol{H}\boldsymbol{\Sigma}^{\natural 2} - \boldsymbol{G}\boldsymbol{U}^\natural\right)_{m,\cdot}\right\|_2 \\
&\leq \underbrace{\frac{1}{\sigma_r^{\natural 2}}\left\|\boldsymbol{U}_{m,\cdot}\left(\boldsymbol{H}\boldsymbol{\Sigma}^{\natural 2} - \boldsymbol{\Sigma}^2\boldsymbol{H}\right)\right\|_2}_{=:\beta_1} + \underbrace{\frac{1}{\sigma_r^{\natural 2}}\left\|\left(\boldsymbol{U}\boldsymbol{\Sigma}^2\boldsymbol{H} - \boldsymbol{G}\boldsymbol{U}^\natural\right)_{m,\cdot}\right\|_2}_{=:\beta_2}.
\end{aligned}$$

  We first bound β_1, where we can use the triangle inequality to achieve

$$\begin{aligned}
\beta_1 &\leq \frac{1}{\sigma_r^{\natural 2}}\left\|\boldsymbol{U}_{m,\cdot}\left(\boldsymbol{R}_{\boldsymbol{U}}\boldsymbol{\Sigma}^{\natural 2} - \boldsymbol{\Sigma}^2\boldsymbol{R}_{\boldsymbol{U}}\right)\right\|_2 + \frac{1}{\sigma_r^{\natural 2}}\left\|\boldsymbol{U}_{m,\cdot}\left(\boldsymbol{H} - \boldsymbol{R}_{\boldsymbol{U}}\right)\boldsymbol{\Sigma}^{\natural 2}\right\|_2 + \frac{1}{\sigma_r^{\natural 2}}\left\|\boldsymbol{U}_{m,\cdot}\boldsymbol{\Sigma}^2\left(\boldsymbol{H} - \boldsymbol{R}_{\boldsymbol{U}}\right)\right\|_2 \\
&\leq \frac{1}{\sigma_r^{\natural 2}}\left\|\boldsymbol{U}_{m,\cdot}\boldsymbol{R}_{\boldsymbol{U}}\left(\boldsymbol{\Sigma}^{\natural 2} - \boldsymbol{R}_{\boldsymbol{U}}^\top\boldsymbol{\Sigma}^2\boldsymbol{R}_{\boldsymbol{U}}\right)\right\|_2 + \frac{1}{\sigma_r^{\natural 2}}\left\|\boldsymbol{U}_{m,\cdot}\right\|_2\left\|\boldsymbol{H} - \boldsymbol{R}\right\|\left(\left\|\boldsymbol{\Sigma}^{\natural 2}\right\| + \left\|\boldsymbol{\Sigma}^2\right\|\right) \\
&\lesssim \frac{1}{\sigma_r^{\natural 2}}\left\|\boldsymbol{U}_{m,\cdot}\right\|_2\left(\left\|\boldsymbol{\Sigma}^{\natural 2} - \boldsymbol{R}_{\boldsymbol{U}}^\top\boldsymbol{\Sigma}^2\boldsymbol{R}_{\boldsymbol{U}}\right\| + \left\|\boldsymbol{\Sigma}^{\natural 2}\right\|\left\|\boldsymbol{H} - \boldsymbol{R}\right\| + \left\|\boldsymbol{\Sigma}^{\natural 2} - \boldsymbol{R}_{\boldsymbol{U}}^\top\boldsymbol{\Sigma}^2\boldsymbol{R}_{\boldsymbol{U}}\right\|\left\|\boldsymbol{H} - \boldsymbol{R}\right\|\right) \\
&\lesssim \frac{1}{\sigma_r^{\natural 2}}\left\|\boldsymbol{U}_{m,\cdot}\boldsymbol{H}\right\|_2\left(\kappa^{\natural 2}\sqrt{\frac{\mu^\natural r}{n_1}}\zeta_{\mathsf{op}} + \kappa^{\natural 2}\frac{\zeta_{\mathsf{op}}^2}{\sigma_r^{\natural 2}}\right) \\
&\lesssim \left(\left\|\boldsymbol{U}^\natural_{m,\cdot}\right\|_2 + \left\|\boldsymbol{U}_{m,\cdot}\boldsymbol{H} - \boldsymbol{U}^\natural_{m,\cdot}\right\|_2\right)\left(\kappa^{\natural 2}\sqrt{\frac{\mu^\natural r}{n_1}}\frac{\zeta_{\mathsf{op}}}{\sigma_r^{\natural 2}} + \kappa^{\natural 2}\frac{\zeta_{\mathsf{op}}^2}{\sigma_r^{\natural 4}}\right).
\end{aligned}$$

  Here the penultimate line follows from Lemma 2, Lemma 5, and the assumption that $\zeta_{\mathsf{op}} \lesssim \sigma_r^{\natural 2}$. Additionally, Lemma 4 tells us that

$$\begin{aligned}
\beta_2 \lesssim\ & \frac{\zeta_{\mathsf{op},m}}{\sigma_r^{\natural 2}}\left(\kappa^{\natural 2}\frac{\zeta_{\mathsf{op}}}{\sigma_r^{\natural 2}}\sqrt{\frac{\mu^\natural r}{n_1}} + \left\|\boldsymbol{U}\boldsymbol{H} - \boldsymbol{U}^\natural\right\|_{2,\infty}\right) + \kappa^{\natural 2}\frac{\zeta_{\mathsf{op}}^2}{\sigma_r^{\natural 4}}\left\|\boldsymbol{U}^\natural_{m,\cdot}\right\|_2 \\
&+ \kappa^{\natural 2}\frac{\zeta_{\mathsf{op}}}{\sigma_r^{\natural 2}}\left\|\boldsymbol{U}^\natural_{m,\cdot}\right\|_2\left\|\boldsymbol{U}_{m,\cdot}\boldsymbol{H} - \boldsymbol{U}^\natural_{m,\cdot}\right\|_2 + \kappa^{\natural 2}\frac{\zeta_{\mathsf{op}}}{\sigma_r^{\natural 2}}\left\|\boldsymbol{U}_{m,\cdot}\boldsymbol{H} - \boldsymbol{U}^\natural_{m,\cdot}\right\|_2^2.
\end{aligned}$$

  Therefore, for all $m \in [n_1]$ we have

$$\begin{aligned}
\alpha_1 &\lesssim \beta_1 + \beta_2 \\
&\lesssim \left(\left\|\boldsymbol{U}^\natural_{m,\cdot}\right\|_2 + \left\|\boldsymbol{U}_{m,\cdot}\boldsymbol{H} - \boldsymbol{U}^\natural_{m,\cdot}\right\|_2\right)\left(\kappa^{\natural 2}\sqrt{\frac{\mu^\natural r}{n_1}}\frac{\zeta_{\mathsf{op}}}{\sigma_r^{\natural 2}} + \kappa^{\natural 2}\frac{\zeta_{\mathsf{op}}^2}{\sigma_r^{\natural 4}}\right) \\
&\quad + \frac{\zeta_{\mathsf{op},m}}{\sigma_r^{\natural 2}}\left(\kappa^{\natural 2}\frac{\zeta_{\mathsf{op}}}{\sigma_r^{\natural 2}}\sqrt{\frac{\mu^\natural r}{n_1}} + \left\|\boldsymbol{U}\boldsymbol{H} - \boldsymbol{U}^\natural\right\|_{2,\infty}\right) \\
&\quad + \kappa^{\natural 2}\frac{\zeta_{\mathsf{op}}}{\sigma_r^{\natural 2}}\left\|\boldsymbol{U}^\natural_{m,\cdot}\right\|_2\left\|\boldsymbol{U}_{m,\cdot}\boldsymbol{H} - \boldsymbol{U}^\natural_{m,\cdot}\right\|_2 + \kappa^{\natural 2}\frac{\zeta_{\mathsf{op}}}{\sigma_r^{\natural 2}}\left\|\boldsymbol{U}_{m,\cdot}\boldsymbol{H} - \boldsymbol{U}^\natural_{m,\cdot}\right\|_2^2.
\end{aligned}$$

- In view of Lemma 3, the second term α_2 can be bounded by

$$\begin{aligned}
\alpha_2 &\leq \frac{1}{\sigma_r^{\natural 2}} \left\| \left(\boldsymbol{G} - \boldsymbol{G}^{\natural}\right)_{m,\cdot} \boldsymbol{U}^{\natural} \right\|_2 \\
&\lesssim \frac{\zeta_{\mathsf{op},m}}{\sigma_r^{\natural 2}} \sqrt{\frac{\mu^{\natural} r}{n_1}} + \frac{1}{\sigma_r^{\natural 2}} \sigma \sigma_1^{\natural} \sqrt{\mu^{\natural} r \log n} \left\| \boldsymbol{U}_{m,\cdot}^{\natural} \right\|_2 + \kappa^{\natural 2} \frac{\zeta_{\mathsf{op}}}{\sigma_r^{\natural 2}} \left\| \boldsymbol{U}_{m,\cdot}^{\natural} \right\|_2^2 \\
&\quad + \kappa^{\natural 2} \frac{\zeta_{\mathsf{op}}}{\sigma_r^{\natural 2}} \left\| \boldsymbol{U}_{m,\cdot} \boldsymbol{H} - \boldsymbol{U}_{m,\cdot}^{\natural} \right\|_2 \left\| \boldsymbol{U}_{m,\cdot}^{\natural} \right\|_2 .
\end{aligned}$$

The preceding bounds taken together allow us to conclude that

$$\begin{aligned}
&\left\| \boldsymbol{U}_{m,\cdot} \boldsymbol{H} - \boldsymbol{U}_{m,\cdot}^{\natural} \right\|_2 \leq \alpha_1 + \alpha_2 \\
&\qquad \lesssim \frac{\zeta_{\mathsf{op},m}}{\sigma_r^{\natural 2}} \left(\sqrt{\frac{\mu^{\natural} r}{n_1}} + \left\| \boldsymbol{U}\boldsymbol{H} - \boldsymbol{U}^{\natural} \right\|_{2,\infty} \right) + \left\| \boldsymbol{U}_{m,\cdot}^{\natural} \right\|_2 \left(\kappa^{\natural 2} \sqrt{\frac{\mu^{\natural} r}{n_1}} \frac{\zeta_{\mathsf{op}}}{\sigma_r^{\natural 2}} + \kappa^{\natural 2} \frac{\zeta_{\mathsf{op}}^2}{\sigma_r^{\natural 4}} \right) \\
&\qquad\quad + \left\| \boldsymbol{U}_{m,\cdot} \boldsymbol{H} - \boldsymbol{U}_{m,\cdot}^{\natural} \right\|_2 \left(\kappa^{\natural 2} \sqrt{\frac{\mu^{\natural} r}{n_1}} \frac{\zeta_{\mathsf{op}}}{\sigma_r^{\natural 2}} + \kappa^{\natural 2} \frac{\zeta_{\mathsf{op}}^2}{\sigma_r^{\natural 4}} \right) + \kappa^{\natural 2} \frac{\zeta_{\mathsf{op}}}{\sigma_r^{\natural 2}} \left\| \boldsymbol{U}_{m,\cdot} \boldsymbol{H} - \boldsymbol{U}_{m,\cdot}^{\natural} \right\|_2^2, \qquad \text{(C.48)}
\end{aligned}$$

provided that $\zeta_{\mathsf{op}} \lesssim \sigma_r^{\natural 2} / \kappa^{\natural 2}$and $n_1 \gtrsim \mu^{\natural} r$. By taking supremum over $m \in [n_1]$, we have

$$\begin{aligned}
\left\| \boldsymbol{U}\boldsymbol{H} - \boldsymbol{U}^{\natural} \right\|_{2,\infty} &\lesssim \frac{\zeta_{\mathsf{op}}}{\sigma_r^{\natural 2}} \left(\sqrt{\frac{\mu^{\natural} r}{n_1}} + \left\| \boldsymbol{U}\boldsymbol{H} - \boldsymbol{U}^{\natural} \right\|_{2,\infty} \right) + \left\| \boldsymbol{U}^{\natural} \right\|_{2,\infty} \left(\kappa^{\natural 2} \sqrt{\frac{\mu^{\natural} r}{n_1}} \frac{\zeta_{\mathsf{op}}}{\sigma_r^{\natural 2}} + \kappa^{\natural 2} \frac{\zeta_{\mathsf{op}}^2}{\sigma_r^{\natural 4}} \right) \\
&\quad + \left\| \boldsymbol{U}\boldsymbol{H} - \boldsymbol{U}^{\natural} \right\|_{2,\infty} \left(\kappa^{\natural 2} \sqrt{\frac{\mu^{\natural} r}{n_1}} \frac{\zeta_{\mathsf{op}}}{\sigma_r^{\natural 2}} + \kappa^{\natural 2} \frac{\zeta_{\mathsf{op}}^2}{\sigma_r^{\natural 4}} \right) + \kappa^{\natural 2} \frac{\zeta_{\mathsf{op}}}{\sigma_r^{\natural 2}} \left\| \boldsymbol{U}\boldsymbol{H} - \boldsymbol{U}^{\natural} \right\|_{2,\infty}^2 \\
&\lesssim \frac{\zeta_{\mathsf{op}}}{\sigma_r^{\natural 2}} \sqrt{\frac{\mu^{\natural} r}{n_1}} + \frac{\zeta_{\mathsf{op}}}{\sigma_r^{\natural 2}} \left\| \boldsymbol{U}\boldsymbol{H} - \boldsymbol{U}^{\natural} \right\|_{2,\infty} + \kappa^{\natural 2} \frac{\zeta_{\mathsf{op}}}{\sigma_r^{\natural 2}} \left\| \boldsymbol{U}\boldsymbol{H} - \boldsymbol{U}^{\natural} \right\|_{2,\infty}^2 ,
\end{aligned}$$

provided that $\zeta_{\mathsf{op}} \lesssim \sigma_r^{\natural 2} / \kappa^{\natural 2}$ and $n_1 \gtrsim \kappa^{\natural 4} \mu^{\natural} r$. Rearrange terms to show that

$$\left\| \boldsymbol{U}\boldsymbol{H} - \boldsymbol{U}^{\natural} \right\|_{2,\infty} \lesssim \frac{\zeta_{\mathsf{op}}}{\sigma_r^{\natural 2}} \sqrt{\frac{\mu^{\natural} r}{n_1}}, \tag{C.49}$$

provided that $\zeta_{\mathsf{op}} \ll \sigma_r^{\natural 2} / \kappa^{\natural 2}$. Then we can use (C.49) to refine the bound (C.48) as

$$\begin{aligned}
\left\| \boldsymbol{U}_{m,\cdot} \boldsymbol{H} - \boldsymbol{U}_{m,\cdot}^{\natural} \right\|_2 &\lesssim \frac{\zeta_{\mathsf{op},m}}{\sigma_r^{\natural 2}} \sqrt{\frac{\mu^{\natural} r}{n_1}} + \left\| \boldsymbol{U}_{m,\cdot}^{\natural} \right\|_2 \left(\kappa^{\natural 2} \sqrt{\frac{\mu^{\natural} r}{n_1}} \frac{\zeta_{\mathsf{op}}}{\sigma_r^{\natural 2}} + \kappa^{\natural 2} \frac{\zeta_{\mathsf{op}}^2}{\sigma_r^{\natural 4}} \right) \\
&\quad + \left\| \boldsymbol{U}_{m,\cdot} \boldsymbol{H} - \boldsymbol{U}_{m,\cdot}^{\natural} \right\|_2 \left(\kappa^{\natural 2} \sqrt{\frac{\mu^{\natural} r}{n_1}} \frac{\zeta_{\mathsf{op}}}{\sigma_r^{\natural 2}} + \kappa^{\natural 2} \frac{\zeta_{\mathsf{op}}^2}{\sigma_r^{\natural 4}} \right) + \kappa^{\natural 2} \frac{\zeta_{\mathsf{op}}}{\sigma_r^{\natural 2}} \left\| \boldsymbol{U}_{m,\cdot} \boldsymbol{H} - \boldsymbol{U}_{m,\cdot}^{\natural} \right\|_2^2,
\end{aligned}$$

provided that $\zeta_{\mathsf{op}} \lesssim \sigma_r^{\natural 2} / \kappa^{\natural 2}$. Again, we can rearrange terms to achieve

$$\left\| \boldsymbol{U}_{m,\cdot} \boldsymbol{H} - \boldsymbol{U}_{m,\cdot}^{\natural} \right\|_2 \lesssim \frac{\zeta_{\mathsf{op},m}}{\sigma_r^{\natural 2}} \sqrt{\frac{\mu^{\natural} r}{n_1}} + \left\| \boldsymbol{U}_{m,\cdot}^{\natural} \right\|_2 \left(\kappa^{\natural 2} \sqrt{\frac{\mu^{\natural} r}{n_1}} \frac{\zeta_{\mathsf{op}}}{\sigma_r^{\natural 2}} + \kappa^{\natural 2} \frac{\zeta_{\mathsf{op}}^2}{\sigma_r^{\natural 4}} \right), \tag{C.50}$$

provided that $\zeta_{\mathsf{op}} \ll \sigma_r^{\natural 2} / \kappa^{\natural 2}$ and $n_1 \gtrsim \mu^{\natural} r$. Combine (C.49) and Lemma 2 gives

$$\left\| \boldsymbol{U}\boldsymbol{R}_{\boldsymbol{U}} - \boldsymbol{U}^{\natural} \right\|_{2,\infty} \leq \left\| \boldsymbol{U}\boldsymbol{H} - \boldsymbol{U}^{\natural} \right\|_{2,\infty} + \left\| \boldsymbol{U} \left(\boldsymbol{H} - \boldsymbol{R}_{\boldsymbol{U}} \right) \right\|_{2,\infty} \lesssim \left\| \boldsymbol{U}\boldsymbol{H} - \boldsymbol{U}^{\natural} \right\|_{2,\infty} + \left\| \boldsymbol{U} \right\|_{2,\infty} \left\| \boldsymbol{H} - \boldsymbol{R}_{\boldsymbol{U}} \right\|$$

$$\lesssim \left\|\boldsymbol{U}\boldsymbol{H}-\boldsymbol{U}^{\natural}\right\|_{2,\infty}+\left(\left\|\boldsymbol{U}^{\natural}\right\|_{2,\infty}+\left\|\boldsymbol{U}\boldsymbol{R}_{\boldsymbol{U}}-\boldsymbol{U}^{\natural}\right\|_{2,\infty}\right)\left\|\boldsymbol{H}-\boldsymbol{R}_{\boldsymbol{U}}\right\|$$
$$\lesssim \kappa^{\natural 2}\frac{\zeta_{\mathsf{op}}}{\sigma_r^{\natural 2}}\sqrt{\frac{\mu^{\natural}r}{n_1}}+\frac{\zeta_{\mathsf{op}}^2}{\sigma_r^{\natural 4}}\sqrt{\frac{\mu^{\natural}r}{n_1}}+\frac{\zeta_{\mathsf{op}}^2}{\sigma_r^{\natural 4}}\left\|\boldsymbol{U}\boldsymbol{R}_{\boldsymbol{U}}-\boldsymbol{U}^{\natural}\right\|_{2,\infty}.$$

Once again, when $\zeta_{\mathsf{op}}\ll\sigma_r^{\natural 2}$, one can rearrange terms to yield

$$\left\|\boldsymbol{U}\boldsymbol{R}_{\boldsymbol{U}}-\boldsymbol{U}^{\natural}\right\|_{2,\infty}\lesssim\frac{\zeta_{\mathsf{op}}}{\sigma_r^{\natural 2}}\sqrt{\frac{\mu^{\natural}r}{n_1}}.$$

Similarly, we can combine (C.50) and Lemma 2 to achieve

$$\left\|\boldsymbol{U}_{m,\cdot}\boldsymbol{R}_{\boldsymbol{U}}-\boldsymbol{U}_{m,\cdot}^{\natural}\right\|_2\lesssim\frac{\zeta_{\mathsf{op},m}}{\sigma_r^{\natural 2}}\sqrt{\frac{\mu^{\natural}r}{n_1}}+\left\|\boldsymbol{U}_{m,\cdot}^{\natural}\right\|_2\left(\kappa^{\natural 2}\sqrt{\frac{\mu^{\natural}r}{n_1}}\frac{\zeta_{\mathsf{op}}}{\sigma_r^{\natural 2}}+\kappa^{\natural 2}\frac{\zeta_{\mathsf{op}}^2}{\sigma_r^{\natural 4}}\right).$$

C.3.8 Proof of Lemma 7

We start with $\|\boldsymbol{G}^{(m)}-\boldsymbol{G}\|$, for which the triangle inequality yields

$$\left\|\boldsymbol{G}^{(m)}-\boldsymbol{G}\right\|\leq\left\|\mathcal{P}_{\textsf{off-diag}}\left(\boldsymbol{G}^{(m)}-\boldsymbol{G}\right)\right\|+\left\|\mathcal{P}_{\textsf{diag}}\left(\boldsymbol{G}^{(m)}-\boldsymbol{G}\right)\right\|. \tag{C.51}$$

- Regarding the first term on the right-hand side of (C.51), it is observed that $\mathcal{P}_{\textsf{off-diag}}(\boldsymbol{G}^{(m)}-\boldsymbol{G})$ in the current paper is the same as the matrix $\boldsymbol{G}^{(m)}-\boldsymbol{G}$ in Cai et al. (2021) (due to the diagonal deletion strategy employed therein). One can then apply Cai et al. (2021, Lemma 6) to show that

$$\begin{aligned}\left\|\mathcal{P}_{\textsf{off-diag}}\left(\boldsymbol{G}^{(m)}-\boldsymbol{G}\right)\right\|&\lesssim\sigma\sqrt{n_2}\left(\sigma\sqrt{n_1}+\left\|\boldsymbol{M}^{\natural\top}\right\|_{2,\infty}\right)\sqrt{\log n}\\&\lesssim\sigma^2\sqrt{n_1n_2\log n}+\sigma\sqrt{n_2}\sqrt{\frac{\mu^{\natural}r}{n_2}}\sigma_1^{\natural}\sqrt{\log n}\\&\asymp\sigma^2\sqrt{n_1n_2\log n}+\sigma\sigma_1^{\natural}\sqrt{\mu^{\natural}r\log n}\end{aligned}$$

  with probability exceeding $1-O(n^{-11})$, where the second inequality relies on (C.39). Here, we have replaced σ_{col} (resp. σ_{row}) in Cai et al. (2021, Lemma 6) with $\sigma\sqrt{n_1}$ (resp. $\sigma\sqrt{n_2}$) under our setting.

- Recalling that the diagonal of $\boldsymbol{G}^{(m)}$ coincides with the true diagonal of $\boldsymbol{G}^{\natural}$, we can invoke Lemma 1 to bound the second term on the right-hand side of (C.51) as follows

$$\left\|\mathcal{P}_{\textsf{diag}}\left(\boldsymbol{G}^{(m)}-\boldsymbol{G}\right)\right\|=\left\|\mathcal{P}_{\textsf{diag}}\left(\boldsymbol{G}^{\natural}-\boldsymbol{G}\right)\right\|\lesssim\kappa^{\natural 2}\sqrt{\frac{\mu^{\natural}r}{n_1}}\zeta_{\mathsf{op}}.$$

Combining the above two bounds and invoke the union bound lead to

$$\begin{aligned}\left\|\boldsymbol{G}^{(m)}-\boldsymbol{G}\right\|&\leq\left\|\mathcal{P}_{\textsf{off-diag}}\left(\boldsymbol{G}^{(m)}-\boldsymbol{G}\right)\right\|+\left\|\mathcal{P}_{\textsf{diag}}\left(\boldsymbol{G}^{(m)}-\boldsymbol{G}\right)\right\|\\&\lesssim\left\{\sigma^2\sqrt{n_1n_2\log n}+\sigma\sigma_1^{\natural}\sqrt{\mu^{\natural}r\log n}\right\}+\kappa^{\natural 2}\sqrt{\frac{\mu^{\natural}r}{n_1}}\zeta_{\mathsf{op}}\\&\lesssim\sigma^2\sqrt{n_1n_2\log n}+\sigma\sigma_1^{\natural}\sqrt{\mu^{\natural}r\log n}+\kappa^{\natural 2}\sqrt{\frac{\mu^{\natural}r}{n_1}}\left(\sigma^2\sqrt{n_1n_2}\log n+\sigma\sigma_1^{\natural}\sqrt{n_1\log n}\right)\\&\lesssim\sigma^2\sqrt{n_1n_2}\log n+\kappa^{\natural 2}\sigma\sigma_1^{\natural}\sqrt{\mu^{\natural}r\log n}\end{aligned}$$

simultaneously for all $1\leq m\leq n_1$, as long as $n_1\gtrsim\kappa^{\natural 4}\mu^{\natural}r$. Here, the penultimate inequality relies on the definition of ζ_{op}.

To finish up, taking the above inequality and Lemma 1 collectively yields

$$\left\|\boldsymbol{G}^{(m)}-\boldsymbol{G}^{\natural}\right\| \le \left\|\boldsymbol{G}^{(m)}-\boldsymbol{G}\right\| + \left\|\boldsymbol{G}-\boldsymbol{G}^{\natural}\right\| \lesssim \sigma^{2}\sqrt{n_{1}n_{2}\log n} + \kappa^{\natural 2}\sigma\sigma_{1}^{\natural}\sqrt{\mu^{\natural}r\log n} + \zeta_{\mathsf{op}} \asymp \zeta_{\mathsf{op}},$$

with the proviso that $n_1 \gtrsim \kappa^{\natural 4}\mu^{\natural}r$.

C.3.9 Proof of Lemma 8

The proof of Lemma 8 can be directly adapted from the proof of Cai et al. (2021, Lemma 9) (see Cai et al. (2021, Appendix C.5)). Two observations are crucial:

- The off-diagonal part of $\boldsymbol{G}^{(m)}$ (resp. $\boldsymbol{G}^{(m,l)}$) in this paper is the same as that of $\boldsymbol{G}^{(m)}$ (resp. $\boldsymbol{G}^{(m,l)}$) defined in Cai et al. (2021).
- Regarding the diagonal, this paper imputes the diagonals of both $\boldsymbol{G}^{(m)}$ and $\boldsymbol{G}^{(m,l)}$ with the diagonal of the ground truth $\boldsymbol{G}^{\natural}$, while the diagonal entries of both $\boldsymbol{G}^{(m)}$ and $\boldsymbol{G}^{(m,l)}$ in Cai et al. (2021) are all zeros.

Therefore, in both the current paper and Cai et al. (2021), we end up dealing with the same matrix $\boldsymbol{G}^{(m)} - \boldsymbol{G}^{(m,l)}$, and consequently, the bound on $\|\boldsymbol{G}^{(m)} - \boldsymbol{G}^{(m,l)}\|$ established in Cai et al. (2021, Appendix C.5.1) remains valid when it comes to our setting.

In addition, we can easily check that the bound on $\|(\boldsymbol{G}^{(m)} - \boldsymbol{G}^{(m,l)})\boldsymbol{U}^{(m,l)}\|$ derived in Cai et al. (2021, Appendix C.5.2) also holds under our setting; this is simply because the analysis in Cai et al. (2021, Appendix C.5.2) remains valid if $\boldsymbol{G}$ and $\boldsymbol{G}^{(l)}$ have the same deterministic diagonal. By replacing σ_{col} with $\sigma\sqrt{n_1}$ in Cai et al. (2021, Lemma 9), we arrive at the result claimed in Lemma 8.

C.3.10 Proof of Lemma 9

For the sake of brevity, we shall only focus on proving (C.16). The proof of (C.17) is similar to — and in fact, simpler than — the proof of (C.16), and can also be directly adapted from the proof of Cai et al. (2021, Lemma 7).

For notational simplicity, we denote $\boldsymbol{B} := \mathcal{P}_{-m,\cdot}(\boldsymbol{M})$. For any $l \in [n_2]$, we can write

$$\begin{aligned}
&\left\|\boldsymbol{e}_l^{\top}\left[\mathcal{P}_{-m,\cdot}\left(\boldsymbol{M}\right)\right]^{\top}\left(\boldsymbol{U}^{(m)}\boldsymbol{H}^{(m)}-\boldsymbol{U}^{\natural}\right)\right\|_2 = \left\|\boldsymbol{B}_{\cdot,l}^{\top}\left(\boldsymbol{U}^{(m)}\boldsymbol{H}^{(m)}-\boldsymbol{U}^{\natural}\right)\right\|_2 \\
&\quad\le \underbrace{\left\|\mathbb{E}\left(\boldsymbol{B}_{\cdot,l}\right)^{\top}\left(\boldsymbol{U}^{(m)}\boldsymbol{H}^{(m)}-\boldsymbol{U}^{\natural}\right)\right\|_2}_{=:\alpha_1} + \underbrace{\left\|\left[\boldsymbol{B}_{\cdot,l}-\mathbb{E}\left(\boldsymbol{B}_{\cdot,l}\right)\right]^{\top}\left(\boldsymbol{U}^{(m)}\boldsymbol{H}^{(m)}-\boldsymbol{U}^{(m,l)}\boldsymbol{H}^{(m,l)}\right)\right\|_2}_{=:\alpha_2} \\
&\quad\quad + \underbrace{\left\|\left[\boldsymbol{B}_{\cdot,l}-\mathbb{E}\left(\boldsymbol{B}_{\cdot,l}\right)\right]^{\top}\left(\boldsymbol{U}^{(m,l)}\boldsymbol{H}^{(m,l)}-\boldsymbol{U}^{\natural}\right)\right\|_2}_{=:\alpha_3}.
\end{aligned}$$

Therefore, we seek to bound α_1, α_2 and α_3 separately.

- Let us begin with the quantity α_1. It is straightforward to see that

$$\begin{aligned}
\alpha_1 &= \left\|\boldsymbol{M}_{\cdot,l}^{\natural\top}\left(\boldsymbol{U}^{(m)}\boldsymbol{H}^{(m)}-\boldsymbol{U}^{\natural}\right)\right\|_2 = \left\|\boldsymbol{M}_{\cdot,l}^{\natural\top}\boldsymbol{U}^{\natural}\boldsymbol{U}^{\natural\top}\left(\boldsymbol{U}^{(m)}\boldsymbol{H}^{(m)}-\boldsymbol{U}^{\natural}\right)\right\|_2 \\
&\le \left\|\boldsymbol{M}^{\natural\top}\right\|_{2,\infty}\left\|\boldsymbol{U}^{\natural\top}\left(\boldsymbol{U}^{(m)}\boldsymbol{H}^{(m)}-\boldsymbol{U}^{\natural}\right)\right\| \\
&= \left\|\boldsymbol{M}^{\natural\top}\right\|_{2,\infty}\left\|\boldsymbol{H}^{(m)\top}\boldsymbol{H}^{(m)}-\boldsymbol{I}_r\right\|,
\end{aligned} \tag{C.52}$$

where the last identity arises from the following relation

$$\left\|\boldsymbol{U}^{\natural\top}\left(\boldsymbol{U}^{(m)}\boldsymbol{H}^{(m)}-\boldsymbol{U}^{\natural}\right)\right\| = \left\|\boldsymbol{U}^{\natural\top}\boldsymbol{U}^{(m)}\boldsymbol{U}^{(m)\top}\boldsymbol{U}^{\natural}-\boldsymbol{I}_r\right\| = \left\|\boldsymbol{H}^{(m)\top}\boldsymbol{H}^{(m)}-\boldsymbol{I}_r\right\|.$$

Let us write the SVD of $\boldsymbol{H}^{(m)} = \boldsymbol{U}^{(m)\top}\boldsymbol{U}^{\natural}$ as $\boldsymbol{X}(\cos\boldsymbol{\Theta})\boldsymbol{Y}^{\top}$, where $\boldsymbol{X},\boldsymbol{Y} \in \mathbb{R}^{r\times r}$ are square orthonormal matrices and $\boldsymbol{\Theta}$ is a diagonal matrix composed of the principal angles between $\boldsymbol{U}^{(m)}$ and $\boldsymbol{U}^{\natural}$. This allows us to deduce that

$$\left\|\boldsymbol{H}^{(m)\top}\boldsymbol{H}^{(m)}-\boldsymbol{I}_r\right\| = \left\|\boldsymbol{Y}\left(\boldsymbol{I}_r-\cos^2\boldsymbol{\Theta}\right)\boldsymbol{Y}^{\top}\right\| = \left\|\boldsymbol{I}_r-\cos^2\boldsymbol{\Theta}\right\| = \left\|\sin^2\boldsymbol{\Theta}\right\| = \left\|\sin\boldsymbol{\Theta}\right\|^2.$$

In view of Davis-Kahan's $\sin\boldsymbol{\Theta}$ Theorem (Chen et al., 2021, Theorem 2.2.1), we have

$$\|\sin\boldsymbol{\Theta}\| \le \frac{\left\|\boldsymbol{G}^{(m)}-\boldsymbol{G}^{\natural}\right\|}{\lambda_r\left(\boldsymbol{G}^{(m)}\right)-\lambda_{r+1}\left(\boldsymbol{G}^{\natural}\right)} \lesssim \frac{\zeta_{\mathsf{op}}}{\sigma_r^{\natural 2}},$$

where the last inequality follows from Lemma 7 and an application of Weyl's inequality:

$$\lambda_r\left(\boldsymbol{G}^{(m)}\right) \ge \lambda_r\left(\boldsymbol{G}^{\natural}\right) - \left\|\boldsymbol{G}^{(m)}-\boldsymbol{G}^{\natural}\right\| \overset{\text{(i)}}{\ge} \sigma_r^{\natural 2} - \widetilde{C}\zeta_{\mathsf{op}} \overset{\text{(ii)}}{\ge} \frac{1}{2}\sigma_r^{\natural 2},$$

with $\widetilde{C}>0$ representing some absolute constant. Here, (i) results from Lemma 7, while (ii) holds provided that $\zeta_{\mathsf{op}} \ll \sigma_r^{\natural 2}$.. Therefore, we arrive at

$$\left\|\boldsymbol{H}^{(m)\top}\boldsymbol{H}^{(m)}-\boldsymbol{I}_r\right\| \le \|\sin\boldsymbol{\Theta}\|^2 \lesssim \frac{\zeta_{\mathsf{op}}^2}{\sigma_r^{\natural 4}}.$$

Substitution into (C.52) yields

$$\alpha_1 \lesssim \frac{\zeta_{\mathsf{op}}^2}{\sigma_r^{\natural 4}}\left\|\boldsymbol{M}^{\natural\top}\right\|_{2,\infty}.$$

- Regarding α_2, it is observed that

$$\begin{aligned}
\alpha_2 &\le \left(\|\boldsymbol{B}_{\cdot,l}\|_2 + \|\mathbb{E}\left(\boldsymbol{B}_{\cdot,l}\right)\|_2\right)\left\|\boldsymbol{U}^{(m)}\boldsymbol{H}^{(m)}-\boldsymbol{U}^{(m,l)}\boldsymbol{H}^{(m,l)}\right\| \\
&\le \left(\|\boldsymbol{M}_{\cdot,l}\|_2 + \left\|\boldsymbol{M}_{\cdot,l}^{\natural}\right\|_2\right)\left\|\left(\boldsymbol{U}^{(m)}\boldsymbol{U}^{(m)\top}-\boldsymbol{U}^{(m,l)}\boldsymbol{U}^{(m,l)\top}\right)\boldsymbol{U}^{\natural}\right\| \\
&\le \left(\left\|\boldsymbol{M}^{\natural\top}\right\|_{2,\infty} + B\sqrt{\log n} + \sigma\sqrt{n_1}\right)\left\|\boldsymbol{U}^{(m)}\boldsymbol{U}^{(m)\top}-\boldsymbol{U}^{(m,l)}\boldsymbol{U}^{(m,l)\top}\right\|.
\end{aligned}$$

  Here, the last inequality arises from Cai et al. (2021, Lemma 12).

- We are now left with the quantity α_3, which can be expressed as

$$\alpha_3 = \left\|\sum_{i\in[n_1]\setminus\{m\}} E_{i,l}\left(\boldsymbol{U}^{(m,l)}\boldsymbol{H}^{(m,l)}-\boldsymbol{U}^{\natural}\right)_{i,\cdot}\right\|_2.$$

  Conditional on $\boldsymbol{U}^{(m,l)}\boldsymbol{H}^{(m,l)}-\boldsymbol{U}^{\natural}$, this term can be viewed as the spectral norm of a sum of independent mean-zero random vectors (where the randomness comes from $\{E_{i,l}\}_{i\in[n_1]\setminus\{m\}}$). To control this term, we first calculate

$$\begin{aligned}
L &\coloneqq \max_{i\in[n_1]\setminus\{m\}}\left\|E_{i,l}\left(\boldsymbol{U}^{(m,l)}\boldsymbol{H}^{(m,l)}-\boldsymbol{U}^{\natural}\right)_{i,\cdot}\right\| \le B\left\|\boldsymbol{U}^{(m,l)}\boldsymbol{H}^{(m,l)}-\boldsymbol{U}^{\natural}\right\|_{2,\infty}, \\
V &\coloneqq \sum_{i\in[n_1]\setminus\{m\}}\mathbb{E}\left[E_{i,l}^2\right]\left\|\left(\boldsymbol{U}^{(m,l)}\boldsymbol{H}^{(m,l)}-\boldsymbol{U}^{\natural}\right)_{i,\cdot}\right\|_2^2 \le \sigma^2 n_1\left\|\boldsymbol{U}^{(m,l)}\boldsymbol{H}^{(m,l)}-\boldsymbol{U}^{\natural}\right\|_{2,\infty}^2.
\end{aligned}$$

  In view of the matrix Bernstein inequality (Tropp, 2015, Theorem 6.1.1),

$$\begin{aligned}
\alpha_3 &= \left\|\sum_{i\in[n_1]\setminus\{m\}} E_{i,l}\left(\boldsymbol{U}^{(m,l)}\boldsymbol{H}^{(m,l)}-\boldsymbol{U}^{\natural}\right)_{i,\cdot}\right\| \lesssim \sqrt{V\log n} + L\log n \\
&\lesssim \left(\sigma\sqrt{n_1\log n} + B\log n\right)\left\|\boldsymbol{U}^{(m,l)}\boldsymbol{H}^{(m,l)}-\boldsymbol{U}^{\natural}\right\|_{2,\infty}
\end{aligned}$$

  with probability exceeding $1-O(n^{-10})$. In addition, the triangle inequality gives

$$\left\|\boldsymbol{U}^{(m,l)}\boldsymbol{H}^{(m,l)}-\boldsymbol{U}^{\natural}\right\|_{2,\infty} \le \left\|\boldsymbol{U}^{(m)}\boldsymbol{H}^{(m)}-\boldsymbol{U}^{\natural}\right\|_{2,\infty} + \left\|\boldsymbol{U}^{(m,l)}\boldsymbol{H}^{(m,l)}-\boldsymbol{U}^{(m)}\boldsymbol{H}^{(m)}\right\|_{2,\infty}$$

$$\begin{aligned}&\le \left\|\boldsymbol{U}^{(m)}\boldsymbol{H}^{(m)}-\boldsymbol{U}^{\natural}\right\|_{2,\infty}+\left\|\boldsymbol{U}^{(m,l)}\boldsymbol{H}^{(m,l)}-\boldsymbol{U}^{(m)}\boldsymbol{H}^{(m)}\right\|\\&\le \left\|\boldsymbol{U}^{(m)}\boldsymbol{H}^{(m)}-\boldsymbol{U}^{\natural}\right\|_{2,\infty}+\left\|\boldsymbol{U}^{(m)}\boldsymbol{U}^{(m)\top}-\boldsymbol{U}^{(m,l)}\boldsymbol{U}^{(m,l)\top}\right\|,\end{aligned}$$

and therefore,

$$\alpha_3 \lesssim \left(\sigma\sqrt{n_1\log n}+B\log n\right)\left(\left\|\boldsymbol{U}^{(m)}\boldsymbol{H}^{(m)}-\boldsymbol{U}^{\natural}\right\|_{2,\infty}+\left\|\boldsymbol{U}^{(m)}\boldsymbol{U}^{(m)\top}-\boldsymbol{U}^{(m,l)}\boldsymbol{U}^{(m,l)\top}\right\|\right).$$

Combine the preceding bounds on α_1, α_2 and α_3 to arrive at

$$\begin{aligned}&\left\|\boldsymbol{e}_l^\top\left[\mathcal{P}_{-m,\cdot}(\boldsymbol{M})\right]^\top\left(\boldsymbol{U}^{(m)}\boldsymbol{H}^{(m)}-\boldsymbol{U}^{\natural}\right)\right\|_2 \le \alpha_1+\alpha_2+\alpha_3\\&\lesssim \frac{\zeta_{\mathsf{op}}^2}{\sigma_r^{\natural 4}}\left\|\boldsymbol{M}^{\natural\top}\right\|_{2,\infty}+\left(\sigma\sqrt{n_1\log n}+B\log n\right)\left\|\boldsymbol{U}^{(m)}\boldsymbol{H}^{(m)}-\boldsymbol{U}^{\natural}\right\|_{2,\infty}\\&\quad+\left(\left\|\boldsymbol{M}^{\natural\top}\right\|_{2,\infty}+B\log n+\sigma\sqrt{n_1\log n}\right)\left\|\boldsymbol{U}^{(m)}\boldsymbol{U}^{(m)\top}-\boldsymbol{U}^{(m,l)}\boldsymbol{U}^{(m,l)\top}\right\|\end{aligned}$$

as claimed.

C.3.11 Proof of Lemma 10

In this subsection, for the sake of brevity, we shall only focus on proving (C.18). The proof of (C.19) is similar to that of (C.18), and can also be easily adapted from the proof of Cai et al. (2021, Lemma 7).

For notational simplicity, we denote $\boldsymbol{B}\coloneqq\mathcal{P}_{-m,\cdot}(\boldsymbol{M})$, allowing us to express

$$\boldsymbol{E}_{m,\cdot}\left[\mathcal{P}_{-m,\cdot}(\boldsymbol{M})\right]^\top\left(\boldsymbol{U}^{(m)}\boldsymbol{H}^{(m)}-\boldsymbol{U}^{\natural}\right)=\sum_{j=1}^{n_2}E_{m,j}\left[\boldsymbol{B}^\top\left(\boldsymbol{U}^{(m)}\boldsymbol{H}^{(m)}-\boldsymbol{U}^{\natural}\right)\right]_{j,\cdot}.$$

Conditional on $\boldsymbol{B}$ and $\boldsymbol{U}^{(m)}\boldsymbol{H}^{(m)}-\boldsymbol{U}^{\natural}$, the above term can be viewed as a sum of independent zero-mean random vectors, where the randomness comes from $\{E_{m,j}\}_{j\in[n_2]}$. We can calculate

$$\begin{aligned}L&\coloneqq\max_{j\in[n_2]}\left\|E_{m,j}\left[\boldsymbol{B}^\top\left(\boldsymbol{U}^{(m)}\boldsymbol{H}^{(m)}-\boldsymbol{U}^{\natural}\right)\right]_{j,\cdot}\right\|\le B_m\left\|\boldsymbol{B}^\top\left(\boldsymbol{U}^{(m)}\boldsymbol{H}^{(m)}-\boldsymbol{U}^{\natural}\right)\right\|_{2,\infty},\\V&\coloneqq\sum_{j\in[n_2]}\mathbb{E}\left(E_{m,j}^2\right)\left\|\boldsymbol{B}^\top\left(\boldsymbol{U}^{(m)}\boldsymbol{H}^{(m)}-\boldsymbol{U}^{\natural}\right)_{j,\cdot}\right\|_2^2\le\sigma_m^2\left\|\boldsymbol{B}^\top\left(\boldsymbol{U}^{(m)}\boldsymbol{H}^{(m)}-\boldsymbol{U}^{\natural}\right)\right\|_{\mathrm{F}}^2.\end{aligned}$$

In view of the matrix Bernstein inequality (Tropp, 2015, Theorem 6.1.1), with probability exceeding $1-O(n^{-11})$

$$\begin{aligned}&\left\|\sum_{j=1}^{n_2}E_{m,j}\left[\boldsymbol{B}^\top\left(\boldsymbol{U}^{(m)}\boldsymbol{H}^{(m)}-\boldsymbol{U}^{\natural}\right)\right]_{j,\cdot}\right\|_2\lesssim\sqrt{V\log n}+L\log n\\&\lesssim\underbrace{\sigma_m\left\|\boldsymbol{B}^\top\left(\boldsymbol{U}^{(m)}\boldsymbol{H}^{(m)}-\boldsymbol{U}^{\natural}\right)\right\|_{\mathrm{F}}\sqrt{\log n}}_{=:\alpha_1}+\underbrace{B_m\left\|\boldsymbol{B}^\top\left(\boldsymbol{U}^{(m)}\boldsymbol{H}^{(m)}-\boldsymbol{U}^{\natural}\right)\right\|_{2,\infty}\log n}_{=:\alpha_2}.\end{aligned}$$

For the first term α_1, with probability exceeding $1-O(n^{-11})$ we have

$$\begin{aligned}\left\|\boldsymbol{B}^\top\left(\boldsymbol{U}^{(m)}\boldsymbol{H}^{(m)}-\boldsymbol{U}^{\natural}\right)\right\|_{\mathrm{F}}&\le\|\boldsymbol{B}\|\left\|\boldsymbol{U}^{(m)}\boldsymbol{H}^{(m)}-\boldsymbol{U}^{\natural}\right\|_{\mathrm{F}}\le\|\boldsymbol{M}\|\left\|\boldsymbol{U}^{(m)}\boldsymbol{H}^{(m)}-\boldsymbol{U}^{\natural}\right\|_{\mathrm{F}}\\&\le\left(\|\boldsymbol{M}^{\natural}\|+\|\boldsymbol{E}\|\right)\left\|\boldsymbol{U}^{(m)}\boldsymbol{H}^{(m)}-\boldsymbol{U}^{\natural}\right\|_{\mathrm{F}}\\&\lesssim\left(\sigma_1^{\natural}+\sigma\sqrt{n}\right)\left\|\boldsymbol{U}^{(m)}\boldsymbol{H}^{(m)}-\boldsymbol{U}^{\natural}\right\|_{\mathrm{F}}\end{aligned}$$

$$\asymp \left(\sigma_1^\natural + \sigma\sqrt{n_2}\right)\left\|\boldsymbol{U}^{(m)}\boldsymbol{H}^{(m)} - \boldsymbol{U}^\natural\right\|_{\mathrm{F}}.$$

Here, the penultimate inequality uses $\|\boldsymbol{E}\| \lesssim \sigma\sqrt{n}$ with probability exceeding $1 - O(n^{-11})$, which follows from standard matrix tail bounds (e.g., Chen et al. (2021, Theorem 3.1.4)); and the last relation holds since $n = \max\{n_1, n_2\}$ and $\sigma\sqrt{n_1} \lesssim \zeta_{\mathsf{op}}/\sigma_1^\natural \ll \sigma_1^\natural$. As a result, we reach

$$\begin{aligned}
\alpha_1 &\lesssim \sigma_m\left(\sigma_1^\natural + \sigma\sqrt{n_2}\right)\left\|\boldsymbol{U}^{(m)}\boldsymbol{H}^{(m)} - \boldsymbol{U}^\natural\right\|_{\mathrm{F}}\sqrt{\log n} \\
&\lesssim \left(\sigma_m\sigma_1^\natural\sqrt{n_1\log n} + \sigma\sigma_m\sqrt{n_1 n_2\log n}\right)\left\|\boldsymbol{U}^{(m)}\boldsymbol{H}^{(m)} - \boldsymbol{U}^\natural\right\|_{2,\infty}.
\end{aligned}$$

Regarding the second term α_2, we know from (C.16) in Lemma 9 that with probability exceeding $1 - O(n^{-10})$,

$$\begin{aligned}
\alpha_2 \le{} & \underbrace{(B_m\log n)\frac{\zeta_{\mathsf{op}}^2}{\sigma_r^{\natural 4}}\left\|\boldsymbol{M}^{\natural\top}\right\|_{2,\infty}}_{=:\beta_1} + \underbrace{(B_m\log n)\left(\sigma\sqrt{n_1\log n} + B\log n\right)\left\|\boldsymbol{U}^{(m)}\boldsymbol{H}^{(m)} - \boldsymbol{U}^\natural\right\|_{2,\infty}}_{=:\beta_2} \\
&+ \underbrace{(B_m\log n)\left(\left\|\boldsymbol{M}^{\natural\top}\right\|_{2,\infty} + B\log n + \sigma\sqrt{n_1\log n}\right)\left\|\boldsymbol{U}^{(m)}\boldsymbol{U}^{(m)\top} - \boldsymbol{U}^{(m,l)}\boldsymbol{U}^{(m,l)\top}\right\|}_{=:\beta_3}
\end{aligned}$$

holds for all $m \in [n_1]$. In what follows, we shall bound β_1, β_2 and β_3 respectively.

- Regarding β_1, we first observe that

$$(B_m\log n)\left\|\boldsymbol{M}^{\natural\top}\right\|_{2,\infty} \le (B_m\log n)\sqrt{\frac{\mu^\natural r}{n_2}}\sigma_1^\natural \lesssim \sigma_m\sqrt{n_2\log n}\sqrt{\frac{\mu^\natural r}{n_2}}\sigma_1^\natural \lesssim \sigma_m\sigma_1^\natural\sqrt{\mu^\natural r\log n}, \qquad \text{(C.53)}$$

  where we have used the noise condition $B_m \lesssim \sigma_m\sqrt{n_2/\log n}$. Therefore

$$\beta_1 \lesssim \sigma_m\sigma_1^\natural\sqrt{n_1\log n}\frac{\zeta_{\mathsf{op}}^2}{\sigma_r^{\natural 4}}\sqrt{\frac{\mu^\natural r}{n_1}}.$$

- When it comes to β_2, we notice that

$$(B_m\log n)\left(\sigma\sqrt{n_1\log n} + B\log n\right) \lesssim \sigma_m\sigma\sqrt{n_1 n_2}\log n. \qquad \text{(C.54)}$$

  Here, we have used the noise assumption $B \lesssim \sigma\sqrt[4]{n_1 n_2}/\sqrt{\log n}$ and $B_m \lesssim \sigma_m\min\{\sqrt[4]{n_1 n_2}, \sqrt{n_2}\}/\sqrt{\log n}$. This in turn leads us to

$$\beta_2 \le \sigma_m\sigma\sqrt{n_1 n_2}\log n\left\|\boldsymbol{U}^{(m)}\boldsymbol{H}^{(m)} - \boldsymbol{U}^\natural\right\|_{2,\infty}.$$

- We are left with the term β_3. From Lemma 8, we see that with probability exceeding $1 - O(n^{-10})$,

$$\begin{aligned}
&(B_m\log n)\left\|\boldsymbol{U}^{(m)}\boldsymbol{U}^{(m)\top} - \boldsymbol{U}^{(m,l)}\boldsymbol{U}^{(m,l)\top}\right\| \\
&\lesssim \frac{B_m\log n}{\sigma_r^{\natural 2}}\left[\left(B\log n + \sigma\sqrt{n_1\log n}\right)^2\left\|\boldsymbol{U}^{(m)}\boldsymbol{H}^{(m)}\right\|_{2,\infty} + \sigma^2\right] + \frac{B_m\log n}{\sigma_r^{\natural 2}}\left(B\log n + \sigma\sqrt{n_1\log n}\right)\left\|\boldsymbol{M}^{\natural\top}\right\|_{2,\infty} \\
&\lesssim \frac{B_m\log n}{\sigma_r^{\natural 2}}\left[\zeta_{\mathsf{op}}\left\|\boldsymbol{U}^{(m)}\boldsymbol{H}^{(m)}\right\|_{2,\infty} + \sigma^2\right] + \left(B\log n + \sigma\sqrt{n_1\log n}\right)\frac{\sigma_m\sigma_1^\natural}{\sigma_r^{\natural 2}}\sqrt{\mu^\natural r\log n}
\end{aligned}$$

  simultaneously for all m and l, where the penultimate inequality follows from (C.53) and

$$\left(\sigma\sqrt{n_1\log n} + B\log n\right)^2 \asymp \sigma^2 n_1\log n + B^2\log^2 d \lesssim \sigma^2 n_1\log n + \sigma^2\sqrt{n_1 n_2}\log n \lesssim \zeta_{\mathsf{op}}. \qquad \text{(C.55)}$$

Therefore, with probability exceeding $1-O(n^{-10})$, we have

$$
\begin{aligned}
\beta_3 &\lesssim \left(\left\|\boldsymbol{M}^{\natural\top}\right\|_{2,\infty}+B\log n+\sigma\sqrt{n_1\log n}\right)\frac{B_m\log n}{\sigma_r^{\natural 2}}\left[\zeta_{\mathsf{op}}\left\|\boldsymbol{U}^{(m)}\boldsymbol{H}^{(m)}\right\|_{2,\infty}+\sigma^2\right] \\
&\quad+\left(\left\|\boldsymbol{M}^{\natural\top}\right\|_{2,\infty}+B\log n+\sigma\sqrt{n_1\log n}\right)\left(B\log n+\sigma\sqrt{n_1\log n}\right)\frac{\sigma_m\sigma_1^{\natural}}{\sigma_r^{\natural 2}}\sqrt{\mu^{\natural}r\log n} \\
&\lesssim \left(\left\|\boldsymbol{M}^{\natural\top}\right\|_{2,\infty}\frac{B_m\log n}{\sigma_r^{\natural 2}}+\frac{B_m\log n\left(B\log n+\sigma\sqrt{n_1\log n}\right)}{\sigma_r^{\natural 2}}\right)\left[\zeta_{\mathsf{op}}\left\|\boldsymbol{U}^{(m)}\boldsymbol{H}^{(m)}\right\|_{2,\infty}+\sigma^2\right] \\
&\quad+\left\|\boldsymbol{M}^{\natural\top}\right\|_{2,\infty}\left(B\log n+\sigma\sqrt{n_1\log n}\right)\frac{\sigma_m\sigma_1^{\natural}}{\sigma_r^{\natural 2}}\sqrt{\mu^{\natural}r\log n}+\left(B\log n+\sigma\sqrt{n_1\log n}\right)^2\frac{\sigma_m\sigma_1^{\natural}}{\sigma_r^{\natural 2}}\sqrt{\mu^{\natural}r\log n} \\
&\overset{\text{(i)}}{\lesssim} \left(\sigma_m\sigma_1^{\natural}\sqrt{\mu^{\natural}r\log n}+\sigma_m\sigma\sqrt{n_1n_2}\log n\right)\left(\frac{\zeta_{\mathsf{op}}}{\sigma_r^{\natural 2}}\left\|\boldsymbol{U}^{(m)}\boldsymbol{H}^{(m)}\right\|_{2,\infty}+\frac{\sigma^2}{\sigma_r^{\natural 2}}\right) \\
&\quad+\left\|\boldsymbol{M}^{\natural\top}\right\|_{2,\infty}\left(B\log n+\sigma\sqrt{n_1\log n}\right)\frac{\sigma_m\sigma_1^{\natural}}{\sigma_r^{\natural 2}}\sqrt{\mu^{\natural}r\log n}+\frac{\zeta_{\mathsf{op}}}{\sigma_r^{\natural 2}}\sigma_m\sigma_1^{\natural}\sqrt{n_1\log n}\sqrt{\frac{\mu^{\natural}r}{n_1}} \\
&\overset{\text{(ii)}}{\lesssim} \left(\sigma_m\sigma_1^{\natural}\sqrt{\mu^{\natural}r\log n}+\sigma_m\sigma\sqrt{n_1n_2}\log n\right)\left(\frac{\zeta_{\mathsf{op}}}{\sigma_r^{\natural 2}}\left\|\boldsymbol{U}^{(m)}\boldsymbol{H}^{(m)}\right\|_{2,\infty}+\frac{\sigma^2}{\sigma_r^{\natural 2}}\right) \\
&\quad+\left(\frac{\mu^{\natural}r}{n_1}+\frac{\mu^{\natural}r}{\sqrt{n_1n_2}}\right)\frac{\sigma\sigma_1^{\natural}\sqrt{n_1\log n}\sigma_m\sigma_1^{\natural}\sqrt{n_1\log n}}{\sigma_r^{\natural 2}}+\frac{\zeta_{\mathsf{op}}}{\sigma_r^{\natural 2}}\sigma_m\sigma_1^{\natural}\sqrt{n_1\log n}\sqrt{\frac{\mu^{\natural}r}{n_1}} \\
&\overset{\text{(iii)}}{\lesssim} \left(\sigma_m\sigma_1^{\natural}\sqrt{\mu^{\natural}r\log n}+\sigma_m\sigma\sqrt{n_1n_2}\log n\right)\frac{\zeta_{\mathsf{op}}}{\sigma_r^{\natural 2}}\left\|\boldsymbol{U}^{(m)}\boldsymbol{H}^{(m)}\right\|_{2,\infty}+\frac{\zeta_{\mathsf{op}}}{\sigma_r^{\natural 2}}\sigma_m\sigma_1^{\natural}\sqrt{n_1\log n}\sqrt{\frac{\mu^{\natural}r}{n_1}}
\end{aligned}
$$

simultaneously for all $m\in[n_1]$, where (i) uses (C.53), (C.54) and (C.55); (ii) comes from

$$
\begin{aligned}
\left\|\boldsymbol{M}^{\natural\top}\right\|_{2,\infty}\left(B\log n+\sigma\sqrt{n_1\log n}\right)\frac{\sigma_m\sigma_1^{\natural}}{\sigma_r^{\natural 2}}\sqrt{\mu^{\natural}r\log n} &\lesssim \sqrt{\frac{\mu^{\natural}r}{n_2}}\sigma_1^{\natural}\sigma\sqrt{n\log n}\frac{\sigma_m\sigma_1^{\natural}}{\sigma_r^{\natural 2}}\sqrt{\mu^{\natural}r\log n} \\
&\lesssim \left(\frac{\mu^{\natural}r}{n_1}+\frac{\mu^{\natural}r}{\sqrt{n_1n_2}}\right)\frac{\sigma\sigma_1^{\natural}\sqrt{n_1\log n}\sigma_m\sigma_1^{\natural}\sqrt{n_1\log n}}{\sigma_r^{\natural 2}}
\end{aligned}
$$

where we use $B\lesssim\sigma\sqrt{n_2/\log n}$; and (iii) holds provided that $\zeta_{\mathsf{op}}\ll\sigma_r^{\natural 2}$, $n_1\gtrsim\mu^{\natural}r$ and $n_2\gtrsim\mu^{\natural}r$.

Putting the above pieces together, we conclude that

$$
\begin{aligned}
&\left\|\boldsymbol{E}_{m,\cdot}\left[\mathcal{P}_{-m,\cdot}\left(\boldsymbol{M}\right)\right]^{\top}\left(\boldsymbol{U}^{(m)}\boldsymbol{H}^{(m)}-\boldsymbol{U}^{\natural}\right)\right\|_2\le\alpha_1+\alpha_2\le\alpha_1+\beta_1+\beta_2+\beta_3 \\
&\quad\lesssim\left(\sigma_m\sigma_1^{\natural}\sqrt{n_1\log n}+\sigma_m\sigma\sqrt{n_1n_2}\log n\right)\left(\left\|\boldsymbol{U}^{(m)}\boldsymbol{H}^{(m)}-\boldsymbol{U}^{\natural}\right\|_{2,\infty}+\frac{\zeta_{\mathsf{op}}}{\sigma_r^{\natural 2}}\left\|\boldsymbol{U}^{(m)}\boldsymbol{H}^{(m)}\right\|_{2,\infty}\right) \\
&\quad+\frac{\zeta_{\mathsf{op}}}{\sigma_r^{\natural 2}}\sigma_m\sigma_1^{\natural}\sqrt{n_1\log n}\sqrt{\frac{\mu^{\natural}r}{n_1}}
\end{aligned}
$$

simultaneously for all $m\in[n_1]$, provided that $\zeta_{\mathsf{op}}\lesssim\sigma_r^{\natural 2}$.

C.3.12 Proof of Lemma 11

In view of the Davis-Kahan $\sin\boldsymbol{\Theta}$ Theorem (Chen et al., 2021, Theorem 2.2.1), we have

$$
\begin{aligned}
\left\|\boldsymbol{U}^{(m)}\boldsymbol{U}^{(m)\top}-\boldsymbol{U}\boldsymbol{U}^{\top}\right\| &\le\frac{\left\|\left(\boldsymbol{G}-\boldsymbol{G}^{(m)}\right)\boldsymbol{U}^{(m)}\right\|}{\lambda_r\left(\boldsymbol{G}^{(m)}\right)-\lambda_{r+1}\left(\boldsymbol{G}\right)}\le\frac{2\left\|\left(\boldsymbol{G}-\boldsymbol{G}^{(m)}\right)\boldsymbol{U}^{(m)}\right\|}{\sigma_r^{\natural 2}} \\
&\le\underbrace{\frac{2\left\|\mathcal{P}_{\mathsf{diag}}\left(\boldsymbol{G}-\boldsymbol{G}^{(m)}\right)\boldsymbol{U}^{(m)}\right\|}{\sigma_r^{\natural 2}}}_{=:\alpha_1}+\underbrace{\frac{2\left\|\mathcal{P}_{\mathsf{off\text{-}diag}}\left(\boldsymbol{G}-\boldsymbol{G}^{(m)}\right)\boldsymbol{U}^{(m)}\right\|}{\sigma_r^{\natural 2}}}_{=:\alpha_2},
\end{aligned}
$$

where the penultimate inequality follows from Weyl's inequality:

$$\lambda_r\left(\boldsymbol{G}^{(m)}\right) \geq \lambda_r\left(\boldsymbol{G}^{\natural}\right) - \left\|\boldsymbol{G}^{(m)} - \boldsymbol{G}^{\natural}\right\| \overset{\text{(i)}}{\geq} \sigma_r^{\natural 2} - \widetilde{C}\zeta_{\mathsf{op}} \overset{\text{(ii)}}{\geq} \frac{3}{4}\sigma_r^{\natural 2},$$

$$\lambda_{r+1}\left(\boldsymbol{G}\right) \leq \lambda_{r+1}\left(\boldsymbol{G}^{\natural}\right) + \left\|\boldsymbol{G} - \boldsymbol{G}^{\natural}\right\| \overset{\text{(iii)}}{\leq} \widetilde{C}\zeta_{\mathsf{op}} \overset{\text{(iv)}}{\leq} \frac{1}{4}\sigma_r^{\natural 2},$$

with $\widetilde{C} > 0$ some absolute constant. Here, (i) comes from Lemma 7; (iii) comes from Lemma 1 and the fact that $\lambda_{r+1}(\boldsymbol{G}^{\natural}) = 0$; (ii) and (iv) are valid provided that $\zeta_{\mathsf{op}} \ll \sigma_r^{\natural 2}$.

Bounding α_1. Recalling that $\mathcal{P}_{\mathsf{diag}}(\boldsymbol{G}^{(m)}) = \mathcal{P}_{\mathsf{diag}}(\boldsymbol{G}^{\natural})$, we obtain

$$\alpha_1 = \frac{2\left\|\mathcal{P}_{\mathsf{diag}}\left(\boldsymbol{G} - \boldsymbol{G}^{\natural}\right)\right\|\left\|\boldsymbol{U}^{(m)}\right\|}{\sigma_r^{\natural 2}} \lesssim \kappa^{\natural 2}\sqrt{\frac{\mu^{\natural} r}{n_1}}\frac{\zeta_{\mathsf{op}}}{\sigma_r^{\natural 2}},$$

where the last inequality follows from Lemma 1.

Bounding α_2. Observe that the symmetric matrix $\mathcal{P}_{\mathsf{off\text{-}diag}}(\boldsymbol{G} - \boldsymbol{G}^{(m)})$ is supported on the m-th row and the m-th column, and

$$\left(\boldsymbol{G} - \boldsymbol{G}^{(m)}\right)_{m,i} = \left(\boldsymbol{G} - \boldsymbol{G}^{(m)}\right)_{i,m} = \boldsymbol{E}_{m,\cdot}\boldsymbol{M}_{i,\cdot}^{\top}, \qquad \forall i \in [n_1]. \tag{C.56}$$

Therefore, we can derive

$$\begin{aligned}
\left\|\mathcal{P}_{\mathsf{off\text{-}diag}}\left(\boldsymbol{G} - \boldsymbol{G}^{(m)}\right)\boldsymbol{U}^{(m)}\right\| &\leq \left\|\mathcal{P}_{\mathsf{off\text{-}diag}}\left(\boldsymbol{G} - \boldsymbol{G}^{(m)}\right)\boldsymbol{U}^{(m)}\right\|_{\mathrm{F}} \overset{\text{(i)}}{\leq} 2\left\|\mathcal{P}_{\mathsf{off\text{-}diag}}\left(\boldsymbol{G} - \boldsymbol{G}^{(m)}\right)\boldsymbol{U}^{(m)}\boldsymbol{H}^{(m)}\right\|_{\mathrm{F}} \\
&\leq 2\left\|\mathcal{P}_{m,\cdot}\left(\boldsymbol{G} - \boldsymbol{G}^{(m)}\right)\boldsymbol{U}^{(m)}\boldsymbol{H}^{(m)}\right\|_{\mathrm{F}} + 2\left\|\mathcal{P}_{\cdot,m}\left(\boldsymbol{G} - \boldsymbol{G}^{(m)}\right)\boldsymbol{U}^{(m)}\boldsymbol{H}^{(m)}\right\|_{\mathrm{F}} \\
&\overset{\text{(ii)}}{=} 2\underbrace{\left\|\boldsymbol{E}_{m,\cdot}\left[\mathcal{P}_{-m,\cdot}\left(\boldsymbol{M}\right)\right]^{\top}\boldsymbol{U}^{(m)}\boldsymbol{H}^{(m)}\right\|_{\mathrm{F}}}_{=:\alpha_{2,1}} + 2\underbrace{\left\|\left(\boldsymbol{G} - \boldsymbol{G}^{(m)}\right)_{m,\cdot}\right\|_2\left\|\boldsymbol{U}_{m,\cdot}^{(m)}\boldsymbol{H}^{(m)}\right\|_2}_{=:\alpha_{2,2}}.
\end{aligned}$$

Here (i) follows from Lemma 2, and (ii) follows from (C.56).

- Regarding α_1, we can invoke (C.19) in Lemma 10 to achieve that with probability exceeding $1 - O(n^{-11})$,

$$\alpha_{2,1} \lesssim \zeta_{\mathsf{op}}\left(\left\|\boldsymbol{U}^{(m)}\boldsymbol{H}^{(m)}\right\|_{2,\infty} + \sqrt{\frac{\mu^{\natural} r}{n_1}}\right).$$

- With regards to α_2, we can invoke Lemma 7 to show that with probability exceeding $1 - O(n^{-11})$,

$$\alpha_{2,2} \leq \left\|\boldsymbol{G} - \boldsymbol{G}^{(m)}\right\|_2\left\|\boldsymbol{U}^{(m)}\boldsymbol{H}^{(m)}\right\|_{2,\infty} \leq \left(\sigma^2\sqrt{n_1 n_2}\log n + \kappa^{\natural 2}\sigma\sigma_1^{\natural}\sqrt{\mu^{\natural} r\log n}\right)\left\|\boldsymbol{U}^{(m)}\boldsymbol{H}^{(m)}\right\|_{2,\infty}.$$

Taking the above bounds on $\alpha_{2,1}$ and $\alpha_{2,2}$ collectively yields

$$\begin{aligned}
&\left\|\mathcal{P}_{\mathsf{off\text{-}diag}}\left(\boldsymbol{G} - \boldsymbol{G}^{(m)}\right)\boldsymbol{U}^{(m)}\right\| \lesssim \alpha_{2,1} + \alpha_{2,2} \\
&\quad\lesssim \zeta_{\mathsf{op}}\left(\left\|\boldsymbol{U}^{(m)}\boldsymbol{H}^{(m)}\right\|_{2,\infty} + \sqrt{\frac{\mu^{\natural} r}{n_1}}\right) + \left(\sigma^2\sqrt{n_1 n_2}\log n + \kappa^{\natural 2}\sigma\sigma_1^{\natural}\sqrt{\mu^{\natural} r\log n}\right)\left\|\boldsymbol{U}^{(m)}\boldsymbol{H}^{(m)}\right\|_{2,\infty} \\
&\quad\lesssim \zeta_{\mathsf{op}}\left(\left\|\boldsymbol{U}^{(m)}\boldsymbol{H}^{(m)}\right\|_{2,\infty} + \sqrt{\frac{\mu^{\natural} r}{n_1}}\right),
\end{aligned}$$

provided that $n_1 \gtrsim \kappa^{\natural 4}\mu^{\natural} r$. Therefore, we arrive at

$$\alpha_2 \lesssim \frac{\zeta_{\mathsf{op}}}{\sigma_r^{\natural 2}}\left(\left\|\boldsymbol{U}^{(m)}\boldsymbol{H}^{(m)}\right\|_{2,\infty} + \sqrt{\frac{\mu^{\natural} r}{n_1}}\right).$$

Combining bounds on α_1 and α_2. The preceding bounds on α_1 and α_2 combined allow one to derive

$$\left\|\boldsymbol{U}^{(m)}\boldsymbol{U}^{(m)\top}-\boldsymbol{U}\boldsymbol{U}^{\top}\right\|\leq\alpha_1+\alpha_2\lesssim\frac{\zeta_{\mathsf{op}}}{\sigma_r^{\natural 2}}\left(\left\|\boldsymbol{U}^{(m)}\boldsymbol{H}^{(m)}\right\|_{2,\infty}+\sqrt{\frac{\mu^{\natural}r}{n_1}}\right)+\kappa^{\natural 2}\sqrt{\frac{\mu^{\natural}r}{n_1}}\frac{\zeta_{\mathsf{op}}}{\sigma_r^{\natural 2}}$$
$$\lesssim\frac{\zeta_{\mathsf{op}}}{\sigma_r^{\natural 2}}\left\|\boldsymbol{U}^{(m)}\boldsymbol{H}^{(m)}\right\|_{2,\infty}+\kappa^{\natural 2}\sqrt{\frac{\mu^{\natural}r}{n_1}}\frac{\zeta_{\mathsf{op}}}{\sigma_r^{\natural 2}}$$

with probability at least $1-O(n^{-11})$. This together with the union bound concludes the proof.

D Analysis for PCA: the approach based on HeteroPCA

In this section, we establish our theoretical guarantees for the approach based on $\mathsf{HeteroPCA}$ — the ones presented in Section 3.2. . We shall establish our inference guarantees by connecting the PCA model with the subspace estimation problem studied in Section 6.

Before embarking on the analysis, we claim that: without loss of generality, we can work with the assumption that

$$\max_{j\in[n]}|\eta_{l,j}|\leq C_{\mathsf{noise}}\omega_l^{\star}\sqrt{\log(n+d)}\qquad\text{for all }l\in[d],\tag{D.1}$$

for some absolute constant $C_{\mathsf{noise}}>0$, in addition to the noise condition we have already imposed in Section 1.1. To see why this is valid, we note that by applying Lemma 31 (with $\delta=C_\delta(n+d)^{-100}$ for some sufficiently small constant $C_\delta>0$) to $\{\eta_{l,j}\}$, we can produce a set of auxiliary random variables $\{\widetilde{\eta}_{l,j}\}$ such that

- the $\widetilde{\eta}_{l,j}$'s are independent sub-Gaussian random variables satisfying
$$\mathbb{E}[\widetilde{\eta}_{l,j}]=0,\qquad\widetilde{\omega}_l^{\star 2}:=\mathbb{E}[\widetilde{\eta}_{l,j}^2]=\left[1+O\left((n+d)^{-50}\right)\right]\omega_l^{\star 2},\qquad\|\widetilde{\eta}_{l,j}\|_{\psi_2}\lesssim\omega_l^{\star},\qquad|\widetilde{\eta}_{l,j}|\lesssim\omega_l^{\star}\sqrt{\log(n+d)};$$
- these auxiliary variables satisfy
$$\mathbb{P}\left(\eta_{l,j}=\widetilde{\eta}_{l,j}\text{ for all }l\in[d]\text{ and }j\in[n]\right)\geq1-O\left((n+d)^{-98}\right).\tag{D.2}$$

The above properties suggest that: if we replace the noise matrix $\boldsymbol{N}=[\eta_{l,j}]_{l\in[d],j\in[n]}$ with $\widetilde{\boldsymbol{N}}=[\widetilde{\eta}_{l,j}]_{l\in[d],j\in[n]}$, then the observations remain unchanged (i.e. $\mathcal{P}_{\Omega}(\boldsymbol{X}+\boldsymbol{N})=\mathcal{P}_{\Omega}(\boldsymbol{X}+\widetilde{\boldsymbol{N}})$) with high probability, and consequently, the resulting estimates are also unchanged. In light of this, our analysis proceeds with the following steps.

- We shall start by proving Theorems 11-14 under the additional assumption (D.1); if this can be accomplished, then the results are clearly valid if we replace $\boldsymbol{N}$ (resp. $\{\omega_l^{\star}\}_{l\in[d]}$) with $\widetilde{\boldsymbol{N}}$ (resp. $\{\widetilde{\omega}_l^{\star}\}_{l=1}^{d}$).
- Then we can invoke the proximity of $\widetilde{\omega}_l^{\star}$ and $\omega_l^{\star}$ as well as (D.2) to establish these theorems without assuming (D.1).

D.1 Connection between PCA and subspace estimation

In order to invoke our theoretical guarantees for subspace estimation (i.e., Theorem 5) to assist in understanding PCA, we need to establish an explicit connection between these two models. This forms the main content of this subsection.

Specification of $\boldsymbol{M}^{\natural}$, $\boldsymbol{M}$ and $\boldsymbol{E}$. Recall that in our PCA model in Section 1.1, we assume the covariance matrix $\boldsymbol{S}^{\star}$ admits the eigen-decomposition $\boldsymbol{U}^{\star}\boldsymbol{\Sigma}^{\star 2}\boldsymbol{U}^{\star\top}$. As a result, the matrix $\boldsymbol{X}$ (cf. (2.1a)) can be equivalently expressed as

$$\boldsymbol{X}=\boldsymbol{U}^{\star}\boldsymbol{\Sigma}^{\star}\left[\boldsymbol{f}_1,\ldots,\boldsymbol{f}_n\right]=\boldsymbol{U}^{\star}\boldsymbol{\Sigma}^{\star}\boldsymbol{F},\qquad\text{where}\qquad\boldsymbol{f}_i\overset{\text{i.i.d.}}{\sim}\mathcal{N}\left(\boldsymbol{0},\boldsymbol{I}_r\right).$$

Recalling that our observation matrix is $\boldsymbol{Y} = \mathcal{P}_\Omega(\boldsymbol{X}+\boldsymbol{N})$, we can specify the matrices $\boldsymbol{M}^\natural$, $\boldsymbol{M}$ and $\boldsymbol{E}$ as defined in Appendix 6.1 as follows:

$$\boldsymbol{M} \coloneqq \frac{1}{\sqrt{np}}\boldsymbol{Y}, \qquad \boldsymbol{M}^\natural \coloneqq \mathbb{E}\left[\boldsymbol{M}\middle|\boldsymbol{F}\right] = \frac{1}{\sqrt{n}}\boldsymbol{U}^\star\boldsymbol{\Sigma}^\star\boldsymbol{F} \qquad \text{and} \qquad \boldsymbol{E} \coloneqq \boldsymbol{M}-\boldsymbol{M}^\natural. \tag{D.3}$$

As usual, we shall let the SVD of $\boldsymbol{M}^\natural$ be $\boldsymbol{M}^\natural = \boldsymbol{U}^\natural\boldsymbol{\Sigma}^\natural\boldsymbol{V}^{\natural\top}$, and use $\kappa^\natural, \mu^\natural, \sigma_r^\natural, \sigma_1^\natural$ to denote the conditional number, the incoherence parameter, the minimum and maximum singular value of $\boldsymbol{M}^\natural$, respectively. Here and below, we shall focus on the randomness of noise and missing data while treating $\boldsymbol{F}$ as given (even though it is generated randomly).

We shall also specify several useful relations between $(\boldsymbol{U}^\natural, \boldsymbol{V}^\natural, \boldsymbol{R}_{\boldsymbol{U}})$ and $(\boldsymbol{U}^\star, \boldsymbol{V}^\star, \boldsymbol{R})$, where the rotation matrix $\boldsymbol{R}$ is defined as $\boldsymbol{R} = \arg\min_{\boldsymbol{O}\in\mathcal{O}^{r\times r}} \|\boldsymbol{U}\boldsymbol{O}-\boldsymbol{U}^\star\|_{\mathrm{F}}$. First, observe that

$$\boldsymbol{V}^\natural = \boldsymbol{M}^{\natural\top}\boldsymbol{U}^\natural\left(\boldsymbol{\Sigma}^\natural\right)^{-1} = \frac{1}{\sqrt{n}}\boldsymbol{F}^\top\underbrace{\boldsymbol{\Sigma}^\star\boldsymbol{U}^{\star\top}\boldsymbol{U}^\natural\left(\boldsymbol{\Sigma}^\natural\right)^{-1}}_{=:\,\boldsymbol{J}}. \tag{D.4}$$

Given that $\boldsymbol{U}^\star$ and $\boldsymbol{U}^\natural$ represent the same column space, there exists $\boldsymbol{Q}\in\mathcal{O}^{r\times r}$ such that

$$\boldsymbol{U}^\natural = \boldsymbol{U}^\star\boldsymbol{Q}, \tag{D.5}$$

thus leading to the expression

$$\boldsymbol{J} = \boldsymbol{\Sigma}^\star\boldsymbol{Q}(\boldsymbol{\Sigma}^\natural)^{-1}. \tag{D.6}$$

The definition of $\boldsymbol{R}$ together with $\boldsymbol{U}^\natural = \boldsymbol{U}^\star\boldsymbol{Q}$ also allows one to derive

$$\boldsymbol{R}_{\boldsymbol{U}} = \arg\min_{\boldsymbol{O}\in\mathcal{O}^{r\times r}} \|\boldsymbol{U}\boldsymbol{O}-\boldsymbol{U}^\natural\|_{\mathrm{F}} = \arg\min_{\boldsymbol{O}\in\mathcal{O}^{r\times r}} \|\boldsymbol{U}\boldsymbol{O}\boldsymbol{Q}^\top-\boldsymbol{U}^\star\|_{\mathrm{F}} = \boldsymbol{R}\boldsymbol{Q}. \tag{D.7}$$

Statistical properties of $\boldsymbol{E} = [E_{i,j}]$. We now describe the statistical properties of the perturbation matrix $\boldsymbol{E}$. To begin with, it is readily seen from the definition of $\boldsymbol{E}$ that $\mathbb{E}[\boldsymbol{E}\mid\boldsymbol{F}] = \boldsymbol{0}$. Moreover, from the definition of $\boldsymbol{M}$ and $\boldsymbol{M}^\natural$, it is observed that

$$\boldsymbol{E} = \boldsymbol{M}-\boldsymbol{M}^{\natural\top} = \frac{1}{\sqrt{n}}\left[\frac{1}{p}\mathcal{P}_\Omega\left(\boldsymbol{U}^\star\boldsymbol{\Sigma}^\star\boldsymbol{F}+\boldsymbol{N}\right)-\boldsymbol{U}^\star\boldsymbol{\Sigma}^\star\boldsymbol{F}\right] = \frac{1}{\sqrt{n}}\left[\frac{1}{p}\mathcal{P}_\Omega\left(\boldsymbol{U}^\star\boldsymbol{\Sigma}^\star\boldsymbol{F}\right)-\boldsymbol{U}^\star\boldsymbol{\Sigma}^\star\boldsymbol{F}\right]+\frac{1}{\sqrt{n}}\frac{1}{p}\mathcal{P}_\Omega\left(\boldsymbol{N}\right), \tag{D.8}$$

which is clearly a zero-mean matrix conditional on $\boldsymbol{F}$. In addition, for location (i,j), we have

$$M_{i,j}^\natural = \frac{1}{\sqrt{n}}\left(\boldsymbol{U}^\star\boldsymbol{\Sigma}^\star\boldsymbol{F}\right)_{i,j} = \frac{1}{\sqrt{n}}\left(\boldsymbol{U}^\star\boldsymbol{\Sigma}^\star\boldsymbol{F}^\star\right)_{i,j} = \frac{1}{\sqrt{n}}\boldsymbol{U}_{i,\cdot}^\star\boldsymbol{\Sigma}^\star\boldsymbol{f}_j,$$

and hence the variance of $E_{i,j}$ can be calculated as

$$\sigma_{i,j}^2 \coloneqq \mathsf{var}\left(E_{i,j}|\boldsymbol{F}\right) = \mathbb{E}\left[E_{i,j}^2|\boldsymbol{F}\right] = \frac{1-p}{np}\left(\boldsymbol{U}^\star\boldsymbol{\Sigma}^\star\boldsymbol{F}\right)_{i,j}^2+\frac{\omega_i^{\star 2}}{np} = \frac{1-p}{np}\left(\boldsymbol{U}_{i,\cdot}^\star\boldsymbol{\Sigma}^\star\boldsymbol{f}_j\right)^2+\frac{\omega_i^{\star 2}}{np}. \tag{D.9}$$

A good event $\mathcal{E}_{\mathsf{good}}$. Finally, the lemma below defines a high-probability event $\mathcal{E}_{\mathsf{good}}$ under which the random quantities defined above enjoy appealing properties. The proof can be found in Appendix E.1.

Lemma 14. *There is an event $\mathcal{E}_{\mathsf{good}}$ with $\mathbb{P}(\mathcal{E}_{\mathsf{good}}) \geq 1-O\big((n+d)^{-10}\big)$, on which the following properties hold.*

- *$\mathcal{E}_{\mathsf{good}}$ is $\sigma(\boldsymbol{F})$-measurable, where $\sigma(\boldsymbol{F})$ is the σ-algebra generated by $\boldsymbol{F}$.*
- *If $n \gg \kappa^2(r+\log(n+d))$, then one has*

$$\left\|\frac{1}{n}\boldsymbol{F}\boldsymbol{F}^\top-\boldsymbol{I}_r\right\| \lesssim \sqrt{\frac{r+\log\left(n+d\right)}{n}}, \tag{D.10}$$

$$\left\|\boldsymbol{\Sigma}^{\natural}-\boldsymbol{\Sigma}^{\star}\right\| \lesssim \kappa\sqrt{\frac{r+\log(n+d)}{n}}\sigma_r^{\star}, \tag{D.11}$$

$$\left\|\boldsymbol{\Sigma}^{\natural 2}-\boldsymbol{\Sigma}^{\star 2}\right\| \lesssim \sqrt{\frac{r+\log(n+d)}{n}}\sigma_1^{\star 2}, \tag{D.12}$$

$$\sigma_r^{\natural} \asymp \sigma_r^{\star} \qquad \text{and} \qquad \sigma_1^{\natural} \asymp \sigma_1^{\star}. \tag{D.13}$$

- *The conditional number $\kappa^{\natural}$ and the incoherence parameter $\mu^{\natural}$ of $\boldsymbol{M}^{\natural}$ obey*

$$\kappa^{\natural} \asymp \sqrt{\kappa}, \tag{D.14}$$

$$\mu^{\natural} \lesssim \kappa\mu\log(n+d). \tag{D.15}$$

In addition, we have the following $\ell_{2,\infty}$ norm bound for $\boldsymbol{U}^{\natural}$ and $\boldsymbol{V}^{\natural}$:

$$\left\|\boldsymbol{U}^{\natural}\right\|_{2,\infty} \leq \sqrt{\frac{\mu r}{d}} \qquad \text{and} \qquad \left\|\boldsymbol{V}^{\natural}\right\|_{2,\infty} \lesssim \sqrt{\frac{r\log(n+d)}{n}}. \tag{D.16}$$

- *The noise levels $\{\sigma_{i,j}\}$ are upper bounded by*

$$\sigma^2 \coloneqq \max_{i\in[d],j\in[n]}\sigma_{i,j}^2 \lesssim \frac{\mu r\log(n+d)}{ndp}\sigma_1^{\star 2}+\frac{\omega_{\max}^2}{np} \eqqcolon \sigma_{\mathsf{ub}}^2, \tag{D.17}$$

$$\max_{i\in[d],j\in[n]}\left|E_{i,j}\right| \lesssim \frac{1}{p}\sqrt{\frac{\mu r\log(n+d)}{nd}}\sigma_1^{\star}+\frac{\omega_{\max}}{p}\sqrt{\frac{\log(n+d)}{n}} \eqqcolon B. \tag{D.18}$$

In addition, for each $i\in[d]$,

$$\sigma_i^2 \coloneqq \max_{j\in[n]}\sigma_{i,j}^2 \lesssim \frac{\log(n+d)}{np}\left\|\boldsymbol{U}_{i,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2^2+\frac{\omega_i^{\star 2}}{np} \coloneqq \sigma_{\mathsf{ub},i}^2$$

and

$$\max_{j\in[n]}\left|E_{i,j}\right| \lesssim \frac{1}{p}\sqrt{\frac{\log(n+d)}{n}}\left\|\boldsymbol{U}_{i,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2+\frac{\omega_l^{\star}}{p}\sqrt{\frac{\log(n+d)}{n}} \eqqcolon B_i.$$

- *The matrices $\boldsymbol{J}$ and $\boldsymbol{Q}$ are close in the sense that*

$$\left\|\boldsymbol{Q}-\boldsymbol{J}\right\| \lesssim \frac{1}{\sigma_r^{\star}}\left\|\boldsymbol{Q}\boldsymbol{\Sigma}^{\natural}-\boldsymbol{\Sigma}^{\star}\boldsymbol{Q}\right\| \lesssim \kappa\sqrt{\frac{r+\log(n+d)}{n}}. \tag{D.19}$$

- *For each $i\in[d]$, one has*

$$\max_{j\in[n]}\left|\boldsymbol{U}_{i,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\boldsymbol{f}_j\right| \lesssim \left\|\boldsymbol{U}_{i,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2\sqrt{\log(n+d)}; \tag{D.20}$$

for each $i,l\in[d]$, we have

$$\left\|\frac{1}{n}\sum_{j=1}^{n}\left(\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\boldsymbol{f}_j\right)^2\boldsymbol{f}_j\boldsymbol{f}_j^{\top}-\left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2^2\boldsymbol{I}_r-2\boldsymbol{\Sigma}^{\star}\boldsymbol{U}_{l,\cdot}^{\star\top}\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\| \lesssim \sqrt{\frac{r\log^3(n+d)}{n}}\left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2^2, \tag{D.21a}$$

$$\left|\frac{1}{n}\sum_{j=1}^{n}\left(\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\boldsymbol{f}_j\right)^2\left(\boldsymbol{U}_{i,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\boldsymbol{f}_j\right)^2-\left(S_{l,l}^{\star}S_{i,i}^{\star}+2S_{i,l}^{\star 2}\right)\right| \lesssim \sqrt{\frac{\log(n+d)^3}{n}}\left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2^2\left\|\boldsymbol{U}_{i,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2^2 \tag{D.21b}$$

$$\text{and} \qquad \left|\frac{1}{n}\sum_{j=1}^{n}\left(\boldsymbol{U}_{i,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\boldsymbol{f}_j\right)^2-S_{i,i}^{\star}\right| \lesssim \sqrt{\frac{\log(n+d)}{n}}\left\|\boldsymbol{U}_{i,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2^2. \tag{D.21c}$$

Proof. See Appendix E.1. □

D.2 Distributional characterization for principal subspace (Proof of Theorem 11)

With the above connection between the subspace estimation model and PCA in place, we can readily move on to invoke Theorem 5 to establish our distributional characterization of $\boldsymbol{U}\boldsymbol{R}-\boldsymbol{U}^{\star}$ for HeteroPCA (as stated in Theorem 11).

Step 1: first- and second-order approximation and the tightness. As a starting point, Theorem 5 taken together with the explicit connection between the subspace estimation model and PCA allows one to approximate $\boldsymbol{U}\boldsymbol{R}-\boldsymbol{U}^{\star}$ in a concise form, which is a crucial first step that enables our subsequent development of the distributional theory. The proof can be found in Appendix E.2.1.

Lemma 15. *Assume that $d \gtrsim \kappa^3\mu^2 r\log^4(n+d)$, $n \gtrsim r\log^4(n+d)$, $ndp^2 \gg \kappa^4\mu^2r^2\log^4(n+d)$, $np \gg \kappa^4\mu r\log^2(n+d)$,*

$$\frac{\omega_{\max}^2}{p\sigma_r^{\star 2}}\sqrt{\frac{d}{n}} \ll \frac{1}{\kappa\log(n+d)} \qquad \text{and} \qquad \frac{\omega_{\max}}{\sigma_r^{\star}}\sqrt{\frac{d}{np}} \ll \frac{1}{\sqrt{\kappa^3\log(n+d)}}.$$

Then for each $l \in [d]$, we have

$$\mathbb{P}\left(\left\|(\boldsymbol{U}\boldsymbol{R}-\boldsymbol{U}^{\star}-\boldsymbol{Z})_{l,\cdot}\right\|_2 \mathbb{1}_{\mathcal{E}_{\mathsf{good}}} \lesssim \zeta_{\mathsf{2nd},l} \,\middle|\, \boldsymbol{F}\right) \ge 1-O\left((n+d)^{-10}\right)$$

almost surely, where

$$\boldsymbol{Z} \coloneqq \left[\boldsymbol{E}\boldsymbol{M}^{\natural\top} + \mathcal{P}_{\mathsf{off\text{-}diag}}\left(\boldsymbol{E}\boldsymbol{E}^{\top}\right)\right]\boldsymbol{U}^{\natural}\left(\boldsymbol{\Sigma}^{\natural}\right)^{-2}\boldsymbol{Q}^{\top} \tag{D.22}$$

and

$$\zeta_{\mathsf{2nd},l} \coloneqq \left\|\boldsymbol{U}_{l,\cdot}^{\star}\right\|_2\left(\sqrt{\frac{\kappa^3\mu r\log(n+d)}{d}}\frac{\zeta_{\mathsf{1st}}}{\sigma_r^{\star 2}} + \kappa\frac{\zeta_{\mathsf{1st}}^2}{\sigma_r^{\star 4}}\right) + \frac{\zeta_{\mathsf{1st}}\zeta_{\mathsf{1st},l}}{\sigma_r^{\star 4}}\sqrt{\frac{\kappa^3\mu r\log(n+d)}{d}}, \tag{D.23a}$$

$$\zeta_{\mathsf{1st}} \coloneqq \frac{\mu r\log^2(n+d)}{\sqrt{ndp}}\sigma_1^{\star 2} + \frac{\omega_{\max}^2}{p}\sqrt{\frac{d}{n}}\log(n+d) + \sigma_1^{\star 2}\sqrt{\frac{\mu r}{np}}\log(n+d) + \sigma_1^{\star}\omega_{\max}\sqrt{\frac{d\log(n+d)}{np}}, \tag{D.23b}$$

$$\begin{aligned}\zeta_{\mathsf{1st},l} \coloneqq &\sqrt{\frac{\mu r\log^4(n+d)}{np^2}}\sigma_1^{\star}\left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2 + \frac{\omega_l^{\star}\omega_{\max}}{p}\sqrt{\frac{d}{n}}\log(n+d) + \sigma_1^{\star}\sqrt{\frac{d}{np}}\left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2\log(n+d)\\ &+ \sigma_1^{\star}\omega_l^{\star}\sqrt{\frac{d\log(n+d)}{np}} + \frac{\omega_{\max}}{p}\sqrt{\frac{d}{n}}\log^{3/2}(n+d)\left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2 + \frac{\omega_l^{\star}}{p}\sqrt{\frac{\mu r\log^3(n+d)}{n}}\sigma_1^{\star}.\end{aligned} \tag{D.23c}$$

In particular, we expect each $\zeta_{\mathsf{2nd},l}$ to be negligible, so that the approximation $\boldsymbol{U}\boldsymbol{R}-\boldsymbol{U}^{\star} \approx \boldsymbol{Z}$ is nearly tight. It is worth noting that the approximation $\boldsymbol{Z}$ consists of both linear and second-order effects of the perturbation matrix $\boldsymbol{E}$.

Step 2: computing the covariance of the first- and second-order approximation. In order to pin down the distribution of $\boldsymbol{Z}$, an important step lies in characterizing its covariance. To be precise, observe that the l-th row of $\boldsymbol{Z}$ $(1 \le l \le d)$ defined in (D.22) satisfies

$$\begin{aligned}\boldsymbol{Z}_{l,\cdot} &= \boldsymbol{e}_l^{\top}\left[\boldsymbol{E}\boldsymbol{M}^{\natural\top} + \mathcal{P}_{\mathsf{off\text{-}diag}}\left(\boldsymbol{E}\boldsymbol{E}^{\top}\right)\right]\boldsymbol{U}^{\natural}(\boldsymbol{\Sigma}^{\natural})^{-2}\boldsymbol{Q}^{\top}\\ &= \boldsymbol{E}_{l,\cdot}\boldsymbol{V}^{\natural}(\boldsymbol{\Sigma}^{\natural})^{-1} + \boldsymbol{E}_{l,\cdot}\left[\mathcal{P}_{-l,\cdot}(\boldsymbol{E})\right]^{\top}\boldsymbol{U}^{\natural}(\boldsymbol{\Sigma}^{\natural})^{-2}\boldsymbol{Q}^{\top}\\ &= \sum_{j=1}^{n}E_{l,j}\left[\boldsymbol{V}_{j,\cdot}^{\natural}(\boldsymbol{\Sigma}^{\natural})^{-1} + \left[\mathcal{P}_{-l,\cdot}(\boldsymbol{E}_{\cdot,j})\right]^{\top}\boldsymbol{U}^{\natural}(\boldsymbol{\Sigma}^{\natural})^{-2}\right]\boldsymbol{Q}^{\top},\end{aligned} \tag{D.24}$$

where we recall that $\mathcal{P}_{-l,\cdot}(\boldsymbol{E})$ is obtained by zeroing out the l-th row of $\boldsymbol{E}$. Conditional on $\boldsymbol{F}$, the covariance matrix of this zero-mean random vector $\boldsymbol{Z}_{l,\cdot}$ can thus be calculated as

$$\begin{aligned}\widetilde{\boldsymbol{\Sigma}}_l &\coloneqq \boldsymbol{Q}(\boldsymbol{\Sigma}^\natural)^{-1}\boldsymbol{V}^{\natural\top}\mathsf{diag}\left\{\sigma_{l,1}^2,\dots,\sigma_{l,n}^2\right\}\boldsymbol{V}^\natural(\boldsymbol{\Sigma}^\natural)^{-1}\boldsymbol{Q}^\top \\ &\quad + \boldsymbol{Q}(\boldsymbol{\Sigma}^\natural)^{-2}\boldsymbol{U}^{\natural\top}\mathsf{diag}\Bigg\{\sum_{j:j\neq l}\sigma_{l,j}^2\sigma_{1,j}^2,\dots,\sum_{j:j\neq l}\sigma_{l,j}^2\sigma_{d,j}^2\Bigg\}\boldsymbol{U}^\natural(\boldsymbol{\Sigma}^\natural)^{-2}\boldsymbol{Q}^\top. \end{aligned}\tag{D.25}$$

Given that the above expression of $\widetilde{\boldsymbol{\Sigma}}_l$ contains components like $\boldsymbol{\Sigma}^\natural$ and $\boldsymbol{V}^\natural$ (which are introduced in order to use the subspace estimation model), it is natural to see whether one can express $\widetilde{\boldsymbol{\Sigma}}_l$ directly in terms of the corresponding quantities introduced for the PCA model. To do so, we first single out a deterministic matrix as follows

$$\boldsymbol{\Sigma}_{U,l}^\star \coloneqq \left(\frac{1-p}{np}\left\|\boldsymbol{U}_{l,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2^2+\frac{\omega_l^{\star 2}}{np}\right)(\boldsymbol{\Sigma}^\star)^{-2}+\frac{2\,(1-p)}{np}\boldsymbol{U}_{l,\cdot}^{\star\top}\boldsymbol{U}_{l,\cdot}^\star+(\boldsymbol{\Sigma}^\star)^{-2}\boldsymbol{U}^{\star\top}\mathsf{diag}\left\{\left[d_{l,i}^\star\right]_{i=1}^d\right\}\boldsymbol{U}^\star(\boldsymbol{\Sigma}^\star)^{-2},\tag{D.26}$$

where we define

$$d_{l,i}^\star \coloneqq \frac{1}{np^2}\left[\omega_l^{\star 2}+(1-p)\left\|\boldsymbol{U}_{l,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2^2\right]\left[\omega_i^{\star 2}+(1-p)\left\|\boldsymbol{U}_{i,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2^2\right]+\frac{2\,(1-p)^2}{np^2}S_{i,l}^{\star 2}.\tag{D.27}$$

In view of the following lemma, $\boldsymbol{\Sigma}_{U,l}^\star$ approximates $\widetilde{\boldsymbol{\Sigma}}_l$ in a reasonably well fashion.

Lemma 16. *Suppose that $n \gg \kappa^8\mu^2r^3\kappa_\omega^2\log^3(n+d)$. On the event $\mathcal{E}_{\mathsf{good}}$ (cf. Lemma 14) , we have*

$$\begin{aligned}\left\|\widetilde{\boldsymbol{\Sigma}}_l-\boldsymbol{\Sigma}_{U,l}^\star\right\| &\lesssim \sqrt{\frac{\kappa^8\mu^2r^3\kappa_\omega^2\log^3(n+d)}{n}}\lambda_{\min}\left(\boldsymbol{\Sigma}_{U,l}^\star\right), \\ \max\left\{\lambda_{\max}\left(\widetilde{\boldsymbol{\Sigma}}_l\right),\lambda_{\max}\left(\boldsymbol{\Sigma}_{U,l}^\star\right)\right\} &\lesssim \frac{1-p}{np\sigma_r^{\star 2}}\left\|\boldsymbol{U}_{l,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2^2+\frac{\omega_l^{\star 2}}{np\sigma_r^{\star 2}}+\frac{\kappa\mu r\,(1-p)^2}{ndp^2\sigma_r^{\star 2}}\left\|\boldsymbol{U}_{l,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2^2+\frac{\kappa\mu r\,(1-p)}{ndp^2\sigma_r^{\star 2}}\omega_l^{\star 2} \\ &\quad+\frac{1-p}{np^2\sigma_r^{\star 4}}\omega_{\max}^2\left\|\boldsymbol{U}_{l,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2^2+\frac{\omega_l^{\star 2}\omega_{\max}^2}{np^2\sigma_r^{\star 4}}, \\ \min\left\{\lambda_{\min}\left(\widetilde{\boldsymbol{\Sigma}}_l\right),\lambda_{\min}\left(\boldsymbol{\Sigma}_{U,l}^\star\right)\right\} &\gtrsim \frac{1-p}{np\sigma_1^{\star 2}}\left\|\boldsymbol{U}_{l,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2^2+\frac{\omega_l^{\star 2}}{np\sigma_1^{\star 2}}+\frac{(1-p)^2}{ndp^2\kappa\sigma_1^{\star 2}}\left\|\boldsymbol{U}_{l,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2^2+\frac{1-p}{ndp^2\kappa\sigma_1^{\star 2}}\omega_l^{\star 2} \\ &\quad+\frac{1-p}{np^2\sigma_1^{\star 4}}\omega_{\min}^2\left\|\boldsymbol{U}_{l,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2^2+\frac{\omega_l^{\star 2}\omega_{\min}^2}{np^2\sigma_1^{\star 4}}.\end{aligned}$$

In addition, the condition number of $\widetilde{\boldsymbol{\Sigma}}_l$ is bounded above by $O(\kappa^3\mu r\kappa_\omega)$.

Proof. See Appendix E.2.2. ∎

Step 3: establishing distributional guarantees for $(\boldsymbol{U}\boldsymbol{R}-\boldsymbol{U}^\star)_{l,\cdot}$. By virtue of the decomposition (D.24), each row $\boldsymbol{Z}_{l,\cdot}$ can be viewed as the sum of a collection of independent zero-mean random variables/vectors. This suggests that $\boldsymbol{Z}_{l,\cdot}$ might be well approximated by certain multivariate Gaussian distributions. If so, then the zero-mean nature of $\boldsymbol{Z}_{l,\cdot}$ in conjunction with the above-mentioned covariance formula allows us to pin down the distribution of each row of $\boldsymbol{U}\boldsymbol{R}-\boldsymbol{U}^\star$ approximately.

Lemma 17 (Gaussian approximation of $\boldsymbol{Z}_{l,\cdot}$). *Suppose that the following conditions hold:*

$$n \gtrsim \kappa^8\mu^2r^4\kappa_\omega^2\log^4(n+d), \qquad d \gtrsim \kappa^7\mu^3r^{7/2}\kappa_\omega^2\log^5(n+d),$$
$$np \gtrsim \kappa^9\mu^3r^{11/2}\kappa_\omega^2\log^7(n+d), \qquad ndp^2 \gtrsim \kappa^9\mu^4r^{13/2}\kappa_\omega^2\log^9(n+d),$$

and

$$\frac{\omega_{\max}}{\sigma_r^\star}\sqrt{\frac{d}{np}} \lesssim \frac{1}{\kappa^{3/2}\mu r^{5/4}\kappa_\omega^{1/2}\log^3(n+d)}, \qquad \frac{\omega_{\max}^2}{p\sigma_r^{\star 2}}\sqrt{\frac{d}{n}} \lesssim \frac{1}{\kappa^4\mu^{3/2}r^{11/4}\kappa_\omega\log^{7/2}(n+d)}.$$

Then it is guaranteed that

$$\sup_{\mathcal{C}\in\mathscr{C}^r}\left|\mathbb{P}\left((\boldsymbol{U}\boldsymbol{R}-\boldsymbol{U}^\star)_{l,\cdot}\in\mathcal{C}\right)-\mathcal{N}\left(\boldsymbol{0},\boldsymbol{\Sigma}_{U,l}^\star\right)\{\mathcal{C}\}\right|\lesssim\frac{1}{\sqrt{\log\left(n+d\right)}}=o\left(1\right),$$

where $\mathscr{C}^r$ denotes the set of all convex sets in $\mathbb{R}^r$.

Proof. See Appendix E.2.3. □

Once Lemma 17 is established, we have solidified the advertised Gaussian approximation of each row of $\boldsymbol{U}\boldsymbol{R}-\boldsymbol{U}^\star$, thus concluding the proof of Theorem 11.

D.3 Validity of confidence regions (Proof of Theorem 12)

Our distributional characterization of $(\boldsymbol{U}\boldsymbol{R}-\boldsymbol{U}^\star)_{l,\cdot}$ hints at the possibility of constructing valid confidence region for $\boldsymbol{U}^\star$, provided that the covariance matrix $\boldsymbol{\Sigma}_{U,l}^\star$ (cf. (D.26)) can be reliably estimated. Similar to the SVD-based approach, we attempt to estimate $\boldsymbol{\Sigma}_{U,l}^\star$ by means of the following plug-in estimator:

$$\boldsymbol{\Sigma}_{U,l}\coloneqq\left(\frac{1-p}{np}\left\|\boldsymbol{U}_{l,\cdot}\boldsymbol{\Sigma}\right\|_2^2+\frac{\omega_l^2}{np}\right)\boldsymbol{\Sigma}^{-2}+\frac{2\left(1-p\right)}{np}\boldsymbol{U}_{l,\cdot}^\top\boldsymbol{U}_{l,\cdot}+\left(\boldsymbol{\Sigma}\right)^{-2}\boldsymbol{U}^\top\mathsf{diag}\left\{\left[d_{l,i}\right]_{1\leq i\leq d}\right\}\boldsymbol{U}\left(\boldsymbol{\Sigma}\right)^{-2},\tag{D.28}$$

where for each $i\in[d]$, we define

$$d_{l,i}\coloneqq\frac{1}{np^2}\left[\omega_l^2+(1-p)\left\|\boldsymbol{U}_{l,\cdot}\boldsymbol{\Sigma}\right\|_2^2\right]\left[\omega_i^2+(1-p)\left\|\boldsymbol{U}_{i,\cdot}\boldsymbol{\Sigma}\right\|_2^2\right]+\frac{2\left(1-p\right)^2}{np^2}S_{l,i}^2,\tag{D.29}$$

$$\omega_i^2\coloneqq\frac{\sum_{j=1}^n y_{i,j}^2\,\mathbb{1}_{(i,j)\in\Omega}}{\sum_{j=1}^n\mathbb{1}_{(i,j)\in\Omega}}-S_{i,i}.\tag{D.30}$$

Step 1: fine-grained estimation guarantees for $\boldsymbol{U}$ and $\boldsymbol{\Sigma}$. To begin with, we need to show that the components in the plug-in estimator are all reliable estimates of their deterministic counterpart (after proper rotation).

Lemma 18. *Recall the definition of ζ_{1st} and $\zeta_{\mathsf{2nd},l}$ in (D.23). Assume that $\zeta_{\mathsf{1st}}/\sigma_r^{\star2}\lesssim1/\sqrt{\kappa^3\mu}$, $n\gg r+\log(n+d)$ and $d\gtrsim\kappa^4\mu^2r\log(n+d)$. Let*

$$\theta\coloneqq\sqrt{\frac{\kappa r\log^2\left(n+d\right)}{np}}\left(1+\frac{\sigma_{\mathsf{ub}}}{\sigma_r^\star}\sqrt{n}\right)\tag{D.31}$$

Then with probability exceeding $1-O((n+d)^{-10})$, we have

$$\left\|(\boldsymbol{U}\boldsymbol{R}-\boldsymbol{U}^\star)_{l,\cdot}\right\|_2\lesssim\frac{\theta}{\sqrt{\kappa}\sigma_r^\star}\left(\left\|\boldsymbol{U}_{l,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2+\omega_l^\star\right)+\zeta_{\mathsf{2nd},l},\tag{D.32a}$$

$$\left\|(\boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{R}-\boldsymbol{U}^\star\boldsymbol{\Sigma}^\star)_{l,\cdot}\right\|_2\lesssim\theta\left(\left\|\boldsymbol{U}_{l,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2+\omega_l^\star\right)+\left\|\boldsymbol{U}_{l,\cdot}^\star\right\|_2\sqrt{\frac{\kappa^2\left(r+\log\left(n+d\right)\right)}{n}}\sigma_1^\star+\zeta_{\mathsf{2nd},l}\sigma_1^\star,\tag{D.32b}$$

$$\left\|\boldsymbol{R}\left(\boldsymbol{\Sigma}^\star\right)^{-2}\boldsymbol{R}^\top-\boldsymbol{\Sigma}^{-2}\right\|\lesssim\sqrt{\frac{\kappa^3\mu r\log\left(n+d\right)}{d}}\frac{\zeta_{\mathsf{1st}}}{\sigma_r^{\star4}}+\kappa\frac{\zeta_{\mathsf{1st}}^2}{\sigma_r^{\star6}}+\sqrt{\frac{\kappa^3\left(r+\log\left(n+d\right)\right)}{n}}\frac{1}{\sigma_r^{\star2}},\tag{D.32c}$$

$$\left\|\boldsymbol{U}\boldsymbol{\Sigma}^{-2}\boldsymbol{R}-\boldsymbol{U}^\star\left(\boldsymbol{\Sigma}^\star\right)^{-2}\right\|\lesssim\frac{\zeta_{\mathsf{1st}}}{\sigma_r^{\star4}}+\sqrt{\frac{\kappa^3\left(r+\log\left(n+d\right)\right)}{n}}\frac{1}{\sigma_r^{\star2}},\tag{D.32d}$$

$$\left\|\boldsymbol{U}\boldsymbol{R}-\boldsymbol{U}^\star\right\|_{2,\infty}\lesssim\frac{\zeta_{\mathsf{1st}}}{\sigma_r^{\star2}}\sqrt{\frac{r\log\left(n+d\right)}{d}}.\tag{D.32e}$$

Proof. See Appendix E.3.1. □

Step 2: faithfulness of the plug-in estimator. With Lemma 18 in hand, we move forward to show that $\boldsymbol{S}$ and $\{\omega_i^2\}_{i=1}^d$ are reliable estimators of the covariance matrix $\boldsymbol{S}^\star$ and the noise levels $\{\omega_i^{\star 2}\}_{i=1}^d$, respectively.

Lemma 19. *Suppose that the conditions of Lemma 18 hold. In addition, assume that* $n \gtrsim \kappa^3 r \log(n+d)$, $\zeta_{\mathsf{1st}}/\sigma_r^{\star 2} \lesssim 1/\sqrt{\log(n+d)}$. *Then with probability exceeding* $1 - O((n+d)^{-10})$, *we have*

$$\left\|\boldsymbol{S}-\boldsymbol{S}^\star\right\|_\infty \lesssim \left(\frac{\zeta_{\mathsf{1st}}}{\sigma_r^\star}\sqrt{\frac{\kappa r \log(n+d)}{d}} + \sqrt{\frac{\kappa^2 \mu r^2 \log(n+d)}{nd}}\sigma_1^\star\right)\sqrt{\frac{\mu r}{d}}\sigma_1^\star \tag{D.33}$$

and, for each $i, j \in [d]$,

$$\begin{aligned}\left|S_{i,j}-S_{i,j}^\star\right| \lesssim & \left(\theta + \sqrt{\frac{\kappa^3 r\log(n+d)}{n}}\right)\left\|\boldsymbol{U}_{i,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2\left\|\boldsymbol{U}_{j,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2 + \theta\left(\omega_i^\star\left\|\boldsymbol{U}_{j,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2 + \omega_j^\star\left\|\boldsymbol{U}_{i,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2\right) \\ & + \sigma_1^\star\left(\zeta_{\mathsf{2nd},i}\left\|\boldsymbol{U}_{j,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2 + \zeta_{\mathsf{2nd},j}\left\|\boldsymbol{U}_{i,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2\right) + \theta^2\omega_i^\star\omega_j^\star + \zeta_{\mathsf{2nd},i}\zeta_{\mathsf{2nd},j}\sigma_1^{\star 2}.\end{aligned} \tag{D.34}$$

Here, the quantity θ *is defined in (D.31). In addition, with probability exceeding* $1 - O((n+d)^{-10})$, *we have*

$$\left|\omega_i^2 - \omega_i^{\star 2}\right| \lesssim \sqrt{\frac{\log^2(n+d)}{np}}\omega_i^{\star 2} + \zeta_{\mathsf{1st}}\frac{\sqrt{\kappa^2\mu r^2\log(n+d)}}{d} + \sqrt{\frac{\kappa^2\mu^2 r^3\log(n+d)}{nd^2}}\sigma_1^{\star 2} \tag{D.35}$$

for all $i \in [d]$, *and*

$$\begin{aligned}\left|\omega_l^2 - \omega_l^{\star 2}\right| \lesssim & \left(\sqrt{\frac{\log^2(n+d)}{np}} + \theta^2\right)\omega_l^{\star 2} + \left(\theta + \sqrt{\frac{\kappa^3 r\log(n+d)}{n}}\right)\left\|\boldsymbol{U}_{l,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2^2 \\ & + \left(\theta\omega_l^\star + \zeta_{\mathsf{2nd},l}\sigma_1^\star\right)\left\|\boldsymbol{U}_{l,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2 + \zeta_{\mathsf{2nd},l}^2\sigma_1^{\star 2}.\end{aligned} \tag{D.36}$$

Proof. See Appendix E.3.2. □

The above two lemmas taken together allow us to demonstrate that our plug-in estimator $\boldsymbol{\Sigma}_{U,l}$ is a faithful estimate of $\boldsymbol{\Sigma}_{U,l}^\star$, as stated below.

Lemma 20. *Suppose that the conditions of Lemma 17, Lemma 18 and Lemma 20 hold. Consider any* $\delta \in (0,1)$, *and we further suppose that* $n \gtrsim \delta^{-2}\kappa^9\mu^2 r^3\log(n+d)$, $d \gtrsim \kappa^3\mu r\log(n+d)$,

$$ndp^2 \gtrsim \delta^{-2}\kappa^8\mu^4 r^4\kappa_\omega^2\log^5(n+d), \qquad np \gtrsim \delta^{-2}\kappa^8\mu^3 r^3\kappa_\omega^2\log^3(n+d),$$

$$\frac{\omega_{\max}^2}{p\sigma_r^{\star 2}}\sqrt{\frac{d}{n}} \lesssim \frac{\delta}{\kappa^3\mu r\kappa_\omega\log^{3/2}(n+d)} \qquad \text{and} \qquad \frac{\omega_{\max}}{\sigma_r^\star}\sqrt{\frac{d}{np}} \lesssim \frac{\delta}{\kappa^{7/2}\mu r\kappa_\omega\log(n+d)}.$$

Then with probability exceeding $1 - O((n+d)^{-10})$, *we have*

$$\left\|\boldsymbol{\Sigma}_{U,l} - \boldsymbol{R}\boldsymbol{\Sigma}_{U,l}^\star\boldsymbol{R}^\top\right\| \lesssim \delta\lambda_{\min}\left(\boldsymbol{\Sigma}_{U,l}^\star\right).$$

Proof. See Appendix E.3.3. □

Step 3: validity of the constructed confidence regions. Finally, we can combine the Gaussian approximation of $\boldsymbol{U}_{l,\cdot}\boldsymbol{R} - \boldsymbol{U}^\star$ established in Theorem 11 and the estimation guarantee of the covariance matrix established in Lemma 20 to justify the validity of the constructed confidence region $\mathsf{CR}_{U,l}^{1-\alpha}$ (cf. Algorithm 3).

Lemma 21. *Suppose that the conditions of Theorem 11 hold. Suppose that* $n \gtrsim \kappa^{12}\mu^3 r^{11/2}\kappa_\omega\log^5(n+d)$, $d \gtrsim \kappa^3\mu r\log(n+d)$,

$$ndp^2 \gtrsim \kappa^{11}\mu^5 r^{13/2}\kappa_\omega^3\log^9(n+d), \qquad np \gtrsim \kappa^{11}\mu^4 r^{11/2}\kappa_\omega^3\log^7(n+d),$$

and

$$\frac{\omega_{\max}^2}{p\sigma_r^{\star 2}}\sqrt{\frac{d}{n}} \lesssim \frac{1}{\kappa^{9/2}\mu^{3/2}r^{9/4}\kappa_\omega^{3/2}\log^{7/2}(n+d)}, \qquad \frac{\omega_{\max}}{\sigma_r^\star}\sqrt{\frac{d}{np}} \lesssim \frac{1}{\kappa^5\mu^{3/2}r^{9/4}\kappa_\omega^{3/2}\log^3(n+d)}.$$

Then it holds that

$$\mathbb{P}\left(\boldsymbol{U}_{l,\cdot}^\star \boldsymbol{R}^\top \in \mathsf{CR}_{U,l}^{1-\alpha}\right) = 1-\alpha+O\left(\frac{1}{\sqrt{\log(n+d)}}\right) = 1-\alpha+o(1).$$

Proof. See Appendix E.3.4. □

With this lemma in place, we have concluded the proof of Theorem 12.

D.4 Entrywise distributional characterization for $\boldsymbol{S}^\star$ (Proof of Theorem 13)

The preceding distributional guarantees for principal subspace in turn allow one to perform inference on the covariance matrix $\boldsymbol{S}^\star$ of the noiseless data vectors $\{\boldsymbol{x}_j\}_{1\le j\le n}$. In this subsection, we demonstrate how to establish the advertised distributional characterization for the matrix $\boldsymbol{S}$ returned by Algorithm 4. Towards this end, we begin with the following decomposition

$$\boldsymbol{S}-\boldsymbol{S}^\star = \underbrace{\boldsymbol{S}-\boldsymbol{M}^\natural\boldsymbol{M}^{\natural\top}}_{=:\boldsymbol{W}} + \underbrace{\boldsymbol{M}^\natural\boldsymbol{M}^{\natural\top}-\boldsymbol{S}^\star}_{=:\boldsymbol{A}},$$

and we shall write $\boldsymbol{W}=[W_{i,j}]_{1\le i,j\le d}$ and $\boldsymbol{A}=[A_{i,j}]$ from now on. In what follows, our proof consists of the following main steps:

1. Show that conditional on $\boldsymbol{F}$, each entry $W_{i,j}$ is approximately a zero-mean Gaussian, whose variance concentrates around some deterministic quantity $\widetilde{v}_{i,j}$.
2. Show that each entry $A_{i,j}$ is approximately Gaussian with mean zero and variance $\overline{v}_{i,j}$.
3. Utilize the (near) independence of $W_{i,j}$ and $A_{i,j}$ to demonstrate that $S_{i,j}-S_{i,j}^\star$ is approximately Gaussian with mean zero and variance $\widetilde{v}_{i,j}+\overline{v}_{i,j}$.

Step 1: first- and second-order approximation of $\boldsymbol{W}$. To begin with, Lemma 15 allows one to derive a reasonably accurate first- and second-order approximation for $\boldsymbol{W}=\boldsymbol{S}-\boldsymbol{M}^\natural\boldsymbol{M}^{\natural\top}$. This is formalized in the following lemma, providing an explicit form of this approximation and the goodness of the approximation.

Lemma 22. *Suppose that the assumptions of Lemma 15 hold. Then one can write*

$$\boldsymbol{W}=\boldsymbol{S}-\boldsymbol{M}^\natural\boldsymbol{M}^{\natural\top}=\boldsymbol{X}+\boldsymbol{\Phi},$$

where

$$\boldsymbol{X} \coloneqq \boldsymbol{E}\boldsymbol{M}^{\natural\top}+\boldsymbol{M}^\natural\boldsymbol{E}^\top+\mathcal{P}_{\textsf{off-diag}}\left(\boldsymbol{E}\boldsymbol{E}^\top\right)\boldsymbol{U}^\natural\boldsymbol{U}^{\natural\top}+\boldsymbol{U}^\natural\boldsymbol{U}^{\natural\top}\mathcal{P}_{\textsf{off-diag}}\left(\boldsymbol{E}\boldsymbol{E}^\top\right) \tag{D.37}$$

and the residual matrix $\boldsymbol{\Phi}$ satisfies: conditional on $\boldsymbol{F}$ and on the $\sigma(\boldsymbol{F})$-measurable event $\mathcal{E}_{\textsf{good}}$ (see Lemma 14),

$$\begin{aligned}|\Phi_{i,j}| &\lesssim \theta^2\left(\left\|\boldsymbol{U}_{i,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2+\omega_i^\star\right)\left(\left\|\boldsymbol{U}_{j,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2+\omega_j^\star\right)+\sigma_1^{\star 2}\zeta_{\textsf{2nd},i}\zeta_{\textsf{2nd},j}\\ &\quad+\sigma_1^{\star 2}\zeta_{\textsf{2nd},i}\left(\left\|\boldsymbol{U}_{j,\cdot}^\star\right\|_2+\theta\frac{\omega_j^\star}{\sigma_1^\star}\right)+\zeta_{\textsf{2nd},j}\sigma_1^{\star 2}\left(\left\|\boldsymbol{U}_{i,\cdot}^\star\right\|_2+\theta\frac{\omega_i^\star}{\sigma_1^\star}\right) =: \zeta_{i,j}\end{aligned} \tag{D.38}$$

holds for any $i,j\in[d]$ with probability exceeding $1-O((n+d)^{-10})$. Here, $\zeta_{\textsf{2nd},i}$ and $\zeta_{\textsf{2nd},j}$ are defined in Lemma 15, and θ is defined in (D.31).

Proof. See Appendix E.4.1. □

Step 2: computing the entrywise variance of our approximation. We can check that the (i,j)-th entry of the matrix $\boldsymbol{X}$ (cf. (D.37)) is given by

$$
\begin{aligned}
X_{i,j} &= \left[\boldsymbol{E}\boldsymbol{M}^{\natural\top} + \boldsymbol{M}^{\natural}\boldsymbol{E}^{\top} + \mathcal{P}_{\mathsf{off\text{-}diag}}\left(\boldsymbol{E}\boldsymbol{E}^{\top}\right)\boldsymbol{U}^{\natural}\boldsymbol{U}^{\natural\top} + \boldsymbol{U}^{\natural}\boldsymbol{U}^{\natural\top}\mathcal{P}_{\mathsf{off\text{-}diag}}\left(\boldsymbol{E}\boldsymbol{E}^{\top}\right)\right]_{i,j} \\
&= \sum_{l=1}^{n}\left\{M_{j,l}^{\natural}E_{i,l} + M_{i,l}^{\natural}E_{j,l} + E_{i,l}\left[\mathcal{P}_{-i,\cdot}\left(\boldsymbol{E}_{\cdot,l}\right)\right]^{\top}\boldsymbol{U}^{\star}\left(\boldsymbol{U}_{j,\cdot}^{\star}\right)^{\top} + E_{j,l}\left[\mathcal{P}_{-j,\cdot}\left(\boldsymbol{E}_{\cdot,l}\right)\right]^{\top}\boldsymbol{U}^{\star}\left(\boldsymbol{U}_{i,\cdot}^{\star}\right)^{\top}\right\} \\
&= \sum_{l=1}^{n}\left[M_{j,l}^{\natural}E_{i,l} + M_{i,l}^{\natural}E_{j,l} + \sum_{k:k\neq i}E_{i,l}E_{k,l}\left(\boldsymbol{U}_{k,\cdot}^{\star}\left(\boldsymbol{U}_{j,\cdot}^{\star}\right)^{\top}\right) + \sum_{k:k\neq j}E_{j,l}E_{k,l}\left(\boldsymbol{U}_{k,\cdot}^{\star}\left(\boldsymbol{U}_{i,\cdot}^{\star}\right)^{\top}\right)\right].
\end{aligned} \tag{D.39}
$$

Here, the penultimate step uses the fact that $\boldsymbol{U}^{\natural} = \boldsymbol{U}^{\star}\boldsymbol{Q}$ for some orthonormal matrix $\boldsymbol{Q}$ (see (D.5)). This allows one to calculate the variance of $X_{i,j}$ conditional on $\boldsymbol{F}$: when $i \neq j$,

$$
\mathsf{var}\left(X_{i,j}|\boldsymbol{F}\right) = \sum_{l=1}^{n}M_{j,l}^{\natural 2}\sigma_{i,l}^{2} + \sum_{l=1}^{n}M_{i,l}^{\natural 2}\sigma_{j,l}^{2} + \sum_{l=1}^{n}\sum_{k:k\neq i}\sigma_{i,l}^{2}\sigma_{k,l}^{2}\left(\boldsymbol{U}_{k,\cdot}^{\star}\boldsymbol{U}_{j,\cdot}^{\star\top}\right)^{2} + \sum_{l=1}^{n}\sum_{k:k\neq j}\sigma_{j,l}^{2}\sigma_{k,l}^{2}\left(\boldsymbol{U}_{k,\cdot}^{\star}\boldsymbol{U}_{i,\cdot}^{\star\top}\right)^{2}; \tag{D.40}
$$

when $i = j$, we have

$$
\begin{aligned}
\mathsf{var}\left(X_{i,i}|\boldsymbol{F}\right) &= 4\sum_{l=1}^{n}M_{i,l}^{\natural 2}\sigma_{i,l}^{2} + 4\sum_{l=1}^{n}\sum_{k:k\neq i}\sigma_{i,l}^{2}\sigma_{k,l}^{2}\left(\boldsymbol{U}_{k,\cdot}^{\natural}\boldsymbol{U}_{i,\cdot}^{\natural\top}\right)^{2} \\
&= 4\sum_{l=1}^{n}M_{i,l}^{\natural 2}\sigma_{i,l}^{2} + 4\sum_{l=1}^{n}\sum_{k:k\neq i}\sigma_{i,l}^{2}\sigma_{k,l}^{2}\left(\boldsymbol{U}_{k,\cdot}^{\star}\boldsymbol{U}_{i,\cdot}^{\star\top}\right)^{2}.
\end{aligned}
$$

The next lemma states that $\mathsf{var}(X_{i,j}|\boldsymbol{F})$ concentrates around some deterministic quantity $\widetilde{v}_{i,j}$ defined as follows:

$$
\begin{aligned}
\widetilde{v}_{i,j} &\coloneqq \frac{2\left(1-p\right)}{np}\left(S_{i,i}^{\star}S_{j,j}^{\star} + 2S_{i,j}^{\star 2}\right) + \frac{1}{np}\left(\omega_{i}^{\star 2}S_{j,j}^{\star} + \omega_{j}^{\star 2}S_{i,i}^{\star}\right) \\
&\quad + \frac{1}{np^{2}}\sum_{k=1}^{d}\left\{\left[\omega_{i}^{\star 2} + \left(1-p\right)S_{i,i}^{\star}\right]\left[\omega_{k}^{\star 2} + \left(1-p\right)S_{k,k}^{\star}\right] + 2\left(1-p\right)^{2}S_{i,k}^{\star 2}\right\}\left(\boldsymbol{U}_{k,\cdot}^{\star}\boldsymbol{U}_{j,\cdot}^{\star\top}\right)^{2} \\
&\quad + \frac{1}{np^{2}}\sum_{k=1}^{d}\left\{\left[\omega_{j}^{\star 2} + \left(1-p\right)S_{j,j}^{\star}\right]\left[\omega_{k}^{\star 2} + \left(1-p\right)S_{k,k}^{\star}\right] + 2\left(1-p\right)^{2}S_{j,k}^{\star 2}\right\}\left(\boldsymbol{U}_{k,\cdot}^{\star}\boldsymbol{U}_{i,\cdot}^{\star\top}\right)^{2}
\end{aligned} \tag{D.41a}
$$

for any $i \neq j$, and

$$
\begin{aligned}
\widetilde{v}_{i,i} &\coloneqq \frac{12\left(1-p\right)}{np}S_{i,i}^{\star 2} + \frac{4}{np}\omega_{i}^{\star 2}S_{i,i}^{\star} \\
&\quad + \frac{4}{np^{2}}\sum_{k=1}^{d}\left\{\left[\omega_{i}^{\star 2} + \left(1-p\right)S_{i,i}^{\star}\right]\left[\omega_{k}^{\star 2} + \left(1-p\right)S_{k,k}^{\star}\right] + 2\left(1-p\right)^{2}S_{i,k}^{\star 2}\right\}\left(\boldsymbol{U}_{k,\cdot}^{\star}\boldsymbol{U}_{i,\cdot}^{\star\top}\right)^{2}
\end{aligned} \tag{D.41b}
$$

for any $i \in [d]$.

Lemma 23. *Suppose that $n \gg \log^{3}(n+d)$, and recall the definition of $\widetilde{v}_{i,j}$ in (D.41). On the event $\mathcal{E}_{\mathsf{good}}$ (cf. Lemma 14), we have*

$$
\mathsf{var}\left(X_{i,j}|\boldsymbol{F}\right) = \widetilde{v}_{i,j} + O\left(\sqrt{\frac{\log^{3}\left(n+d\right)}{n}} + \frac{\kappa\mu^{2}r^{2} + \kappa_{\omega}\mu r}{d}\right)\widetilde{v}_{i,j} \tag{D.42}
$$

for any $i,j \in [d]$. In addition, for any $i,j \in [d]$ it holds that

$$
\widetilde{v}_{i,j} \gtrsim \frac{1}{ndp^{2}\kappa \wedge np}\left\|\boldsymbol{U}_{i,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_{2}^{2}\left\|\boldsymbol{U}_{j,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_{2}^{2} + \left(\frac{\sigma_{r}^{\star 2}}{ndp^{2} \wedge np} + \frac{\omega_{\min}^{2}}{np^{2}}\right)\left(\omega_{j}^{\star 2}\left\|\boldsymbol{U}_{i,\cdot}^{\star}\right\|_{2}^{2} + \omega_{i}^{\star 2}\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_{2}^{2}\right). \tag{D.43}
$$

Proof. See Appendix E.4.2. □

Step 3: establishing approximate Gaussianity of $W_{i,j}$. We are now ready to invoke the Berry-Esseen Theorem to show that $W_{i,j}$ is approximately Gaussian with mean zero and variance $\widetilde{v}_{i,j}$, as stated in the next lemma.

Lemma 24. *Suppose that $d \gtrsim \kappa^8\mu^3r^3\kappa_\omega^2\log^5(n+d)$,*

$$ndp^2 \gtrsim \kappa^{10}\mu^4r^4\kappa_\omega^2\log^9(n+d), \qquad np \gtrsim \kappa^{10}\mu^3r^3\kappa_\omega^2\log^7(n+d),$$

$$\frac{\omega_{\max}^2}{p\sigma_r^{\star 2}}\sqrt{\frac{d}{n}} \lesssim \frac{1}{\sqrt{\kappa^7\mu^2r^2\kappa_\omega\log^7(n+d)}}, \qquad \frac{\omega_{\max}}{\sigma_r^\star}\sqrt{\frac{d}{np}} \lesssim \frac{1}{\sqrt{\kappa^8\mu^2r^2\kappa_\omega\log^6(n+d)}},$$

and

$$\left\|\boldsymbol{U}_{i,\cdot}^\star\right\|_2 + \left\|\boldsymbol{U}_{j,\cdot}^\star\right\|_2 \gtrsim \kappa r\log^{5/2}(n+d)\left[\frac{\omega_{\max}}{\sigma_r^\star}\sqrt{\frac{d}{np}} + \kappa_\omega\frac{\omega_{\max}^2}{p\sigma_r^{\star 2}}\sqrt{\frac{d}{n}} + \frac{\kappa\mu r\kappa_\omega\log(n+d)}{\sqrt{ndp^2}}\right]\sqrt{\frac{1}{d}}.$$

Then it holds that

$$\sup_{t\in\mathbb{R}}\left|\mathbb{P}\left(\left(\widetilde{v}_{i,j}\right)^{-1/2}W_{i,j} \le t \,\middle|\, \boldsymbol{F}\right) - \Phi(t)\right| \lesssim \frac{1}{\sqrt{\log(n+d)}} = o(1).$$

Proof. See Appendix E.4.3. □

An important observation that one should bear in mind is that: conditional on $\boldsymbol{F}$, the distribution of $W_{i,j}$ is approximately $\mathcal{N}(0,\widetilde{v}_{i,j})$, where $\widetilde{v}_{i,j}$ does not depend on $\boldsymbol{F}$. This suggests that $W_{i,j}$ is nearly independent of the σ-algebra $\sigma(\boldsymbol{F})$.

Step 4: establishing approximate Gaussianity of $A_{i,j}$. We now move on to the matrix $\boldsymbol{A}$. It follows from (D.3) that

$$A_{i,j} = \left(\boldsymbol{M}^\natural\boldsymbol{M}^{\natural\top} - \boldsymbol{S}^\star\right)_{i,j} = \frac{1}{n}\sum_{l=1}^n\left(\boldsymbol{U}_{i,\cdot}^\star\boldsymbol{\Sigma}^\star\boldsymbol{f}_l\right)\left(\boldsymbol{U}_{j,\cdot}^\star\boldsymbol{\Sigma}^\star\boldsymbol{f}_l\right) - \boldsymbol{U}_{i,\cdot}^\star\boldsymbol{\Sigma}^{\star 2}\boldsymbol{U}_{j,\cdot}^{\star\top}. \tag{D.44}$$

In view of the independence of $\{\boldsymbol{f}_l\}_{1\le l\le n}$, the variance of $A_{i,j}$ can be calculated as follows

$$\begin{aligned}
\overline{v}_{i,j} &\coloneqq \mathsf{var}\left(A_{i,j}\right) = \frac{1}{n}\mathsf{var}\left[\left(\boldsymbol{U}_{i,\cdot}^\star\boldsymbol{\Sigma}^\star\boldsymbol{f}_1\right)\left(\boldsymbol{U}_{j,\cdot}^\star\boldsymbol{\Sigma}^\star\boldsymbol{f}_1\right)\right] \\
&= \frac{1}{n}\left[\left\|\boldsymbol{U}_{i,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2^2\left\|\boldsymbol{U}_{j,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2^2 + 2\left(\boldsymbol{U}_{i,\cdot}^\star\boldsymbol{\Sigma}^{\star 2}\boldsymbol{U}_{j,\cdot}^{\star\top}\right)^2\right] - \frac{1}{n}\left(\boldsymbol{U}_{i,\cdot}^\star\boldsymbol{\Sigma}^{\star 2}\boldsymbol{U}_{j,\cdot}^{\star\top}\right)^2 \\
&= \frac{1}{n}\left[\left\|\boldsymbol{U}_{i,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2^2\left\|\boldsymbol{U}_{j,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2^2 + \left(\boldsymbol{U}_{i,\cdot}^\star\boldsymbol{\Sigma}^{\star 2}\boldsymbol{U}_{j,\cdot}^{\star\top}\right)^2\right] = \frac{1}{n}\left(S_{i,i}^\star S_{j,j}^\star + S_{i,j}^{\star 2}\right).
\end{aligned} \tag{D.45}$$

With the assistance of the Berry-Esseen Theorem, the lemma below demonstrates that $A_{i,j}$ is approximately Gaussian.

Lemma 25. *It holds that*

$$\sup_{z\in\mathbb{R}}\left|\mathbb{P}\left(\left(\overline{v}_{i,j}\right)^{-1/2}A_{i,j} \le z\right) - \Phi(z)\right| \lesssim \frac{1}{\sqrt{n}}.$$

Proof. See Appendix E.4.4. □

Note that $A_{i,j}$ is $\sigma(\boldsymbol{F})$-measurable. This fact taken collectively with the near independence between $W_{i,j}$ and $\sigma(\boldsymbol{F})$ implies that $W_{i,j}$ and $A_{i,j}$ are nearly statistically independent.

Step 5: distributional characterization of $S_{i,j} - S^{\star}_{i,j}$. The approximate Gaussianity of $W_{i,j}$ and $A_{i,j}$, as well as the near independence between them, leads to the conjecture that $S_{i,j} - S^{\star}_{i,j} = W_{i,j} + A_{i,j}$ is approximately distributed as $\mathcal{N}(0, v^{\star}_{i,j})$ with

$$v^{\star}_{ij} \coloneqq \widetilde{v}_{i,j} + \overline{v}_{i,j}.$$

The following lemma rigorizes this conjecture, which in turn concludes the proof of Theorem 13.

Lemma 26. *Suppose that the assumptions of Lemma 24 hold, and suppose that $n \gtrsim \log(n+d)$. Then we have*

$$\sup_{t\in\mathbb{R}} \left| \mathbb{P}\left(\left(S_{i,j} - S^{\star}_{i,j}\right) / \sqrt{v^{\star}_{i,j}} \le t \right) - \Phi\left(t\right) \right| \lesssim \frac{1}{\sqrt{\log\left(n+d\right)}} = o\left(1\right).$$

Proof. See Appendix E.4.5. □

D.5 Validity of confidence intervals (Proof of Theorem 14)

Armed with the above entrywise distributional theory for $\boldsymbol{S} - \boldsymbol{S}^{\star}$, we hope to construct valid confidence interval for each $S^{\star}_{i,j}$ based on the estimate $\boldsymbol{S}$ returned by HeteroPCA. This requires a faithful estimate of the variance $v^{\star}_{i,j}$. Towards this, we define the following plug-in estimator: if $i \neq j$, let

$$\begin{aligned} v_{i,j} = {} & \frac{2-p}{np} S_{i,i} S_{j,j} + \frac{4-3p}{np} S_{i,j}^2 + \frac{1}{np}\left(\omega_i^2 S_{j,j} + \omega_j^2 S_{i,i}\right) \\ & + \frac{1}{np^2} \sum_{k=1}^{d} \left\{ \left[\omega_i^2 + \left(1-p\right) S_{i,i}\right] \left[\omega_k^2 + \left(1-p\right) S_{k,k}\right] + 2\left(1-p\right)^2 S_{i,k}^2 \right\} \left(\boldsymbol{U}_{k,\cdot} \boldsymbol{U}_{j,\cdot}^{\top}\right)^2 \\ & + \frac{1}{np^2} \sum_{k=1}^{d} \left\{ \left[\omega_j^2 + \left(1-p\right) S_{j,j}\right] \left[\omega_k^2 + \left(1-p\right) S_{k,k}\right] + 2\left(1-p\right)^2 S_{j,k}^2 \right\} \left(\boldsymbol{U}_{k,\cdot} \boldsymbol{U}_{i,\cdot}^{\top}\right)^2; \end{aligned} \tag{D.46a}$$

otherwise, let

$$\begin{aligned} v_{i,i} = {} & \frac{12-9p}{np} S_{i,i}^2 + \frac{4}{np}\omega_i^2 S_{i,i} \\ & + \frac{4}{np^2} \sum_{k=1}^{d} \left\{ \left[\omega_i^2 + \left(1-p\right) S_{i,i}\right] \left[\omega_k^2 + \left(1-p\right) S_{k,k}\right] + 2\left(1-p\right)^2 S_{i,k}^2 \right\} \left(\boldsymbol{U}_{k,\cdot} \boldsymbol{U}_{i,\cdot}^{\top}\right)^2. \end{aligned} \tag{D.46b}$$

Step 1: faithfulness of the plug-in estimator. With the fine-grained estimation guarantees in Lemma 18 and Lemma 19 in place, we can demonstrate that $v_{i,j}$ is a reliable estimate of $v^{\star}_{i,j}$, as formally stated below.

Lemma 27. *Suppose that the conditions of Lemma 24 hold. For any $\delta \in (0,1)$, we further assume that $n \gtrsim \delta^{-2}\kappa^7\mu^2 r^3 \kappa_\omega^2 \log(n+d)$, $d \gtrsim \kappa^2\mu\log(n+d)$,*

$$ndp^2 \gtrsim \delta^{-2}\kappa^6\mu^4 r^6 \kappa_\omega^2 \log^5\left(n+d\right), \qquad np \gtrsim \delta^{-2}\kappa^6\mu^3 r^5\kappa_\omega^2 \log^3\left(n+d\right),$$
$$\frac{\omega_{\max}^2}{p\sigma_r^{\star 2}}\sqrt{\frac{d}{n}} \lesssim \frac{\delta}{\kappa^2\mu r^2\kappa_\omega \log^{3/2}\left(n+d\right)}, \qquad \frac{\omega_{\max}}{\sigma_r^{\star}}\sqrt{\frac{d}{np}} \lesssim \frac{\delta}{\kappa^{5/2}\mu r^2 \kappa_\omega \log\left(n+d\right)},$$

and

$$\left\|\boldsymbol{U}^{\star}_{i,\cdot}\right\|_2 + \left\|\boldsymbol{U}^{\star}_{j,\cdot}\right\|_2 \gtrsim \delta^{-1}\kappa\mu r^2\kappa_\omega \log\left(n+d\right) \left[\frac{\kappa\mu r\log\left(n+d\right)}{\sqrt{ndp}} + \frac{\omega_{\max}^2}{p\sigma_r^{\star 2}}\sqrt{\frac{d}{n}} + \frac{\omega_{\max}}{\sigma_r^{\star}}\sqrt{\frac{d}{np}} \right] \sqrt{\frac{r}{d}}.$$

Then with probability exceeding $1 - O((n+d)^{-10})$, the quantity $v_{i,j}$ defined in (D.46) obeys

$$\left|v^{\star}_{i,j} - v_{i,j}\right| \lesssim \delta v^{\star}_{i,j}.$$

Proof. See Appendix E.5.1. □

Step 2: validity of the constructed confidence intervals. Armed with the Gaussian approximation in Lemma 26 as well as the faithfulness of $v_{i,j}$ as an estimate of $v_{i,j}^{\star}$ in the previous lemma, we show in the next lemma that the confidence interval constructed in Algorithm 4 is valid and nearly accurate.

Lemma 28. *Suppose that the conditions of Lemma 24 hold. Further suppose that $n \gtrsim \kappa^9 \mu^3 r^4 \kappa_\omega^3 \log^4(n+d)$, $d \gtrsim \kappa^2 \mu \log(n+d)$,*

$$ndp^2 \gtrsim \kappa^8 \mu^5 r^7 \kappa_\omega^3 \log^8(n+d), \qquad np \gtrsim \kappa^8 \mu^4 r^6 \kappa_\omega^3 \log^6(n+d),$$

$$\frac{\omega_{\max}^2}{p\sigma_r^{\star 2}}\sqrt{\frac{d}{n}} \lesssim \frac{1}{\kappa^3 \mu^{3/2} r^{5/2} \kappa_\omega^{3/2} \log^3(n+d)}, \qquad \frac{\omega_{\max}}{\sigma_r^{\star}}\sqrt{\frac{d}{np}} \lesssim \frac{1}{\kappa^{7/2} \mu^{3/2} r^{5/2} \kappa_\omega^{3/2} \log^{5/2}(n+d)},$$

and

$$\left\|\boldsymbol{U}_{i,\cdot}^{\star}\right\|_2 + \left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2 \gtrsim \kappa^2 \mu^{3/2} r^{5/2} \kappa_\omega^{3/2} \log^{5/2}(n+d)\left[\frac{\kappa\mu r \log(n+d)}{\sqrt{ndp}} + \frac{\omega_{\max}^2}{p\sigma_r^{\star 2}}\sqrt{\frac{d}{n}} + \frac{\omega_{\max}}{\sigma_r^{\star}}\sqrt{\frac{d}{np}}\right]\sqrt{\frac{r}{d}}.$$

Then the confidence region $\mathsf{CI}_{i,j}^{1-\alpha}$ returned from Algorithm 4 satisfies

$$\mathbb{P}\left(S_{i,j}^{\star} \in \mathsf{CI}_{i,j}^{1-\alpha}\right) = 1 - \alpha + O\left(\frac{1}{\sqrt{\log(n+d)}}\right) = 1 - \alpha + o(1).$$

Proof. See Appendix E.5.2. □

E Auxiliary lemmas: the approach based on HeteroPCA

E.1 Proof of Lemma 14

In this section we will establish a couple of useful properties that occur with high probability.

1. In view of standard results on Gaussian random matrices, we have

$$\left\|\frac{1}{n}\boldsymbol{F}\boldsymbol{F}^{\top} - \boldsymbol{I}_r\right\| \lesssim \sqrt{\frac{r}{n}} + \sqrt{\frac{\log(n+d)}{n}} + \frac{r}{n} + \frac{\log(n+d)}{n} \lesssim \sqrt{\frac{r + \log(n+d)}{n}} \tag{E.1}$$

with probability exceeding $1 - O((n+d)^{-100})$, provided that $n \gtrsim r + \log(n+d)$. As an immediate consequence, it is seen from the definition (D.3) that

$$\begin{aligned}\left\|\boldsymbol{M}^{\natural}\boldsymbol{M}^{\natural\top} - \boldsymbol{S}^{\star}\right\| &= \left\|\boldsymbol{U}^{\star}\boldsymbol{\Sigma}^{\star}\left(\frac{1}{n}\boldsymbol{F}\boldsymbol{F}^{\top} - \boldsymbol{I}_r\right)\boldsymbol{\Sigma}^{\star}\boldsymbol{U}^{\star\top}\right\| \leq \left\|\boldsymbol{U}^{\star}\right\|^2 \left\|\boldsymbol{\Sigma}^{\star}\right\|^2 \left\|\frac{1}{n}\boldsymbol{F}\boldsymbol{F}^{\top} - \boldsymbol{I}_r\right\| \\ &\lesssim \sqrt{\frac{r + \log(n+d)}{n}}\sigma_1^{\star 2}. \end{aligned} \tag{E.2}$$

Weyl's inequality then tells us that

$$\left\|\boldsymbol{\Sigma}^{\natural 2} - \boldsymbol{\Sigma}^{\star 2}\right\| \leq \left\|\boldsymbol{M}^{\natural}\boldsymbol{M}^{\natural\top} - \boldsymbol{S}^{\star}\right\| \lesssim \sqrt{\frac{r + \log(n+d)}{n}}\sigma_1^{\star 2}, \tag{E.3}$$

thus indicating that

$$\sigma_r^{\natural} \asymp \sigma_r^{\star} \qquad \text{and} \qquad \sigma_1^{\natural} \asymp \sigma_1^{\star} \tag{E.4}$$

as long as $n \gg \frac{\sigma_1^{\star 4}}{\sigma_r^{\star 4}}\big(r + \log(n+d)\big) = \kappa^2\big(r + \log(n+d)\big)$ (see the definition of κ in (3.1)). Consequently,

$$\left\|\boldsymbol{\Sigma}^{\natural} - \boldsymbol{\Sigma}^{\star}\right\| \leq \frac{\left\|\boldsymbol{\Sigma}^{\natural 2} - \boldsymbol{\Sigma}^{\star 2}\right\|}{\sigma_r^{\star}} \lesssim \sqrt{\frac{r + \log(n+d)}{n}}\frac{\sigma_1^{\star 2}}{\sigma_r^{\star}} \asymp \kappa\sqrt{\frac{(r + \log(n+d))}{n}}\sigma_r^{\star}. \tag{E.5}$$

2. Given the calculation (D.9), let us define

$$\sigma^2 := \max_{i\in[d],j\in[n]} \sigma_{i,j}^2 \asymp \frac{1-p}{np} \max_{i\in[d],j\in[n]} \left(\boldsymbol{U}_{i,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\boldsymbol{f}_j\right)^2 + \frac{\omega_i^{\star 2}}{np},$$

and

$$\sigma_i^2 := \max_{j\in[n]} \sigma_{i,j}^2 \asymp \frac{1-p}{np} \max_{j\in[n]} \left(\boldsymbol{U}_{i,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\boldsymbol{f}_j\right)^2 + \frac{\omega_i^{\star 2}}{np}$$

for each $i \in [d]$. Note that for each $i \in [d]$ and $j \in [n]$, one has $\boldsymbol{U}_{i,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\boldsymbol{f}_j \sim \mathcal{N}(0, \|\boldsymbol{U}_{i,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\|_2^2)$, thus revealing that

$$\max_{j\in[n]} \left|\boldsymbol{U}_{i,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\boldsymbol{f}_j\right| \lesssim \left\|\boldsymbol{U}_{i,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2 \sqrt{\log(n+d)} \tag{E.6}$$

with probability exceeding $1 - O((n+d)^{-100})$. As a result, taking the union bound gives

$$\sigma^2 \lesssim \frac{\|\boldsymbol{U}^{\star}\boldsymbol{\Sigma}^{\star}\|_{2,\infty}^2 \log(n+d) + \omega_{\max}^2}{np} \lesssim \frac{\mu r \log(n+d)}{ndp}\sigma_1^{\star 2} + \frac{\omega_{\max}^2}{np} =: \sigma_{\mathsf{ub}}^2$$

and

$$\sigma_i^2 \lesssim \frac{\log(n+d)}{np} \left\|\boldsymbol{U}_{i,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2^2 + \frac{\omega_i^{\star 2}}{np} := \sigma_{\mathsf{ub},i}^2.$$

3. In addition, it follows from the expression (D.8) and the property (E.6) that

$$\begin{aligned}
\max_{i\in[d],j\in[n]} |E_{i,j}| &\leq \frac{1}{\sqrt{np}} \max_{i\in[d],j\in[n]} \left|\boldsymbol{U}_{i,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\boldsymbol{f}_j\right| + \frac{1}{\sqrt{np}} \max_{i\in[d],j\in[n]} |N_{i,j}| \\
&\lesssim \frac{1}{\sqrt{np}} \|\boldsymbol{U}^{\star}\boldsymbol{\Sigma}^{\star}\|_{2,\infty} \sqrt{\log(n+d)} + \frac{\omega_{\max}\sqrt{\log(n+d)}}{\sqrt{np}} \\
&\lesssim \frac{1}{p}\sqrt{\frac{\mu r \log(n+d)}{nd}}\sigma_1^{\star} + \frac{\omega_{\max}}{p}\sqrt{\frac{\log(n+d)}{n}}
\end{aligned}$$

occurs with probability exceeding $1 - O((n+d)^{-100})$. Therefore, we shall take

$$B := \frac{1}{p}\sqrt{\frac{\mu r \log(n+d)}{nd}}\sigma_1^{\star} + \frac{\omega_{\max}}{p}\sqrt{\frac{\log(n+d)}{n}} \asymp \sigma_{\mathsf{ub}}\sqrt{\frac{\log(n+d)}{p}}$$

as an upper bound for $\{|E_{i,j}| : i \in [d], j \in [n]\}$. Similarly, for each $i \in [d]$, we can take

$$B_i := \frac{1}{p}\sqrt{\frac{\log(n+d)}{n}} \left\|\boldsymbol{U}_{i,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2 + \frac{\omega_l^{\star}}{p}\sqrt{\frac{\log(n+d)}{n}}$$

as an upper bound for $\{|E_{i,j}| : j \in [n]\}$.

4. Recall that the top-r eigen-decomposition of $\boldsymbol{M}^{\natural}\boldsymbol{M}^{\natural\top}$ and $\boldsymbol{S}^{\star}$ are denoted by $\boldsymbol{U}^{\natural}\boldsymbol{\Sigma}^{\natural 2}\boldsymbol{U}^{\natural\top}$ and $\boldsymbol{U}^{\star}\boldsymbol{\Sigma}^{\star}\boldsymbol{U}^{\star\top}$, respectively, and that $\boldsymbol{Q}$ is a rotation matrix such that $\boldsymbol{U}^{\natural} = \boldsymbol{U}^{\star}\boldsymbol{Q}$. Therefore, the matrix $\boldsymbol{J}$ defined in (D.6) obeys

$$\begin{aligned}
\|\boldsymbol{Q} - \boldsymbol{J}\| &= \left\|\boldsymbol{Q} - \boldsymbol{\Sigma}^{\star}\boldsymbol{Q}\left(\boldsymbol{\Sigma}^{\natural}\right)^{-1}\right\| \leq \frac{1}{\sigma_r^{\natural}} \left\|\boldsymbol{Q}\boldsymbol{\Sigma}^{\natural} - \boldsymbol{\Sigma}^{\star}\boldsymbol{Q}\right\| = \frac{1}{\sigma_r^{\natural}} \left\|\boldsymbol{Q}\boldsymbol{\Sigma}^{\natural}\boldsymbol{Q}^{\top} - \boldsymbol{\Sigma}^{\star}\right\| \\
&= \frac{1}{\sigma_r^{\natural}} \left\|\boldsymbol{U}^{\star}\left(\boldsymbol{Q}\boldsymbol{\Sigma}^{\natural}\boldsymbol{Q}^{\top} - \boldsymbol{\Sigma}^{\star}\right)\boldsymbol{U}^{\star\top}\right\| = \frac{1}{\sigma_r^{\natural}} \left\|\boldsymbol{U}^{\natural}\boldsymbol{\Sigma}^{\natural}\boldsymbol{U}^{\natural\top} - \boldsymbol{U}^{\star}\boldsymbol{\Sigma}^{\star}\boldsymbol{U}^{\star\top}\right\|.
\end{aligned}$$

Invoke the perturbation bound for matrix square roots (Schmitt, 1992, Lemma 2.1) to derive

$$\left\|\boldsymbol{U}^{\natural}\boldsymbol{\Sigma}^{\natural}\boldsymbol{U}^{\natural\top} - \boldsymbol{U}^{\star}\boldsymbol{\Sigma}^{\star}\boldsymbol{U}^{\star\top}\right\| \lesssim \frac{1}{\sigma_r^{\natural} + \sigma_r^{\star}} \left\|\boldsymbol{U}^{\natural}\boldsymbol{\Sigma}^{\natural 2}\boldsymbol{U}^{\natural\top} - \boldsymbol{U}^{\star}\boldsymbol{\Sigma}^{\star 2}\boldsymbol{U}^{\star\top}\right\|$$

$$\asymp \frac{1}{\sigma_r^\star} \left\| \boldsymbol{M}^\natural \boldsymbol{M}^{\natural\top} - \boldsymbol{S}^\star \right\|,$$

where the last step follows from (E.4). To summarize, with probability exceeding $1 - O((n+d)^{-100})$ one has

$$\|\boldsymbol{Q} - \boldsymbol{J}\| \lesssim \frac{1}{\sigma_r^\star} \left\| \boldsymbol{Q}\boldsymbol{\Sigma}^\natural - \boldsymbol{\Sigma}^\star \boldsymbol{Q} \right\| \lesssim \frac{1}{\sigma_r^{\star 2}} \left\| \boldsymbol{M}^\natural \boldsymbol{M}^{\natural\top} - \boldsymbol{S}^\star \right\| \lesssim \kappa \sqrt{\frac{r + \log(n+d)}{n}}, \tag{E.7}$$

where we have made used of the properties (E.4) and (E.2).

5. In view of (E.4), we know that with probability exceeding $1 - O((n+d)^{-100})$, the conditional number of $\boldsymbol{M}^\natural$ satisfies

$$\kappa^\natural \asymp \sqrt{\kappa}. \tag{E.8}$$

Recalling that $\boldsymbol{U}^\natural = \boldsymbol{U}^\star \boldsymbol{Q}$, we can see from the incoherence assumption that

$$\left\| \boldsymbol{U}^\natural \right\|_{2,\infty} = \|\boldsymbol{U}^\star \boldsymbol{Q}\|_{2,\infty} = \|\boldsymbol{U}^\star\|_{2,\infty} \leq \sqrt{\frac{\mu r}{d}}. \tag{E.9}$$

In addition, it is readily seen from (D.4) that

$$\left\| \boldsymbol{V}^\natural \right\|_{2,\infty} = \left\| \frac{1}{\sqrt{n}} \boldsymbol{F}^\top \boldsymbol{J} \right\|_{2,\infty} \overset{\text{(i)}}{\lesssim} \sqrt{\frac{1}{n}} \left\| \boldsymbol{F}^\top \right\|_{2,\infty} \overset{\text{(ii)}}{\lesssim} \sqrt{\frac{r \log(n+d)}{n}} \tag{E.10}$$

with probability exceeding $1 - O((n+d)^{-100})$. Here, (i) holds since, according to (E.7),

$$\|\boldsymbol{J}\| \leq \|\boldsymbol{Q}\| + \|\boldsymbol{J} - \boldsymbol{Q}\| \leq 1 + \kappa \sqrt{\frac{r + \log(n+d)}{n}} \leq 2 \tag{E.11}$$

holds as long as $n \gg \kappa^2(r + \log(n+d))$; and (ii) follows from the standard Gaussian concentration inequality. Combine (E.9) , (E.10) and

$$\begin{aligned}
\left\| \boldsymbol{M}^\natural \right\|_\infty &= \max_{i,j} \left| \boldsymbol{U}_{i,\cdot}^\natural \boldsymbol{\Sigma}^\natural \boldsymbol{V}_{j,\cdot}^{\natural\top} \right| \leq \sigma_1^\natural \left\| \boldsymbol{U}^\natural \right\|_{2,\infty} \left\| \boldsymbol{V}^\natural \right\|_{2,\infty} \overset{\text{(i)}}{\lesssim} \sqrt{\frac{\mu \log(n+d)}{nd}} \sigma_1^\natural \sqrt{r} \\
&\lesssim \sqrt{\frac{\mu \log(n+d)}{nd}} \kappa^\natural \left\| \boldsymbol{M}^\natural \right\|_{\mathrm{F}} \overset{\text{(ii)}}{\asymp} \sqrt{\frac{\kappa \mu \log(n+d)}{nd}} \left\| \boldsymbol{M}^\natural \right\|_{\mathrm{F}}.
\end{aligned}$$

to reach

$$\mu^\natural \lesssim \kappa \mu \log(n+d)$$

with probability exceeding $1 - O((n+d)^{-100})$. Here, (i) follows from (E.10), whereas (ii) arises from (E.8).

6. Invoke Chen et al. (2019a, Lemma 14) to show that, for all $l \in [d]$,

$$\left\| \frac{1}{n} \sum_{j=1}^n \left(\boldsymbol{U}_{l,\cdot}^\star \boldsymbol{\Sigma}^\star \boldsymbol{f}_j \right)^2 \boldsymbol{f}_j \boldsymbol{f}_j^\top - \left\| \boldsymbol{U}_{l,\cdot}^\star \boldsymbol{\Sigma}^\star \right\|_2^2 \boldsymbol{I}_r - 2\boldsymbol{\Sigma}^\star \boldsymbol{U}_{l,\cdot}^{\star\top} \boldsymbol{U}_{l,\cdot}^\star \boldsymbol{\Sigma}^\star \right\| \lesssim \sqrt{\frac{r \log^3(n+d)}{n}} \left\| \boldsymbol{U}_{l,\cdot}^\star \boldsymbol{\Sigma}^\star \right\|_2^2 \tag{E.12}$$

holds with probability exceeding $1 - O((n+d)^{-100})$, provided that $n \gg r \log^3(n+d)$. In addition, we make note of the following lemmas.

Lemma 29. *Assume that $n \gg \log(n+d)$. For any fixed vector $\boldsymbol{u} \in \mathbb{R}^r$, with probability exceeding $1 - O((n+d)^{-100})$ we have*

$$\left| \frac{1}{n} \sum_{j=1}^n \left(\boldsymbol{u}^\top \boldsymbol{f}_j \right)^2 - \|\boldsymbol{u}\|_2^2 \right| \lesssim \sqrt{\frac{\log(n+d)}{n}} \|\boldsymbol{u}\|_2^2.$$

Lemma 30. *Assume that $n \gg \log(n+d)$. For any fixed unit vectors $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^r$, with probability exceeding $1 - O((n+d)^{-100})$ we have*

$$\left| \frac{1}{n} \sum_{j=1}^{n} \left(\boldsymbol{u}^\top \boldsymbol{f}_j\right)^2 \left(\boldsymbol{v}^\top \boldsymbol{f}_j\right)^2 - \|\boldsymbol{u}\|_2^2 \|\boldsymbol{v}\|_2^2 - 2\left(\boldsymbol{u}^\top \boldsymbol{v}\right)^2 \right| \lesssim \sqrt{\frac{\log(n+d)^3}{n}} \|\boldsymbol{u}\|_2^2 \|\boldsymbol{v}\|_2^2 .$$

In view of Lemma 29 and Lemma 30, we know that for each $i, l \in [d]$

$$\left| \frac{1}{n} \sum_{j=1}^{n} \left(\boldsymbol{U}_{l,\cdot}^\star \boldsymbol{\Sigma}^\star \boldsymbol{f}_j\right)^2 \left(\boldsymbol{U}_{i,\cdot}^\star \boldsymbol{\Sigma}^\star \boldsymbol{f}_j\right)^2 - \left[\left\|\boldsymbol{U}_{l,\cdot}^\star \boldsymbol{\Sigma}^\star\right\|_2^2 \left\|\boldsymbol{U}_{i,\cdot}^\star \boldsymbol{\Sigma}^\star\right\|_2^2 + 2S_{i,l}^{\star 2} \right] \right| \lesssim \sqrt{\frac{\log(n+d)^3}{n}} \left\|\boldsymbol{U}_{l,\cdot}^\star \boldsymbol{\Sigma}^\star\right\|_2^2 \left\|\boldsymbol{U}_{i,\cdot}^\star \boldsymbol{\Sigma}^\star\right\|_2^2 \tag{E.13a}$$

$$\text{and} \qquad \left| \frac{1}{n} \sum_{j=1}^{n} \left(\boldsymbol{U}_{i,\cdot}^\star \boldsymbol{\Sigma}^\star \boldsymbol{f}_j\right)^2 - \left\|\boldsymbol{U}_{i,\cdot}^\star \boldsymbol{\Sigma}^\star\right\|_2^2 \right| \lesssim \sqrt{\frac{\log(n+d)}{n}} \left\|\boldsymbol{U}_{i,\cdot}^\star \boldsymbol{\Sigma}^\star\right\|_2^2 \tag{E.13b}$$

hold with probability exceeding $1 - O((n+d)^{-100})$.

7. With the above properties in place, we can now formally define the "good" event $\mathcal{E}_{\mathsf{good}}$ as follows

$$\mathcal{E}_{\mathsf{good}} \coloneqq \{\text{All equations from (E.1) to (E.13b) hold}\} .$$

It is immediately seen from the above analysis that

$$\mathbb{P}\left(\mathcal{E}_{\mathsf{good}}\right) \geq 1 - O\left((n+d)^{-100}\right) .$$

By construction, it is self-evident that the event $\mathcal{E}_{\mathsf{good}}$ is $\sigma(\boldsymbol{F})$-measurable.

Proof of Lemma 29. Recognizing that $\boldsymbol{u}^\top \boldsymbol{f}_j \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, \|\boldsymbol{u}\|_2^2)$, we have

$$\frac{1}{n} \sum_{j=1}^{n} \left(\boldsymbol{u}^\top \boldsymbol{f}_j\right)^2 / \|\boldsymbol{u}\|_2^2 \sim \chi_n^2,$$

where χ_n^2 denotes the chi-square distribution with n degrees of freedom. One can then apply the tail bound for chi-square random variables (see, e.g., Wainwright (2019, Example 2.5)) to establish the desired result. $\square$

Proof of Lemma 30. Given that $\boldsymbol{u}^\top \boldsymbol{f}_j \sim \mathcal{N}(0, \|\boldsymbol{u}\|_2^2)$ for all $j \in [n]$, we have, with probability exceeding $1 - O((n+d)^{-100})$, that

$$\max_{j \in [n]} \left|\boldsymbol{u}^\top \boldsymbol{f}_j\right| \leq C_1 \|\boldsymbol{u}\|_2 \sqrt{\log(n+d)}$$

for some sufficiently large constant $C_1 > 0$. For each $j \in [n]$, let $X_j = (\boldsymbol{u}^\top \boldsymbol{f}_j)^2$ and $Y_j = (\boldsymbol{v}^\top \boldsymbol{f}_j)^2$ and define the event

$$\mathcal{A}_j \coloneqq \left\{ \left|\boldsymbol{u}^\top \boldsymbol{f}_j\right| \leq C_1 \|\boldsymbol{u}\|_2 \sqrt{\log(n+d)} \right\} .$$

Then with probability exceeding $1 - O((n+d)^{-100})$, it holds that

$$\frac{1}{n} \sum_{j=1}^{n} X_j Y_j = \frac{1}{n} \sum_{j=1}^{n} X_j Y_j \mathbb{1}_{\mathcal{A}_j},$$

which motivates us to decompose

$$\left| \frac{1}{n} \sum_{j=1}^{n} \left(X_j Y_j - \mathbb{E}\left[X_j Y_j\right]\right) \right| \leq \underbrace{\left| \frac{1}{n} \sum_{j=1}^{n} \left\{ X_j Y_j \mathbb{1}_{\mathcal{A}_j} - \mathbb{E}\left[X_j Y_j \mathbb{1}_{\mathcal{A}_j}\right] \right\} \right|}_{=:\alpha_1} + \underbrace{\left| \mathbb{E}\left[X_1 Y_1 \mathbb{1}_{\mathcal{A}_1^c}\right] \right|}_{=:\alpha_2} .$$

- Let us first bound α_2. It is straightforward to derive that

$$\begin{aligned}
\alpha_2 &= \mathbb{E}\left[\left(\boldsymbol{u}^\top \boldsymbol{f}_1\right)^2 \left(\boldsymbol{v}^\top \boldsymbol{f}_1\right)^2 \mathbb{1}_{|\boldsymbol{u}^\top \boldsymbol{f}_1|>C_1\|\boldsymbol{u}\|_2\sqrt{\log(n+d)}}\right] \\
&\overset{\text{(i)}}{\leq} \left(\mathbb{E}\left[\left(\boldsymbol{u}^\top \boldsymbol{f}_1\right)^6\right]\right)^{\frac{1}{3}} \left(\mathbb{E}\left[\left(\boldsymbol{v}^\top \boldsymbol{f}_1\right)^6\right]\right)^{\frac{1}{3}} \left[\mathbb{P}\left(\left|\boldsymbol{u}^\top \boldsymbol{f}_1\right| > C_1 \|\boldsymbol{u}\|_2 \sqrt{\log(n+d)}\right)\right]^{1/3} \\
&\lesssim \|\boldsymbol{u}\|_2^2 \|\boldsymbol{v}\|_2^2 \exp\left(-\frac{C_1^2 \log(n+d)}{6}\right) \overset{\text{(ii)}}{\lesssim} \|\boldsymbol{u}\|_2^2 \|\boldsymbol{v}\|_2^2 (n+d)^{-100},
\end{aligned}$$

where (i) comes from HÃ¶lder's inequality, and (ii) holds for C_1 large enough.

- Next, we shall apply the Bernstein inequality (Vershynin, 2018, Theorem 2.8.2) to bound α_1. Note that for each $j \in [n]$

$$\left\|X_j Y_j \mathbb{1}_{\mathcal{A}_j} - \mathbb{E}\left[X_j Y_j \mathbb{1}_{\mathcal{A}_j}\right]\right\|_{\psi_1} \leq \left\|X_j Y_j \mathbb{1}_{\mathcal{A}_j}\right\|_{\psi_1} + \left\|\mathbb{E}\left[X_j Y_j \mathbb{1}_{\mathcal{A}_j}\right]\right\|_{\psi_1}, \tag{E.14}$$

where $\|\cdot\|_{\psi_1}$ denotes the sub-exponential norm (Vershynin, 2012). The first term of (E.14) obeys

$$\begin{aligned}
\left\|X_j Y_j \mathbb{1}_{\mathcal{A}_j}\right\|_{\psi_1} &\leq C_1^2 \|\boldsymbol{u}\|_2^2 \log(n+d) \|Y_j\|_{\psi_1} \lesssim C_1^2 \|\boldsymbol{u}\|_2^2 \log(n+d) \left\|\boldsymbol{v}^\top \boldsymbol{f}_j\right\|_{\psi_2}^2 \\
&\lesssim C_1^2 \log(n+d) \|\boldsymbol{u}\|_2^2 \|\boldsymbol{v}\|_2^2,
\end{aligned}$$

where $\|\cdot\|_{\psi_2}$ denotes the sub-Gaussian norm (Vershynin, 2012). Turning to the second term of (E.14), we have

$$\begin{aligned}
\left\|\mathbb{E}\left[X_j Y_j \mathbb{1}_{\mathcal{A}_j}\right]\right\|_{\psi_1} &\lesssim \mathbb{E}\left[X_j Y_j \mathbb{1}_{\mathcal{A}_j}\right] \leq \mathbb{E}\left[\left(\boldsymbol{u}^\top \boldsymbol{f}_j\right)^2 \left(\boldsymbol{v}^\top \boldsymbol{f}_j\right)^2\right] \\
&= \|\boldsymbol{u}\|_2^2 \|\boldsymbol{v}\|_2^2 + 2\left(\boldsymbol{u}^\top \boldsymbol{v}\right)^2 \leq 3\|\boldsymbol{u}\|_2^2 \|\boldsymbol{v}\|_2^2,
\end{aligned}$$

where the last step arises from Cauchy-Schwarz. The above results taken collectively give

$$\left\|X_j Y_j \mathbb{1}_{\mathcal{A}_j} - \mathbb{E}\left[X_j Y_j \mathbb{1}_{\mathcal{A}_j}\right]\right\|_{\psi_1} \lesssim C_1^2 \log(n+d) \|\boldsymbol{u}\|_2^2 \|\boldsymbol{v}\|_2^2.$$

Applying the Bernstein inequality (Vershynin, 2018, Theorem 2.8.2) then yields

$$\alpha_1 \lesssim \sqrt{\frac{\log(n+d)^3}{n}} \|\boldsymbol{u}\|_2^2 \|\boldsymbol{v}\|_2^2 + \frac{\log^2(n+d)}{n} \|\boldsymbol{u}\|_2^2 \|\boldsymbol{v}\|_2^2 \lesssim \sqrt{\frac{\log(n+d)^3}{n}} \|\boldsymbol{u}\|_2^2 \|\boldsymbol{v}\|_2^2$$

with probability exceeding $1 - O((n+d)^{-100})$, provided that $n \gg \log(n+d)$.

Combine the preceding bounds on α_1 and α_2 to achieve

$$\begin{aligned}
\left|\frac{1}{n}\sum_{j=1}^n \left[X_j Y_j - \mathbb{E}\left[X_j Y_j\right]\right]\right| &\leq \alpha_1 + \alpha_2 \lesssim \sqrt{\frac{\log(n+d)^3}{n}} \|\boldsymbol{u}\|_2^2 \|\boldsymbol{v}\|_2^2 + (n+d)^{-10} \|\boldsymbol{u}\|_2^2 \|\boldsymbol{v}\|_2^2 \\
&\lesssim \sqrt{\frac{\log(n+d)^3}{n}} \|\boldsymbol{u}\|_2^2 \|\boldsymbol{v}\|_2^2
\end{aligned}$$

with probability exceeding $1 - O((n+d)^{-100})$. It is straightforward to verify that

$$\mathbb{E}\left[X_j Y_j\right] = \|\boldsymbol{u}\|_2^2 \|\boldsymbol{v}\|_2^2 + 2\left(\boldsymbol{u}^\top \boldsymbol{v}\right)^2$$

for each $j \in [n]$, thus concluding the proof. □

E.2 Auxiliary lemmas for Theorem 11

E.2.1 Proof of Lemma 15

To begin with, we remind the reader that $\boldsymbol{R} \in \mathcal{O}^{r\times r}$ represents the rotation matrix that best aligns $\boldsymbol{U}$ and $\boldsymbol{U}^\star$, and the rotation matrix $\boldsymbol{Q}$ is chosen to satisfy $\boldsymbol{U}^\star\boldsymbol{Q} = \boldsymbol{U}^\natural$. In addition, we have also shown in (D.7) that $\boldsymbol{RQ} \in \mathcal{O}^{r\times r}$ is the rotation matrix that best aligns $\boldsymbol{U}$ and $\boldsymbol{U}^\natural$. Suppose for the moment that the assumptions of Theorem 5 are satisfied (which we shall verify shortly). Then conditional on $\boldsymbol{F}$, invoking Theorem 5 leads to

$$\boldsymbol{URQ} - \boldsymbol{U}^\natural = \left[\boldsymbol{E}\boldsymbol{M}^{\natural\top} + \mathcal{P}_{\mathsf{off\text{-}diag}}\left(\boldsymbol{E}\boldsymbol{E}^\top\right)\right]\boldsymbol{U}^\natural\left(\boldsymbol{\Sigma}^\natural\right)^{-2} + \boldsymbol{\Psi}, \tag{E.15}$$

where the residual matrix $\boldsymbol{\Psi}$ can be controlled with high probability as follows: for each $l \in [d]$

$$\mathbb{P}\left(\left\|\boldsymbol{\Psi}_{l,\cdot}\right\|_2 \lesssim \zeta_{\mathsf{2nd},l}\left(\boldsymbol{F}\right) \,\middle|\, \boldsymbol{F}\right) \geq 1 - O\left((n+d)^{-10}\right).$$

Here, the quantity $\zeta_{\mathsf{2nd}}\left(\boldsymbol{F}\right)$ is defined as

$$\zeta_{\mathsf{2nd},l}\left(\boldsymbol{F}\right) \coloneqq \left\|\boldsymbol{U}^\natural_{l,\cdot}\right\|_2\left(\kappa^{\natural2}\sqrt{\frac{\mu^\natural r}{d}}\frac{\zeta_{\mathsf{1st}}\left(\boldsymbol{F}\right)}{\sigma_r^{\natural2}} + \kappa^{\natural2}\frac{\zeta^2_{\mathsf{1st}}\left(\boldsymbol{F}\right)}{\sigma_r^{\natural4}}\right) + \kappa^{\natural2}\frac{\zeta_{\mathsf{1st}}\left(\boldsymbol{F}\right)\zeta_{\mathsf{1st},l}\left(\boldsymbol{F}\right)}{\sigma_r^{\natural4}}\sqrt{\frac{\mu^\natural r}{d}}$$

with the quantities $\zeta_{\mathsf{1st}}\left(\boldsymbol{F}\right)$ and $\zeta_{\mathsf{1st},l}\left(\boldsymbol{F}\right)$ given by

$$\begin{aligned}
\zeta_{\mathsf{1st}}\left(\boldsymbol{F}\right) &\coloneqq \sigma^2\sqrt{nd}\log\left(n+d\right) + \sigma\sigma_1^\natural\sqrt{d\log\left(n+d\right)},\\
\zeta_{\mathsf{1st},l}\left(\boldsymbol{F}\right) &\coloneqq \sigma\sigma_l\sqrt{nd}\log\left(n+d\right) + \sigma_l\sigma_1^\natural\sqrt{d\log\left(n+d\right)}.
\end{aligned}$$

Given that $\boldsymbol{U}^\star = \boldsymbol{U}^\natural\boldsymbol{Q}^\top$, the above decomposition (E.15) can alternatively be written as

$$\boldsymbol{UR} - \boldsymbol{U}^\star = \underbrace{\left[\boldsymbol{E}\boldsymbol{M}^{\natural\top} + \mathcal{P}_{\mathsf{off\text{-}diag}}\left(\boldsymbol{E}\boldsymbol{E}^\top\right)\right]\boldsymbol{U}^\natural\left(\boldsymbol{\Sigma}^\natural\right)^{-2}\boldsymbol{Q}^\top}_{=:\boldsymbol{Z}} + \boldsymbol{\Psi}\boldsymbol{Q}^\top.$$

When the event $\mathcal{E}_{\mathsf{good}}$ occurs, we can see from (D.13), (D.14), (D.15) and (D.17) that

$$\begin{aligned}
\zeta_{\mathsf{1st},l}\left(\boldsymbol{F}\right) &\lesssim \sigma_{\mathsf{ub}}\sigma_{\mathsf{ub},l}\sqrt{nd}\log\left(n+d\right) + \sigma_{\mathsf{ub},l}\sigma_1^\star\sqrt{d\log\left(n+d\right)}\\
&\asymp \sqrt{\frac{\mu r\log^4\left(n+d\right)}{np^2}}\sigma_1^\star\left\|\boldsymbol{U}^\star_{l,\cdot}\boldsymbol{\Sigma}^\star\right\|_2 + \frac{\omega_l^\star\omega_{\max}}{p}\sqrt{\frac{d}{n}}\log\left(n+d\right) + \sigma_1^\star\sqrt{\frac{d}{np}}\left\|\boldsymbol{U}^\star_{l,\cdot}\boldsymbol{\Sigma}^\star\right\|_2\log\left(n+d\right)\\
&\quad + \sigma_1^\star\omega_l^\star\sqrt{\frac{d\log\left(n+d\right)}{np}} + \frac{\omega_{\max}}{p}\sqrt{\frac{d}{n}}\log^{3/2}\left(n+d\right)\left\|\boldsymbol{U}^\star_{l,\cdot}\boldsymbol{\Sigma}^\star\right\|_2 + \frac{\omega_l^\star}{p}\sqrt{\frac{\mu r\log^3\left(n+d\right)}{n}}\sigma_1^\star\\
&=: \zeta_{\mathsf{1st},l},
\end{aligned}$$

$$\begin{aligned}
\zeta_{\mathsf{1st}}\left(\boldsymbol{F}\right) &\lesssim \sigma_{\mathsf{ub}}^2\sqrt{nd}\log\left(n+d\right) + \sigma_{\mathsf{ub}}\sigma_1^\natural\sqrt{d\log\left(n+d\right)}\\
&\asymp \frac{\mu r\log^2\left(n+d\right)}{\sqrt{nd}p}\sigma_1^{\star2} + \frac{\omega_{\max}^2}{p}\sqrt{\frac{d}{n}}\log\left(n+d\right) + \sigma_1^{\star2}\sqrt{\frac{\mu r}{np}}\log\left(n+d\right) + \sigma_1^\star\omega_{\max}\sqrt{\frac{d\log\left(n+d\right)}{np}}\\
&=: \zeta_{\mathsf{1st}},
\end{aligned}$$

and

$$\begin{aligned}
\zeta_{\mathsf{2nd},l}\left(\boldsymbol{F}\right) &\lesssim \left\|\boldsymbol{U}^\star_{l,\cdot}\right\|_2\left(\sqrt{\frac{\kappa^3\mu r\log\left(n+d\right)}{d}}\frac{\zeta_{\mathsf{1st}}}{\sigma_r^{\star2}} + \kappa\frac{\zeta_{\mathsf{1st}}^2}{\sigma_r^{\star4}}\right) + \frac{\zeta_{\mathsf{1st}}\zeta_{\mathsf{1st},l}}{\sigma_r^{\star4}}\sqrt{\frac{\kappa^3\mu r\log\left(n+d\right)}{d}}\\
&=: \zeta_{\mathsf{2nd},l}.
\end{aligned}$$

These bounds taken together imply that

$$\mathbb{P}\left(\left\|\boldsymbol{\Psi}_{l,\cdot}\right\|_2\mathbb{1}_{\mathcal{E}_{\mathsf{good}}} \lesssim \zeta_{\mathsf{2nd},l} \,\middle|\, \boldsymbol{F}\right) \geq 1 - O\left((n+d)^{-10}\right).$$

Additionally, in view of the facts that $\mathcal{E}_{\mathsf{good}}$ is $\sigma(\boldsymbol{F})$-measurable and

$$\left\|\boldsymbol{\Psi}_{l,\cdot}\right\|_2 = \left\|\boldsymbol{\Psi}_{l,\cdot}\boldsymbol{Q}^\top\right\|_2 = \left\|(\boldsymbol{U}\boldsymbol{R} - \boldsymbol{U}^\star - \boldsymbol{Z})_{l,\cdot}\right\|_{2,\infty},$$

one can readily demonstrate that

$$\mathbb{P}\left(\left\|(\boldsymbol{U}\boldsymbol{R} - \boldsymbol{U}^\star - \boldsymbol{Z})_{l,\cdot}\right\|_2 \lesssim \zeta_{\mathsf{2nd},l} \,\middle|\, \boldsymbol{F}\right) \geq 1 - O\left((n+d)^{-10}\right)$$

on the high-probability event $\mathcal{E}_{\mathsf{good}}$.

It remains to verify the assumptions of Theorem 5, which requires

$$d \gtrsim \kappa^{\natural 4}\mu^\natural r + \mu^{\natural 2} r \log^2(n+d), \quad n \gtrsim r\log^4(n+d), \quad B \lesssim \frac{\sigma_{\mathsf{ub}}\min\left\{\sqrt{n_2}, \sqrt[4]{n_1 n_2}\right\}}{\sqrt{\log n}}, \quad \zeta_{\mathsf{1st}} \ll \frac{\sigma_r^{\natural 2}}{\kappa^{\natural 2}} \tag{E.16}$$

whenever $\mathcal{E}_{\mathsf{good}}$ occurs. According to (D.13), (D.14), (D.15), (D.18) and the definition of ζ_{2nd}, the conditions in (E.16) are guaranteed to hold as long as

$$d \gtrsim \kappa^3\mu^2 r\log^4(n+d), \qquad n \gtrsim r\log^4(n+d), \qquad \text{and} \qquad \zeta_{\mathsf{1st}} \ll \frac{\sigma_r^{\star 2}}{\kappa}.$$

Finally, we would like to take a closer inspection on the condition $\zeta_{\mathsf{1st}} \ll \sigma_r^{\star 2}/\kappa$; in fact, we seek to derive sufficient conditions which guarantee that $\zeta_{\mathsf{1st}}/\sigma_r^{\star 2} \ll \delta$ for any given $\delta > 0$, which will also be useful in subsequent analysis. It is straightforward to check that $\zeta_{\mathsf{1st}}/\sigma_r^{\star 2} \ll \delta$ is equivalent to

$$ndp^2 \gg \delta^{-2}\kappa^2\mu^2 r^2\log^4(n+d), \qquad np \gg \delta^{-2}\kappa^2\mu r\log^2(n+d), \tag{E.17a}$$

$$\text{and} \qquad \frac{\omega_{\max}^2}{p\sigma_r^{\star 2}}\sqrt{\frac{d}{n}} \ll \frac{\delta}{\log(n+d)}, \qquad \frac{\omega_{\max}}{\sigma_r^\star}\sqrt{\frac{d}{np}} \ll \frac{\delta}{\sqrt{\kappa\log(n+d)}}. \tag{E.17b}$$

By taking $\delta \coloneqq 1/\kappa$, we see that $\zeta_{\mathsf{1st}} \ll \sigma_r^{\star 2}/\kappa$ is guaranteed as long as the following conditions hold:

$$ndp^2 \gg \kappa^4\mu^2 r^2\log^4(n+d), \qquad np \gg \kappa^4\mu r\log^2(n+d)$$

$$\text{and} \qquad \frac{\omega_{\max}^2}{p\sigma_r^{\star 2}}\sqrt{\frac{d}{n}} \ll \frac{1}{\kappa\log(n+d)}, \qquad \frac{\omega_{\max}}{\sigma_r^\star}\sqrt{\frac{d}{np}} \ll \frac{1}{\sqrt{\kappa^3\log(n+d)}}.$$

This concludes the proof.

E.2.2 Proof of Lemma 16

Let us write $\widetilde{\boldsymbol{\Sigma}}_l$ as the superposition of two components:

$$\begin{aligned}\widetilde{\boldsymbol{\Sigma}}_l = &\underbrace{\boldsymbol{Q}(\boldsymbol{\Sigma}^\natural)^{-1}\boldsymbol{V}^{\natural\top}\mathsf{diag}\left\{\sigma_{l,1}^2,\dots,\sigma_{l,n}^2\right\}\boldsymbol{V}^\natural(\boldsymbol{\Sigma}^\natural)^{-1}\boldsymbol{Q}^\top}_{=:\widetilde{\boldsymbol{\Sigma}}_{l,1}} \\ &+ \underbrace{\boldsymbol{Q}(\boldsymbol{\Sigma}^\natural)^{-2}\boldsymbol{U}^{\natural\top}\mathsf{diag}\left\{\sum_{j:j\neq l}\sigma_{l,j}^2\sigma_{1,j}^2,\dots,\sum_{j:j\neq l}\sigma_{l,j}^2\sigma_{d,j}^2\right\}\boldsymbol{U}^\natural(\boldsymbol{\Sigma}^\natural)^{-2}\boldsymbol{Q}^\top}_{=:\widetilde{\boldsymbol{\Sigma}}_{l,2}}.\end{aligned}$$

We shall control $\widetilde{\boldsymbol{\Sigma}}_{l,1}$ and $\widetilde{\boldsymbol{\Sigma}}_{l,2}$ separately. Throughout this subsection we assume that $\mathcal{E}_{\mathsf{good}}$ happens.

Step 1: identifying a good approximation of $\widetilde{\boldsymbol{\Sigma}}_{l,1}$. We start with the concentration of the first component $\widetilde{\boldsymbol{\Sigma}}_{l,1}$. The matrices in the middle part satisfy

$$\boldsymbol{V}^{\natural\top}\mathsf{diag}\left\{\sigma_{l,1}^2,\dots,\sigma_{l,n}^2\right\}\boldsymbol{V}^{\natural}=\frac{1}{n}\boldsymbol{J}^\top\boldsymbol{F}\mathsf{diag}\left\{\sigma_{l,1}^2,\dots,\sigma_{l,n}^2\right\}\boldsymbol{F}^\top\boldsymbol{J}=\boldsymbol{J}^\top\left(\frac{1}{n}\sum_{j=1}^n\sigma_{l,j}^2\boldsymbol{f}_j\boldsymbol{f}_j^\top\right)\boldsymbol{J},$$

where we have made use of the identity (D.4). To control the term within the parentheses of the above identity, we can make use of the variance calculation in (D.9) to obtain

$$\frac{1}{n}\sum_{j=1}^n\sigma_{l,j}^2\boldsymbol{f}_j\boldsymbol{f}_j^\top=\frac{1-p}{n^2p}\sum_{j=1}^n\left(\boldsymbol{U}_{l,\cdot}^\star\boldsymbol{\Sigma}^\star\boldsymbol{f}_j\right)^2\boldsymbol{f}_j\boldsymbol{f}_j^\top+\frac{\omega_l^{\star2}}{n^2p}\sum_{j=1}^n\boldsymbol{f}_j\boldsymbol{f}_j^\top.$$

In addition, it is seen from (D.21a) and (D.10) that

$$\left\|\frac{1}{n}\sum_{j=1}^n\left(\boldsymbol{U}_{l,\cdot}^\star\boldsymbol{\Sigma}^\star\boldsymbol{f}_j\right)^2\boldsymbol{f}_j\boldsymbol{f}_j^\top-\left\|\boldsymbol{U}_{l,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2^2\boldsymbol{I}_r-2\boldsymbol{\Sigma}^\star\boldsymbol{U}_{l,\cdot}^{\star\top}\boldsymbol{U}_{l,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|\lesssim\sqrt{\frac{r\log^3(n+d)}{n}}\left\|\boldsymbol{U}_{l,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2^2$$

$$\text{and}\qquad\left\|\frac{1}{n}\sum_{j=1}^n\boldsymbol{f}_i\boldsymbol{f}_i^\top-\boldsymbol{I}_r\right\|=\left\|\frac{1}{n}\boldsymbol{F}\boldsymbol{F}^\top-\boldsymbol{I}_r\right\|\lesssim\sqrt{\frac{r+\log(n+d)}{n}}.$$

These bounds taken together allow us to express

$$\frac{1}{n}\sum_{j=1}^n\sigma_{l,j}^2\boldsymbol{f}_j\boldsymbol{f}_j^\top=\frac{1-p}{np}\left(\left\|\boldsymbol{U}_{l,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2^2\boldsymbol{I}_r+2\boldsymbol{\Sigma}^\star\boldsymbol{U}_{l,\cdot}^{\star\top}\boldsymbol{U}_{l,\cdot}^\star\boldsymbol{\Sigma}^\star\right)+\frac{\omega_l^{\star2}}{np}\boldsymbol{I}_r+\boldsymbol{R}_{1,1}$$

for some residual matrix $\boldsymbol{R}_{1,1}$ satisfying

$$\left\|\boldsymbol{R}_{1,1}\right\|\lesssim\frac{1-p}{np}\sqrt{\frac{r\log^3(n+d)}{n}}\left\|\boldsymbol{U}_{l,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2^2+\frac{\omega_l^{\star2}}{np}\sqrt{\frac{r+\log(n+d)}{n}}.$$

Putting the above pieces together, we arrive at

$$\begin{aligned}\widetilde{\boldsymbol{\Sigma}}_{l,1}&=\boldsymbol{Q}(\boldsymbol{\Sigma}^\natural)^{-1}\boldsymbol{J}^\top\left(\frac{1}{n}\sum_{j=1}^n\sigma_{l,j}^2\boldsymbol{f}_j^\star\boldsymbol{f}_j^{\star\top}\right)\boldsymbol{J}(\boldsymbol{\Sigma}^\natural)^{-1}\boldsymbol{Q}^\top\\&=\underbrace{\left(\frac{1-p}{np}\left\|\boldsymbol{U}_{l,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2^2+\frac{\omega_l^{\star2}}{np}\right)\boldsymbol{Q}(\boldsymbol{\Sigma}^\natural)^{-1}\boldsymbol{J}^\top\boldsymbol{J}(\boldsymbol{\Sigma}^\natural)^{-1}\boldsymbol{Q}^\top}_{=:\boldsymbol{\Sigma}_{1,1}}\\&\quad+\underbrace{\frac{2(1-p)}{np}\boldsymbol{Q}(\boldsymbol{\Sigma}^\natural)^{-1}\boldsymbol{J}^\top\boldsymbol{\Sigma}^\star\boldsymbol{U}_{l,\cdot}^{\star\top}\boldsymbol{U}_{l,\cdot}^\star\boldsymbol{\Sigma}^\star\boldsymbol{J}(\boldsymbol{\Sigma}^\natural)^{-1}\boldsymbol{Q}^\top}_{=:\boldsymbol{\Sigma}_{1,2}}+\boldsymbol{R}_{1,2}\end{aligned}$$

for some residual matrix

$$\boldsymbol{R}_{1,2}=\boldsymbol{Q}(\boldsymbol{\Sigma}^\natural)^{-1}\boldsymbol{J}^\top\boldsymbol{R}_{1,1}\boldsymbol{J}(\boldsymbol{\Sigma}^\natural)^{-1}\boldsymbol{Q}^\top.$$

This motivates us to look at $\boldsymbol{\Sigma}_{1,1}$ and $\boldsymbol{\Sigma}_{1,2}$ separately.

- Regarding the matrix $\boldsymbol{\Sigma}_{1,1}$, we make the observation that

$$\begin{aligned}\left\|(\boldsymbol{\Sigma}^\natural)^{-1}\boldsymbol{J}^\top\boldsymbol{J}(\boldsymbol{\Sigma}^\natural)^{-1}-(\boldsymbol{\Sigma}^\natural)^{-2}\right\|&\overset{\text{(i)}}{=}\left\|(\boldsymbol{\Sigma}^\natural)^{-1}\left(\boldsymbol{J}^\top\boldsymbol{J}-\boldsymbol{Q}^\top\boldsymbol{Q}\right)(\boldsymbol{\Sigma}^\natural)^{-1}\right\|\\&\leq\left\|(\boldsymbol{\Sigma}^\natural)^{-1}(\boldsymbol{J}-\boldsymbol{Q})^\top\boldsymbol{J}(\boldsymbol{\Sigma}^\natural)^{-1}\right\|+\left\|(\boldsymbol{\Sigma}^\natural)^{-1}\boldsymbol{Q}^\top(\boldsymbol{J}-\boldsymbol{Q})(\boldsymbol{\Sigma}^\natural)^{-1}\right\|\end{aligned}$$

$$\overset{\text{(ii)}}{\lesssim} \frac{1}{\sigma_r^{\star 2}} \|\boldsymbol{J}-\boldsymbol{Q}\| \left(\|\boldsymbol{Q}\|+\|\boldsymbol{J}\|\right) \overset{\text{(iii)}}{\lesssim} \frac{\kappa}{\sigma_r^{\star 2}} \sqrt{\frac{r+\log(n+d)}{n}}. \tag{E.18}$$

Here, (i) comes from the fact that $\boldsymbol{Q}$ is a orthonormal matrix, (ii) follows from the property (D.13), whereas (iii) utilizes the property (D.19) and its direct application

$$\|\boldsymbol{J}\| \le \|\boldsymbol{Q}\| + \|\boldsymbol{J}-\boldsymbol{Q}\| \le 1+\kappa\sqrt{\frac{r+\log(n+d)}{n}} \le 2, \tag{E.19}$$

provided that $n \gg \kappa^2(r+\log(n+d))$. In addition, from (D.19) and (D.13) we know that

$$\begin{aligned}\left\|\boldsymbol{Q}(\boldsymbol{\Sigma}^\natural)^{-1} - (\boldsymbol{\Sigma}^\star)^{-1}\boldsymbol{Q}\right\| &\le \left\|(\boldsymbol{\Sigma}^\star)^{-1}\left(\boldsymbol{\Sigma}^\star\boldsymbol{Q}-\boldsymbol{Q}\boldsymbol{\Sigma}^\natural\right)(\boldsymbol{\Sigma}^\natural)^{-1}\right\| \lesssim \frac{1}{\sigma_r^{\star 2}}\left\|\boldsymbol{\Sigma}^\star\boldsymbol{Q}-\boldsymbol{Q}\boldsymbol{\Sigma}^\natural\right\| \\ &\lesssim \frac{\kappa}{\sigma_r^\star}\sqrt{\frac{r+\log(n+d)}{n}}. \end{aligned}\tag{E.20}$$

This immediately leads to

$$\begin{aligned}\left\|\boldsymbol{Q}(\boldsymbol{\Sigma}^\natural)^{-2}\boldsymbol{Q}^\top - (\boldsymbol{\Sigma}^\star)^{-2}\right\| &= \left\|\boldsymbol{Q}(\boldsymbol{\Sigma}^\natural)^{-1}(\boldsymbol{\Sigma}^\natural)^{-1}\boldsymbol{Q}^\top - (\boldsymbol{\Sigma}^\star)^{-1}\boldsymbol{Q}\boldsymbol{Q}^\top(\boldsymbol{\Sigma}^\star)^{-1}\right\| \\ &\le \left\|\left[\boldsymbol{Q}(\boldsymbol{\Sigma}^\natural)^{-1} - (\boldsymbol{\Sigma}^\star)^{-1}\boldsymbol{Q}\right](\boldsymbol{\Sigma}^\natural)^{-1}\boldsymbol{Q}^\top\right\| + \left\|(\boldsymbol{\Sigma}^\star)^{-1}\boldsymbol{Q}\left[\boldsymbol{Q}(\boldsymbol{\Sigma}^\natural)^{-1} - (\boldsymbol{\Sigma}^\star)^{-1}\boldsymbol{Q}\right]^\top\right\| \\ &\le \left(\left\|(\boldsymbol{\Sigma}^\natural)^{-1}\right\| + \left\|(\boldsymbol{\Sigma}^\star)^{-1}\right\|\right)\|\boldsymbol{Q}\|\left\|\boldsymbol{Q}(\boldsymbol{\Sigma}^\natural)^{-1} - (\boldsymbol{\Sigma}^\star)^{-1}\boldsymbol{Q}\right\| \\ &\lesssim \frac{\kappa}{\sigma_r^{\star 2}}\sqrt{\frac{r+\log(n+d)}{n}}, \end{aligned}\tag{E.21}$$

where we have again used (D.13). Taking (E.18) and (E.21) collectively gives

$$\begin{aligned}&\left\|\boldsymbol{Q}(\boldsymbol{\Sigma}^\natural)^{-1}\boldsymbol{J}^\top\boldsymbol{J}(\boldsymbol{\Sigma}^\natural)^{-1}\boldsymbol{Q}^\top - (\boldsymbol{\Sigma}^\star)^{-2}\right\| \\ &\quad\le \left\|\boldsymbol{Q}\left[(\boldsymbol{\Sigma}^\natural)^{-1}\boldsymbol{J}^\top\boldsymbol{J}(\boldsymbol{\Sigma}^\natural)^{-1} - (\boldsymbol{\Sigma}^\natural)^{-2}\right]\boldsymbol{Q}^\top\right\| + \left\|\boldsymbol{Q}(\boldsymbol{\Sigma}^\natural)^{-2}\boldsymbol{Q}^\top - (\boldsymbol{\Sigma}^\star)^{-2}\right\| \\ &\quad\le \left\|(\boldsymbol{\Sigma}^\natural)^{-1}\boldsymbol{J}^\top\boldsymbol{J}(\boldsymbol{\Sigma}^\natural)^{-1} - (\boldsymbol{\Sigma}^\natural)^{-2}\right\| + \left\|\boldsymbol{Q}(\boldsymbol{\Sigma}^\natural)^{-2}\boldsymbol{Q}^\top - (\boldsymbol{\Sigma}^\star)^{-2}\right\| \\ &\quad\lesssim \frac{\kappa}{\sigma_r^{\star 2}}\sqrt{\frac{r+\log(n+d)}{n}}. \end{aligned}$$

Substitution into the definition of $\boldsymbol{\Sigma}_{1,1}$ allows us to conclude that

$$\left\|\boldsymbol{\Sigma}_{1,1} - \left(\frac{1-p}{np}\left\|\boldsymbol{U}_{l,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2^2 + \frac{\sigma_l^2}{np}\right)(\boldsymbol{\Sigma}^\star)^{-2}\right\| \lesssim \left(\frac{1-p}{np}\left\|\boldsymbol{U}_{l,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2^2 + \frac{\omega_l^{\star 2}}{np}\right)\frac{\kappa}{\sigma_r^{\star 2}}\sqrt{\frac{r+\log(n+d)}{n}}. \tag{E.22}$$

- When it comes to the remaining term $\boldsymbol{\Sigma}_{1,2}$, we first notice that

$$\begin{aligned}&\left\|\boldsymbol{U}_{l,\cdot}^\star\boldsymbol{\Sigma}^\star\boldsymbol{J}(\boldsymbol{\Sigma}^\natural)^{-1} - \boldsymbol{U}_{l,\cdot}^\natural\right\|_2 \le \left\|\boldsymbol{U}_{l,\cdot}^\star\boldsymbol{\Sigma}^\star\boldsymbol{Q}(\boldsymbol{\Sigma}^\natural)^{-1} - \boldsymbol{U}_{l,\cdot}^\natural\right\|_2 + \left\|\boldsymbol{U}_{l,\cdot}^\star\boldsymbol{\Sigma}^\star\left(\boldsymbol{J}-\boldsymbol{Q}\right)(\boldsymbol{\Sigma}^\natural)^{-1}\right\|_2 \\ &\quad\le \left\|\boldsymbol{U}_{l,\cdot}^\star\left(\boldsymbol{\Sigma}^\star\boldsymbol{Q}-\boldsymbol{Q}\boldsymbol{\Sigma}^\natural\right)(\boldsymbol{\Sigma}^\natural)^{-1}\right\|_2 + \left\|\boldsymbol{U}_{l,\cdot}^\star\boldsymbol{Q} - \boldsymbol{U}_{l,\cdot}^\natural\right\|_2 + \left\|\boldsymbol{U}_{l,\cdot}^\star\boldsymbol{\Sigma}^\star\left(\boldsymbol{J}-\boldsymbol{Q}\right)(\boldsymbol{\Sigma}^\natural)^{-1}\right\|_2 \\ &\quad\lesssim \frac{1}{\sigma_r^\star}\left\|\boldsymbol{U}_{l,\cdot}^\star\right\|_2\left\|\boldsymbol{\Sigma}^\star\boldsymbol{Q}-\boldsymbol{Q}\boldsymbol{\Sigma}^\natural\right\| + \frac{\sigma_1^\star}{\sigma_r^\star}\left\|\boldsymbol{U}_{l,\cdot}^\star\right\|_2\left\|\boldsymbol{J}-\boldsymbol{Q}\right\| \\ &\quad\lesssim \kappa^{3/2}\sqrt{\frac{r+\log(n+d)}{n}}\left\|\boldsymbol{U}_{l,\cdot}^\star\right\|_2, \end{aligned}$$

where the penultimate line uses (D.13) and fact that $\boldsymbol{U}^\star\boldsymbol{Q} = \boldsymbol{U}^\natural$, and the last line relies on the property (D.19). An immediate consequence is that

$$\left\|\boldsymbol{U}_{l,\cdot}^\star\boldsymbol{\Sigma}^\star\boldsymbol{J}(\boldsymbol{\Sigma}^\natural)^{-1}\right\|_2 \le \left\|\boldsymbol{U}_{l,\cdot}^\natural\right\|_2 + \left\|\boldsymbol{U}_{l,\cdot}^\star\boldsymbol{\Sigma}^\star\boldsymbol{J}(\boldsymbol{\Sigma}^\natural)^{-1} - \boldsymbol{U}_{l,\cdot}^\natural\right\|_2 \lesssim \left\|\boldsymbol{U}_{l,\cdot}^\star\right\|_2,$$

provided that $n \gg \kappa^3 r + \kappa^3 \log(n+d)$. This combined with the fact $\boldsymbol{U}^\star \boldsymbol{Q} = \boldsymbol{U}^\natural$ immediately yields

$$
\begin{aligned}
&\left\|\boldsymbol{Q}(\boldsymbol{\Sigma}^\natural)^{-1}\boldsymbol{J}^\top \boldsymbol{\Sigma}^\star \boldsymbol{U}_{l,\cdot}^{\star\top} \boldsymbol{U}_{l,\cdot}^\star \boldsymbol{\Sigma}^\star \boldsymbol{J}(\boldsymbol{\Sigma}^\natural)^{-1}\boldsymbol{Q}^\top - \boldsymbol{U}_{l,\cdot}^{\star\top}\boldsymbol{U}_{l,\cdot}^\star\right\| \\
&\quad = \left\|(\boldsymbol{\Sigma}^\natural)^{-1}\boldsymbol{J}^\top \boldsymbol{\Sigma}^\star \boldsymbol{U}_{l,\cdot}^{\star\top} \boldsymbol{U}_{l,\cdot}^\star \boldsymbol{\Sigma}^\star \boldsymbol{J}(\boldsymbol{\Sigma}^\natural)^{-1} - \boldsymbol{U}_{l,\cdot}^{\natural\top}\boldsymbol{U}_{l,\cdot}^\natural\right\| \\
&\quad \le \left\|(\boldsymbol{\Sigma}^\natural)^{-1}\boldsymbol{J}^\top \boldsymbol{\Sigma}^\star \boldsymbol{U}_{l,\cdot}^{\star\top}\left(\boldsymbol{U}_{l,\cdot}^\star \boldsymbol{\Sigma}^\star \boldsymbol{J}(\boldsymbol{\Sigma}^\natural)^{-1} - \boldsymbol{U}_{l,\cdot}^\natural\right)\right\| + \left\|\left(\boldsymbol{U}_{l,\cdot}^\star \boldsymbol{\Sigma}^\star \boldsymbol{J}(\boldsymbol{\Sigma}^\natural)^{-1} - \boldsymbol{U}_{l,\cdot}^\natural\right)^\top \boldsymbol{U}_{l,\cdot}^\natural\right\| \\
&\quad \le \left\|\boldsymbol{U}_{l,\cdot}^\star \boldsymbol{\Sigma}^\star \boldsymbol{J}(\boldsymbol{\Sigma}^\natural)^{-1} - \boldsymbol{U}_{l,\cdot}^\natural\right\|_2 \left(\left\|\boldsymbol{U}_{l,\cdot}^\star \boldsymbol{\Sigma}^\star \boldsymbol{J}(\boldsymbol{\Sigma}^\natural)^{-1}\right\|_2 + \left\|\boldsymbol{U}_{l,\cdot}^\natural\right\|_2\right) \\
&\quad \lesssim \kappa^{3/2}\sqrt{\frac{r+\log(n+d)}{n}}\left\|\boldsymbol{U}_{l,\cdot}^\star\right\|_2^2,
\end{aligned}
$$

and as a result,

$$
\left\|\boldsymbol{\Sigma}_{1,2} - \frac{2(1-p)}{np}\boldsymbol{U}_{l,\cdot}^{\star\top}\boldsymbol{U}_{l,\cdot}^\star\right\| \lesssim \frac{1-p}{np}\kappa^{3/2}\sqrt{\frac{r+\log(n+d)}{n}}\left\|\boldsymbol{U}_{l,\cdot}^\star\right\|_2^2. \tag{E.23}
$$

Combining the above bounds (E.22) and (E.23), we can demonstrate that

$$
\widetilde{\boldsymbol{\Sigma}}_{l,1} = \boldsymbol{\Sigma}_{1,1} + \boldsymbol{\Sigma}_{1,2} + \boldsymbol{R}_{1,2} = \underbrace{\left(\frac{1-p}{np}\left\|\boldsymbol{U}_{l,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2^2 + \frac{\omega_l^{\star 2}}{np}\right)(\boldsymbol{\Sigma}^\star)^{-2} + \frac{2(1-p)}{np}\boldsymbol{U}_{l,\cdot}^{\star\top}\boldsymbol{U}_{l,\cdot}^\star}_{=:\,\boldsymbol{\Sigma}_{l,1}^\star} + \boldsymbol{R}_1
$$

holds for some residual matrix $\boldsymbol{R}_l$ satisfying

$$
\begin{aligned}
\|\boldsymbol{R}_1\| &\le \|\boldsymbol{R}_{1,2}\| + \left\|\boldsymbol{\Sigma}_{1,1} - \left(\frac{1-p}{np}\left\|\boldsymbol{U}_{l,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2^2 + \frac{\omega_l^{\star 2}}{np}\right)(\boldsymbol{\Sigma}^\star)^{-2}\right\| + \left\|\boldsymbol{\Sigma}_{1,2} - \frac{2(1-p)}{np}\boldsymbol{U}_{l,\cdot}^{\star\top}\boldsymbol{U}_{l,\cdot}^\star\right\| \\
&\overset{\text{(i)}}{\lesssim} \frac{1}{\sigma_r^{\star 2}}\left(\frac{1-p}{np}\sqrt{\frac{r\log^3(n+d)}{n}}\left\|\boldsymbol{U}_{l,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2^2 + \frac{\omega_l^{\star 2}}{np}\sqrt{\frac{r+\log(n+d)}{n}}\right) \\
&\quad + \left(\frac{1-p}{np}\left\|\boldsymbol{U}_{l,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2^2 + \frac{\omega_l^{\star 2}}{np}\right)\frac{\kappa}{\sigma_r^{\star 2}}\sqrt{\frac{r+\log(n+d)}{n}} + \frac{1-p}{np}\kappa^{3/2}\sqrt{\frac{r+\log(n+d)}{n}}\left\|\boldsymbol{U}_{l,\cdot}^\star\right\|_2^2 \\
&\lesssim \frac{1-p}{np\sigma_r^{\star 2}}\left(\sqrt{\frac{r\log^3(n+d)}{n}} + \kappa^{3/2}\sqrt{\frac{r+\log(n+d)}{n}}\right)\left\|\boldsymbol{U}_{l,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2^2 + \frac{\omega_l^{\star 2}}{np\sigma_r^{\star 2}}\kappa\sqrt{\frac{r+\log(n+d)}{n}}.
\end{aligned}
$$

Here, (i) has made use of (D.13) and the property (E.19).

Step 2: controlling the spectrum of $\widetilde{\boldsymbol{\Sigma}}_{l,1}$. We first observe that

$$
\left(\frac{1-p}{np}\left\|\boldsymbol{U}_{l,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2^2 + \frac{\omega_l^{\star 2}}{np}\right)(\boldsymbol{\Sigma}^\star)^{-2} + \frac{2(1-p)}{np}\boldsymbol{U}_{l,\cdot}^{\star\top}\boldsymbol{U}_{l,\cdot}^\star \succeq \boldsymbol{\Sigma}_{l,1}^\star \succeq \left(\frac{1-p}{np}\left\|\boldsymbol{U}_{l,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2^2 + \frac{\omega_l^{\star 2}}{np}\right)(\boldsymbol{\Sigma}^\star)^{-2},
$$

and as a result,

$$
\lambda_{\max}\left(\boldsymbol{\Sigma}_{l,1}^\star\right) \le \frac{3(1-p)}{np\sigma_r^{\star 2}}\left\|\boldsymbol{U}_{l,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2^2 + \frac{\omega_l^{\star 2}}{np\sigma_r^{\star 2}}, \tag{E.24a}
$$

$$
\lambda_{\min}\left(\boldsymbol{\Sigma}_{l,1}^\star\right) \ge \frac{1-p}{np\sigma_1^{\star 2}}\left\|\boldsymbol{U}_{l,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2^2 + \frac{\omega_l^{\star 2}}{np\sigma_1^{\star 2}}. \tag{E.24b}
$$

This immediately implies that the condition number of $\boldsymbol{\Sigma}_{l,1}^\star$ is at most 3κ. In addition, it follows from the preceding bounds that

$$
\frac{\|\boldsymbol{R}_1\|}{\lambda_{\min}\left(\boldsymbol{\Sigma}_{U,l}^\star\right)} \lesssim \sqrt{\frac{\kappa^2 r\log^3(n+d) + \kappa^5 r + \kappa^5\log(n+d)}{n}} \asymp \sqrt{\frac{\kappa^5 r\log^3(n+d)}{n}}.
$$

This means that $\|\boldsymbol{R}_l\| \ll \lambda_{\min}(\boldsymbol{\Sigma}_{l,1}^\star)$ holds as long as $n \gg \kappa^5 r \log^3(n+d)$, and as a consequence,

$$\lambda_{\min}\big(\widetilde{\boldsymbol{\Sigma}}_{l,1}\big) \in \left[\lambda_{\min}\left(\boldsymbol{\Sigma}_{l,1}^\star\right) - \|\boldsymbol{R}_1\|, \lambda_{\min}\left(\boldsymbol{\Sigma}_{l,1}^\star\right) + \|\boldsymbol{R}_1\|\right] \quad \Longrightarrow \quad \lambda_{\min}\big(\widetilde{\boldsymbol{\Sigma}}_{l}\big) \asymp \lambda_{\min}\left(\boldsymbol{\Sigma}_{U,l}^\star\right).$$

Therefore, one can conclude that

$$\left\|\widetilde{\boldsymbol{\Sigma}}_{l,1} - \boldsymbol{\Sigma}_{l,1}^\star\right\| = \|\boldsymbol{R}_1\| \lesssim \sqrt{\frac{\kappa^5 r \log^3(n+d)}{n}} \lambda_{\min}\left(\boldsymbol{\Sigma}_{U,l}^\star\right).$$

Step 3: identifying a good approximation of $\widetilde{\boldsymbol{\Sigma}}_{l,2}$. We now turn attention to approximating $\widetilde{\boldsymbol{\Sigma}}_{l,2}$, and it suffices to study $\sum_{j\neq l}^n \sigma_{l,j}^2 \sigma_{i,j}^2$ for each $i \in [d]$. In view of the expression (D.9), we have

$$\begin{aligned}
\sum_{j:j\neq l} \sigma_{l,j}^2\sigma_{i,j}^2 &= \sum_{j:j\neq l}\left[\frac{1-p}{np}\left(\boldsymbol{U}_{l,\cdot}^\star\boldsymbol{\Sigma}^\star\boldsymbol{f}_j\right)^2 + \frac{\omega_l^{\star 2}}{np}\right]\left[\frac{1-p}{np}\left(\boldsymbol{U}_{i,\cdot}^\star\boldsymbol{\Sigma}^\star\boldsymbol{f}_j\right)^2 + \frac{\omega_i^{\star 2}}{np}\right] \\
&= \left(\frac{1-p}{np}\right)^2 \underbrace{\sum_{j:j\neq l}\left(\boldsymbol{U}_{l,\cdot}^\star\boldsymbol{\Sigma}^\star\boldsymbol{f}_j\right)^2\left(\boldsymbol{U}_{i,\cdot}^\star\boldsymbol{\Sigma}^\star\boldsymbol{f}_j\right)^2}_{=:\alpha_1} + \frac{(1-p)\,\omega_i^{\star 2}}{n^2p^2}\underbrace{\sum_{j:j\neq l}\left(\boldsymbol{U}_{l,\cdot}^\star\boldsymbol{\Sigma}^\star\boldsymbol{f}_j\right)^2}_{=:\alpha_2} \\
&\quad + \frac{(1-p)\,\omega_l^{\star 2}}{n^2p^2}\underbrace{\sum_{j:j\neq l}\left(\boldsymbol{U}_{i,\cdot}^\star\boldsymbol{\Sigma}^\star\boldsymbol{f}_j\right)^2}_{=:\alpha_3} + \frac{(n-1)\,\omega_l^{\star 2}\omega_i^{\star 2}}{n^2p^2}.
\end{aligned} \tag{E.25}$$

To proceed, we can see from (D.21b), (D.21c) and (D.20) that

$$\begin{aligned}
\left|\frac{1}{n}\alpha_1 - \left[\left\|\boldsymbol{U}_{l,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2^2\left\|\boldsymbol{U}_{i,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2^2 + 2S_{l,i}^{\star 2}\right]\right| &\lesssim \sqrt{\frac{\log^3(n+d)}{n}}\left\|\boldsymbol{U}_{l,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2^2\left\|\boldsymbol{U}_{i,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2^2, \\
\left|\frac{1}{n}\alpha_2 - \left\|\boldsymbol{U}_{l,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2^2\right| &\lesssim \sqrt{\frac{\log(n+d)}{n}}\left\|\boldsymbol{U}_{l,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2^2, \\
\left|\frac{1}{n}\alpha_3 - \left\|\boldsymbol{U}_{i,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2^2\right| &\lesssim \sqrt{\frac{\log(n+d)}{n}}\left\|\boldsymbol{U}_{i,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2^2.
\end{aligned}$$

Substitution into (E.25) yields

$$\sum_{j:j\neq l}\sigma_{l,j}^2\sigma_{i,j}^2 = \underbrace{\frac{(1-p)^2}{np^2}\left(\left\|\boldsymbol{U}_{l,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2^2\left\|\boldsymbol{U}_{i,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2^2 + 2S_{l,i}^{\star 2}\right) + \frac{1-p}{np^2}\left(\omega_i^{\star 2}\left\|\boldsymbol{U}_{l,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2^2 + \omega_l^{\star 2}\left\|\boldsymbol{U}_{i,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2^2\right) + \frac{\omega_l^{\star 2}\omega_i^{\star 2}}{np^2}}_{=:\,d_{l,i}} + r_i \tag{E.26}$$

for some residual term r_i satisfying

$$|r_i| \lesssim \sqrt{\frac{\log^3(n+d)}{n}}\, d_{l,i}. \tag{E.27}$$

As a consequence, the above calculations together with the definition of $\widetilde{\boldsymbol{\Sigma}}_{l,2}$ allow one to write

$$\widetilde{\boldsymbol{\Sigma}}_{l,2} = \boldsymbol{Q}(\boldsymbol{\Sigma}^\natural)^{-2}\boldsymbol{U}^{\natural\top}\mathsf{diag}\left\{d_{l,i}\right\}_{i=1}^d\boldsymbol{U}^\natural(\boldsymbol{\Sigma}^\natural)^{-2}\boldsymbol{Q}^\top + \boldsymbol{R}_2 \tag{E.28}$$

for some residual matrix $\boldsymbol{R}_2$ satisfying

$$\begin{aligned}
\|\boldsymbol{R}_2\| &\le \left\|\boldsymbol{Q}(\boldsymbol{\Sigma}^\natural)^{-2}\boldsymbol{U}^{\natural\top}\mathsf{diag}\left\{|r_1|,\dots,|r_d|\right\}\boldsymbol{U}^\natural(\boldsymbol{\Sigma}^\natural)^{-2}\boldsymbol{Q}^\top\right\| \\
&\lesssim \frac{1}{\left(\sigma_r(\boldsymbol{\Sigma}^\natural)\right)^4}\max_{1\le i\le d}|r_i| \lesssim \sqrt{\frac{\log^3(n+d)}{n}}\max_{1\le i\le d}\frac{1}{\sigma_r^{\star 4}}d_{l,i},
\end{aligned}$$

where the last inequality makes use of (D.13) and (E.27).

Further, it turns out that the term $\boldsymbol{Q}(\boldsymbol{\Sigma}^{\natural})^{-2}\boldsymbol{U}^{\natural\top}$ used in (E.28) can be well approximated by $(\boldsymbol{\Sigma}^{\star})^{-2}\boldsymbol{Q}$. To see this, we note that

$$\begin{aligned}\left\|\boldsymbol{Q}(\boldsymbol{\Sigma}^{\natural})^{-2}-(\boldsymbol{\Sigma}^{\star})^{-2}\boldsymbol{Q}\right\| &\leq \left\|\left[\boldsymbol{Q}(\boldsymbol{\Sigma}^{\natural})^{-1}-(\boldsymbol{\Sigma}^{\star})^{-1}\boldsymbol{Q}\right](\boldsymbol{\Sigma}^{\natural})^{-1}\right\|+\left\|(\boldsymbol{\Sigma}^{\star})^{-1}\left[\boldsymbol{Q}(\boldsymbol{\Sigma}^{\natural})^{-1}-(\boldsymbol{\Sigma}^{\star})^{-1}\boldsymbol{Q}\right]\right\| \\ &\overset{\text{(i)}}{\lesssim} \frac{1}{\sigma_r^{\star}}\left\|\boldsymbol{Q}(\boldsymbol{\Sigma}^{\natural})^{-1}-(\boldsymbol{\Sigma}^{\star})^{-1}\boldsymbol{Q}\right\| \overset{\text{(ii)}}{\lesssim} \frac{\kappa}{\sigma_r^{\star2}}\sqrt{\frac{r+\log(n+d)}{n}},\end{aligned}$$

where (i) comes from (D.13), and (ii) results from (E.20). This combined with the identity $\boldsymbol{U}^{\star}=\boldsymbol{U}^{\natural}\boldsymbol{Q}^{\top}$ gives

$$\begin{aligned}\left\|\boldsymbol{Q}(\boldsymbol{\Sigma}^{\natural})^{-2}\boldsymbol{U}^{\natural\top}-(\boldsymbol{\Sigma}^{\star})^{-2}\boldsymbol{U}^{\star\top}\right\| &= \left\|\left[\boldsymbol{Q}(\boldsymbol{\Sigma}^{\natural})^{-2}-(\boldsymbol{\Sigma}^{\star})^{-2}\boldsymbol{Q}\right]\boldsymbol{U}^{\natural\top}\right\| \\ &\leq \left\|\boldsymbol{Q}(\boldsymbol{\Sigma}^{\natural})^{-2}-(\boldsymbol{\Sigma}^{\star})^{-2}\boldsymbol{Q}\right\| \\ &\lesssim \frac{\kappa}{\sigma_r^{\star2}}\sqrt{\frac{r+\log(n+d)}{n}}.\end{aligned}$$

Therefore, substituting this into (E.28) reveals that

$$\begin{aligned}&\left\|\widetilde{\boldsymbol{\Sigma}}_{l,2}-(\boldsymbol{\Sigma}^{\star})^{-2}\boldsymbol{U}^{\star\top}\mathsf{diag}\{d_{l,i}\}_{i=1}^{d}\boldsymbol{U}^{\star}(\boldsymbol{\Sigma}^{\star})^{-2}\right\| \\ &\quad\leq \left\|\boldsymbol{Q}(\boldsymbol{\Sigma}^{\natural})^{-2}\boldsymbol{U}^{\natural\top}\mathsf{diag}\{d_{l,i}\}_{i=1}^{d}\boldsymbol{U}^{\natural}(\boldsymbol{\Sigma}^{\natural})^{-2}\boldsymbol{Q}^{\top}-(\boldsymbol{\Sigma}^{\star})^{-2}\boldsymbol{U}^{\star\top}\mathsf{diag}\{d_{l,i}\}_{i=1}^{d}\boldsymbol{U}^{\star}(\boldsymbol{\Sigma}^{\star})^{-2}\right\|+\|\boldsymbol{R}_2\| \\ &\quad\leq \left\|\left(\boldsymbol{Q}(\boldsymbol{\Sigma}^{\natural})^{-2}\boldsymbol{U}^{\natural\top}-(\boldsymbol{\Sigma}^{\star})^{-2}\boldsymbol{U}^{\star\top}\right)\mathsf{diag}\{d_{l,i}\}_{i=1}^{d}\boldsymbol{U}^{\natural}(\boldsymbol{\Sigma}^{\natural})^{-2}\boldsymbol{Q}^{\top}\right\| \\ &\qquad+\left\|(\boldsymbol{\Sigma}^{\star})^{-2}\boldsymbol{U}^{\star\top}\mathsf{diag}\{d_{l,i}\}_{i=1}^{d}\left(\boldsymbol{Q}(\boldsymbol{\Sigma}^{\natural})^{-2}\boldsymbol{U}^{\natural\top}-(\boldsymbol{\Sigma}^{\star})^{-2}\boldsymbol{U}^{\star\top}\right)^{\top}\right\|+\|\boldsymbol{R}_2\| \\ &\quad\lesssim \frac{\kappa}{\sigma_r^{\star4}}\sqrt{\frac{r+\log(n+d)}{n}}\max_{1\leq i\leq d}d_{l,i}+\|\boldsymbol{R}_2\|,\end{aligned}$$

where the last relation relies on (D.13). These bounds taken collectively allow one to express

$$\widetilde{\boldsymbol{\Sigma}}_{l,2}=\boldsymbol{\Sigma}_{l,2}^{\star}+\boldsymbol{R}_3, \tag{E.29}$$

where

$$\boldsymbol{\Sigma}_{l,2}^{\star}\coloneqq(\boldsymbol{\Sigma}^{\star})^{-2}\boldsymbol{U}^{\star\top}\mathsf{diag}\{d_{l,i}\}_{i=1}^{d}\boldsymbol{U}^{\star}(\boldsymbol{\Sigma}^{\star})^{-2} \tag{E.30}$$

and $\boldsymbol{R}_3$ is some residual matrix satisfying

$$\|\boldsymbol{R}_3\|\lesssim\|\boldsymbol{R}_2\|+\frac{\kappa}{\sigma_r^{\star4}}\sqrt{\frac{r+\log(n+d)}{n}}\max_{1\leq i\leq d}d_{l,i}\lesssim\sqrt{\frac{\kappa^2r\log^3(n+d)}{n}}\frac{1}{\sigma_r^{\star4}}\max_{1\leq i\leq d}d_{l,i}.$$

Step 4: controlling the spectrum of $\widetilde{\boldsymbol{\Sigma}}_{l,2}$. Next, we will investigate the spectrum of $\widetilde{\boldsymbol{\Sigma}}_{l,2}$. It is straightforward to show that the matrix $\boldsymbol{\Sigma}_{l,2}^{\star}$ defined in (E.30) obeys

$$\left\|\boldsymbol{\Sigma}_{l,2}^{\star}\right\|\leq\frac{1}{\sigma_r^{\star4}}\max_{1\leq i\leq d}|d_{l,i}|\leq\frac{3\kappa\mu r(1-p)^2}{ndp^2\sigma_r^{\star2}}\left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2^2+\frac{\kappa\mu r(1-p)}{ndp^2\sigma_r^{\star2}}\omega_l^{\star2}+\frac{1-p}{np^2\sigma_r^{\star4}}\omega_{\max}^2\left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2^2+\frac{\omega_l^{\star2}\omega_{\max}^{\star2}}{np^2\sigma_r^{\star4}}. \tag{E.31}$$

In addition, we claim that the following inequality is valid.

***Claim* 1.** It holds that

$$\boldsymbol{U}^{\star\top}\mathsf{diag}\left\{\left\|\boldsymbol{U}_{i,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2^2\right\}_{i=1}^{d}\boldsymbol{U}^{\star}\succeq\frac{\sigma_r^{\star2}}{4d}\boldsymbol{I}_r. \tag{E.32}$$

With this claim in place, we can combine it with the definition (E.26) of $d_{l,i}$ to bound

$$
\begin{aligned}
\boldsymbol{\Sigma}_{l,2}^{\star} &= \left(\boldsymbol{\Sigma}^{\star}\right)^{-2} \boldsymbol{U}^{\star\top}\mathsf{diag}\left\{d_{l,i}\right\}_{i=1}^{d} \boldsymbol{U}^{\star}\left(\boldsymbol{\Sigma}^{\star}\right)^{-2} \\
&\succeq \frac{(1-p)^2}{np^2}\left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2^2 \left(\boldsymbol{\Sigma}^{\star}\right)^{-2} \boldsymbol{U}^{\star\top}\mathsf{diag}\left\{\left\|\boldsymbol{U}_{i,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2^2\right\}_{i=1}^{d} \boldsymbol{U}^{\star}\left(\boldsymbol{\Sigma}^{\star}\right)^{-2} \\
&\quad + \frac{1-p}{np^2}\omega_l^{\star 2}\left(\boldsymbol{\Sigma}^{\star}\right)^{-2} \boldsymbol{U}^{\star\top}\mathsf{diag}\left\{\left\|\boldsymbol{U}_{i,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2^2\right\}_{i=1}^{d} \boldsymbol{U}^{\star}\left(\boldsymbol{\Sigma}^{\star}\right)^{-2} + \frac{1-p}{np^2}\omega_{\min}^2\left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2^2\left(\boldsymbol{\Sigma}^{\star}\right)^{-4} \\
&\quad + \frac{1}{np^2}\omega_l^{\star 2}\left(\boldsymbol{\Sigma}^{\star}\right)^{-2} \boldsymbol{U}^{\star\top}\mathsf{diag}\left\{\omega_i^{\star 2}\right\}_{i=1}^{d} \boldsymbol{U}^{\star}\left(\boldsymbol{\Sigma}^{\star}\right)^{-2} \\
&\succeq \left[\frac{(1-p)^2}{4ndp^2}\sigma_r^{\star 2}\left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2^2 + \frac{1-p}{4ndp^2}\sigma_r^{\star 2}\omega_l^{\star 2} + \frac{1-p}{np^2}\omega_{\min}^2\left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2^2 + \frac{\omega_l^{\star 2}\omega_{\min}^2}{np^2}\right]\left(\boldsymbol{\Sigma}^{\star}\right)^{-4},
\end{aligned}
$$

which in turn leads to

$$
\begin{aligned}
\lambda_{\min}\left(\boldsymbol{\Sigma}_{l,2}^{\star}\right) &\geq \frac{1}{\sigma_1^{\star 4}}\left[\frac{(1-p)^2}{4ndp^2}\sigma_r^{\star 2}\left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2^2 + \frac{1-p}{4ndp^2}\sigma_r^{\star 2}\omega_l^{\star 2} + \frac{1-p}{np^2}\omega_{\min}^2\left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2^2 + \frac{\omega_l^{\star 2}\omega_{\min}^2}{np^2}\right] \\
&= \frac{(1-p)^2}{4ndp^2\kappa\sigma_1^{\star 2}}\left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2^2 + \frac{1-p}{4ndp^2\kappa\sigma_1^{\star 2}}\omega_l^{\star 2} + \frac{1-p}{np^2\sigma_1^{\star 4}}\omega_{\min}^2\left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2^2 + \frac{\omega_l^{\star 2}\omega_{\min}^2}{np^2\sigma_1^{\star 4}}. \qquad \text{(E.33)}
\end{aligned}
$$

The above bounds (E.31) and (E.33) imply that the condition number of $\boldsymbol{\Sigma}_{l,2}^{\star}$ — denoted by $\kappa(\boldsymbol{\Sigma}_{l,2}^{\star})$ — is upper bounded by

$$
\kappa(\boldsymbol{\Sigma}_{l,2}^{\star}) = \frac{\left\|\boldsymbol{\Sigma}_{l,2}^{\star}\right\|}{\lambda_{\min}\left(\boldsymbol{\Sigma}_{l,2}^{\star}\right)} \lesssim \kappa^3\mu r\kappa_{\omega}.
$$

Moreover, we can also obtain

$$
\begin{aligned}
\left\|\boldsymbol{R}_3\right\| &\lesssim \sqrt{\frac{\kappa^2 r\log^3(n+d)}{n}}\frac{1}{\sigma_r^{\star 4}}\max_{1\leq i\leq d}\left|d_{l,i}\right| \\
&\lesssim \sqrt{\frac{\kappa^2 r\log^3(n+d)}{n}}\kappa^3\mu r\kappa_{\omega}\lambda_{\min}\left(\boldsymbol{\Sigma}_{l,2}^{\star}\right) \\
&\asymp \sqrt{\frac{\kappa^8\mu^2 r^3\kappa_{\omega}^2\log^3(n+d)}{n}}\lambda_{\min}\left(\boldsymbol{\Sigma}_{l,2}^{\star}\right).
\end{aligned}
$$

Step 5: putting everything together. Combining the above results, we are allowed to express

$$
\widetilde{\boldsymbol{\Sigma}}_l = \boldsymbol{\Sigma}_{U,l}^{\star} + \boldsymbol{R}_l
$$

where $\boldsymbol{\Sigma}_{U,l}^{\star} = \boldsymbol{\Sigma}_{l,1}^{\star} + \boldsymbol{\Sigma}_{l,2}^{\star}$ and the residual matrix $\boldsymbol{R}_l$ satisfies

$$
\begin{aligned}
\left\|\boldsymbol{R}_l\right\| &\leq \left\|\boldsymbol{R}_1\right\| + \left\|\boldsymbol{R}_3\right\| \\
&\lesssim \sqrt{\frac{\kappa^5 r\log^3(n+d)}{n}}\lambda_{\min}\left(\boldsymbol{\Sigma}_{l,1}^{\star}\right) + \sqrt{\frac{\kappa^8\mu^2 r^3\kappa_{\omega}^2\log^3(n+d)}{n}}\lambda_{\min}\left(\boldsymbol{\Sigma}_{l,2}^{\star}\right) \\
&\lesssim \sqrt{\frac{\kappa^8\mu^2 r^3\kappa_{\omega}^2\log^3(n+d)}{n}}\lambda_{\min}\left(\boldsymbol{\Sigma}_{U,l}^{\star}\right).
\end{aligned}
$$

An immediate consequence is that the conditional number of $\boldsymbol{\Sigma}_{U,l}^{\star}$ is upper bounded by the maximum of the conditional numbers of $\boldsymbol{\Sigma}_{l,1}^{\star}$ and $\boldsymbol{\Sigma}_{l,2}^{\star}$, which is at most $O(\kappa^3\mu r\kappa_{\omega})$. Consequently, when $n \gg \kappa^8\mu^2 r^3\kappa_{\omega}^2\log^3(n+d)$, the conditional number of $\widetilde{\boldsymbol{\Sigma}}_l$ is also upper bounded by $O(\kappa^3\mu r\kappa_{\omega})$. This finishes the proof, as long as Claim 1 can be justified.

Proof of Claim 1. For any $c \in (0,1)$, we define an index set

$$\mathcal{I}_c := \{i \in [d] : \|\boldsymbol{U}^\star_{i,\cdot}\|_2 \geq \sqrt{c/d}\}.$$

Then for any $\boldsymbol{v} \in \mathbb{R}^d$, one has

$$\begin{aligned}
\boldsymbol{v}^\top \boldsymbol{U}^{\star\top} \mathsf{diag}\left\{\left\|\boldsymbol{U}^\star_{i,\cdot}\boldsymbol{\Sigma}^\star\right\|_2^2\right\}_{i=1}^d \boldsymbol{U}^\star \boldsymbol{v} &= \sum_{i=1}^d \left(\boldsymbol{v}^\top \boldsymbol{U}^\star_{i,\cdot}\right)^2 \left\|\boldsymbol{U}^\star_{i,\cdot}\boldsymbol{\Sigma}^\star\right\|_2^2 \geq \sigma_r^{\star 2} \sum_{i=1}^d \left(\boldsymbol{v}^\top \boldsymbol{U}^\star_{i,\cdot}\right)^2 \left\|\boldsymbol{U}^\star_{i,\cdot}\right\|_2^2 \\
&\geq \sigma_r^{\star 2} \sum_{i\in\mathcal{I}_c} \left(\boldsymbol{v}^\top \boldsymbol{U}^\star_{i,\cdot}\right)^2 \left\|\boldsymbol{U}^\star_{i,\cdot}\right\|_2^2 \geq \frac{c\sigma_r^{\star 2}}{d} \sum_{i\in\mathcal{I}_c} \left(\boldsymbol{v}^\top \boldsymbol{U}^\star_{i,\cdot}\right)^2 \\
&= \frac{c\sigma_r^{\star 2}}{d}\left[\|\boldsymbol{v}\|_2^2 - \sum_{i\in[d]\setminus\mathcal{I}_c} \left(\boldsymbol{v}^\top \boldsymbol{U}^\star_{i,\cdot}\right)^2\right] \\
&\geq \frac{c\sigma_r^{\star 2}}{d}\left[\|\boldsymbol{v}\|_2^2 - \sum_{i\in[d]\setminus\mathcal{I}_c} \|\boldsymbol{v}\|_2^2 \left\|\boldsymbol{U}^\star_{i,\cdot}\right\|_2^2\right] \\
&\geq \frac{c\sigma_r^{\star 2}}{d}\left[\|\boldsymbol{v}\|_2^2 - \sum_{i\in[d]\setminus\mathcal{I}_c} \|\boldsymbol{v}\|_2^2 \frac{c}{d}\right] \\
&\geq \frac{c\sigma_r^{\star 2}}{d}\left[\|\boldsymbol{v}\|_2^2 - c\|\boldsymbol{v}\|_2^2\right] \geq \frac{c(1-c)\,\sigma_r^{\star 2}}{d}\|\boldsymbol{v}\|_2^2.
\end{aligned}$$

Since the above inequality holds for arbitrary $c \in (0,1)$ and any $\boldsymbol{v} \in \mathbb{R}^d$, taking $c = 1/2$ leads to

$$\inf_{\boldsymbol{v}\in\mathbb{R}^d} \boldsymbol{v}^\top \boldsymbol{U}^{\star\top} \mathsf{diag}\left\{\left\|\boldsymbol{U}^\star_{i,\cdot}\boldsymbol{\Sigma}^\star\right\|_2^2\right\}_{i=1}^d \boldsymbol{U}^\star \boldsymbol{v} \geq \frac{\sigma_r^{\star 2}}{4d}\|\boldsymbol{v}\|_2^2.$$

Therefore, we can conclude that

$$\boldsymbol{U}^{\star\top} \mathsf{diag}\left\{\left\|\boldsymbol{U}^\star_{i,\cdot}\boldsymbol{\Sigma}^\star\right\|_2^2\right\}_{i=1}^d \boldsymbol{U}^\star \succeq \frac{\sigma_r^{\star 2}}{4d}\boldsymbol{I}_r.$$

□

E.2.3 Proof of Lemma 17

Step 1: Gaussian approximation of $\boldsymbol{Z}_{l,\cdot}$ using the Berry-Esseen Theorem. It is first seen that

$$\boldsymbol{Z}_{l,\cdot} = \sum_{j=1}^n \boldsymbol{Y}_j, \qquad \text{where} \quad \boldsymbol{Y}_j = E_{l,j}\left[\boldsymbol{V}^\natural_{j,\cdot}(\boldsymbol{\Sigma}^\natural)^{-1} + \left[\mathcal{P}_{-l,\cdot}\left(\boldsymbol{E}_{\cdot,j}\right)\right]^\top \boldsymbol{U}^\natural(\boldsymbol{\Sigma}^\natural)^{-2}\right]\boldsymbol{Q}^\top.$$

Invoking the Berry-Esseen theorem (see Theorem 16) then yields

$$\sup_{\mathcal{C}\in\mathscr{C}^r}\left|\mathbb{P}\left(\boldsymbol{Z}_{l,\cdot}\in\mathcal{C}\,\middle|\,\boldsymbol{F}\right) - \mathbb{P}\left(\mathcal{N}\big(\boldsymbol{0},\widetilde{\boldsymbol{\Sigma}}_l\big)\in\mathcal{C}\,\middle|\,\boldsymbol{F}\right)\right| \lesssim r^{1/4}\gamma\left(\boldsymbol{F}\right), \tag{E.34}$$

where $\widetilde{\boldsymbol{\Sigma}}_l$ is the covariance matrix of $\boldsymbol{Z}_{l,\cdot}$ that has been calculated in (D.25), and $\gamma\left(\boldsymbol{F}\right)$ is defined to be

$$\gamma\left(\boldsymbol{F}\right) := \sum_{i=1}^n \mathbb{E}\left[\left\|\boldsymbol{Y}_j\widetilde{\boldsymbol{\Sigma}}_l^{-1/2}\right\|_2^3\,\middle|\,\boldsymbol{F}\right].$$

In the sequel, let us develop an upper bound on $\gamma\left(\boldsymbol{F}\right)$. Towards this, we find it helpful to define the quantity B_l such that

$$\max_{j\in[n]}|E_{l,j}| \leq B_l.$$

- Towards this, we first provide a (conditional) high probability bound for each $\|\boldsymbol{Y}_j\|$. In view of Cai et al. (2021, Lemma 12), with probability exceeding $1 - O((n+d)^{-101})$ we have

$$\left\|\left[\mathcal{P}_{-l,\cdot}\left(\boldsymbol{E}_{\cdot,j}\right)\right]^{\top}\boldsymbol{U}^{\natural}\right\|_2 = \left\|\sum_{i:i\neq l} E_{i,j}\boldsymbol{U}^{\natural}_{i,\cdot}\right\|_2 \lesssim \left(B\log\left(n+d\right) + \sigma\sqrt{d\log\left(n+d\right)}\right)\left\|\boldsymbol{U}^{\natural}\right\|_{2,\infty}$$
$$\asymp \left(B\log\left(n+d\right) + \sigma\sqrt{d\log\left(n+d\right)}\right)\left\|\boldsymbol{U}^{\star}\right\|_{2,\infty}$$

holds for all $j \in [n]$, where the last relation makes use of the fact that $\boldsymbol{U}^{\natural} = \boldsymbol{U}^{\star}\boldsymbol{Q}$ for some orthonormal matrix $\boldsymbol{Q}$ (see (D.5)). Consequently, with probability exceeding $1 - O((n+d)^{-101})$,

$$\|\boldsymbol{Y}_j\|_2 \leq \frac{1}{\sigma_r^{\natural}} B_l \left\|\boldsymbol{V}^{\natural}\right\|_{2,\infty} + \frac{1}{\sigma_r^{\natural 2}} B_l \left(B\log\left(n+d\right) + \sigma\sqrt{d\log\left(n+d\right)}\right)\left\|\boldsymbol{U}^{\star}\right\|_{2,\infty}$$

holds for each $j \in [n]$. If we define

$$C_{\mathsf{prob}} \coloneqq \widetilde{C}_1 B_l \left[\frac{1}{\sigma_r^{\natural}}\left\|\boldsymbol{V}^{\natural}\right\|_{2,\infty} + \frac{1}{\sigma_r^{\natural 2}}\left(B\log\left(n+d\right) + \sigma\sqrt{d\log\left(n+d\right)}\right)\left\|\boldsymbol{U}^{\star}\right\|_{2,\infty}\right]$$

for some sufficiently large constant $\widetilde{C}_1 > 0$, then the union bound over $1 \leq j \leq n$ guarantees that

$$\mathbb{P}\left(\max_{1\leq j\leq n}\|\boldsymbol{Y}_j\|_2 \leq C_{\mathsf{prob}}\middle|\boldsymbol{F}\right) \geq 1 - O\left(\left(n+d\right)^{-100}\right). \tag{E.35}$$

- In addition, we also know that $\|\boldsymbol{Y}_j\|_2$ admits a trivial deterministic upper bound as follows

$$\|\boldsymbol{Y}_j\|_2 \lesssim \frac{1}{\sigma_r^{\natural}} B_l \left\|\boldsymbol{V}^{\natural}\right\|_{2,\infty} + \frac{1}{\sigma_r^{\natural 2}} B_l \|\boldsymbol{E}_{\cdot,j}\|_2 \lesssim \frac{1}{\sigma_r^{\natural}} B_l \left\|\boldsymbol{V}^{\natural}\right\|_{2,\infty} + \frac{1}{\sigma_r^{\natural 2}} B_l B\sqrt{d}.$$

This means that by defining the quantity

$$C_{\mathsf{det}} \coloneqq \widetilde{C}_2 B_l \left(\frac{1}{\sigma_r^{\natural}}\left\|\boldsymbol{V}^{\natural}\right\|_{2,\infty} + \frac{1}{\sigma_r^{\natural 2}} B\sqrt{d}\right)$$

for some sufficiently large constant $\widetilde{C}_2 > 0$, we can ensure that

$$\max_{1\leq j\leq n}\|\boldsymbol{Y}_j\|_2 \leq C_{\mathsf{det}}. \tag{E.36}$$

The above arguments taken together allow one to bound $\gamma\left(\boldsymbol{F}\right)$ as follows

$$\begin{aligned}
\gamma\left(\boldsymbol{F}\right) &\leq \lambda_{\min}^{-3/2}\big(\widetilde{\boldsymbol{\Sigma}}_l\big)\sum_{j=1}^{n}\mathbb{E}\left[\|\boldsymbol{Y}_j\|_2^3\,\middle|\boldsymbol{F}\right] \\
&= \lambda_{\min}^{-3/2}\big(\widetilde{\boldsymbol{\Sigma}}_l\big)\sum_{j=1}^{n}\mathbb{E}\left[\|\boldsymbol{Y}_j\|_2^3 \mathbb{1}_{\|\boldsymbol{Y}_j\|_2\leq C_{\mathsf{prob}}}\,\middle|\boldsymbol{F}\right] + \lambda_{\min}^{-3/2}\big(\widetilde{\boldsymbol{\Sigma}}_l\big)\sum_{j=1}^{n}\mathbb{E}\left[\|\boldsymbol{Y}_j\|_2^3 \mathbb{1}_{\|\boldsymbol{Y}_j\|_2> C_{\mathsf{prob}}}\,\middle|\boldsymbol{F}\right] \\
&\overset{\text{(i)}}{\lesssim} \lambda_{\min}^{-3/2}\big(\widetilde{\boldsymbol{\Sigma}}_l\big) C_{\mathsf{prob}}\sum_{j=1}^{n}\mathbb{E}\left[\|\boldsymbol{Y}_j\|_2^2\,\middle|\boldsymbol{F}\right] + \lambda_{\min}^{-3/2}\big(\widetilde{\boldsymbol{\Sigma}}_l\big)\sum_{j=1}^{n} C_{\mathsf{det}}^3\mathbb{P}\left(\max_{1\leq j\leq n}\|\boldsymbol{Y}_j\|_2 \leq C_{\mathsf{prob}}\middle|\boldsymbol{F}\right) \\
&\overset{\text{(ii)}}{\lesssim} \lambda_{\min}^{-3/2}\big(\widetilde{\boldsymbol{\Sigma}}_l\big)\mathsf{tr}\big(\widetilde{\boldsymbol{\Sigma}}_l\big) C_{\mathsf{prob}} + \lambda_{\min}^{-3/2}\big(\widetilde{\boldsymbol{\Sigma}}_l\big) C_{\mathsf{det}}^3\left(n+d\right)^{-99}.
\end{aligned}$$

Here (i) relies on (E.36); (ii) follows from (E.35) as well as

$$\sum_{j=1}^{n}\mathbb{E}\left[\|\boldsymbol{Y}_j\|_2^2\,\middle|\boldsymbol{F}\right] = \sum_{j=1}^{n}\mathbb{E}\left[\boldsymbol{Y}_j\boldsymbol{Y}_j^{\top}\middle|\boldsymbol{F}\right] = \sum_{j=1}^{n}\mathbb{E}\left[\mathsf{tr}\left(\boldsymbol{Y}_j^{\top}\boldsymbol{Y}_j\right)\middle|\boldsymbol{F}\right] = \mathsf{tr}\left[\sum_{j=1}^{n}\mathbb{E}\left(\boldsymbol{Y}_j^{\top}\boldsymbol{Y}_j\middle|\boldsymbol{F}\right)\right]$$

$$= \mathsf{tr}\left(\mathbb{E}\left[\boldsymbol{Z}_{l,\cdot}^{\top}\boldsymbol{Z}_{l,\cdot}\big|\boldsymbol{F}\right]\right) = \mathsf{tr}\big(\widetilde{\boldsymbol{\Sigma}}_l\big). \tag{E.37}$$

Substitution into (E.34) then yields

$$\begin{aligned}&\sup_{\mathcal{C}\in\mathscr{C}^r}\left|\mathbb{P}\left(\boldsymbol{Z}_{l,\cdot}\in\mathcal{C}\big|\boldsymbol{F}\right)-\mathbb{P}\left(\mathcal{N}\big(\boldsymbol{0},\widetilde{\boldsymbol{\Sigma}}_l\big)\in\mathcal{C}\big|\boldsymbol{F}\right)\right|\\&\qquad\lesssim\underbrace{r^{1/4}\lambda_{\min}^{-3/2}\big(\widetilde{\boldsymbol{\Sigma}}_l\big)\mathsf{tr}\big(\widetilde{\boldsymbol{\Sigma}}_l\big)C_{\mathsf{prob}}}_{=:\alpha}+\underbrace{r^{1/4}\lambda_{\min}^{-3/2}\big(\widetilde{\boldsymbol{\Sigma}}_l\big)C_{\mathsf{det}}^3\left(n+d\right)^{-99}}_{=:\beta}.\end{aligned}\tag{E.38}$$

When the high-probability event $\mathcal{E}_{\mathsf{good}}$ occurs, we know from Lemma 16 that the condition number of $\widetilde{\boldsymbol{\Sigma}}_l$ is bounded by $O(\kappa^3\mu r\kappa_\omega)$. This implies that

$$\frac{\mathsf{tr}\big(\widetilde{\boldsymbol{\Sigma}}_l\big)}{\lambda_{\min}^{3/2}\big(\widetilde{\boldsymbol{\Sigma}}_l\big)}\leq\frac{r\big\|\widetilde{\boldsymbol{\Sigma}}_l\big\|}{\lambda_{\min}^{3/2}\big(\widetilde{\boldsymbol{\Sigma}}_l\big)}\lesssim\frac{\kappa^3\mu r^2\kappa_\omega}{\lambda_{\min}^{1/2}\big(\widetilde{\boldsymbol{\Sigma}}_l\big)}.$$

and as a consequence,

$$\begin{aligned}\alpha&\lesssim\frac{\kappa^3\mu r^{9/4}\kappa_\omega}{\lambda_{\min}^{1/2}\big(\widetilde{\boldsymbol{\Sigma}}_l\big)}C_{\mathsf{prob}}\\&\asymp\underbrace{\frac{\kappa^3\mu r^{9/4}\kappa_\omega}{\lambda_{\min}^{1/2}\big(\widetilde{\boldsymbol{\Sigma}}_l\big)}\frac{B_l}{\sigma_r^{\natural}}\left\|\boldsymbol{V}^{\natural}\right\|_{2,\infty}}_{=:\alpha_1}+\underbrace{\frac{\kappa^3\mu r^{9/4}\kappa_\omega\log\left(n+d\right)}{\lambda_{\min}^{1/2}\big(\widetilde{\boldsymbol{\Sigma}}_l\big)}\frac{B_lB}{\sigma_r^{\natural 2}}\left\|\boldsymbol{U}^{\star}\right\|_{2,\infty}}_{=:\alpha_2}\\&\quad+\underbrace{\frac{\kappa^3\mu r^{9/4}\kappa_\omega\log^{3/2}\left(n+d\right)}{\lambda_{\min}^{1/2}\big(\widetilde{\boldsymbol{\Sigma}}_l\big)}\frac{\sigma B_l}{\sigma_r^{\natural 2}}\sqrt{d}\left\|\boldsymbol{U}^{\star}\right\|_{2,\infty}}_{=:\alpha_3}.\end{aligned}$$

All of these terms involve $\lambda_{\min}^{1/2}\big(\widetilde{\boldsymbol{\Sigma}}_l\big)$, which can be lower bounded using Lemma 16 as follows

$$\begin{aligned}\lambda_{\min}^{1/2}\big(\widetilde{\boldsymbol{\Sigma}}_l\big)&\gtrsim\frac{1}{\sqrt{np}\sigma_1^{\star}}\left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2+\frac{\omega_l^{\star}}{\sqrt{np}\sigma_1^{\star}}+\frac{1}{\sqrt{ndp^2\kappa}\sigma_1^{\star}}\left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2+\frac{1}{\sqrt{ndp^2\kappa}\sigma_1^{\star}}\omega_l^{\star}\\&\quad+\frac{1}{\sqrt{np^2\sigma_1^{\star 2}}}\omega_{\min}\left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2+\frac{\omega_l^{\star}\omega_{\min}}{\sqrt{np^2\sigma_1^{\star 2}}}.\end{aligned}\tag{E.39}$$

Moreover, it is seen from (D.18) that

$$B_l\lesssim\frac{1}{p}\sqrt{\frac{\log\left(n+d\right)}{n}}\left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2+\frac{\omega_l^{\star}}{p}\sqrt{\frac{\log\left(n+d\right)}{n}}\asymp\frac{1}{p}\sqrt{\frac{\log\left(n+d\right)}{n}}\left(\left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2+\omega_l^{\star}\right).\tag{E.40}$$

In the sequel, we shall bound α_1, α_2, and α_3 separately.

- We start with bounding α_1, where we have

$$\begin{aligned}\alpha_1&\overset{\text{(i)}}{\lesssim}\frac{\kappa^3\mu r^{9/4}\kappa_\omega}{\lambda_{\min}^{1/2}\big(\widetilde{\boldsymbol{\Sigma}}_l\big)}\frac{1}{\sigma_r^{\star}}\frac{1}{p}\sqrt{\frac{\log\left(n+d\right)}{n}}\left(\left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2+\omega_l^{\star}\right)\sqrt{\frac{r\log\left(n+d\right)}{n}}\\&\lesssim\frac{1}{\lambda_{\min}^{1/2}\big(\widetilde{\boldsymbol{\Sigma}}_l\big)}\frac{\kappa^3\mu r^{11/4}\kappa_\omega}{np\sigma_r^{\star}}\log\left(n+d\right)\left(\left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2+\omega_l^{\star}\right)\\&\overset{\text{(ii)}}{\lesssim}\frac{\kappa^3\mu r^{11/4}\kappa_\omega}{np\sigma_r^{\star}}\log\left(n+d\right)\left(\left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2+\omega_l^{\star}\right)\left[\frac{1}{\sqrt{np}\sigma_1^{\star}}\left(\left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2+\omega_l^{\star}\right)\right]^{-1}\\&\asymp\frac{\kappa^{7/2}\mu r^{11/4}\kappa_\omega}{\sqrt{np}}\log\left(n+d\right)\overset{\text{(iii)}}{\lesssim}\frac{1}{\sqrt{\log\left(n+d\right)}}.\end{aligned}$$

  Here (i) follows from (D.13), (D.16), (D.18) and (E.40); (ii) makes use of (E.39); and (iii) is valid with the proviso that $np\gtrsim\kappa^7\mu^2r^{11/2}\kappa_\omega^2\log^3(n+d)$.

- Regarding α_2, we know that

$$
\begin{aligned}
\alpha_2 &\overset{\text{(i)}}{\lesssim} \frac{\kappa^3 \mu r^{9/4} \kappa_\omega \log^2(n+d)}{\lambda_{\min}^{1/2}(\widetilde{\boldsymbol{\Sigma}}_l)} \frac{1}{\sigma_r^{\star 2}} \left(\frac{1}{p^2} \sqrt{\frac{\mu r}{n^2 d}} \sigma_1^\star + \frac{\omega_{\max}}{np^2} \right) \left(\left\| \boldsymbol{U}_{l,\cdot}^\star \boldsymbol{\Sigma}^\star \right\|_2 + \omega_l^\star \right) \| \boldsymbol{U}^\star \|_{2,\infty} \\
&\lesssim \frac{1}{\lambda_{\min}^{1/2}(\widetilde{\boldsymbol{\Sigma}}_l)} \frac{1}{\sigma_r^{\star 2}} \left(\frac{\kappa^3 \mu^2 r^{13/4} \kappa_\omega}{ndp^2} \sigma_1^\star + \frac{\kappa^3 \mu^{3/2} r^{11/4} \kappa_\omega}{n\sqrt{d} p^2} \omega_{\max} \right) \log^2(n+d) \left(\left\| \boldsymbol{U}_{l,\cdot}^\star \boldsymbol{\Sigma}^\star \right\|_2 + \omega_l^\star \right) \\
&\overset{\text{(ii)}}{\lesssim} \frac{1}{\sigma_r^{\star 2}} \left(\frac{\kappa^3 \mu^2 r^{13/4} \kappa_\omega}{ndp^2} \sigma_1^\star + \frac{\kappa^3 \mu^{3/2} r^{11/4} \kappa_\omega}{n\sqrt{d} p^2} \omega_{\max} \right) \log^2(n+d) \\
&\quad \cdot \left(\left\| \boldsymbol{U}_{l,\cdot}^\star \boldsymbol{\Sigma}^\star \right\|_2 + \omega_l^\star \right) \left[\frac{1}{\sqrt{ndp^2 \kappa} \sigma_1^\star} \left(\left\| \boldsymbol{U}_{l,\cdot}^\star \boldsymbol{\Sigma}^\star \right\|_2 + \omega_l^\star \right) \right]^{-1} \\
&\lesssim \left(\frac{\kappa^{9/2} \mu^2 r^{13/4} \kappa_\omega}{\sqrt{ndp^2}} + \frac{\kappa^4 \mu^{3/2} r^{11/4} \kappa_\omega}{\sqrt{np^2}} \frac{\omega_{\max}}{\sigma_r^\star} \right) \log^2(n+d) \overset{\text{(iii)}}{\lesssim} \frac{1}{\sqrt{\log(n+d)}}.
\end{aligned}
$$

Here, (i) follows from (D.13), (D.18) and (E.40); (ii) makes use of (E.39); and (iii) holds provided that $ndp^2 \gtrsim \kappa^9 \mu^4 r^{13/2} \kappa_\omega^2 \log^5(n+d)$ and

$$
\frac{\omega_{\max}}{\sigma_r^\star \sqrt{np^2}} \lesssim \frac{\omega_{\max}^2}{p \sigma_r^{\star 2}} \sqrt{\frac{d}{n}} + \frac{1}{\sqrt{ndp}} \lesssim \frac{1}{\kappa^4 \mu^{3/2} r^{11/4} \kappa_\omega \log^{5/2}(n+d)},
$$

which can be guaranteed by $ndp^2 \gtrsim \kappa^8 \mu^3 r^{11/2} \kappa_\omega^2 \log^5(n+d)$ and

$$
\frac{\omega_{\max}^2}{p \sigma_r^{\star 2}} \sqrt{\frac{d}{n}} \lesssim \frac{1}{\kappa^4 \mu^{3/2} r^{11/4} \kappa_\omega \log^{5/2}(n+d)}.
$$

- When it comes to α_3, we have

$$
\begin{aligned}
\alpha_3 &\overset{\text{(i)}}{\lesssim} \frac{\kappa^3 \mu^{3/2} r^{11/4} \kappa_\omega \log^2(n+d)}{\lambda_{\min}^{1/2}(\widetilde{\boldsymbol{\Sigma}}_l)} \frac{1}{\sigma_r^{\star 2}} \left(\sqrt{\frac{\mu r \log(n+d)}{ndp}} \sigma_1^\star + \frac{\omega_{\max}}{\sqrt{np}} \right) \sqrt{\frac{1}{np^2}} \left(\left\| \boldsymbol{U}_{l,\cdot}^\star \boldsymbol{\Sigma}^\star \right\|_2 + \omega_l^\star \right) \\
&\lesssim \frac{\kappa^3 \mu^{3/2} r^{11/4} \kappa_\omega \log^2(n+d)}{\lambda_{\min}^{1/2}(\widetilde{\boldsymbol{\Sigma}}_l)} \frac{1}{\sigma_r^{\star 2}} \left(\sqrt{\frac{\mu r \log(n+d)}{n^2 dp^3}} \sigma_1^\star + \frac{\omega_{\max}}{\sqrt{n^2 p^3}} \right) \left(\left\| \boldsymbol{U}_{l,\cdot}^\star \boldsymbol{\Sigma}^\star \right\|_2 + \omega_l^\star \right) \\
&\overset{\text{(ii)}}{\lesssim} \frac{1}{\sigma_r^{\star 2}} \left(\frac{\kappa^3 \mu^2 r^{13/4} \kappa_\omega \log^{5/2}(n+d)}{\sqrt{n^2 dp^3}} \sigma_1^\star + \omega_{\max} \frac{\kappa^3 \mu^{3/2} r^{11/4} \kappa_\omega \log^2(n+d)}{\sqrt{n^2 p^3}} \right) \\
&\quad \cdot \left(\left\| \boldsymbol{U}_{l,\cdot}^\star \boldsymbol{\Sigma}^\star \right\|_2 + \omega_l^\star \right) \left[\frac{1}{\sqrt{np} \sigma_1^\star} \left(\left\| \boldsymbol{U}_{l,\cdot}^\star \boldsymbol{\Sigma}^\star \right\|_2 + \omega_l^\star \right) \right]^{-1} \\
&\lesssim \frac{\kappa^4 \mu^2 r^{13/4} \log^{5/2}(n+d)}{\sqrt{ndp^2}} + \frac{\omega_{\max}}{\sigma_r^\star} \frac{\kappa^{7/2} \mu^{3/2} r^{11/4} \kappa_\omega \log^2(n+d)}{\sqrt{np^2}} \overset{\text{(iii)}}{\lesssim} \frac{1}{\sqrt{\log(n+d)}}.
\end{aligned}
$$

Here, (i) follows from (D.17), (D.18) and (E.40); (ii) makes use of (E.39); and (iii) holds provided that $ndp^2 \gtrsim \kappa^8 \mu^4 r^{13/2} \kappa_\omega^2 \log^6(n+d)$ and

$$
\frac{\omega_{\max}}{\sigma_r^\star \sqrt{np^2}} \lesssim \frac{\omega_{\max}^2}{p \sigma_r^{\star 2}} \sqrt{\frac{d}{n}} + \frac{1}{\sqrt{ndp}} \lesssim \frac{1}{\kappa^{7/2} \mu^{3/2} r^{11/4} \kappa_\omega \log^{5/2}(n+d)},
$$

which can be guaranteed by $ndp^2 \gtrsim \kappa^7 \mu^3 r^{11/2} \kappa_\omega^2 \log^5(n+d)$ and

$$
\frac{\omega_{\max}^2}{p \sigma_r^{\star 2}} \sqrt{\frac{d}{n}} \lesssim \frac{1}{\kappa^{7/2} \mu^{3/2} r^{11/4} \kappa_\omega \log^{5/2}(n+d)}.
$$

In addition, we have learned from (E.39) and (E.40) that

$$B_l \lesssim \frac{1}{p}\sqrt{\frac{\log(n+d)}{n}}\sqrt{np}\sigma_1^\star \lambda_{\min}^{1/2}(\widetilde{\boldsymbol{\Sigma}}_l) \asymp \sqrt{\frac{\log(n+d)}{p}}\sigma_1^\star \lambda_{\min}^{1/2}(\widetilde{\boldsymbol{\Sigma}}_l). \tag{E.41}$$

Therefore the term β defined in (E.38) can be bounded by

$$\begin{aligned}
\beta &\overset{\text{(i)}}{\lesssim} r^{1/4}\lambda_{\min}^{-3/2}(\widetilde{\boldsymbol{\Sigma}}_l) B_l^3 \left(\frac{1}{\sigma_r^\star}\sqrt{\frac{r\log(n+d)}{n}} + \frac{1}{\sigma_r^{\star 2}}B\sqrt{d}\right)^3 (n+d)^{-99} \\
&\overset{\text{(ii)}}{\lesssim} \kappa^{3/2} r^{1/4}\frac{\log^{3/2}(n+d)}{p^{3/2}}\left(1+\frac{1}{\sigma_r^\star}B\sqrt{d}\right)^3 (n+d)^{-99} \\
&\overset{\text{(iii)}}{\lesssim} \kappa^{3/2} r^{1/4}\frac{\log^{3/2}(n+d)}{p^{3/2}} d^{3/2}(n+d)^{-99} \overset{\text{(iv)}}{\lesssim} (n+d)^{-50}.
\end{aligned}$$

Here (i) utilizes (D.13) and (D.16); (ii) follows from (E.41); (iii) holds as long as $B \lesssim \sigma_r^\star$, which can be guaranteed by $ndp^2 \gtrsim \kappa\mu r\log(n+d)$ and

$$\frac{\omega_{\max}}{p\sigma_r^\star}\sqrt{\frac{1}{n}} \lesssim \frac{\omega_{\max}^2}{p\sigma_r^{\star 2}}\sqrt{\frac{d}{n}} + \frac{1}{\sqrt{ndp}} \lesssim \frac{1}{\sqrt{\log(n+d)}}; \tag{E.42}$$

and (iv) is valid provided that $np \gtrsim \kappa r\log(n+d)$.

Therefore, the preceding bounds allow one to conclude that

$$\sup_{\mathcal{C}\in\mathscr{C}^r}\left|\mathbb{P}\left(\boldsymbol{Z}_{l,\cdot}\in\mathcal{C}\middle|\boldsymbol{F}\right) - \mathbb{P}\left(\mathcal{N}\left(\boldsymbol{0},\widetilde{\boldsymbol{\Sigma}}_l\right)\in\mathcal{C}\middle|\boldsymbol{F}\right)\right| \lesssim \alpha_1+\alpha_2+\alpha_3+\beta \lesssim \frac{1}{\sqrt{\log(n+d)}}, \tag{E.43}$$

provided that $np \gtrsim \kappa^7\mu^2 r^{11/2}\kappa_\omega^2\log^3(n+d)$, $ndp^2 \gtrsim \kappa^9\mu^4 r^{13/2}\kappa_\omega^2\log^6(n+d)$ and

$$\frac{\omega_{\max}^2}{p\sigma_r^{\star 2}}\sqrt{\frac{d}{n}} \lesssim \frac{1}{\kappa^4\mu^{3/2}r^{11/4}\kappa_\omega\log^{5/2}(n+d)}.$$

Step 2: bounding TV distance between Gaussian distributions. In view of Lemma 16, we know that

$$\left\|\widetilde{\boldsymbol{\Sigma}}_l - \boldsymbol{\Sigma}_{U,l}^\star\right\| \lesssim \sqrt{\frac{\kappa^8\mu^2 r^3\kappa_\omega^2\log^3(n+d)}{n}}\lambda_{\min}\left(\boldsymbol{\Sigma}_{U,l}^\star\right) \tag{E.44}$$

holds on the event $\mathcal{E}_{\mathsf{good}}$. In addition, $\boldsymbol{\Sigma}_{U,l}^\star$ is assumed to be non-singular. As a result, conditional on $\boldsymbol{F}$ and the event $\mathcal{E}_{\mathsf{good}}$, one can invoke Theorem 17 and (E.44) to arrive at

$$\begin{aligned}
&\sup_{\mathcal{C}\in\mathscr{C}^r}\left|\mathbb{P}\left(\mathcal{N}(\boldsymbol{0},\widetilde{\boldsymbol{\Sigma}}_l)\in\mathcal{C}\mid\boldsymbol{F}\right) - \mathbb{P}\left(\mathcal{N}\left(\boldsymbol{0},\boldsymbol{\Sigma}_{U,l}^\star\right)\in\mathcal{C}\right)\right| \le \mathsf{TV}\left(\mathcal{N}(\boldsymbol{0},\widetilde{\boldsymbol{\Sigma}}_l),\mathcal{N}\left(\boldsymbol{0},\boldsymbol{\Sigma}_{U,l}^\star\right)\right) \\
&\asymp \left\|\left(\boldsymbol{\Sigma}_{U,l}^\star\right)^{-1/2}\widetilde{\boldsymbol{\Sigma}}_l\left(\boldsymbol{\Sigma}_{U,l}^\star\right)^{-1/2} - \boldsymbol{I}_d\right\|_{\mathrm{F}} = \left\|\left(\boldsymbol{\Sigma}_{U,l}^\star\right)^{-1/2}\left(\widetilde{\boldsymbol{\Sigma}}_l - \boldsymbol{\Sigma}_{U,l}^\star\right)\left(\boldsymbol{\Sigma}_{U,l}^\star\right)^{-1/2}\right\|_{\mathrm{F}} \\
&\lesssim \left\|\left(\boldsymbol{\Sigma}_{U,l}^\star\right)^{-1/2}\right\|\left\|\widetilde{\boldsymbol{\Sigma}}_l - \boldsymbol{\Sigma}_{U,l}^\star\right\|_{\mathrm{F}}\left\|\left(\boldsymbol{\Sigma}_{U,l}^\star\right)^{-1/2}\right\| \lesssim \sqrt{r}\left\|\left(\boldsymbol{\Sigma}_{U,l}^\star\right)^{-1}\right\|\left\|\widetilde{\boldsymbol{\Sigma}}_l - \boldsymbol{\Sigma}_{U,l}^\star\right\| \\
&\lesssim \frac{1}{\lambda_{\min}\left(\boldsymbol{\Sigma}_{U,l}^\star\right)}\cdot\sqrt{\frac{\kappa^8\mu^2 r^4\kappa_\omega^2\log^3(n+d)}{n}}\lambda_{\min}\left(\boldsymbol{\Sigma}_{U,l}^\star\right) \\
&\lesssim \sqrt{\frac{\kappa^8\mu^2 r^4\kappa_\omega^2\log^3(n+d)}{n}}.
\end{aligned}$$

where $\mathsf{TV}(\cdot,\cdot)$ represents the total-variation distance between two distributions (Tsybakov and Zaiats, 2009). The above inequality taken together with (E.43) yields

$$\sup_{\mathcal{C}\in\mathscr{C}^r}\left|\mathbb{P}\left(\boldsymbol{Z}_{l,\cdot}\in\mathcal{C}\middle|\boldsymbol{F}\right) - \mathbb{P}\left(\mathcal{N}\left(\boldsymbol{0},\boldsymbol{\Sigma}_l^\star\right)\in\mathcal{C}\right)\right| \lesssim \frac{1}{\sqrt{\log(n+d)}} \tag{E.45}$$

on the event $\mathcal{E}_{\mathsf{good}}$, provided that $n \gtrsim \kappa^8\mu^2 r^4\kappa_\omega^2\log^4(n+d)$.

Step 3: accounting for higher-order errors. Let us assume that the high-probability event $\mathcal{E}_{\mathsf{good}}$ happens. Lemma 15 tells us that

$$\mathbb{P}\left(\left\|(\boldsymbol{U}\boldsymbol{R}-\boldsymbol{U}^{\star}-\boldsymbol{Z})_{l,\cdot}\right\|_{2}\lesssim\zeta_{\mathsf{2nd},l}\,\middle|\,\boldsymbol{F}\right)\geq1-O\left((n+d)^{-10}\right),$$

which immediately gives

$$\mathbb{P}\left(\left\|\left(\boldsymbol{\Sigma}_{U,l}^{\star}\right)^{-1/2}(\boldsymbol{U}\boldsymbol{R}-\boldsymbol{U}^{\star}-\boldsymbol{Z})\right\|_{2,\infty}\leq\zeta\,\middle|\,\boldsymbol{F}\right)\geq1-O\left((n+d)^{-10}\right).\tag{E.46}$$

Here, the quantity ζ is defined as

$$\zeta\coloneqq c_{\zeta}\zeta_{\mathsf{2nd},l}\left(\lambda_{\min}\left(\boldsymbol{\Sigma}_{U,l}^{\star}\right)\right)^{-\frac{1}{2}}\tag{E.47}$$

for some sufficiently large constant $c_{\zeta}>0$.

For any convex set $\mathcal{C}\in\mathscr{C}^{r}$ and any ε, recalling the definition of $\mathcal{C}^{\varepsilon}$ in (A.3) in Appendix A, we have

$$\begin{aligned}\mathbb{P}\left(\left(\boldsymbol{\Sigma}_{U,l}^{\star}\right)^{-1/2}\boldsymbol{Z}_{l,\cdot}\in\mathcal{C}^{-\zeta}\,\middle|\,\boldsymbol{F}\right)&=\mathbb{P}\left(\left(\boldsymbol{\Sigma}_{U,l}^{\star}\right)^{-1/2}\boldsymbol{Z}_{l,\cdot}\in\mathcal{C}^{-\zeta},\left\|\left(\boldsymbol{\Sigma}_{U,l}^{\star}\right)^{-1/2}(\boldsymbol{U}\boldsymbol{R}-\boldsymbol{U}^{\star}-\boldsymbol{Z})\right\|_{2,\infty}\leq\zeta\,\middle|\,\boldsymbol{F}\right)\\&\quad+\mathbb{P}\left(\left(\boldsymbol{\Sigma}_{U,l}^{\star}\right)^{-1/2}\boldsymbol{Z}_{l,\cdot}\in\mathcal{C}^{-\zeta},\left\|\left(\boldsymbol{\Sigma}_{U,l}^{\star}\right)^{-1/2}(\boldsymbol{U}\boldsymbol{R}-\boldsymbol{U}^{\star}-\boldsymbol{Z})\right\|_{2,\infty}>\zeta\,\middle|\,\boldsymbol{F}\right)\\&\leq\mathbb{P}\left(\left(\boldsymbol{\Sigma}_{U,l}^{\star}\right)^{-1/2}(\boldsymbol{U}\boldsymbol{R}-\boldsymbol{U}^{\star})_{l,\cdot}\in\mathcal{C}\,\middle|\,\boldsymbol{F}\right)+\mathbb{P}\left(\left\|\left(\boldsymbol{\Sigma}_{U,l}^{\star}\right)^{-1/2}(\boldsymbol{U}\boldsymbol{R}-\boldsymbol{U}^{\star}-\boldsymbol{Z})\right\|_{2,\infty}>\zeta\,\middle|\,\boldsymbol{F}\right)\\&\leq\mathbb{P}\left(\left(\boldsymbol{\Sigma}_{U,l}^{\star}\right)^{-1/2}(\boldsymbol{U}\boldsymbol{R}-\boldsymbol{U}^{\star})_{l,\cdot}\in\mathcal{C}\,\middle|\,\boldsymbol{F}\right)+O\left((n+d)^{-10}\right),\end{aligned}\tag{E.48}$$

where the first inequality follows from the definition of $\mathcal{C}^{-\zeta}$, and the last inequality makes use of (E.46). Similarly,

$$\mathbb{P}\left(\left(\boldsymbol{\Sigma}_{U,l}^{\star}\right)^{-1/2}(\boldsymbol{U}\boldsymbol{R}-\boldsymbol{U}^{\star})_{l,\cdot}\in\mathcal{C}\,\middle|\,\boldsymbol{F}\right)\leq\mathbb{P}\left(\boldsymbol{\Sigma}_{l}^{-1/2}\boldsymbol{Z}_{l,\cdot}\in\mathcal{C}^{\zeta}\,\middle|\,\boldsymbol{F}\right)+O\left((n+d)^{-10}\right).\tag{E.49}$$

In addition, for any set $\mathcal{X}\subseteq\mathbb{R}^{r}$ and any matrix $\boldsymbol{A}\in\mathbb{R}^{r\times r}$, let us denote by $\boldsymbol{A}\mathcal{X}$ be the set $\{\boldsymbol{A}\boldsymbol{x}:\boldsymbol{x}\in\mathcal{X}\}$. It is easily seen that when $\boldsymbol{A}$ is non-singular, $\boldsymbol{A}\mathcal{C}\subseteq\mathscr{C}^{r}$ holds if and only if $\mathcal{C}\in\mathscr{C}^{r}$. We can then deduce that

$$\begin{aligned}\mathbb{P}\left(\left(\boldsymbol{\Sigma}_{U,l}^{\star}\right)^{-1/2}\boldsymbol{Z}_{l,\cdot}\in\mathcal{C}^{\zeta}\,\middle|\,\boldsymbol{F}\right)&=\mathbb{P}\left(\boldsymbol{Z}_{l,\cdot}\in\left(\boldsymbol{\Sigma}_{U,l}^{\star}\right)^{1/2}\mathcal{C}^{\zeta}\,\middle|\,\boldsymbol{F}\right)\\&\overset{\text{(i)}}{\leq}\mathbb{P}\left(\mathcal{N}\left(\boldsymbol{0},\boldsymbol{\Sigma}_{U,l}^{\star}\right)\in\left(\boldsymbol{\Sigma}_{U,l}^{\star}\right)^{1/2}\mathcal{C}^{\zeta}\right)+O\left(\frac{1}{\sqrt{\log(n+d)}}\right)\\&=\mathbb{P}\left(\mathcal{N}\left(\boldsymbol{0},\boldsymbol{I}_{r}\right)\in\mathcal{C}^{\zeta}\right)+O\left(\frac{1}{\sqrt{\log(n+d)}}\right)\\&\overset{\text{(ii)}}{\leq}\mathbb{P}\left(\mathcal{N}\left(\boldsymbol{0},\boldsymbol{I}_{r}\right)\in\mathcal{C}\right)+\zeta\left(0.59r^{1/4}+0.21\right)+\left(\frac{1}{\sqrt{\log(n+d)}}\right)\\&\overset{\text{(iii)}}{\leq}\mathbb{P}\left(\mathcal{N}\left(\boldsymbol{0},\boldsymbol{I}_{r}\right)\in\mathcal{C}\right)+O\left(\frac{1}{\sqrt{\log(n+d)}}\right),\end{aligned}$$

where (i) uses (E.45), (ii) is a consequence of Theorem 18, and (iii) holds provided that $\zeta\lesssim1/(r^{1/4}\log^{1/2}(n+d))$. Similarly we can show that

$$\mathbb{P}\left(\left(\boldsymbol{\Sigma}_{U,l}^{\star}\right)^{-1/2}\boldsymbol{Z}_{l,\cdot}\in\mathcal{C}^{-\zeta}\,\middle|\,\boldsymbol{F}\right)\geq\mathbb{P}\left(\mathcal{N}\left(\boldsymbol{0},\boldsymbol{I}_{r}\right)\in\mathcal{C}\right)-O\left(\frac{1}{\sqrt{\log(n+d)}}\right).$$

Combine the above two inequalities with (E.48) and (E.49) to achieve

$$\left|\mathbb{P}\left(\left(\boldsymbol{\Sigma}_{U,l}^{\star}\right)^{-1/2}(\boldsymbol{U}\boldsymbol{R}-\boldsymbol{U}^{\star})_{l,\cdot}\in\mathcal{C}\,\middle|\,\boldsymbol{F}\right)-\mathbb{P}\left(\mathcal{N}\left(\boldsymbol{0},\boldsymbol{I}_{r}\right)\in\mathcal{C}\right)\right|\lesssim\frac{1}{\sqrt{\log(n+d)}}.$$

It is worth noting that this inequality holds for all $\mathcal{C} \in \mathscr{C}^r$. As a result, on the event $\mathcal{E}_{\mathsf{good}}$ we can obtain

$$\begin{aligned}
&\sup_{\mathcal{C}\in\mathscr{C}^r}\left|\mathbb{P}\left(\left(\boldsymbol{U}\boldsymbol{R}-\boldsymbol{U}^\star\right)_{l,\cdot}\in\mathcal{C}\,\middle|\,\boldsymbol{F}\right)-\mathbb{P}\left(\mathcal{N}\left(\mathbf{0},\boldsymbol{\Sigma}_l\right)\in\mathcal{C}\right)\right| \\
&\quad=\sup_{\mathcal{C}\in\mathscr{C}^r}\left|\mathbb{P}\left(\left(\boldsymbol{U}\boldsymbol{R}-\boldsymbol{U}^\star\right)_{l,\cdot}\in\boldsymbol{\Sigma}_l^{1/2}\mathcal{C}\right)-\mathbb{P}\left(\mathcal{N}\left(\mathbf{0},\boldsymbol{\Sigma}_l\right)\in\boldsymbol{\Sigma}_l^{1/2}\mathcal{C}\right)\right| \\
&\quad=\sup_{\mathcal{C}\in\mathscr{C}^r}\left|\mathbb{P}\left(\boldsymbol{\Sigma}_l^{-1/2}\left(\boldsymbol{U}\boldsymbol{R}-\boldsymbol{U}^\star\right)_{l,\cdot}\in\mathcal{C}\right)-\mathbb{P}\left(\mathcal{N}\left(\mathbf{0},\boldsymbol{I}_r\right)\in\mathcal{C}\right)\right| \\
&\quad\lesssim\frac{1}{\sqrt{\log\left(n+d\right)}},
\end{aligned}\tag{E.50}$$

where the first identity makes use of the fact that $\mathcal{C}\to(\boldsymbol{\Sigma}_{U,l}^\star)^{1/2}\mathcal{C}$ is a one-to one mapping from $\mathscr{C}^r$ to $\mathscr{C}^r$ (since $\boldsymbol{\Sigma}_{U,l}^\star$ has full rank).

It remains to verify the conditions required to guarantee $\zeta\lesssim 1/(r^{1/4}\log^{1/2}(n+d))$. More generally, we shall check that under what conditions we can guarantee

$$\zeta_{\mathsf{2nd},l}\lesssim\delta\lambda_{\min}^{1/2}\left(\boldsymbol{\Sigma}_{U,l}^\star\right)$$

for some $\delta>0$. Recall from Lemma 15 that

$$\zeta_{\mathsf{2nd},l}\lesssim\underbrace{\left\|\boldsymbol{U}_{l,\cdot}^\star\right\|_2\sqrt{\frac{\kappa^3\mu r\log\left(n+d\right)}{d}}\frac{\zeta_{\mathsf{1st}}}{\sigma_r^{\star2}}}_{=:\gamma_1}+\underbrace{\left\|\boldsymbol{U}_{l,\cdot}^\star\right\|_2\kappa\frac{\zeta_{\mathsf{1st}}^2}{\sigma_r^{\star4}}}_{=:\gamma_2}+\underbrace{\frac{\zeta_{\mathsf{1st}}\zeta_{\mathsf{1st},l}}{\sigma_r^{\star4}}\sqrt{\frac{\kappa^3\mu r\log\left(n+d\right)}{d}}}_{=:\gamma_3},$$

and from Lemma 16 that

$$\begin{aligned}
\lambda_{\min}^{1/2}\left(\boldsymbol{\Sigma}_{U,l}^\star\right)\gtrsim{}&\frac{1}{\sqrt{np}\sigma_1^\star}\left\|\boldsymbol{U}_{l,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2+\frac{\omega_l^\star}{\sqrt{np}\sigma_1^\star}+\frac{1}{\sqrt{ndp^2\kappa}\sigma_1^\star}\left\|\boldsymbol{U}_{l,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2+\frac{1}{\sqrt{ndp^2\kappa}\sigma_1^\star}\omega_l^\star \\
&+\frac{1}{\sqrt{np^2}\sigma_1^{\star2}}\omega_{\min}\left\|\boldsymbol{U}_{l,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2+\frac{\omega_l^\star\omega_{\min}}{\sqrt{np^2}\sigma_1^{\star2}}.
\end{aligned}$$

- Regarding γ_1, we can derive

$$\begin{aligned}
\gamma_1\asymp{}&\underbrace{\sqrt{\frac{\kappa^5\mu^3r^3\log^5\left(n+d\right)}{nd^2p^2}}\left\|\boldsymbol{U}_{l,\cdot}^\star\right\|_2}_{=:\gamma_{1,1}}+\underbrace{\frac{\omega_{\max}^2}{\sigma_r^{\star2}p}\sqrt{\frac{\kappa^3\mu r\log^3\left(n+d\right)}{n}}\left\|\boldsymbol{U}_{l,\cdot}^\star\right\|_2}_{=:\gamma_{1,2}} \\
&+\underbrace{\left\|\boldsymbol{U}_{l,\cdot}^\star\right\|_2\sqrt{\frac{\kappa^5\mu^2r^2\log^3\left(n+d\right)}{ndp}}}_{=:\gamma_{1,3}}+\underbrace{\left\|\boldsymbol{U}_{l,\cdot}^\star\right\|_2\frac{\omega_{\max}}{\sigma_r^\star}\sqrt{\frac{\kappa^3\mu r\log^2\left(n+d\right)}{np}}}_{=:\gamma_{1,4}} \\
\lesssim{}&\delta\lambda_{\min}^{1/2}\left(\boldsymbol{\Sigma}_{U,l}^\star\right),
\end{aligned}$$

where the last line holds since

$$\begin{aligned}
\gamma_{1,1}&\lesssim\delta\frac{1}{\sqrt{ndp^2\kappa}\sigma_1^\star}\left\|\boldsymbol{U}_{l,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2, \\
\gamma_{1,2}&\lesssim\delta\frac{\omega_l^\star\omega_{\min}}{\sqrt{np^2}\sigma_1^{\star2}}, \\
\gamma_{1,3}&\lesssim\delta\frac{1}{\sqrt{np}\sigma_1^\star}\left\|\boldsymbol{U}_{l,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2, \\
\gamma_{1,4}&\lesssim\delta\frac{\omega_{\min}}{\sqrt{np}\sigma_1^\star},
\end{aligned}$$

provided that $d\gtrsim\delta^{-2}\kappa^7\mu^3r^3\kappa_\omega^2\log^4(n+d)$.

- When it comes to γ_2, we observe that

$$\begin{aligned}\gamma_2 &\asymp \underbrace{\frac{\kappa^3\mu^2r^2\log^4(n+d)}{ndp^2}\left\|\boldsymbol{U}_{l,\cdot}^\star\right\|_2}_{=:\gamma_{2,1}} + \underbrace{\kappa\frac{\omega_{\max}^4}{p^2\sigma_r^{\star 4}}\frac{d}{n}\log^2(n+d)\left\|\boldsymbol{U}_{l,\cdot}^\star\right\|_2}_{=:\gamma_{2,2}} \\ &\quad + \underbrace{\frac{\kappa^3\mu r\log^2(n+d)}{np}\left\|\boldsymbol{U}_{l,\cdot}^\star\right\|_2}_{=:\gamma_{2,3}} + \underbrace{\frac{\omega_{\max}^2}{\sigma_r^{\star 2}}\frac{d\kappa^2\log(n+d)}{np}\left\|\boldsymbol{U}_{l,\cdot}^\star\right\|_2}_{=:\gamma_{2,4}} \\ &\lesssim \delta\lambda_{\min}^{1/2}\left(\boldsymbol{\Sigma}_{U,l}^\star\right),\end{aligned}$$

where the last line holds since

$$\begin{aligned}\gamma_{2,1} &\lesssim \delta\frac{1}{\sqrt{ndp^2\kappa}\sigma_1^\star}\left\|\boldsymbol{U}_{l,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2, \\ \gamma_{2,2} &\lesssim \delta\frac{\omega_{\min}\omega_l^\star}{\sqrt{np^2}\sigma_1^{\star 2}}, \\ \gamma_{2,3} &\lesssim \delta\frac{1}{\sqrt{np}\sigma_1^\star}\left\|\boldsymbol{U}_{l,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2, \\ \gamma_{2,4} &\lesssim \delta\frac{\omega_{\min}}{\sqrt{np}\sigma_1^\star},\end{aligned}$$

provided that $ndp^2 \gtrsim \delta^{-2}\kappa^8\mu^4r^4\log^8(n+d)$, $np \gtrsim \delta^{-2}\kappa^7\mu^2r^2\log^4(n+d)$,

$$\frac{\omega_{\max}}{\sigma_r^\star}\sqrt{\frac{d}{np}} \lesssim \frac{\delta}{\kappa^{5/2}\mu^{1/2}r^{1/2}\kappa_\omega^{1/2}\log(n+d)}, \qquad \frac{\omega_{\max}^2}{p\sigma_r^{\star 2}}\sqrt{\frac{d}{n}} \lesssim \frac{\delta}{\kappa^2\mu^{1/2}r^{1/2}\kappa_\omega\log^2(n+d)}.$$

- We are left with γ_3, which can be bounded by

$$\begin{aligned}\gamma_3 &\asymp \underbrace{\sqrt{\frac{\kappa^5\mu^2r^2\log^4(n+d)}{ndp^2}}\left\|\boldsymbol{U}_{l,\cdot}^\star\right\|_2\frac{\zeta_{\mathsf{1st}}}{\sigma_r^{\star 2}}}_{=:\gamma_{3,1}} + \underbrace{\frac{\omega_l^\star\omega_{\max}}{p\sigma_r^{\star 2}}\sqrt{\frac{\kappa^3\mu r\log^3(n+d)}{n}}\frac{\zeta_{\mathsf{1st}}}{\sigma_r^{\star 2}}}_{=:\gamma_{3,2}} + \underbrace{\sqrt{\frac{\kappa^5\mu r\log^3(n+d)}{np}}\left\|\boldsymbol{U}_{l,\cdot}^\star\right\|_2\frac{\zeta_{\mathsf{1st}}}{\sigma_r^{\star 2}}}_{=:\gamma_{3,3}} \\ &\quad + \underbrace{\frac{\omega_l^\star}{\sigma_r^\star}\sqrt{\frac{\kappa^4\mu r\log^2(n+d)}{np}}\frac{\zeta_{\mathsf{1st}}}{\sigma_r^{\star 2}}}_{=:\gamma_{3,4}} + \underbrace{\frac{\omega_{\max}}{p\sigma_r^\star}\sqrt{\frac{\kappa^4\mu r\log^4(n+d)}{n}}\left\|\boldsymbol{U}_{l,\cdot}^\star\right\|_2\frac{\zeta_{\mathsf{1st}}}{\sigma_r^{\star 2}}}_{=:\gamma_{3,5}} + \underbrace{\frac{\omega_l^\star}{p\sigma_r^\star}\sqrt{\frac{\kappa^4\mu^2r^2\log^4(n+d)}{nd}}\frac{\zeta_{\mathsf{1st}}}{\sigma_r^{\star 2}}}_{=:\gamma_{3,6}}.\end{aligned}$$

Similar to γ_1, we can show that $\gamma_{3,1} \lesssim \delta\lambda_{\min}^{1/2}(\boldsymbol{\Sigma}_{U,l}^\star)$, $\gamma_{3,3} \lesssim \delta\lambda_{\min}^{1/2}(\boldsymbol{\Sigma}_{U,l}^\star)$ and $\gamma_{3,5} \lesssim \delta\lambda_{\min}^{1/2}(\boldsymbol{\Sigma}_{U,l}^\star)$, provided that $ndp^2 \gtrsim \delta^{-2}\kappa^9\mu^4r^4\kappa_\omega^2\log^7(n+d)$, $np \gtrsim \delta^{-2}\kappa^9\mu^3r^3\kappa_\omega^2\log^6(n+d)$, and

$$\frac{\omega_{\max}}{p\sigma_r^\star}\sqrt{\frac{1}{n}} \lesssim \delta\sqrt{\frac{1}{\kappa^8\mu^3r^3\kappa_\omega^2\log^7(n+d)}}.$$

The last condition can be guaranteed by $ndp^2 \gtrsim \delta^{-2}\kappa^9\mu^4r^4\kappa_\omega^2\log^7(n+d)$ and

$$\frac{\omega_{\max}^2}{p\sigma_r^{\star 2}}\sqrt{\frac{d}{n}} \lesssim 1.$$

We are left with bounding $\gamma_{3,2}$, $\gamma_{3,4}$ and $\gamma_{3,6}$, where we have

$$\gamma_{3,2} \lesssim \delta\frac{\omega_l^\star\omega_{\min}}{\sqrt{np^2}\sigma_1^{\star 2}} \lesssim \delta\lambda_{\min}^{1/2}\left(\boldsymbol{\Sigma}_{U,l}^\star\right),$$

$$\gamma_{3,4} \lesssim \delta \frac{\omega_l^\star}{\sqrt{np}\sigma_1^\star} \lesssim \delta \lambda_{\min}^{1/2}\left(\boldsymbol{\Sigma}_{U,l}^\star\right),$$
$$\gamma_{3,6} \lesssim \delta \frac{1}{\sqrt{ndp^2\kappa}\sigma_1^\star}\omega_l^\star \lesssim \delta \lambda_{\min}^{1/2}\left(\boldsymbol{\Sigma}_{U,l}^\star\right),$$

provided that

$$\frac{\zeta_{\mathsf{1st}}}{\sigma_r^{\star 2}} \lesssim \frac{\delta}{\sqrt{\kappa^6\mu^2r^2\kappa_\omega \log^4(n+d)}}.$$

In view of (E.17), the above condition is equivalent to

$$ndp^2 \gtrsim \delta^{-2}\kappa^8\mu^4r^4\kappa_\omega \log^8(n+d), \qquad np \gtrsim \delta^{-2}\kappa^8\mu^3r^3\kappa_\omega\log^6(n+d),$$
$$\text{and} \quad \frac{\omega_{\max}^2}{p\sigma_r^{\star 2}}\sqrt{\frac{d}{n}} \lesssim \frac{\delta}{\sqrt{\kappa^6\mu^2r^2\kappa_\omega\log^6(n+d)}}, \qquad \frac{\omega_{\max}}{\sigma_r^\star}\sqrt{\frac{d}{np}} \lesssim \frac{\delta}{\sqrt{\kappa^7\mu^2r^2\kappa_\omega\log^5(n+d)}}.$$

Therefore we have

$$\gamma_3 \lesssim \delta\lambda_{\min}^{1/2}\left(\boldsymbol{\Sigma}_{U,l}^\star\right)$$

under the conditions

$$ndp^2 \gtrsim \delta^{-2}\kappa^9\mu^4r^4\kappa_\omega^2 \log^8(n+d), \qquad np \gtrsim \delta^{-2}\kappa^9\mu^3r^3\kappa_\omega^2\log^6(n+d),$$
$$\text{and} \quad \frac{\omega_{\max}^2}{p\sigma_r^{\star 2}}\sqrt{\frac{d}{n}} \lesssim \frac{\delta}{\sqrt{\kappa^6\mu^2r^2\kappa_\omega\log^6(n+d)}}, \qquad \frac{\omega_{\max}}{\sigma_r^\star}\sqrt{\frac{d}{np}} \lesssim \frac{\delta}{\sqrt{\kappa^7\mu^2r^2\kappa_\omega\log^5(n+d)}}.$$

Therefore, $\zeta_{\mathsf{2nd},l} \lesssim \delta\lambda_{\min}^{1/2}\left(\boldsymbol{\Sigma}_{U,l}^\star\right)$ is guaranteed to hold as long as $d \gtrsim \delta^{-2}\kappa^7\mu^3r^3\kappa_\omega^2\log^4(n+d)$,

$$ndp^2 \gtrsim \delta^{-2}\kappa^9\mu^4r^4\kappa_\omega^2 \log^8(n+d), \qquad np \gtrsim \delta^{-2}\kappa^9\mu^3r^3\kappa_\omega^2\log^6(n+d),$$
$$\text{and} \quad \frac{\omega_{\max}^2}{p\sigma_r^{\star 2}}\sqrt{\frac{d}{n}} \lesssim \frac{\delta}{\sqrt{\kappa^6\mu^2r^2\kappa_\omega\log^6(n+d)}}, \qquad \frac{\omega_{\max}}{\sigma_r^\star}\sqrt{\frac{d}{np}} \lesssim \frac{\delta}{\sqrt{\kappa^7\mu^2r^2\kappa_\omega\log^5(n+d)}}.$$

By taking $\delta = 1/(r^{1/4}\log^{1/2}(n+d))$, we see that $\zeta \lesssim 1/(r^{1/4}\log^{1/2}(n+d))$ holds provided that $d \gtrsim \kappa^7\mu^3r^{7/2}\kappa_\omega^2\log^5(n+d)$,

$$ndp^2 \gtrsim \kappa^9\mu^4r^{9/2}\kappa_\omega^2 \log^9(n+d), \qquad np \gtrsim \kappa^9\mu^3r^{7/2}\kappa_\omega^2\log^7(n+d),$$
$$\text{and} \quad \frac{\omega_{\max}^2}{p\sigma_r^{\star 2}}\sqrt{\frac{d}{n}} \lesssim \frac{1}{\sqrt{\kappa^6\mu^2r^{5/2}\kappa_\omega\log^7(n+d)}}, \qquad \frac{\omega_{\max}}{\sigma_r^\star}\sqrt{\frac{d}{np}} \lesssim \frac{1}{\sqrt{\kappa^7\mu^2r^{5/2}\kappa_\omega\log^6(n+d)}}.$$

Step 4: distributional characterization of $\boldsymbol{Z}_{l,\cdot}$. For any convex set $\mathcal{C} \in \mathscr{C}^r$, it holds that

$$\begin{aligned}
&\left|\mathbb{P}\left((\boldsymbol{U}\boldsymbol{R}-\boldsymbol{U}^\star)_{l,\cdot} \in \mathcal{C}\right) - \mathbb{P}\left(\mathcal{N}\left(\boldsymbol{0},\boldsymbol{\Sigma}_{U,l}^\star\right)\in\mathcal{C}\right)\right| \\
&\quad = \left|\mathbb{E}\left[\mathbb{P}\left((\boldsymbol{U}\boldsymbol{R}-\boldsymbol{U}^\star)_{l,\cdot} \in \mathcal{C}\middle|\boldsymbol{F}\right) - \mathbb{P}\left(\mathcal{N}\left(\boldsymbol{0},\boldsymbol{\Sigma}_{U,l}^\star\right)\in\mathcal{C}\right)\right]\right| \\
&\quad \leq \left|\mathbb{E}\left[\left[\mathbb{P}\left((\boldsymbol{U}\boldsymbol{R}-\boldsymbol{U}^\star)_{l,\cdot} \in \mathcal{C}\middle|\boldsymbol{F}\right) - \mathbb{P}\left(\mathcal{N}\left(\boldsymbol{0},\boldsymbol{\Sigma}_{U,l}^\star\right)\in\mathcal{C}\right)\right]\mathbb{1}_{\mathcal{E}_{\mathsf{good}}}\right]\right| \\
&\qquad + \left|\mathbb{E}\left[\left[\mathbb{P}\left((\boldsymbol{U}\boldsymbol{R}-\boldsymbol{U}^\star)_{l,\cdot} \in \mathcal{C}\middle|\boldsymbol{F}\right) - \mathbb{P}\left(\mathcal{N}\left(\boldsymbol{0},\boldsymbol{\Sigma}_{U,l}^\star\right)\in\mathcal{C}\right)\right]\mathbb{1}_{\mathcal{E}_{\mathsf{good}}^{\mathsf{c}}}\right]\right| \\
&\quad \leq \frac{1}{\sqrt{\log(n+d)}} + 2\mathbb{P}\left(\mathcal{E}_{\mathsf{good}}^{\mathsf{c}}\right) \\
&\quad \lesssim \frac{1}{\sqrt{\log(n+d)}} + \frac{1}{(n+d)^{100}} \lesssim \frac{1}{\sqrt{\log(n+d)}},
\end{aligned}$$

where the penultimate line relies on (E.50). This allows one to conclude that

$$\sup_{\mathcal{C}\in\mathscr{C}^r}\left|\mathbb{P}\left(\boldsymbol{Z}_{l,\cdot}\in\mathcal{C}\right)-\mathbb{P}\left(\mathcal{N}\left(\boldsymbol{0},\boldsymbol{\Sigma}_{U,l}^{\star}\right)\in\mathcal{C}\right)\right|\lesssim\frac{1}{\sqrt{\log\left(n+d\right)}}=o\left(1\right).$$

E.3 Auxiliary lemmas for Theorem 12

E.3.1 Proof of Lemma 18

This lemma is concerned with several different quantities, which we seek to control separately.

Bounding $\|(\boldsymbol{U}\boldsymbol{R}-\boldsymbol{U}^{\star})_{l,\cdot}\|_2$ and $\|\boldsymbol{U}\boldsymbol{R}-\boldsymbol{U}^{\star}\|_{2,\infty}$. In view of Lemma 15, we can begin with the decomposition

$$\begin{aligned}\left\|\boldsymbol{U}_{l,\cdot}\boldsymbol{R}-\boldsymbol{U}_{l,\cdot}^{\star}\right\|_2 &\leq\left\|\boldsymbol{Z}_{l,\cdot}\right\|_2+\zeta_{\mathsf{2nd},l}=\left\|\left[\boldsymbol{E}\boldsymbol{M}^{\natural\top}+\mathcal{P}_{\mathsf{off\text{-}diag}}\left(\boldsymbol{E}\boldsymbol{E}^{\top}\right)\right]_{l,\cdot}\boldsymbol{U}^{\natural}\left(\boldsymbol{\Sigma}^{\natural}\right)^{-2}\boldsymbol{Q}^{\top}\right\|_2+\zeta_{\mathsf{2nd},l}\\ &\leq\underbrace{\left\|\boldsymbol{E}_{l,\cdot}\boldsymbol{V}^{\natural}(\boldsymbol{\Sigma}^{\natural})^{-1}\boldsymbol{Q}^{\top}\right\|_2}_{=:\alpha_1}+\underbrace{\left\|\boldsymbol{E}_{l,\cdot}\left[\mathcal{P}_{-l,\cdot}\left(\boldsymbol{E}\right)\right]^{\top}\boldsymbol{U}^{\natural}\left(\boldsymbol{\Sigma}^{\natural}\right)^{-2}\boldsymbol{Q}^{\top}\right\|_2}_{=:\alpha_2}+\zeta_{\mathsf{2nd},l},\end{aligned}$$

where $\boldsymbol{Z}$ and $\zeta_{\mathsf{2nd},l}$ are defined in (D.22).

We first bound α_1. One can decompose

$$\boldsymbol{E}_{l,\cdot}\boldsymbol{V}^{\natural}(\boldsymbol{\Sigma}^{\natural})^{-1}\boldsymbol{Q}^{\top}=\sum_{j=1}^{n}E_{l,j}\boldsymbol{V}_{j,\cdot}^{\natural}(\boldsymbol{\Sigma}^{\natural})^{-1}\boldsymbol{Q}^{\top},$$

which is an independent sum of random vectors (where the randomness comes from $\{E_{l,j}\}_{1\leq j\leq n}$) conditional on $\boldsymbol{F}$. By carrying out the following calculation (see (D.8) and (D.9))

$$\begin{aligned}L&\coloneqq\max_{1\leq j\leq n}\left\|E_{l,j}\boldsymbol{V}_{j,\cdot}^{\natural}(\boldsymbol{\Sigma}^{\natural})^{-1}\boldsymbol{Q}^{\top}\right\|_2\leq\left\{\max_{1\leq j\leq n}\left|E_{l,j}\right|\right\}\left\|\boldsymbol{V}^{\natural}(\boldsymbol{\Sigma}^{\natural})^{-1}\right\|_{2,\infty}\\ &\lesssim\frac{1}{\sqrt{np}}\left(\max_{1\leq j\leq n}\left|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\boldsymbol{f}_j\right|+\omega_l^{\star}\sqrt{\log\left(n+d\right)}\right)\left\|\boldsymbol{V}^{\natural}(\boldsymbol{\Sigma}^{\natural})^{-1}\right\|_{2,\infty},\\ V&\coloneqq\sum_{j=1}^{n}\mathbb{E}\left[E_{l,j}^2\left\|\boldsymbol{V}_{j,\cdot}^{\natural}(\boldsymbol{\Sigma}^{\natural})^{-1}\boldsymbol{Q}^{\top}\right\|_2^2\right]\leq\sum_{j=1}^{n}\mathbb{E}\left[E_{l,j}^2\right]\left\|\boldsymbol{V}_{j,\cdot}^{\natural}(\boldsymbol{\Sigma}^{\natural})^{-1}\right\|_2^2\leq\max_j\mathbb{E}\left[E_{l,j}^2\right]\left\|\boldsymbol{V}^{\natural}(\boldsymbol{\Sigma}^{\natural})^{-1}\right\|_{\mathrm{F}}^2\\ &\lesssim\frac{1}{np}\left[\max_{1\leq j\leq n}\left(\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\boldsymbol{f}_j\right)^2+\omega_l^{\star2}\right]\left\|\boldsymbol{V}^{\natural}(\boldsymbol{\Sigma}^{\natural})^{-1}\right\|_{\mathrm{F}}^2,\end{aligned}$$

we can invoke the Bernstein inequality (Chen et al., 2021, Corollary 3.1.3) to demonstrate that

$$\begin{aligned}\left\|\boldsymbol{Z}_{l,\cdot}\right\|_2&\lesssim\sqrt{V\log\left(n+d\right)}+L\log\left(n+d\right)\\ &\lesssim\sqrt{\frac{\log\left(n+d\right)}{np}}\left[\max_{1\leq j\leq n}\left|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\boldsymbol{f}_j\right|+\omega_l^{\star}\right]\left\|\boldsymbol{V}^{\natural}(\boldsymbol{\Sigma}^{\natural})^{-1}\right\|_{\mathrm{F}}\\ &\quad+\frac{\log\left(n+d\right)}{\sqrt{np}}\left(\max_{1\leq j\leq n}\left|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\boldsymbol{f}_j\right|+\omega_l^{\star}\sqrt{\log\left(n+d\right)}\right)\left\|\boldsymbol{V}^{\natural}(\boldsymbol{\Sigma}^{\natural})^{-1}\right\|_{2,\infty}\end{aligned}$$

with probability at least $1-O((n+d)^{-10})$. On the event $\mathcal{E}_{\mathsf{good}}$, we can further derive

$$\begin{aligned}\left\|\boldsymbol{Z}_{l,\cdot}\right\|_2&\overset{\text{(i)}}{\lesssim}\frac{1}{\sigma_r^{\star}}\sqrt{\frac{r\log\left(n+d\right)}{np}}\left[\left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2\sqrt{\log\left(n+d\right)}+\omega_l^{\star}\right]\\ &\quad+\frac{1}{\sigma_r^{\star}}\frac{\log\left(n+d\right)}{\sqrt{np}}\left(\left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2\sqrt{\log\left(n+d\right)}+\omega_l^{\star}\sqrt{\log\left(n+d\right)}\right)\sqrt{\frac{\log\left(n+d\right)}{n}}\end{aligned}$$

$$
\asymp \frac{1}{\sigma_r^\star}\left\|\boldsymbol{U}_{l,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2\left(\sqrt{\frac{r\log^2(n+d)}{np}}+\frac{\log^{3/2}(n+d)}{np}\right)+\frac{\omega_l^\star}{\sigma_r^\star}\left(\sqrt{\frac{r\log(n+d)}{np}}+\frac{\log^2(n+d)}{np}\right)
$$
$$
\overset{\text{(ii)}}{\lesssim} \frac{1}{\sigma_r^\star}\left\|\boldsymbol{U}_{l,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2\sqrt{\frac{r\log^2(n+d)}{np}}+\frac{\omega_l^\star}{\sigma_r^\star}\sqrt{\frac{r\log(n+d)}{np}}
$$

where (i) uses (D.13), (D.20) as well as (D.16), and (ii) holds true as long as $np \gtrsim \log^3(n+d)$.

$$
\alpha_1 \lesssim \frac{1}{\sigma_r^\star}\left(\left\|\boldsymbol{U}_{l,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2+\omega_l^\star\right)\sqrt{\frac{r\log^2(n+d)}{np}}.
$$

Next we bound α_2. In order to do so, we know from (D.13) that on the event $\mathcal{E}_{\mathsf{good}}$,

$$
\alpha_2\,\mathbb{1}_{\mathcal{E}_{\mathsf{good}}} \lesssim \frac{1}{\sigma_r^{\star 2}}\left\|\boldsymbol{E}_{l,\cdot}\left[\mathcal{P}_{-l,\cdot}(\boldsymbol{E})\right]^\top\boldsymbol{U}^\natural\right\|_2 \asymp \frac{1}{\sigma_r^{\star 2}}\left\|\boldsymbol{E}_{l,\cdot}\left[\mathcal{P}_{-l,\cdot}(\boldsymbol{E})\right]^\top\boldsymbol{U}^\star\right\|_2, \tag{E.51}
$$

where the last relation arises from the fact $\boldsymbol{U}^\natural=\boldsymbol{U}^\star\boldsymbol{Q}$ for some orthonormal matrix $\boldsymbol{Q}$ (see (D.5)). Therefore, it suffices to bound

$$
\left\|\boldsymbol{E}_{l,\cdot}\left[\mathcal{P}_{-l,\cdot}(\boldsymbol{E})\right]^\top\boldsymbol{U}^\star\right\|_2=\left\|\sum_{j=1}^n E_{l,j}\boldsymbol{C}_{j,\cdot}\boldsymbol{U}^\star\right\|_2=\left\|\sum_{j=1}^n\boldsymbol{X}_j\right\|_2,
$$

where $\boldsymbol{C}=[\mathcal{P}_{-l,\cdot}(\boldsymbol{E})]^\top$ and $\boldsymbol{X}_j=E_{l,j}\boldsymbol{C}_{j,\cdot}\boldsymbol{U}^\star$ for $j\in[n]$. Recognizing that $\sum_{j=1}^n\boldsymbol{X}_j$ is a sum of independent random vectors (conditional on $\boldsymbol{F}$), we can employ the truncated Bernstein inequality (see, e.g., Chen et al. (2021, Theorem 3.1.1)) to bound the above quantity. From now on, we will always assume occurrence of $\mathcal{E}_{\mathsf{good}}$ when bounding $\|(\boldsymbol{U}\boldsymbol{R}-\boldsymbol{U}^\star\|_2$ (recall that $\mathcal{E}_{\mathsf{good}}$ is $\sigma(\boldsymbol{F})$-measurable).

- It is first observed that

$$
\max_{1\le j\le n}\|\boldsymbol{X}_j\|_2 \le \max_{1\le j\le n}|E_{l,j}|\,\|\boldsymbol{C}_{j,\cdot}\boldsymbol{U}^\star\|_2.
$$

Recall that for each $j\in[n]$,

$$
E_{l,j}=\frac{1}{\sqrt{np}}\left[(\delta_{l,j}-1)\,\boldsymbol{U}_{l,\cdot}^\star\boldsymbol{\Sigma}^\star\boldsymbol{f}_j+N_{l,j}\right]
$$

with $\delta_{l,j}=\mathbb{1}_{(l,j)\in\Omega}$. It is seen from (D.20) that

$$
|E_{l,j}| \le \widetilde{C}_1\sqrt{\frac{1}{np^2}}\left(\left\|\boldsymbol{U}_{l,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2+\omega_l^\star\right)\sqrt{\log(n+d)} =: L_E
$$

for some sufficiently large constant $\widetilde{C}_1>0$. In addition, for each $j\in[n]$, we can write $\boldsymbol{C}_{j,\cdot}\boldsymbol{U}^\star=\sum_{i:i\neq l}E_{i,j}\boldsymbol{U}_{i,\cdot}^\star$. It is then straightforward to compute

$$
\begin{aligned}
L_1 &\coloneqq \max_{i:i\neq l}\left\|E_{i,j}\boldsymbol{U}_{i,\cdot}^\star\right\|_2 \lesssim \sqrt{\frac{\log(n+d)}{np^2}}\left(\|\boldsymbol{U}^\star\boldsymbol{\Sigma}^\star\|_{2,\infty}+\omega_{\max}\right)\sqrt{\frac{\mu r}{d}}\\
&\asymp \sqrt{\frac{\mu r\log(n+d)}{ndp^2}}\left(\|\boldsymbol{U}^\star\boldsymbol{\Sigma}^\star\|_{2,\infty}+\omega_{\max}\right) \lesssim \sqrt{\frac{\mu r\log(n+d)}{dp}}\sigma_{\mathsf{ub}},\\
V_1 &\coloneqq \sum_{i:i\neq l}\mathbb{E}\left[E_{i,j}^2\left\|\boldsymbol{U}_{i,\cdot}^\star\right\|_2^2\right] \le \sum_{i:i\neq l}\sigma_{i,j}^2\left\|\boldsymbol{U}_{i,\cdot}^\star\right\|_2^2 \le \sigma_{\mathsf{ub}}^2\|\boldsymbol{U}^\star\|_{\mathrm{F}}^2 = r\sigma_{\mathsf{ub}}^2,
\end{aligned}
$$

where σ_{ub} is defined in (D.17). Apply the matrix Bernstein inequality (Tropp, 2015, Theorem 6.1.1) to achieve

$$
\mathbb{P}\left(\|\boldsymbol{C}_{j,\cdot}\boldsymbol{U}^\star\|_2 \ge t \,\middle|\, \boldsymbol{F}\right) \le (d+1)\exp\left(\frac{-t^2/2}{V_1+L_1t/3}\right).
$$

If we take

$$L_C \coloneqq \widetilde{C}_2 \left[\sqrt{V_1 \log(n+d)} + L_1 \log(n+d) \right].$$

for some sufficiently large constant $\widetilde{C}_2 > 0$, then the above inequality tells us that for each $j \in [n]$

$$\mathbb{P}\left(\left\| \boldsymbol{C}_{j,\cdot} \boldsymbol{U}^\star \right\|_2 \geq L_C \,\middle|\, \boldsymbol{F} \right) \leq (n+d)^{-20}.$$

Consequently, by setting

$$\begin{aligned}
L \coloneqq L_E L_C &\asymp \sqrt{\frac{\log(n+d)}{np^2}} \left(\left\| \boldsymbol{U}_{l,\cdot}^\star \boldsymbol{\Sigma}^\star \right\|_2 + \omega_l^\star \right) \left[\sqrt{V_1 \log(n+d)} + L_1 \log(n+d) \right] \\
&\lesssim \sqrt{\frac{r \log^2(n+d)}{np^2}} \sigma_{\mathsf{ub}} \left(\left\| \boldsymbol{U}_{l,\cdot}^\star \boldsymbol{\Sigma}^\star \right\|_2 + \omega_l^\star \right) \left[1 + \sqrt{\frac{\mu}{dp}} \right],
\end{aligned}$$

we can further derive

$$\mathbb{P}\left(\|\boldsymbol{X}_j\| \geq L \,\middle|\, \boldsymbol{F} \right) \leq \mathbb{P}\left(\left\| \boldsymbol{C}_{j,\cdot} \boldsymbol{U}^\star \right\|_2 \geq L_C \,\middle|\, \boldsymbol{F} \right) \leq (n+d)^{-20} \coloneqq q_0.$$

- Next, we can develop the following upper bound

$$\begin{aligned}
q_1 \coloneqq \left\| \mathbb{E}\left[\boldsymbol{X}_j \mathbb{1}_{\|\boldsymbol{X}_j\| \geq L} \,\middle|\, \boldsymbol{F} \right] \right\| &\overset{\text{(i)}}{\leq} \mathbb{E}\left[\|\boldsymbol{X}_j\| \mathbb{1}_{\|\boldsymbol{X}_j\| \geq L} \,\middle|\, \boldsymbol{F} \right] \\
&\overset{\text{(ii)}}{\leq} L_E \mathbb{E}\left[\left\| \boldsymbol{C}_{j,\cdot} \boldsymbol{U}^\star \right\|_2 \mathbb{1}_{\|\boldsymbol{C}_{j,\cdot} \boldsymbol{U}^\star\|_2 \geq L_C} \,\middle|\, \boldsymbol{F} \right] \\
&= L_E \int_0^\infty \mathbb{P}\left(\left\| \boldsymbol{C}_{j,\cdot} \boldsymbol{U}^\star \right\|_2 \mathbb{1}_{\|\boldsymbol{C}_{j,\cdot} \boldsymbol{U}^\star\|_2 \geq L_C} \geq t \,\middle|\, \boldsymbol{F} \right) \mathrm{d}t \\
&= L_E \int_0^{L_C} \mathbb{P}\left(\left\| \boldsymbol{C}_{j,\cdot} \boldsymbol{U}^\star \right\|_2 \geq L_C \,\middle|\, \boldsymbol{F} \right) \mathrm{d}t + L_E \int_{L_C}^\infty \mathbb{P}\left(\left\| \boldsymbol{C}_{j,\cdot} \boldsymbol{U}^\star \right\|_2 \geq t \,\middle|\, \boldsymbol{F} \right) \mathrm{d}t \\
&\leq L_E L_C (n+d)^{-20} + L_E \int_{L_C}^\infty \mathbb{P}\left(\left\| \boldsymbol{C}_{j,\cdot} \boldsymbol{U}^\star \right\|_2 \geq t \,\middle|\, \boldsymbol{F} \right) \mathrm{d}t.
\end{aligned}$$

Here, (i) makes use of Jensen's inequality, while (ii) holds since the conditions $\|\boldsymbol{X}_j\| \geq L$ and $|E_{l,j}| \leq L_E$ taken together imply that $\|\boldsymbol{C}_{j,\cdot} \boldsymbol{U}^\star\|_2 \geq L_C$. Note that for any $t \geq L_C$, we have $t \gg \sqrt{V_1 \log(n+d)}$ and $t \gg L_1 \log(n+d)$ as long as $\widetilde{C}_2$ is sufficiently large. Therefore, we arrive at

$$\mathbb{P}\left(\left\| \boldsymbol{C}_{j,\cdot} \boldsymbol{U}^\star \right\|_2 \geq t \,\middle|\, \boldsymbol{F} \right) \leq (d+1) \exp\left(-t / \max\left\{ 4\sqrt{V_1 / \log(n+d)}, 4L_1/3 \right\} \right),$$

which immediately gives

$$\begin{aligned}
\int_{L_C}^\infty \mathbb{P}\left(\left\| \boldsymbol{C}_{j,\cdot} \boldsymbol{U}^\star \right\|_2 \geq t \,\middle|\, \boldsymbol{F} \right) \mathrm{d}t &\leq (d+1) \int_{L_C}^\infty \exp\left(-t / \max\left\{ 4\sqrt{V_1 / \log(n+d)}, 4L_1/3 \right\} \right) \mathrm{d}t \\
&\leq (d+1) \max\left\{ 4\sqrt{V_1 / \log(n+d)}, 4L_1/3 \right\} \exp\left(-L_C / \max\left\{ 4\sqrt{V_1 / \log(n+d)}, 4L_1/3 \right\} \right) \\
&\leq 4(d+1) \widetilde{C}_2^{-1} L_C \exp\left(-4\widetilde{C}_2 \log(n+d) \right) \leq L_C (n+d)^{-20},
\end{aligned}$$

provided that $\widetilde{C}_2$ is sufficiently large. As a consequence, we reach

$$q_1 \leq 2 L_E L_C (n+d)^{-20} \leq L (n+d)^{-19}.$$

- Finally, let us calculate the variance statistics as follows

$$v \coloneqq \sum_{j=1}^n \mathbb{E}\left[\|\boldsymbol{X}_j\|_2^2 \,\middle|\, \boldsymbol{F} \right] = \sum_{j=1}^n \mathbb{E}\left[\left\| E_{l,j} \boldsymbol{C}_{j,\cdot} \boldsymbol{U}^\star \right\|_2^2 \,\middle|\, \boldsymbol{F} \right] = \sum_{j=1}^n \sigma_{l,j}^2 \mathbb{E}\left[\left\| \boldsymbol{C}_{j,\cdot} \boldsymbol{U}^\star \right\|_2^2 \,\middle|\, \boldsymbol{F} \right]$$

$$
\begin{aligned}
&\leq \max_{1\leq j\leq n}\sigma_{l,j}^2\mathbb{E}\left[\|\boldsymbol{C}\boldsymbol{U}^\star\|_{\mathrm{F}}^2 \mid \boldsymbol{F}\right] = \max_{1\leq j\leq n}\sigma_{l,j}^2\mathbb{E}\left[\left\|\left[\mathcal{P}_{-l,\cdot}(\boldsymbol{E})\right]^\top\boldsymbol{U}^\star\right\|_{\mathrm{F}}^2 \mid \boldsymbol{F}\right]\\
&\leq \frac{\left\|\boldsymbol{U}_{l,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2^2\log(n+d)+\omega_l^{\star 2}}{np}\mathbb{E}\left[\left\|\left[\mathcal{P}_{-l,\cdot}(\boldsymbol{E})\right]^\top\boldsymbol{U}^\star\right\|_{\mathrm{F}}^2 \mid \boldsymbol{F}\right],
\end{aligned}
$$

where the last inequality results from the following relation

$$
\max_{1\leq j\leq n}\sigma_{l,j}^2 \asymp \frac{1-p}{np}\max_{1\leq j\leq n}\left(\boldsymbol{U}_{l,\cdot}^\star\boldsymbol{\Sigma}^\star\boldsymbol{f}_j\right)^2+\frac{\omega_l^{\star 2}}{np}\lesssim\frac{\left\|\boldsymbol{U}_{l,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2^2\log(n+d)+\omega_l^{\star 2}}{np}.
$$

Notice that

$$
\begin{aligned}
\mathbb{E}\left[\left\|\left[\mathcal{P}_{-l,\cdot}(\boldsymbol{E})\right]^\top\boldsymbol{U}^\star\right\|_{\mathrm{F}}^2 \mid \boldsymbol{F}\right] &= \mathbb{E}\left[\mathsf{tr}\left(\boldsymbol{U}^{\star\top}\mathcal{P}_{-l,\cdot}(\boldsymbol{E})\left[\mathcal{P}_{-l,\cdot}(\boldsymbol{E})\right]^\top\boldsymbol{U}^\star\right) \mid \boldsymbol{F}\right]\\
&= \mathbb{E}\left[\mathsf{tr}\left(\mathcal{P}_{-l,\cdot}(\boldsymbol{E})\left[\mathcal{P}_{-l,\cdot}(\boldsymbol{E})\right]^\top\boldsymbol{U}^\star\boldsymbol{U}^{\star\top}\right) \mid \boldsymbol{F}\right]\\
&= \mathsf{tr}\left\{\boldsymbol{D}\boldsymbol{U}^\star\boldsymbol{U}^{\star\top}\right\},
\end{aligned}
$$

where $\boldsymbol{D}=\mathbb{E}\big[\mathcal{P}_{-l,\cdot}(\boldsymbol{E})[\mathcal{P}_{-l,\cdot}(\boldsymbol{E})]^\top|\boldsymbol{F}\big]\in\mathbb{R}^{d\times d}$ is a diagonal matrix with the i-th diagonal entry given by

$$
D_{i,i}=\begin{cases}\sum_{j=1}^n\sigma_{i,j}^2 & \text{if } i\neq l,\\ 0 & \text{if } i=l.\end{cases}
$$

As a result, we have demonstrated that

$$
\begin{aligned}
\mathbb{E}\left[\left\|\left[\mathcal{P}_{-l,\cdot}(\boldsymbol{E})\right]^\top\boldsymbol{U}^\star\right\|_{\mathrm{F}}^2 \mid \boldsymbol{F}\right] &= \mathsf{tr}\left\{\boldsymbol{D}\boldsymbol{U}^\star\boldsymbol{U}^{\star\top}\right\}=\sum_{i=1}^d D_{i,i}\left\|\boldsymbol{U}_{i,\cdot}^\star\right\|_2^2\\
&\leq\sum_{i=1}^d\left(\sum_{j=1}^n\sigma_{i,j}^2\right)\left\|\boldsymbol{U}_{i,\cdot}^\star\right\|_2^2\leq n\sigma_{\mathsf{ub}}^2\left\|\boldsymbol{U}^\star\right\|_{\mathrm{F}}^2=nr\sigma_{\mathsf{ub}}^2.
\end{aligned}
$$

Taking the above inequalities collectively yields

$$
v\lesssim\frac{r\sigma_{\mathsf{ub}}^2}{p}\left(\left\|\boldsymbol{U}_{l,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2^2\log(n+d)+\omega_l^{\star 2}\right).
$$

Equipped with the above quantities, we are ready to invoke the truncated matrix Bernstein inequality (Chen et al., 2021, Theorem 3.1.1) to show that

$$
\begin{aligned}
\left\|\sum_{j=1}^n\boldsymbol{X}_j\right\|_2 &\lesssim\sqrt{v\log(n+d)}+L\log(n+d)+nq_1\\
&\lesssim\sigma_{\mathsf{ub}}\left(\left\|\boldsymbol{U}_{l,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2\sqrt{\log(n+d)}+\omega_l^\star\right)\sqrt{\frac{r\log(n+d)}{p}}+\sqrt{\frac{r\log^4(n+d)}{np^2}}\sigma_{\mathsf{ub}}\left(\left\|\boldsymbol{U}_{l,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2+\omega_l^\star\right)\left[1+\sqrt{\frac{\mu}{dp}}\right]\\
&\lesssim\sigma_{\mathsf{ub}}\left(\left\|\boldsymbol{U}_{l,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2+\omega_l^\star\right)\sqrt{\frac{r\log^2(n+d)}{p}}
\end{aligned}
$$

with probability exceeding $1-O((n+d)^{-10})$, where the last line holds as long as $np\gtrsim\log^2(n+d)$ and $ndp^2\gtrsim\mu\log^2(n+d)$. Substitution into (E.51) yields

$$
\alpha_2\,\mathbb{1}_{\mathcal{E}_{\mathsf{good}}}\lesssim\frac{1}{\sigma_r^{\star 2}}\sigma_{\mathsf{ub}}\left(\left\|\boldsymbol{U}_{l,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2+\omega_l^\star\right)\sqrt{\frac{r\log^2(n+d)}{p}}.
$$

Putting the above bounds together allows one to conclude that

$$\begin{aligned}
\left\|\boldsymbol{U}_{l,\cdot}\boldsymbol{R}-\boldsymbol{U}_{l,\cdot}^{\star}\right\|_{2} &\leq \alpha_1+\alpha_2+\zeta_{\mathsf{2nd},l} \\
&\lesssim \frac{1}{\sigma_r^{\star}}\left(\left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2+\omega_l^{\star}\right)\sqrt{\frac{r\log^2(n+d)}{np}}+\frac{\sigma_{\mathsf{ub}}}{\sigma_r^{\star 2}}\left(\left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2+\omega_l^{\star}\right)\sqrt{\frac{r\log^2(n+d)}{p}}+\zeta_{\mathsf{2nd},l} \\
&\lesssim \frac{1}{\sigma_r^{\star}}\sqrt{\frac{r\log^2(n+d)}{np}}\left(\left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2+\omega_l^{\star}\right)\left(1+\frac{\sigma_{\mathsf{ub}}}{\sigma_r^{\star}}\sqrt{n}\right)+\zeta_{\mathsf{2nd},l} \\
&\asymp \frac{\theta}{\sqrt{\kappa}\sigma_r^{\star}}\left(\left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2+\omega_l^{\star}\right)+\zeta_{\mathsf{2nd},l},
\end{aligned} \tag{E.52}$$

where we define

$$\theta := \sqrt{\frac{\kappa r\log^2(n+d)}{np}}\left(1+\frac{\sigma_{\mathsf{ub}}}{\sigma_r^{\star}}\sqrt{n}\right)$$

for notational simplicity. By taking the supremum over $l\in[d]$, we further arrive at

$$\begin{aligned}
\left\|\boldsymbol{U}\boldsymbol{R}-\boldsymbol{U}^{\star}\right\|_{2,\infty} &= \max_{l\in[d]}\left\|\boldsymbol{U}_{l,\cdot}\boldsymbol{R}-\boldsymbol{U}_{l,\cdot}^{\star}\right\|_{2} \\
&\lesssim \frac{1}{\sigma_r^{\star}}\sqrt{\frac{r\log^2(n+d)}{np}}\left(\left\|\boldsymbol{U}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_{2,\infty}+\omega_{\max}\right)\left[1+\frac{\sigma_{\mathsf{ub}}}{\sigma_r^{\star}}\sqrt{n}\right]+\max_{l\in[d]}\zeta_{\mathsf{2nd},l} \\
&\lesssim \frac{\sigma_{\mathsf{ub}}}{\sigma_r^{\star}}\left(1+\frac{\sigma_{\mathsf{ub}}}{\sigma_r^{\star}}\sqrt{n}\right)\sqrt{r\log^2(n+d)}+\frac{\kappa^{3/2}\mu r\log^{1/2}(n+d)}{d}\frac{\zeta_{\mathsf{1st}}}{\sigma_r^{\star 2}}+\frac{\zeta_{\mathsf{1st}}^2}{\sigma_r^{\star 4}}\sqrt{\frac{\kappa^3\mu r\log(n+d)}{d}} \\
&\asymp \frac{\zeta_{\mathsf{1st}}}{\sigma_r^{\star 2}}\sqrt{\frac{r\log(n+d)}{d}}+\frac{\kappa^{3/2}\mu r\log^{1/2}(n+d)}{d}\frac{\zeta_{\mathsf{1st}}}{\sigma_r^{\star 2}}+\frac{\zeta_{\mathsf{1st}}^2}{\sigma_r^{\star 4}}\sqrt{\frac{\kappa^3\mu r\log(n+d)}{d}} \\
&\asymp \frac{\zeta_{\mathsf{1st}}}{\sigma_r^{\star 2}}\sqrt{\frac{r\log(n+d)}{d}},
\end{aligned} \tag{E.53}$$

where the last relation holds provided that $d\gtrsim\kappa^3\mu^2 r\log(n+d)$ and $\zeta_{\mathsf{1st}}/\sigma_r^{\star 2}\lesssim 1/\sqrt{\kappa^3\mu}$.

Bounding $\|\boldsymbol{R}^{\top}\boldsymbol{\Sigma}^{-2}\boldsymbol{R}-(\boldsymbol{\Sigma}^{\star})^{-2}\|$. Conditional on $\boldsymbol{F}$, we learn from Lemma 5 that with probability exceeding $1-O((n+d)^{-10})$

$$\begin{aligned}
\left\|\boldsymbol{Q}^{\top}\boldsymbol{R}^{\top}\boldsymbol{\Sigma}^{2}\boldsymbol{R}\boldsymbol{Q}-\boldsymbol{\Sigma}^{\natural 2}\right\| \overset{\text{(i)}}{=} \left\|\boldsymbol{R}_{\boldsymbol{U}}^{\top}\boldsymbol{\Sigma}^{2}\boldsymbol{R}_{\boldsymbol{U}}-\boldsymbol{\Sigma}^{\natural 2}\right\| &\overset{\text{(ii)}}{\lesssim} \kappa^{\natural 2}\sqrt{\frac{\mu^{\natural}r}{d}}\zeta_{\mathsf{1st}}+\kappa^{\natural 2}\frac{\zeta_{\mathsf{1st}}^2}{\sigma_r^{\natural 2}} \\
&\overset{\text{(iii)}}{\lesssim} \sqrt{\frac{\kappa^3\mu r\log(n+d)}{d}}\zeta_{\mathsf{1st}}+\kappa\frac{\zeta_{\mathsf{1st}}^2}{\sigma_r^{\star 2}}.
\end{aligned} \tag{E.54}$$

Here, (i) follows from the fact that $\boldsymbol{R}_{\boldsymbol{U}}=\boldsymbol{R}\boldsymbol{Q}$; (ii) follows from Lemma 5; and (iii) utilizes (D.13), (D.15) and (D.14). We can then readily derive

$$\sigma_1^2 \overset{\text{(i)}}{\leq} \sigma_1^{\natural 2}+\left\|\boldsymbol{Q}^{\top}\boldsymbol{R}^{\top}\boldsymbol{\Sigma}^{2}\boldsymbol{R}\boldsymbol{Q}-\boldsymbol{\Sigma}^{\natural 2}\right\| \overset{\text{(ii)}}{\leq} \sigma_1^{\natural 2}+\sqrt{\frac{\kappa^3\mu r\log(n+d)}{d}}\zeta_{\mathsf{1st}}+\kappa\frac{\zeta_{\mathsf{1st}}^2}{\sigma_r^{\star 2}} \overset{\text{(iii)}}{\leq} 2\sigma_1^{\natural 2} \overset{\text{(iv)}}{\leq} 4\sigma_1^{\star 2}, \tag{E.55}$$

$$\sigma_r^2 \overset{\text{(v)}}{\geq} \sigma_r^{\natural 2}-\left\|\boldsymbol{Q}^{\top}\boldsymbol{R}^{\top}\boldsymbol{\Sigma}^{2}\boldsymbol{R}\boldsymbol{Q}-\boldsymbol{\Sigma}^{\natural 2}\right\| \overset{\text{(vi)}}{\geq} \sigma_r^{\natural 2}-\sqrt{\frac{\kappa^3\mu r\log(n+d)}{d}}\zeta_{\mathsf{1st}}-\kappa\frac{\zeta_{\mathsf{1st}}^2}{\sigma_r^{\star 2}} \overset{\text{(vii)}}{\geq} \frac{1}{2}\sigma_r^{\natural 2} \overset{\text{(viii)}}{\geq} \frac{1}{4}\sigma_r^{\star 2}. \tag{E.56}$$

Here, (i) and (v) follow from Weyl's inequality; (ii) and (vi) rely on (E.54); (iii) and (vii) make use of (D.13) and hold true provided that $\zeta_{\mathsf{1st}}/\sigma_r^{\star 2}\ll 1/\sqrt{\kappa}$ and $d\gtrsim\kappa^2\mu r\log(n+d)$; (iv) and (viii) utilize (E.3) and are valid as long as $n\gg r+\log(n+d)$. On the event $\mathcal{E}_{\mathsf{good}}$, it is seen from (D.19) that

$$\left\|\boldsymbol{Q}\boldsymbol{\Sigma}^{\natural}\boldsymbol{Q}^{\top}-\boldsymbol{\Sigma}^{\star}\right\| = \left\|\boldsymbol{Q}\boldsymbol{\Sigma}^{\natural}-\boldsymbol{\Sigma}^{\star}\boldsymbol{Q}\right\| \lesssim \sqrt{\frac{\kappa(r+\log(n+d))}{n}}\sigma_1^{\star},$$

and as a result,

$$\begin{aligned}\left\|\boldsymbol{Q}\boldsymbol{\Sigma}^{\natural 2}\boldsymbol{Q}^{\top}-\boldsymbol{\Sigma}^{\star 2}\right\| &= \left\|\left(\boldsymbol{Q}\boldsymbol{\Sigma}^{\natural}\boldsymbol{Q}^{\top}-\boldsymbol{\Sigma}^{\star}\right)\boldsymbol{Q}\boldsymbol{\Sigma}^{\natural}\boldsymbol{Q}^{\top}+\boldsymbol{\Sigma}^{\star}\left(\boldsymbol{Q}\boldsymbol{\Sigma}^{\natural}\boldsymbol{Q}^{\top}-\boldsymbol{\Sigma}^{\star}\right)\right\| \\ &\leq \left\|\boldsymbol{Q}\boldsymbol{\Sigma}^{\natural}\boldsymbol{Q}^{\top}-\boldsymbol{\Sigma}^{\star}\right\|\left(\sigma_1^{\natural}+\sigma_1^{\star}\right) \\ &\lesssim \sqrt{\frac{\kappa\left(r+\log\left(n+d\right)\right)}{n}}\sigma_1^{\star 2},\end{aligned} \tag{E.57}$$

where the last line relies on (D.13). Taking (E.54) and (E.57) together yields

$$\begin{aligned}\left\|\boldsymbol{R}^{\top}\boldsymbol{\Sigma}^{2}\boldsymbol{R}-\boldsymbol{\Sigma}^{\star 2}\right\| &\leq \left\|\boldsymbol{R}^{\top}\boldsymbol{\Sigma}^{2}\boldsymbol{R}-\boldsymbol{Q}\boldsymbol{\Sigma}^{\natural 2}\boldsymbol{Q}^{\top}\right\|+\left\|\boldsymbol{Q}\boldsymbol{\Sigma}^{\natural 2}\boldsymbol{Q}^{\top}-\boldsymbol{\Sigma}^{\star 2}\right\| \\ &= \left\|\boldsymbol{Q}^{\top}\boldsymbol{R}^{\top}\boldsymbol{\Sigma}^{2}\boldsymbol{R}\boldsymbol{Q}-\boldsymbol{\Sigma}^{\natural 2}\right\|+\left\|\boldsymbol{Q}\boldsymbol{\Sigma}^{\natural 2}\boldsymbol{Q}^{\top}-\boldsymbol{\Sigma}^{\star 2}\right\| \\ &\lesssim \sqrt{\frac{\kappa^{3}\mu r\log\left(n+d\right)}{d}}\zeta_{\mathsf{1st}}+\kappa\frac{\zeta_{\mathsf{1st}}^{2}}{\sigma_r^{\star 2}}+\sqrt{\frac{\kappa\left(r+\log\left(n+d\right)\right)}{n}}\sigma_1^{\star 2}.\end{aligned}$$

In view of the perturbation bound for matrix square roots (Schmitt, 1992, Lemma 2.2), we obtain

$$\begin{aligned}\left\|\boldsymbol{R}^{\top}\boldsymbol{\Sigma}\boldsymbol{R}-\boldsymbol{\Sigma}^{\star}\right\| &\leq \frac{1}{\sigma_r^{\star}+\sigma_r}\left\|\boldsymbol{R}^{\top}\boldsymbol{\Sigma}^{2}\boldsymbol{R}-\boldsymbol{\Sigma}^{\star 2}\right\| \leq \frac{1}{\sigma_r^{\star}}\left\|\boldsymbol{R}^{\top}\boldsymbol{\Sigma}^{2}\boldsymbol{R}-\boldsymbol{\Sigma}^{\star 2}\right\| \\ &\lesssim \sqrt{\frac{\kappa^{3}\mu r\log\left(n+d\right)}{d}}\frac{\zeta_{\mathsf{1st}}}{\sigma_r^{\star}}+\kappa\frac{\zeta_{\mathsf{1st}}^{2}}{\sigma_r^{\star 3}}+\sqrt{\frac{\kappa^{2}\left(r+\log\left(n+d\right)\right)}{n}}\sigma_1^{\star}.\end{aligned} \tag{E.58}$$

In addition, we also have

$$\begin{aligned}\left\|\boldsymbol{R}^{\top}\boldsymbol{\Sigma}^{-2}\boldsymbol{R}-\left(\boldsymbol{\Sigma}^{\star}\right)^{-2}\right\| &= \left\|\boldsymbol{R}^{\top}\boldsymbol{\Sigma}^{-2}\boldsymbol{R}\left(\boldsymbol{R}^{\top}\boldsymbol{\Sigma}^{2}\boldsymbol{R}-\boldsymbol{\Sigma}^{\star 2}\right)\left(\boldsymbol{\Sigma}^{\star}\right)^{-2}\right\| \\ &\leq \left\|\boldsymbol{R}^{\top}\boldsymbol{\Sigma}^{-2}\boldsymbol{R}\right\|\left\|\boldsymbol{R}^{\top}\boldsymbol{\Sigma}^{2}\boldsymbol{R}-\boldsymbol{\Sigma}^{\star 2}\right\|\left\|\left(\boldsymbol{\Sigma}^{\star}\right)^{-2}\right\| \\ &\lesssim \frac{1}{\sigma_r^{\star 4}}\left\|\boldsymbol{R}^{\top}\boldsymbol{\Sigma}^{2}\boldsymbol{R}-\boldsymbol{\Sigma}^{\star 2}\right\| \\ &\lesssim \sqrt{\frac{\kappa^{3}\mu r\log\left(n+d\right)}{d}}\frac{\zeta_{\mathsf{1st}}}{\sigma_r^{\star 4}}+\kappa\frac{\zeta_{\mathsf{1st}}^{2}}{\sigma_r^{\star 6}}+\sqrt{\frac{\kappa^{3}\left(r+\log\left(n+d\right)\right)}{n}}\frac{1}{\sigma_r^{\star 2}}.\end{aligned} \tag{E.59}$$

Here, the penultimate line comes from (E.55).

Bounding $\|(\boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{R}-\boldsymbol{U}^{\star}\boldsymbol{\Sigma}^{\star})_{l,\cdot}\|_2$. We start by bounding

$$\begin{aligned}\left\|\boldsymbol{U}_{l\cdot}\boldsymbol{\Sigma}\boldsymbol{R}-\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2 &\leq \left\|\left(\boldsymbol{U}_{l\cdot}\boldsymbol{R}-\boldsymbol{U}_{l,\cdot}^{\star}\right)\boldsymbol{R}^{\top}\boldsymbol{\Sigma}\boldsymbol{R}+\boldsymbol{U}_{l,\cdot}^{\star}\left(\boldsymbol{R}^{\top}\boldsymbol{\Sigma}\boldsymbol{R}-\boldsymbol{\Sigma}^{\star}\right)\right\|_2 \\ &\overset{\text{(i)}}{\leq} 2\left\|\boldsymbol{U}_{l\cdot}\boldsymbol{R}-\boldsymbol{U}_{l,\cdot}^{\star}\right\|_2\sigma_1^{\star}+\left\|\boldsymbol{U}_{l,\cdot}^{\star}\right\|_2\left\|\boldsymbol{R}^{\top}\boldsymbol{\Sigma}\boldsymbol{R}-\boldsymbol{\Sigma}^{\star}\right\| \\ &\overset{\text{(ii)}}{\lesssim} \theta\left(\left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2+\omega_l^{\star}\right)+\zeta_{\mathsf{2nd},l}\sigma_1^{\star} \\ &\quad+\left\|\boldsymbol{U}_{l,\cdot}^{\star}\right\|_2\left(\sqrt{\frac{\kappa^{3}\mu r\log\left(n+d\right)}{d}}\frac{\zeta_{\mathsf{1st}}}{\sigma_r^{\star}}+\kappa\frac{\zeta_{\mathsf{1st}}^{2}}{\sigma_r^{\star 3}}+\sqrt{\frac{\kappa^{2}\left(r+\log\left(n+d\right)\right)}{n}}\sigma_1^{\star}\right) \\ &\lesssim \theta\left(\left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2+\omega_l^{\star}\right)+\zeta_{\mathsf{2nd},l}\sigma_1^{\star}+\zeta_{\mathsf{2nd},l}\sigma_r^{\star}+\left\|\boldsymbol{U}_{l,\cdot}^{\star}\right\|_2\sqrt{\frac{\kappa^{2}\left(r+\log\left(n+d\right)\right)}{n}}\sigma_1^{\star} \\ &\asymp \theta\left(\left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2+\omega_l^{\star}\right)+\left\|\boldsymbol{U}_{l,\cdot}^{\star}\right\|_2\sqrt{\frac{\kappa^{2}\left(r+\log\left(n+d\right)\right)}{n}}\sigma_1^{\star}+\zeta_{\mathsf{2nd},l}\sigma_1^{\star}.\end{aligned} \tag{E.60}$$

Here, (i) arises from (E.56), whereas (ii) follows from (E.52) and (E.58). In addition, we observe that

$$\left\|\boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{R}-\boldsymbol{U}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_{2,\infty} \leq \left\|\left(\boldsymbol{U}\boldsymbol{R}-\boldsymbol{U}^{\star}\right)\boldsymbol{R}^{\top}\boldsymbol{\Sigma}\boldsymbol{R}+\boldsymbol{U}^{\star}\left(\boldsymbol{R}^{\top}\boldsymbol{\Sigma}\boldsymbol{R}-\boldsymbol{\Sigma}^{\star}\right)\right\|_{2,\infty}$$

$$
\begin{aligned}
&\overset{\text{(i)}}{\leq} \|\boldsymbol{U}\boldsymbol{R}-\boldsymbol{U}^{\star}\|_{2,\infty}\sigma_1^{\star}+\|\boldsymbol{U}^{\star}\|_{2,\infty}\left\|\boldsymbol{R}^{\top}\boldsymbol{\Sigma}\boldsymbol{R}-\boldsymbol{\Sigma}^{\star}\right\| \\
&\overset{\text{(ii)}}{\lesssim} \frac{\zeta_{\mathsf{1st}}}{\sigma_r^{\star}}\sqrt{\frac{\kappa r\log(n+d)}{d}}+\sqrt{\frac{\mu r}{d}}\left(\sqrt{\frac{\kappa^3\mu r\log(n+d)}{d}}\frac{\zeta_{\mathsf{1st}}}{\sigma_r^{\star}}+\kappa\frac{\zeta_{\mathsf{1st}}^2}{\sigma_r^{\star 3}}+\sqrt{\frac{\kappa^2(r+\log(n+d))}{n}}\sigma_1^{\star}\right) \\
&\overset{\text{(iii)}}{\lesssim} \frac{\zeta_{\mathsf{1st}}}{\sigma_r^{\star}}\sqrt{\frac{\kappa r\log(n+d)}{d}}+\sqrt{\frac{\kappa^2\mu r^2\log(n+d)}{nd}}\sigma_1^{\star}.
\end{aligned}
\tag{E.61}
$$

Here, (i) results from (E.56); (ii) follows from (E.53) and (E.58); (iii) holds true provided that $d\gtrsim\kappa^2\mu^2 r$ and $\zeta_{\mathsf{1st}}/\sigma_r^{\star 2}\lesssim 1/\sqrt{\kappa\mu}$.

Bounding $\|\boldsymbol{U}\boldsymbol{\Sigma}^{-2}\boldsymbol{R}-\boldsymbol{U}^{\star}(\boldsymbol{\Sigma}^{\star})^{-2}\|$. It is first observed that, on the event $\mathcal{E}_{\mathsf{good}}$,

$$
\|\boldsymbol{U}\boldsymbol{R}-\boldsymbol{U}^{\star}\|\overset{\text{(i)}}{=}\left\|\boldsymbol{U}\boldsymbol{R}_{\boldsymbol{U}}-\boldsymbol{U}^{\natural}\right\|\overset{\text{(ii)}}{\lesssim}\frac{\zeta_{\mathsf{1st}}}{\sigma_r^{\natural 2}}\overset{\text{(iii)}}{\lesssim}\frac{\zeta_{\mathsf{1st}}}{\sigma_r^{\star 2}}. \tag{E.62}
$$

Here, (i) comes from the facts that $\boldsymbol{R}_{\boldsymbol{U}}=\boldsymbol{R}\boldsymbol{Q}$ and $\boldsymbol{U}^{\natural}=\boldsymbol{U}^{\star}\boldsymbol{Q}$ for some orthonormal matrix $\boldsymbol{Q}$ (see (D.5)); (ii) follows from Lemma 2; and (iii) is a consequence of (D.13). This immediately gives

$$
\begin{aligned}
\left\|\boldsymbol{U}\boldsymbol{\Sigma}^{-2}\boldsymbol{R}-\boldsymbol{U}^{\star}(\boldsymbol{\Sigma}^{\star})^{-2}\right\| &= \left\|(\boldsymbol{U}\boldsymbol{R}-\boldsymbol{U}^{\star})\boldsymbol{R}^{\top}\boldsymbol{\Sigma}^{-2}\boldsymbol{R}+\boldsymbol{U}^{\star}\left[\boldsymbol{R}^{\top}\boldsymbol{\Sigma}^{-2}\boldsymbol{R}-(\boldsymbol{\Sigma}^{\star})^{-2}\right]\right\| \\
&\leq \frac{1}{\sigma_r^2}\|\boldsymbol{U}\boldsymbol{R}-\boldsymbol{U}^{\star}\|+\|\boldsymbol{U}^{\star}\|\left\|\boldsymbol{R}^{\top}\boldsymbol{\Sigma}^{-2}\boldsymbol{R}-(\boldsymbol{\Sigma}^{\star})^{-2}\right\| \\
&\overset{\text{(i)}}{\lesssim} \frac{\zeta_{\mathsf{1st}}}{\sigma_r^{\star 4}}+\sqrt{\frac{\kappa^3\mu r\log(n+d)}{d}}\frac{\zeta_{\mathsf{1st}}}{\sigma_r^{\star 4}}+\kappa\frac{\zeta_{\mathsf{1st}}^2}{\sigma_r^{\star 6}}+\sqrt{\frac{\kappa^3(r+\log(n+d))}{n}}\frac{1}{\sigma_r^{\star 2}} \\
&\overset{\text{(ii)}}{\asymp} \frac{\zeta_{\mathsf{1st}}}{\sigma_r^{\star 4}}+\sqrt{\frac{\kappa^3(r+\log(n+d))}{n}}\frac{1}{\sigma_r^{\star 2}},
\end{aligned}
$$

where (i) follows from (E.59) and (E.62), and (ii) holds true as long as $d\gtrsim\kappa^3\mu r\log(n+d)$ and $\zeta_{\mathsf{1st}}/\sigma_r^{\star 2}\lesssim 1/\sqrt{\kappa}$.

E.3.2 Proof of Lemma 19

Before proceeding, let us make some useful observations: for all $l\in[d]$,

$$
\begin{aligned}
\|\boldsymbol{U}_{l,\cdot}\boldsymbol{\Sigma}\|_2 &= \|\boldsymbol{U}_{l,\cdot}\boldsymbol{\Sigma}\boldsymbol{R}\|_2 \leq \left\|\boldsymbol{U}_{l\cdot}\boldsymbol{\Sigma}\boldsymbol{R}-\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2+\left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2 \\
&\overset{\text{(i)}}{\lesssim} \theta\left(\left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2+\omega_l^{\star}\right)+\left\|\boldsymbol{U}_{l,\cdot}^{\star}\right\|_2\sqrt{\frac{\kappa^2(r+\log(n+d))}{n}}\sigma_1^{\star}+\zeta_{\mathsf{2nd},l}\sigma_1^{\star}+\left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2 \\
&\lesssim \theta\left(\left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2+\omega_l^{\star}\right)+\frac{1}{\sigma_r^{\star}}\left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2\sqrt{\frac{\kappa^2(r+\log(n+d))}{n}}\sigma_1^{\star}+\zeta_{\mathsf{2nd},l}\sigma_1^{\star}+\left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2 \\
&\overset{\text{(ii)}}{\asymp} \left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2+\theta\omega_l^{\star}+\zeta_{\mathsf{2nd},l}\sigma_1^{\star},
\end{aligned}
\tag{E.63}
$$

where θ is defined in (D.31). Here, (i) relies on (D.32b), while (ii) holds true as long as $n\gtrsim\kappa^3 r\log(n+d)$ and $\theta\ll 1$. Additionally, note that

$$
\begin{aligned}
\theta &\asymp \sqrt{\frac{\kappa r\log^2(n+d)}{np}}\left(1+\sqrt{\frac{\kappa\mu r\log(n+d)}{dp}}+\frac{\omega_{\max}}{\sigma_r^{\star}\sqrt{p}}\right) \\
&\asymp \sqrt{\frac{\kappa r\log^2(n+d)}{np}}+\sqrt{\frac{\kappa^2\mu r^2\log^3(n+d)}{ndp^2}}+\frac{\omega_{\max}}{\sigma_r^{\star}}\sqrt{\frac{\kappa r\log^2(n+d)}{np^2}}.
\end{aligned}
\tag{E.64}
$$

In view of the following relation (which makes use of the AM-GM inequality)

$$
\frac{\zeta_{\mathsf{1st}}}{\sigma_r^{\star 2}}\geq\frac{\kappa\mu r\log^2(n+d)}{\sqrt{ndp}}+\frac{\omega_{\max}^2}{p\sigma_r^{\star 2}}\sqrt{\frac{d}{n}}\log(n+d)\geq\frac{\omega_{\max}}{\sigma_r^{\star}}\sqrt{\frac{\kappa\mu r\log^3(n+d)}{np^2}}, \tag{E.65}
$$

we can see that

$$\theta \lesssim \frac{1}{\sqrt{\mu}} \cdot \frac{\zeta_{\mathsf{1st}}}{\sigma_r^{\star 2}}. \tag{E.66}$$

This means that $\theta \ll 1$ can be guaranteed by the condition $\zeta_{\mathsf{1st}} \ll \sigma_r^{\star 2}$. We are now positioned to embark on the proof.

Step 1: bounding $|S_{i,j} - S_{i,j}^\star|$. We first develop an entrywise upper bound on $\boldsymbol{S} - \boldsymbol{S}^\star$ as follows

$$\begin{aligned}
\|\boldsymbol{S} - \boldsymbol{S}^\star\|_\infty &= \left\| (\boldsymbol{U\Sigma R})(\boldsymbol{U\Sigma R})^\top - \boldsymbol{U}^\star\boldsymbol{\Sigma}^\star(\boldsymbol{U}^\star\boldsymbol{\Sigma}^\star)^\top \right\|_\infty \\
&= \left\| (\boldsymbol{U\Sigma R} - \boldsymbol{U}^\star\boldsymbol{\Sigma}^\star)(\boldsymbol{U\Sigma R})^\top + \boldsymbol{U}^\star\boldsymbol{\Sigma}^\star(\boldsymbol{U\Sigma R} - \boldsymbol{U}^\star\boldsymbol{\Sigma}^\star)^\top \right\|_\infty \\
&\le \|\boldsymbol{U\Sigma R} - \boldsymbol{U}^\star\boldsymbol{\Sigma}^\star\|_{2,\infty} \|\boldsymbol{U\Sigma}\|_{2,\infty} + \|\boldsymbol{U\Sigma R} - \boldsymbol{U}^\star\boldsymbol{\Sigma}^\star\|_{2,\infty} \|\boldsymbol{U}^\star\boldsymbol{\Sigma}^\star\|_{2,\infty} \\
&\lesssim \|\boldsymbol{U\Sigma R} - \boldsymbol{U}^\star\boldsymbol{\Sigma}^\star\|_{2,\infty} \|\boldsymbol{U}^\star\boldsymbol{\Sigma}^\star\|_{2,\infty} + \|\boldsymbol{U\Sigma R} - \boldsymbol{U}^\star\boldsymbol{\Sigma}^\star\|_{2,\infty}^2 \\
&\lesssim \left(\frac{\zeta_{\mathsf{1st}}}{\sigma_r^\star}\sqrt{\frac{\kappa r \log(n+d)}{d}} + \sqrt{\frac{\kappa^2 \mu r^2 \log(n+d)}{nd}}\sigma_1^\star \right)\sqrt{\frac{\mu r}{d}}\sigma_1^\star \\
&\quad + \left(\frac{\zeta_{\mathsf{1st}}}{\sigma_r^\star}\sqrt{\frac{\kappa r \log(n+d)}{d}} + \sqrt{\frac{\kappa^2 \mu r^2 \log(n+d)}{nd}}\sigma_1^\star \right)^2 \\
&\asymp \left(\frac{\zeta_{\mathsf{1st}}}{\sigma_r^\star}\sqrt{\frac{\kappa r \log(n+d)}{d}} + \sqrt{\frac{\kappa^2 \mu r^2 \log(n+d)}{nd}}\sigma_1^\star \right)\sqrt{\frac{\mu r}{d}}\sigma_1^\star.
\end{aligned} \tag{E.67}$$

Here the penultimate relation follows from (E.61), and the last relation holds provided that

$$\frac{\zeta_{\mathsf{1st}}}{\sigma_r^\star}\sqrt{\frac{\kappa r \log(n+d)}{d}} + \sqrt{\frac{\kappa^2 \mu r^2 \log(n+d)}{nd}}\sigma_1^\star \lesssim \sqrt{\frac{\mu r}{d}}\sigma_1^\star,$$

which can be guaranteed by $\zeta_{\mathsf{1st}}/\sigma_r^{\star 2} \lesssim 1/\sqrt{\log(n+d)}$ and $n \gtrsim \kappa^2 r \log(n+d)$. Focusing on the (i,j)-th entry, we can obtain (without loss of generality, assume $\omega_i^\star \le \omega_j^\star$)

$$\begin{aligned}
\left|S_{i,j} - S_{i,j}^\star\right| &= \left| (\boldsymbol{U}_{i,\cdot}\boldsymbol{\Sigma R})(\boldsymbol{U}_{j,\cdot}\boldsymbol{\Sigma R})^\top - \boldsymbol{U}_{i,\cdot}^\star\boldsymbol{\Sigma}^\star(\boldsymbol{U}_{j,\cdot}^\star\boldsymbol{\Sigma}^\star)^\top \right| \\
&\le \left\|\boldsymbol{U}_{i\cdot}\boldsymbol{\Sigma R} - \boldsymbol{U}_{i,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2 \|\boldsymbol{U}_{j,\cdot}\boldsymbol{\Sigma}\|_2 + \left\|\boldsymbol{U}_{j\cdot}\boldsymbol{\Sigma R} - \boldsymbol{U}_{j,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2 \left\|\boldsymbol{U}_{i,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2 \\
&\overset{\text{(i)}}{\lesssim} \left\|\boldsymbol{U}_{i\cdot}\boldsymbol{\Sigma R} - \boldsymbol{U}_{i,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2 \left\|\boldsymbol{U}_{j,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2 + \left\|\boldsymbol{U}_{i\cdot}\boldsymbol{\Sigma R} - \boldsymbol{U}_{i,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2 \left(\theta\omega_j^\star + \zeta_{\mathsf{2nd},j}\sigma_1^\star\right) \\
&\quad + \left\|\boldsymbol{U}_{j\cdot}\boldsymbol{\Sigma R} - \boldsymbol{U}_{j,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2 \left\|\boldsymbol{U}_{i,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2 \\
&\overset{\text{(ii)}}{\lesssim} \left[\theta\left(\left\|\boldsymbol{U}_{i,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2 + \omega_i^\star\right) + \left\|\boldsymbol{U}_{i,\cdot}^\star\right\|_2 \sqrt{\frac{\kappa^2(r + \log(n+d))}{n}}\sigma_1^\star + \zeta_{\mathsf{2nd},i}\sigma_1^\star \right] \left\|\boldsymbol{U}_{j,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2 \\
&\quad + \left[\theta\left(\left\|\boldsymbol{U}_{j,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2 + \omega_j^\star\right) + \left\|\boldsymbol{U}_{j,\cdot}^\star\right\|_2 \sqrt{\frac{\kappa^2(r + \log(n+d))}{n}}\sigma_1^\star + \zeta_{\mathsf{2nd},j}\sigma_1^\star \right] \left\|\boldsymbol{U}_{i,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2 \\
&\quad + \left[\theta\left(\left\|\boldsymbol{U}_{i,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2 + \omega_i^\star\right) + \left\|\boldsymbol{U}_{i,\cdot}^\star\right\|_2 \sqrt{\frac{\kappa^2(r + \log(n+d))}{n}}\sigma_1^\star + \zeta_{\mathsf{2nd},i}\sigma_1^\star \right] \left(\theta\omega_j^\star + \zeta_{\mathsf{2nd},j}\sigma_1^\star\right) \\
&\overset{\text{(iii)}}{\lesssim} \left(\theta + \sqrt{\frac{\kappa^3 r \log(n+d)}{n}}\right) \left\|\boldsymbol{U}_{i,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2 \left\|\boldsymbol{U}_{j,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2 + \theta\left(\omega_i^\star \left\|\boldsymbol{U}_{j,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2 + \omega_j^\star \left\|\boldsymbol{U}_{i,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2\right) \\
&\quad + \sigma_1^\star\left(\zeta_{\mathsf{2nd},i}\left\|\boldsymbol{U}_{j,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2 + \zeta_{\mathsf{2nd},j}\left\|\boldsymbol{U}_{i,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2\right) + \theta^2\omega_i^\star\omega_j^\star + \zeta_{\mathsf{2nd},i}\zeta_{\mathsf{2nd},j}\sigma_1^{\star 2}.
\end{aligned} \tag{E.68}$$

Here, (i) follows from (E.63), (ii) arises from (D.32b) and the fact that $\left\|\boldsymbol{U}_{j,\cdot}^\star\right\|_2 \le \frac{1}{\sigma_r^\star}\left\|\boldsymbol{U}_{j,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2$, while (iii) uses the AM-GM inequality $\zeta_{\mathsf{2nd}}\theta\sigma_1^\star\omega_i^\star \lesssim \theta^2\omega_i^\star\omega_j^\star + \zeta_{\mathsf{2nd}}^2\sigma_1^{\star 2}$, and holds true as long as $\theta \ll 1$ (which is guaranteed when $\zeta_{\mathsf{1st}} \ll \sigma_r^{\star 2}$) and $n \gtrsim \kappa^3 r \log(n+d)$.

Step 2: bounding $|\omega_i^2 - \omega_i^{\star 2}|$. With the above bound on $\left|S_{i,i} - S_{i,i}^\star\right|$ in place, we can move on to control $\left|\omega_i^2 - \omega_i^{\star 2}\right|$. First of all, invoke Chernoff's inequality (see (Vershynin, 2018, Exercise 2.3.5)) to show that

$$\mathbb{P}\left(\sum_{j=1}^n \mathbb{1}_{(i,j)\in\Omega} < \frac{1}{2}np\right) \le \exp\left(-\frac{c}{4}np\right) \le (n+d)^{-10}$$

as long as $np \gg \log(n+d)$. In what follows, we shall define the following event

$$\mathcal{E}_i \coloneqq \left\{\sum_{j=1}^n \mathbb{1}_{(l,j)\in\Omega} \ge np/2\right\},$$

which occurs with probability at least $1 - O((n+d)^{-10})$.

Recall the definition of $y_{i,j}$ in (1.3). Conditional on

$$\Omega_{i,\cdot} = \{(i,j) : (i,j) \in \Omega\},$$

one can verify that $\{y_{i,j} : (i,j) \in \Omega_{l,\cdot}\}$ are independent sub-Gaussian random variables with sub-Gaussian norm bounded above by

$$\|y_{i,j}\|_{\psi_2} = \|x_{i,j}\|_{\psi_2} + \|\eta_{i,j}\|_{\psi_2} \lesssim \sqrt{S_{i,i}^\star} + \omega_i^\star$$

for each $j \in [n]$. As a result, we know that: conditional on $\Omega_{i,\cdot}$, $\{y_{i,j}^2 : (i,j) \in \Omega_{i,\cdot}\}$ are independent sub-exponential random variables obeying

$$K \coloneqq \max_{1\le j\le n} \left\|y_{i,j}^2\right\|_{\psi_1} \le \max_{1\le j\le n} \|y_{i,j}\|_{\psi_2}^2 \lesssim S_{i,i}^\star + \omega_i^{\star 2}.$$

In addition, it is easily observed that $\mathbb{E}[y_{i,j}^2] = S_{i,i}^\star + \omega_i^{\star 2}$. Then we can apply the Bernstein inequality (see (Vershynin, 2018, Theorem 2.8.1)) to demonstrate that: for any $t > 0$,

$$\mathbb{P}\left(\left|\frac{\sum_{j=1}^n y_{i,j}^2 \mathbb{1}_{(i,j)\in\Omega}}{\sum_{j=1}^n \mathbb{1}_{(i,j)\in\Omega}} - \omega_i^{\star 2} - S_{i,i}^\star\right| \ge t \,\middle|\, \Omega_{i,\cdot}\right) \le 2\exp\left[-c\left(\sum_{j=1}^n \mathbb{1}_{(i,j)\in\Omega}\right)\min\left\{\frac{t^2}{K^2}, \frac{t}{K}\right\}\right]$$

for some universal constant $c > 0$. Therefore, when $\mathcal{E}_i$ happens, we have

$$\mathbb{P}\left(\left|\frac{\sum_{j=1}^n y_{i,j}^2 \mathbb{1}_{(i,j)\in\Omega}}{\sum_{j=1}^n \mathbb{1}_{(i,j)\in\Omega}} - \omega_i^{\star 2} - S_{i,i}^\star\right| \ge t \,\middle|\, \Omega_{i,\cdot}\right)\mathbb{1}_{\mathcal{E}_i} \le 2\exp\left[-\frac{c}{2}np\min\left\{\frac{t^2}{K^2}, \frac{t}{K}\right\}\right].$$

Take expectation to further achieve that: for any $t > 0$,

$$\begin{aligned}
&\mathbb{P}\left(\left|\frac{\sum_{j=1}^n y_{i,j}^2 \mathbb{1}_{(i,j)\in\Omega}}{\sum_{j=1}^n \mathbb{1}_{(i,j)\in\Omega}} - \omega_i^{\star 2} - S_{i,i}^\star\right| \ge t\right) = \mathbb{E}\left[\mathbb{P}\left(\left|\frac{\sum_{j=1}^n y_{i,j}^2 \mathbb{1}_{(i,j)\in\Omega}}{\sum_{j=1}^n \mathbb{1}_{(i,j)\in\Omega}} - \omega_i^{\star 2} - S_{i,i}^\star\right| \ge t \,\middle|\, \Omega_{i,\cdot}\right)\right] \\
&\quad = \mathbb{E}\left[\mathbb{P}\left(\left|\frac{\sum_{j=1}^n y_{i,j}^2 \mathbb{1}_{(i,j)\in\Omega}}{\sum_{j=1}^n \mathbb{1}_{(i,j)\in\Omega}} - \omega_i^{\star 2} - S_{i,i}^\star\right| \ge t \,\middle|\, \Omega_{i,\cdot}\right)\mathbb{1}_{\mathcal{E}_i}\right] \\
&\qquad + \mathbb{E}\left[\mathbb{P}\left(\left|\frac{\sum_{j=1}^n y_{i,j}^2 \mathbb{1}_{(i,j)\in\Omega}}{\sum_{j=1}^n \mathbb{1}_{(i,j)\in\Omega}} - \omega_i^{\star 2} - S_{i,i}^\star\right| \ge t \,\middle|\, \Omega_{i,\cdot}\right)\mathbb{1}_{\mathcal{E}_i^\mathsf{c}}\right] \\
&\quad \le 2\exp\left[-\frac{c}{2}np\min\left\{\frac{t^2}{K^2}, \frac{t}{K}\right\}\right] + \mathbb{P}\left(\mathcal{E}_i^\mathsf{c}\right) \lesssim 2\exp\left[-\frac{c}{2}np\min\left\{\frac{t^2}{K^2}, \frac{t}{K}\right\}\right] + (n+d)^{-10}.
\end{aligned}$$

By taking

$$t = \widetilde{C}K\sqrt{\frac{\log^2(n+d)}{np}} \asymp \widetilde{C}\left(\omega_i^{\star 2} + S_{i,i}^\star\right)\sqrt{\frac{\log^2(n+d)}{np}}$$

for some sufficiently large constant $\widetilde{C} > 0$, we see that with probability exceeding $1 - O((n+d)^{-10})$

$$\left| \frac{\sum_{j=1}^{n} y_{i,j}^2 \mathbb{1}_{(i,j)\in\Omega}}{\sum_{j=1}^{n} \mathbb{1}_{(i,j)\in\Omega}} - \omega_i^{\star 2} - S_{i,i}^{\star} \right| \lesssim \sqrt{\frac{\log^2(n+d)}{np}} \left(\omega_i^{\star 2} + S_{i,i}^{\star}\right).$$

A direct consequence is that

$$\begin{aligned} \left|\omega_i^2 - \omega_i^{\star 2}\right| &\le \left| \frac{\sum_{j=1}^{n} y_{i,j}^2 \mathbb{1}_{(i,j)\in\Omega}}{\sum_{j=1}^{n} \mathbb{1}_{(i,j)\in\Omega}} - \omega_i^{\star 2} - S_{i,i}^{\star} \right| + \left|S_{i,i} - S_{i,i}^{\star}\right| \\ &\lesssim \sqrt{\frac{\log^2(n+d)}{np}} \left(\omega_i^{\star 2} + S_{i,i}^{\star}\right) + \left|S_{i,i} - S_{i,i}^{\star}\right|. \end{aligned}$$

In view of (E.67) and the fact that $\|\boldsymbol{S}^{\star}\|_{\infty} = \|\boldsymbol{U}^{\star}\|_{2,\infty}^2 \|\boldsymbol{\Sigma}^{\star 2}\| \le \frac{\mu r}{d}\sigma_1^{\star 2}$, the above inequality implies that, for all $i \in [d]$,

$$\begin{aligned} \left|\omega_i^2 - \omega_i^{\star 2}\right| &\lesssim \sqrt{\frac{\log^2(n+d)}{np}} \left(\omega_i^{\star 2} + \frac{\mu r}{d}\sigma_1^{\star 2}\right) + \left(\frac{\zeta_{\mathsf{1st}}}{\sigma_r^{\star}} \sqrt{\frac{\kappa r \log(n+d)}{d}} + \sqrt{\frac{\kappa^2 \mu r^2 \log(n+d)}{nd}}\sigma_1^{\star}\right) \sqrt{\frac{\mu r}{d}}\sigma_1^{\star} \\ &\asymp \sqrt{\frac{\log^2(n+d)}{np}}\omega_i^{\star 2} + \sqrt{\frac{\log^2(n+d)}{np}}\frac{\mu r}{d}\sigma_1^{\star 2} + \zeta_{\mathsf{1st}}\frac{\sqrt{\kappa^2 \mu r^2 \log(n+d)}}{d} + \sqrt{\frac{\kappa^2 \mu^2 r^3 \log(n+d)}{nd^2}}\sigma_1^{\star 2} \\ &\asymp \sqrt{\frac{\log^2(n+d)}{np}}\omega_i^{\star 2} + \zeta_{\mathsf{1st}}\frac{\sqrt{\kappa^2 \mu r^2 \log(n+d)}}{d} + \sqrt{\frac{\kappa^2 \mu^2 r^3 \log(n+d)}{nd^2}}\sigma_1^{\star 2}, \end{aligned}$$

where the last line makes use of the definition of ζ_{1st} in (D.23b). Additionally, in view of (E.68), we can derive for $i = l$ that

$$\begin{aligned} \left|\omega_l^2 - \omega_l^{\star 2}\right| &\lesssim \sqrt{\frac{\log^2(n+d)}{np}} \left(\omega_l^{\star 2} + S_{l,l}^{\star}\right) + \left(\theta + \sqrt{\frac{\kappa^3 r \log(n+d)}{n}}\right) \left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2^2 \\ &\quad + \left(\theta\omega_l^{\star} + \zeta_{\mathsf{2nd},l}\sigma_1^{\star}\right)\left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2 + \theta^2\omega_l^{\star 2} + \zeta_{\mathsf{2nd},l}^2\sigma_1^{\star 2} \\ &\asymp \left(\sqrt{\frac{\log^2(n+d)}{np}} + \theta^2\right)\omega_l^{\star 2} + \left(\theta + \sqrt{\frac{\kappa^3 r \log(n+d)}{n}}\right) \left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2^2 \\ &\quad + \left(\theta\omega_l^{\star} + \zeta_{\mathsf{2nd},l}\sigma_1^{\star}\right)\left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2 + \zeta_{\mathsf{2nd},l}^2\sigma_1^{\star 2}. \end{aligned}$$

Here, the last relation invokes the following condition

$$\sqrt{\frac{\log^2(n+d)}{np}} S_{l,l}^{\star} \lesssim \theta \left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2^2,$$

which is a direct consequence of (E.64).

E.3.3 Proof of Lemma 20

Before proceeding, let us recall a few facts as follows. To begin with, from the definition (D.26) of $\boldsymbol{\Sigma}_{U,l}^{\star}$, we can decompose

$$\begin{aligned} \boldsymbol{R}\boldsymbol{\Sigma}_{U,l}^{\star}\boldsymbol{R}^{\top} = &\underbrace{\frac{1-p}{np}\left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2^2 \boldsymbol{R}\left(\boldsymbol{\Sigma}^{\star}\right)^{-2}\boldsymbol{R}^{\top}}_{=:\boldsymbol{A}_1} + \underbrace{\frac{\omega_l^{\star 2}}{np}\boldsymbol{R}\left(\boldsymbol{\Sigma}^{\star}\right)^{-2}\boldsymbol{R}^{\top}}_{=:\boldsymbol{A}_2} + \underbrace{\frac{2(1-p)}{np}\boldsymbol{R}\boldsymbol{U}_{l,\cdot}^{\star\top}\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{R}^{\top}}_{=:\boldsymbol{A}_3} \\ &+ \underbrace{\boldsymbol{R}\left(\boldsymbol{\Sigma}^{\star}\right)^{-2}\boldsymbol{U}^{\star\top}\mathsf{diag}\left\{\left[d_{l,i}^{\star}\right]_{i=1}^{d}\right\}\boldsymbol{U}^{\star}\left(\boldsymbol{\Sigma}^{\star}\right)^{-2}\boldsymbol{R}^{\top}}_{=:\boldsymbol{A}_4}, \end{aligned}$$

where we recall that

$$d_{l,i}^{\star} \coloneqq \frac{1}{np^2}\left[\omega_l^{\star 2} + (1-p)\left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2^2\right]\left[\omega_i^{\star 2} + (1-p)\left\|\boldsymbol{U}_{i,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2^2\right] + \frac{2(1-p)^2}{np^2}S_{i,l}^{\star 2}.$$

Regarding $\boldsymbol{\Sigma}_{U,l}$, we remind the reader of its definition in (D.28) as follows

$$\boldsymbol{\Sigma}_{U,l} = \underbrace{\frac{1-p}{np}\left\|\boldsymbol{U}_{l,\cdot}\boldsymbol{\Sigma}\right\|_2^2\boldsymbol{\Sigma}^{-2}}_{=:\boldsymbol{B}_1} + \underbrace{\frac{\omega_l^2}{np}\boldsymbol{\Sigma}^{-2}}_{=:\boldsymbol{B}_2} + \underbrace{\frac{2(1-p)}{np}\boldsymbol{U}_{l,\cdot}^{\top}\boldsymbol{U}_{l,\cdot}}_{=:\boldsymbol{B}_3} + \underbrace{\boldsymbol{\Sigma}^{-2}\boldsymbol{U}^{\top}\mathsf{diag}\left\{\left[d_{l,i}\right]_{1\le i\le d}\right\}\boldsymbol{U}\boldsymbol{\Sigma}^{-2}}_{=:\boldsymbol{B}_4},$$

where we define

$$d_{l,i} \coloneqq \frac{1}{np^2}\left[\omega_l^2 + (1-p)\left\|\boldsymbol{U}_{l,\cdot}\boldsymbol{\Sigma}\right\|_2^2\right]\left[\omega_i^2 + (1-p)\left\|\boldsymbol{U}_{i,\cdot}\boldsymbol{\Sigma}\right\|_2^2\right] + \frac{2(1-p)^2}{np^2}S_{l,i}^2.$$

In addition, we also recall from Lemma 16 that

$$\begin{aligned}\lambda_{\min}\left(\boldsymbol{\Sigma}_{U,l}^{\star}\right) \gtrsim{}& \frac{1}{np\sigma_1^{\star 2}}\left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2^2 + \frac{\omega_l^{\star 2}}{np\sigma_1^{\star 2}} + \frac{1}{ndp^2\kappa\sigma_1^{\star 2}}\left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2^2 + \frac{1}{ndp^2\kappa\sigma_1^{\star 2}}\omega_l^{\star 2}\\ &+ \frac{1}{np^2\sigma_1^{\star 4}}\omega_{\min}^2\left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2^2 + \frac{\omega_l^{\star 2}\omega_{\min}^2}{np^2\sigma_1^{\star 4}}.\end{aligned}$$

We are now ready to present the proof, with the focus of bounding $\|\boldsymbol{A}_i - \boldsymbol{B}_i\|$ for each $1 \le i \le 4$ as well as $\left|d_{l,i}^{\star} - d_{l,i}\right|$.

Step 1: controlling $\|\boldsymbol{A}_1 - \boldsymbol{B}_1\|$. It follows from the triangle inequality and Lemma 18 that

$$\begin{aligned}\left|\left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2 - \left\|\boldsymbol{U}_{l,\cdot}\boldsymbol{\Sigma}\right\|_2\right| &= \left|\left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2 - \left\|\boldsymbol{U}_{l,\cdot}\boldsymbol{\Sigma}\boldsymbol{R}\right\|_2\right| \le \left\|\boldsymbol{U}_{l\cdot}\boldsymbol{\Sigma}\boldsymbol{R} - \boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2\\ &\lesssim \theta\left(\left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2 + \omega_l^{\star}\right) + \left\|\boldsymbol{U}_{l,\cdot}^{\star}\right\|_2\sqrt{\frac{\kappa^2\left(r+\log\left(n+d\right)\right)}{n}}\sigma_1^{\star} + \zeta_{\mathsf{2nd},l}\sigma_1^{\star}.\end{aligned}$$

Here, the last line follows from (D.32b). This immediately gives

$$\begin{aligned}\left|\left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2^2 - \left\|\boldsymbol{U}_{l,\cdot}\boldsymbol{\Sigma}\right\|_2^2\right| &\lesssim \left|\left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2 - \left\|\boldsymbol{U}_{l,\cdot}\boldsymbol{\Sigma}\right\|_2\right|\left|\left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2 + \left\|\boldsymbol{U}_{l,\cdot}\boldsymbol{\Sigma}\right\|_2\right|\\ &\asymp \left|\left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2 - \left\|\boldsymbol{U}_{l,\cdot}\boldsymbol{\Sigma}\right\|_2\right|\left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2 + \left|\left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2 - \left\|\boldsymbol{U}_{l,\cdot}\boldsymbol{\Sigma}\right\|_2\right|^2\\ &\lesssim \theta\left(\left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2 + \omega_l^{\star}\right)\left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2 + \left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2\left\|\boldsymbol{U}_{l,\cdot}^{\star}\right\|_2\sqrt{\frac{\kappa^2\left(r+\log\left(n+d\right)\right)}{n}}\sigma_1^{\star}\\ &\quad + \zeta_{\mathsf{2nd},l}\sigma_1^{\star}\left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2 + \theta^2\left(\left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2^2 + \omega_l^{\star 2}\right) + \left\|\boldsymbol{U}_{l,\cdot}^{\star}\right\|_2^2\frac{\kappa^2\left(r+\log\left(n+d\right)\right)}{n}\sigma_1^{\star 2} + \zeta_{\mathsf{2nd},l}^2\sigma_1^{\star 2}\\ &\lesssim \left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2^2\left(\theta + \sqrt{\frac{\kappa^3\left(r+\log\left(n+d\right)\right)}{n}}\right) + \theta\omega_l^{\star}\left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2\\ &\quad + \zeta_{\mathsf{2nd},l}\sigma_1^{\star}\left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2 + \theta^2\omega_l^{\star 2} + \zeta_{\mathsf{2nd},l}^2\sigma_1^{\star 2}.\end{aligned}\tag{E.69}$$

Here the last relation holds provided that $\theta \ll 1$ and $n \gtrsim \kappa^3 r\log(n+d)$. In addition, from (D.32c) in Lemma 18, we know that

$$\left\|\boldsymbol{R}\left(\boldsymbol{\Sigma}^{\star}\right)^{-2}\boldsymbol{R}^{\top} - \boldsymbol{\Sigma}^{-2}\right\| \lesssim \sqrt{\frac{\kappa^3\mu r\log\left(n+d\right)}{d}}\frac{\zeta_{\mathsf{1st}}}{\sigma_r^{\star 4}} + \kappa\frac{\zeta_{\mathsf{1st}}^2}{\sigma_r^{\star 6}} + \sqrt{\frac{\kappa^3\left(r+\log\left(n+d\right)\right)}{n}}\frac{1}{\sigma_r^{\star 2}}.\tag{E.70}$$

These two inequalities help us derive

$$\left\|\boldsymbol{A}_1 - \boldsymbol{B}_1\right\| \le \frac{1-p}{np}\left|\left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2^2 - \left\|\boldsymbol{U}_{l,\cdot}\boldsymbol{\Sigma}\right\|_2^2\right|\left\|\boldsymbol{R}\boldsymbol{\Sigma}^{-2}\boldsymbol{R}^{\top}\right\| + \frac{1-p}{np}\left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2^2\left\|\boldsymbol{R}\left(\boldsymbol{\Sigma}^{\star}\right)^{-2}\boldsymbol{R}^{\top} - \boldsymbol{\Sigma}^{-2}\right\|$$

$$\overset{\text{(i)}}{\lesssim} \underbrace{\frac{1}{np\sigma_r^{\star 2}} \left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2^2 \left(\theta + \sqrt{\frac{\kappa^3\left(r+\log\left(n+d\right)\right)}{n}}\right)}_{=:\alpha_{1,1}} + \underbrace{\frac{1}{np\sigma_r^{\star 2}}\theta\omega_l^{\star}\left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2}_{=:\alpha_{1,2}}$$
$$+ \frac{1}{np\sigma_r^{\star 2}}\zeta_{\mathsf{2nd},l}\sigma_1^{\star}\left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2 + \underbrace{\frac{1}{np\sigma_r^{\star 2}}\theta^2\omega_l^{\star 2}}_{=:\alpha_{1,3}} + \frac{1}{np\sigma_r^{\star 2}}\zeta_{\mathsf{2nd},l}^2\sigma_1^{\star 2}$$
$$+ \underbrace{\frac{1}{np}\left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2^2\sqrt{\frac{\kappa^3\mu r\log\left(n+d\right)}{d}}\frac{\zeta_{\mathsf{1st}}}{\sigma_r^{\star 4}}}_{=:\alpha_{1,4}} + \underbrace{\frac{1}{np}\left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2^2\kappa\frac{\zeta_{\mathsf{1st}}^2}{\sigma_r^{\star 6}}}_{=:\alpha_{1,5}}$$
$$\overset{\text{(ii)}}{\lesssim} \delta\lambda_{\min}\left(\boldsymbol{\Sigma}_{U,l}^{\star}\right) + \frac{\sqrt{\kappa}}{np\sigma_r^{\star}}\zeta_{\mathsf{2nd},l}\left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2 + \frac{\kappa}{np}\zeta_{\mathsf{2nd},l}^2.$$

Here, (i) results from (E.55), (E.69) and (E.70); (ii) holds true due to the following inequalities:

$$\alpha_{1,1}+\alpha_{1,4}+\alpha_{1,5} \lesssim \delta\frac{1}{np\sigma_1^{\star 2}}\left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2^2 \lesssim \delta\lambda_{\min}\left(\boldsymbol{\Sigma}_{U,l}^{\star}\right),$$
$$\alpha_{1,2} \lesssim \delta\frac{1}{np\sigma_1^{\star 2}}\omega_l^{\star}\left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2 \lesssim \delta\frac{1}{np\sigma_1^{\star 2}}\left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2^2 + \delta\frac{\omega_l^{\star 2}}{np\sigma_1^{\star 2}} \lesssim \delta\lambda_{\min}\left(\boldsymbol{\Sigma}_{U,l}^{\star}\right),$$
$$\alpha_{1,3} \lesssim \delta\frac{\omega_l^{\star 2}}{np\sigma_1^{\star 2}} \lesssim \delta\lambda_{\min}\left(\boldsymbol{\Sigma}_{U,l}^{\star}\right),$$

provided that $\theta \lesssim \delta/\kappa$, $n \gtrsim \delta^{-2}\kappa^5 r\log(n+d)$, $\zeta_{\mathsf{1st}}/\sigma_r^{\star 2} \lesssim 1/\kappa$ and $d \gtrsim \kappa^3\mu r\log(n+d)$.

Step 2: controlling $\|\boldsymbol{A}_2 - \boldsymbol{B}_2\|$. Recall from (D.36) that

$$\left|\omega_l^2 - \omega_l^{\star 2}\right| \lesssim \left(\sqrt{\frac{\log^2\left(n+d\right)}{np}} + \theta^2\right)\omega_l^{\star 2} + \left(\theta + \sqrt{\frac{\kappa^3 r\log\left(n+d\right)}{n}}\right)\left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2^2$$
$$+ \left(\theta\omega_l^{\star} + \zeta_{\mathsf{2nd},l}\sigma_1^{\star}\right)\left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2 + \zeta_{\mathsf{2nd},l}^2\sigma_1^{\star 2}. \tag{E.71}$$

Armed with this inequality, we can derive

$$\left\|\boldsymbol{A}_2 - \boldsymbol{B}_2\right\| = \left\|\frac{\omega_l^{\star 2}}{np}\left[\boldsymbol{R}\left(\boldsymbol{\Sigma}^{\star}\right)^{-2}\boldsymbol{R}^{\top} - \boldsymbol{\Sigma}^{-2}\right] + \frac{\omega_l^{\star 2} - \omega_l^2}{np}\boldsymbol{\Sigma}^{-2}\right\|$$
$$\overset{\text{(i)}}{\lesssim} \frac{\omega_l^{\star 2}}{np}\left\|\boldsymbol{R}\left(\boldsymbol{\Sigma}^{\star}\right)^{-2}\boldsymbol{R}^{\top} - \boldsymbol{\Sigma}^{-2}\right\| + \frac{\left|\omega_l^{\star 2} - \omega_l^2\right|}{np\sigma_r^{\star 2}}$$
$$\overset{\text{(ii)}}{\lesssim} \underbrace{\frac{\omega_l^{\star 2}}{np}\left(\sqrt{\frac{\kappa^3\mu r\log\left(n+d\right)}{d}}\frac{\zeta_{\mathsf{1st}}}{\sigma_r^{\star 4}} + \kappa\frac{\zeta_{\mathsf{1st}}^2}{\sigma_r^{\star 6}}\right)}_{=:\alpha_{2,1}} + \underbrace{\frac{\omega_l^{\star 2}}{np\sigma_r^{\star 2}}\sqrt{\frac{\kappa^3\left(r+\log\left(n+d\right)\right)}{n}}}_{=:\alpha_{2,2}}$$
$$+ \underbrace{\frac{1}{np\sigma_r^{\star 2}}\left(\sqrt{\frac{\log^2\left(n+d\right)}{np}} + \theta^2\right)\omega_l^{\star 2}}_{=:\alpha_{2,3}} + \underbrace{\frac{1}{np\sigma_r^{\star 2}}\left(\theta + \sqrt{\frac{\kappa^3 r\log\left(n+d\right)}{n}}\right)\left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2^2}_{=:\alpha_{2,4}}$$
$$+ \underbrace{\frac{1}{np\sigma_r^{\star 2}}\theta\omega_l^{\star}\left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2}_{=:\alpha_{2,5}} + \frac{1}{np\sigma_r^{\star 2}}\zeta_{\mathsf{2nd},l}\sigma_1^{\star}\left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2 + \frac{1}{np\sigma_r^{\star 2}}\zeta_{\mathsf{2nd},l}^2\sigma_1^{\star 2}$$
$$\overset{\text{(iii)}}{\lesssim} \delta\lambda_{\min}\left(\boldsymbol{\Sigma}_{U,l}^{\star}\right) + \frac{\sqrt{\kappa}}{np\sigma_r^{\star}}\zeta_{\mathsf{2nd}}\left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2 + \frac{\kappa}{np}\zeta_{\mathsf{2nd}}^2.$$

Here, (i) arises from (E.55); (ii) utilizes (E.70) and (E.71); (iii) follows from the inequalities below:

$$\begin{aligned}
\alpha_{2,1}+\alpha_{2,2}+\alpha_{2,3} &\lesssim \delta\frac{\omega_l^{\star 2}}{np\sigma_1^{\star 2}} \lesssim \delta\lambda_{\min}\left(\boldsymbol{\Sigma}_{U,l}^{\star}\right), \\
\alpha_{2,4} &\lesssim \delta\frac{1}{np\sigma_1^{\star 2}}\left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2^2 \lesssim \delta\lambda_{\min}\left(\boldsymbol{\Sigma}_{U,l}^{\star}\right), \\
\alpha_{2,5} &\lesssim \delta\frac{1}{np\sigma_1^{\star 2}}\omega_l^{\star}\left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2 \lesssim \delta\frac{1}{np\sigma_1^{\star 2}}\left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2^2+\delta\frac{\omega_l^{\star 2}}{np\sigma_1^{\star 2}} \lesssim \delta\lambda_{\min}\left(\boldsymbol{\Sigma}_{U,l}^{\star}\right),
\end{aligned}$$

provided that $\zeta_{\mathsf{1st}}/\sigma_r^{\star 2}\lesssim 1/\kappa$, $d\gtrsim\kappa^3\mu r\log(n+d)$, $n\gtrsim\delta^{-2}\kappa^5 r\log(n+d)$, $np\gtrsim\delta^{-2}\log^2(n+d)$ and $\theta\lesssim\delta/\kappa$.

Step 3: controlling $\|\boldsymbol{A}_3-\boldsymbol{B}_3\|$. We have learned from (D.32a) in Lemma 18 we know that

$$\left\|\boldsymbol{U}_{l,\cdot}\boldsymbol{R}-\boldsymbol{U}_{l,\cdot}^{\star}\right\|_2 \lesssim \frac{\theta}{\sqrt{\kappa}\sigma_r^{\star}}\left(\left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2+\omega_l^{\star}\right)+\zeta_{\mathsf{2nd},l}.$$

This immediately gives

$$\begin{aligned}
\left\|\boldsymbol{U}_{l,\cdot}\right\|_2 = \left\|\boldsymbol{U}_{l,\cdot}\boldsymbol{R}\right\|_2 &\le \left\|\boldsymbol{U}_{l,\cdot}^{\star}\right\|_2+\left\|\boldsymbol{U}_{l,\cdot}\boldsymbol{R}-\boldsymbol{U}_{l,\cdot}^{\star}\right\|_2 \lesssim \left\|\boldsymbol{U}_{l,\cdot}^{\star}\right\|_2+\frac{\theta}{\sqrt{\kappa}\sigma_r^{\star}}\left(\left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2+\omega_l^{\star}\right)+\zeta_{\mathsf{2nd}} \\
&\le \left\|\boldsymbol{U}_{l,\cdot}^{\star}\right\|_2+\frac{\theta}{\sqrt{\kappa}\sigma_r^{\star}}\left(\left\|\boldsymbol{U}_{l,\cdot}^{\star}\right\|_2\sigma_1^{\star}+\omega_l^{\star}\right)+\zeta_{\mathsf{2nd}} \\
&\lesssim \left\|\boldsymbol{U}_{l,\cdot}^{\star}\right\|_2+\frac{\theta}{\sqrt{\kappa}\sigma_r^{\star}}\omega_l^{\star}+\zeta_{\mathsf{2nd},l},
\end{aligned} \tag{E.72}$$

where the last line holds provided that $\theta\ll 1$. As a consequence, we arrive at

$$\begin{aligned}
\left\|\boldsymbol{A}_3-\boldsymbol{B}_3\right\| &= \frac{2(1-p)}{np}\left\|\boldsymbol{R}\boldsymbol{U}_{l,\cdot}^{\star\top}\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{R}^{\top}-\boldsymbol{U}_{l,\cdot}^{\top}\boldsymbol{U}_{l,\cdot}\right\| = \frac{2(1-p)}{np}\left\|\boldsymbol{U}_{l,\cdot}^{\star\top}\boldsymbol{U}_{l,\cdot}^{\star}-\boldsymbol{R}^{\top}\boldsymbol{U}_{l,\cdot}^{\top}\boldsymbol{U}_{l,\cdot}\boldsymbol{R}\right\| \\
&\le \frac{2(1-p)}{np}\left\|\left(\boldsymbol{U}^{\star}-\boldsymbol{U}\boldsymbol{R}\right)_{l,\cdot}^{\top}\boldsymbol{U}_{l,\cdot}^{\star}\right\|+\frac{2(1-p)}{np}\left\|\left(\boldsymbol{U}_{l}\boldsymbol{R}\right)^{\top}\left(\boldsymbol{U}_{l,\cdot}^{\star}-\boldsymbol{U}_{l,\cdot}\boldsymbol{R}\right)\right\| \\
&\lesssim \frac{1}{np}\left\|\boldsymbol{U}_{l,\cdot}^{\star}-\boldsymbol{U}_{l,\cdot}\boldsymbol{R}\right\|_2\left(\left\|\boldsymbol{U}_{l,\cdot}^{\star}\right\|_2+\left\|\boldsymbol{U}_{l,\cdot}\right\|_2\right) \\
&\overset{\text{(i)}}{\lesssim} \frac{1}{np}\left[\frac{\theta}{\sqrt{\kappa}\sigma_r^{\star}}\left(\left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2+\omega_l^{\star}\right)+\zeta_{\mathsf{2nd},l}\right]\left(\left\|\boldsymbol{U}_{l,\cdot}^{\star}\right\|_2+\frac{\theta}{\sqrt{\kappa}\sigma_r^{\star}}\omega_l^{\star}+\zeta_{\mathsf{2nd},l}\right) \\
&\overset{\text{(ii)}}{\lesssim} \underbrace{\frac{1}{np}\frac{\theta}{\sqrt{\kappa}\sigma_r^{\star 2}}\left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2^2}_{=:\alpha_{3,1}}+\underbrace{\frac{1}{np}\frac{\theta}{\sqrt{\kappa}\sigma_r^{\star 2}}\omega_l^{\star}\left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2}_{=:\alpha_{3,2}}+\frac{1}{np}\zeta_{\mathsf{2nd},l}\left\|\boldsymbol{U}_{l,\cdot}^{\star}\right\|_2+\underbrace{\frac{1}{np}\frac{\theta^2}{\kappa\sigma_r^{\star 2}}\omega_l^{\star 2}}_{=:\alpha_{3,3}} \\
&\quad+\frac{1}{np}\zeta_{\mathsf{2nd},l}\frac{\theta}{\sqrt{\kappa}\sigma_r^{\star}}\omega_l^{\star}+\frac{1}{np}\zeta_{\mathsf{2nd},l}^2 \\
&\overset{\text{(iii)}}{\lesssim} \delta\lambda_{\min}\left(\boldsymbol{\Sigma}_{U,l}^{\star}\right)+\frac{1}{np}\zeta_{\mathsf{2nd},l}\left\|\boldsymbol{U}_{l,\cdot}^{\star}\right\|_2+\frac{1}{np}\zeta_{\mathsf{2nd},l}\frac{\theta}{\sqrt{\kappa}\sigma_r^{\star}}\omega_l^{\star}+\frac{1}{np}\zeta_{\mathsf{2nd},l}^2.
\end{aligned}$$

Here, (i) follows from (D.32a) and (E.72); (ii) holds provided that $\theta\ll 1$; and (iii) is valid due to the following facts

$$\begin{aligned}
\alpha_{3,1} &\lesssim \delta\frac{1}{np\sigma_1^{\star 2}}\left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2^2 \lesssim \delta\lambda_{\min}\left(\boldsymbol{\Sigma}_{U,l}^{\star}\right), \\
\alpha_{3,2} &\lesssim \delta\frac{1}{np\sigma_1^{\star 2}}\omega_l^{\star}\left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2 \lesssim \delta\frac{1}{np\sigma_1^{\star 2}}\left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2^2+\delta\frac{\omega_l^{\star 2}}{np\sigma_1^{\star 2}} \lesssim \delta\lambda_{\min}\left(\boldsymbol{\Sigma}_{U,l}^{\star}\right), \\
\alpha_{3,3} &\lesssim \delta\frac{\omega_l^{\star 2}}{np\sigma_1^{\star 2}} \lesssim \delta\lambda_{\min}\left(\boldsymbol{\Sigma}_{U,l}^{\star}\right),
\end{aligned}$$

provided that $\theta\lesssim\delta/\sqrt{\kappa}$.

Step 4: bounding $\|\boldsymbol{A}_4 - \boldsymbol{B}_4\|$. To begin with, we recall from (D.32d) in Lemma 18 that

$$\left\|\boldsymbol{U}^{\star}\left(\boldsymbol{\Sigma}^{\star}\right)^{-2}\boldsymbol{R}^{\top} - \boldsymbol{U}\boldsymbol{\Sigma}^{-2}\right\| \lesssim \frac{\zeta_{\mathsf{1st}}}{\sigma_r^{\star 4}} + \sqrt{\frac{\kappa^3\left(r + \log\left(n+d\right)\right)}{n}}\frac{1}{\sigma_r^{\star 2}}. \tag{E.73}$$

This allows one to upper bound

$$\begin{aligned}
\|\boldsymbol{A}_4 - \boldsymbol{B}_4\| &\leq \left\|\left(\boldsymbol{U}^{\star}\left(\boldsymbol{\Sigma}^{\star}\right)^{-2}\boldsymbol{R}^{\top} - \boldsymbol{U}\boldsymbol{\Sigma}^{-2}\right)^{\top}\mathsf{diag}\left\{\left[d_{l,i}^{\star}\right]_{i=1}^{d}\right\}\boldsymbol{U}^{\star}\left(\boldsymbol{\Sigma}^{\star}\right)^{-2}\boldsymbol{R}^{\top}\right\| \\
&\quad + \left\|\boldsymbol{U}\boldsymbol{\Sigma}^{-2}\mathsf{diag}\left\{\left[d_{l,i}^{\star} - d_{l,i}\right]_{i=1}^{d}\right\}\boldsymbol{U}^{\star}\left(\boldsymbol{\Sigma}^{\star}\right)^{-2}\boldsymbol{R}^{\top}\right\| \\
&\quad + \left\|\boldsymbol{U}\boldsymbol{\Sigma}^{-2}\mathsf{diag}\left\{\left[d_{l,i}\right]_{i=1}^{d}\right\}\left(\boldsymbol{U}^{\star}\left(\boldsymbol{\Sigma}^{\star}\right)^{-2}\boldsymbol{R}^{\top} - \boldsymbol{U}\boldsymbol{\Sigma}^{-2}\right)\right\| \\
&\overset{\text{(i)}}{\lesssim} \frac{1}{\sigma_r^{\star 2}}\left\|\boldsymbol{U}^{\star}\left(\boldsymbol{\Sigma}^{\star}\right)^{-2}\boldsymbol{R}^{\top} - \boldsymbol{U}\boldsymbol{\Sigma}^{-2}\right\|\max_{1\leq i\leq d}\left(d_{l,i}^{\star} + d_{l,i}\right) + \frac{1}{\sigma_r^{\star 4}}\max_{1\leq i\leq d}\left|d_{l,i}^{\star} - d_{l,i}\right| \\
&\overset{\text{(ii)}}{\lesssim} \left(\frac{\zeta_{\mathsf{1st}}}{\sigma_r^{\star 6}} + \frac{1}{\sigma_r^{\star 4}}\sqrt{\frac{\kappa^3\left(r + \log\left(n+d\right)\right)}{n}}\right)\max_{1\leq i\leq d}\left\{d_{l,i}^{\star} + \left|d_{l,i}^{\star} - d_{l,i}\right|\right\} + \frac{1}{\sigma_r^{\star 4}}\max_{1\leq i\leq d}\left|d_{l,i}^{\star} - d_{l,i}\right| \\
&\overset{\text{(iii)}}{\lesssim} \underbrace{\left(\frac{\zeta_{\mathsf{1st}}}{\sigma_r^{\star 6}} + \frac{1}{\sigma_r^{\star 4}}\sqrt{\frac{\kappa^3\left(r + \log\left(n+d\right)\right)}{n}}\right)\max_{1\leq i\leq d}d_{l,i}^{\star}}_{=:\alpha_4} + \underbrace{\frac{1}{\sigma_r^{\star 4}}\max_{1\leq i\leq d}\left|d_{l,i}^{\star} - d_{l,i}\right|}_{=:\beta}.
\end{aligned} \tag{E.74}$$

Here, (i) relies on (E.55); (ii) comes from (E.73); (iii) holds true provided that $\zeta_{\mathsf{1st}} \lesssim \sigma_r^{\star 2}$ and $n \gtrsim \kappa^3 r\log(n+d)$. Note that, for each $i \in [d]$,

$$\begin{aligned}
d_{l,i}^{\star} &= \frac{1}{np^2}\left[\omega_l^{\star 2} + (1-p)\left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2^2\right]\left[\omega_i^{\star 2} + (1-p)\left\|\boldsymbol{U}_{i,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2^2\right] + \frac{2\left(1-p\right)^2}{np^2}\left(\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star 2}\boldsymbol{U}_{i,\cdot}^{\star\top}\right)^2 \\
&\lesssim \frac{1}{np^2}\left[\omega_l^{\star 2} + (1-p)\left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2^2\right]\left[\omega_{\max}^2 + \frac{\mu r}{d}\sigma_1^{\star 2}\right] + \frac{\mu r}{ndp^2}\sigma_1^{\star 2}\left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2^2 \\
&\lesssim \frac{1}{np^2}\omega_l^{\star 2}\omega_{\max}^2 + \frac{\mu r}{ndp^2}\sigma_1^{\star 2}\omega_l^{\star 2} + \frac{1}{np^2}\omega_{\max}^2\left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2^2 + \frac{\mu r}{ndp^2}\sigma_1^{\star 2}\left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2^2,
\end{aligned} \tag{E.75}$$

which in turn results in

$$\begin{aligned}
\alpha_4 &\overset{\text{(i)}}{\lesssim} \underbrace{\left(\frac{\zeta_{\mathsf{1st}}}{\sigma_r^{\star 6}} + \frac{1}{\sigma_r^{\star 4}}\sqrt{\frac{\kappa^3 r\log\left(n+d\right)}{n}}\right)\frac{\omega_l^{\star 2}\omega_{\max}^2}{np^2}}_{=:\alpha_{4,1}} + \underbrace{\frac{\kappa\mu r}{ndp^2}\left(\frac{\zeta_{\mathsf{1st}}}{\sigma_r^{\star 4}} + \frac{1}{\sigma_r^{\star 2}}\sqrt{\frac{\kappa^3 r\log\left(n+d\right)}{n}}\right)\omega_l^{\star 2}}_{=:\alpha_{4,2}} \\
&\quad + \underbrace{\frac{1}{np^2}\left(\frac{\zeta_{\mathsf{1st}}}{\sigma_r^{\star 6}} + \frac{1}{\sigma_r^{\star 4}}\sqrt{\frac{\kappa^3 r\log\left(n+d\right)}{n}}\right)\omega_{\max}^2\left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2^2}_{=:\alpha_{4,3}} \\
&\quad + \underbrace{\frac{\kappa\mu r}{ndp^2}\left(\frac{\zeta_{\mathsf{1st}}}{\sigma_r^{\star 4}} + \frac{1}{\sigma_r^{\star 2}}\sqrt{\frac{\kappa^3 r\log\left(n+d\right)}{n}}\right)\left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2^2}_{=:\alpha_{4,4}} \\
&\overset{\text{(ii)}}{\lesssim} \delta\lambda_{\min}\left(\boldsymbol{\Sigma}_{U,l}^{\star}\right).
\end{aligned}$$

Here, (i) follows from (E.75), while (ii) holds since

$$\begin{aligned}
\alpha_{4,1} &\lesssim \delta\frac{\omega_l^{\star 2}\omega_{\min}^2}{np^2\sigma_1^{\star 4}} \lesssim \delta\lambda_{\min}\left(\boldsymbol{\Sigma}_{U,l}^{\star}\right), \\
\alpha_{4,2} &\lesssim \delta\frac{1}{ndp^2\kappa\sigma_1^{\star 2}}\omega_l^{\star 2} \lesssim \delta\lambda_{\min}\left(\boldsymbol{\Sigma}_{U,l}^{\star}\right),
\end{aligned}$$

$$\alpha_{4,3} \lesssim \delta \frac{1}{np^2\sigma_1^{\star 4}} \omega_{\min}^2 \left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2^2 \lesssim \delta \lambda_{\min}\left(\boldsymbol{\Sigma}_{U,l}^{\star}\right),$$
$$\alpha_{4,4} \lesssim \delta \frac{1}{ndp^2\kappa\sigma_1^{\star 2}} \left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2^2 \lesssim \delta \lambda_{\min}\left(\boldsymbol{\Sigma}_{U,l}^{\star}\right),$$

provided that $\zeta_{\mathsf{1st}}/\sigma_r^{\star 2} \lesssim \delta/(\kappa^3\mu r\kappa_\omega)$, $n \gtrsim \delta^{-2}\kappa^9 r\log(n+d)$ and $n \gtrsim \kappa^7 r\kappa_\omega^2 \log(n+d)$. Note that we still need to bound the term β in (E.74), which we leave to the next step.

Step 5: bounding $|d_{l,i}^{\star} - d_{l,i}|$. For each $i \in [d]$, we can decompose

$$\begin{aligned}
\frac{1}{\sigma_r^{\star 4}}\left|d_{l,i}^{\star} - d_{l,i}\right| &\leq \frac{1}{np^2\sigma_r^{\star 4}}\left|\left[\omega_l^{\star 2} + (1-p)\,S_{l,l}^{\star}\right]\left[\omega_i^{\star 2} + (1-p)\,S_{i,i}^{\star}\right] - \left[\omega_l^{2} + (1-p)\,S_{l,l}\right]\left[\omega_i^{2} + (1-p)\,S_{i,i}\right]\right| \\
&\quad + \frac{2\,(1-p)^2}{np^2\sigma_r^{\star 4}}\left|S_{i,l}^{\star 2} - S_{i,l}^2\right| \\
&\leq \underbrace{\frac{1}{np^2\sigma_r^{\star 4}}\left|\omega_l^{\star 2} + (1-p)\,S_{l,l}^{\star}\right|\left|\omega_i^{\star 2} + (1-p)\,S_{i,i}^{\star} - \omega_i^2 - (1-p)\,S_{i,i}\right|}_{=:\beta_1} \\
&\quad + \underbrace{\frac{1}{np^2\sigma_r^{\star 4}}\left|\omega_l^{\star 2} + (1-p)\,S_{l,l}^{\star} - \omega_l^2 - (1-p)\,S_{l,l}\right|\left|\omega_i^{\star 2} + (1-p)\,S_{i,i}^{\star}\right|}_{=:\beta_2} \\
&\quad + \underbrace{\frac{1}{np^2\sigma_r^{\star 4}}\left|\omega_l^{\star 2} + (1-p)\,S_{l,l}^{\star} - \omega_l^2 - (1-p)\,S_{l,l}\right|\left|\omega_i^{2} + (1-p)\,S_{i,i} - \omega_i^{\star 2} - (1-p)\,S_{i,i}^{\star}\right|}_{=:\beta_3} \\
&\quad + \underbrace{\frac{2\,(1-p)^2}{np^2\sigma_r^{\star 4}}\left|S_{i,l}^{\star 2} - S_{i,l}^2\right|}_{=:\beta_4}.
\end{aligned}$$

Denote $\Delta_i \coloneqq \omega_i^{\star 2} + (1-p)S_{i,i}^{\star} - \omega_i^2 - (1-p)S_{i,i}$. We know from (D.35) and (D.33) in Lemma 19 that for each $i \in [d]$,

$$\begin{aligned}
|\Delta_i| &\leq \left|\omega_i^{\star 2} - \omega_i^2\right| + \left\|\boldsymbol{S} - \boldsymbol{S}^{\star}\right\|_{\infty} \\
&\lesssim \sqrt{\frac{\log^2(n+d)}{np}}\omega_{\max}^{\star 2} + \zeta_{\mathsf{1st}}\frac{\sqrt{\kappa^2\mu r^2\log(n+d)}}{d} + \sqrt{\frac{\kappa^2\mu^2 r^3\log(n+d)}{nd^2}}\sigma_1^{\star 2};
\end{aligned} \tag{E.76}$$

and we have also learned from (D.36) and (D.34) that

$$\begin{aligned}
|\Delta_l| &\leq \left|\omega_l^{\star 2} - \omega_l^2\right| + \left|S_{l,l} - S_{l,l}^{\star}\right| \\
&\lesssim \left(\sqrt{\frac{\log^2(n+d)}{np}} + \theta^2\right)\omega_l^{\star 2} + \left(\theta + \sqrt{\frac{\kappa^3 r\log(n+d)}{n}}\right)\left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2^2 \\
&\quad + \theta\omega_l^{\star}\left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2 + \zeta_{\mathsf{2nd},l}\sigma_1^{\star}\left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2 + \zeta_{\mathsf{2nd},l}^2\sigma_1^{\star 2}.
\end{aligned} \tag{E.77}$$

In addition, it is straightforward to verify that for each $i \in [d]$,

$$\left|\omega_i^{\star 2} + (1-p)\,S_{i,i}^{\star}\right| \leq \omega_{\max}^2 + \frac{\mu r}{d}\sigma_1^{\star 2} \tag{E.78}$$

and

$$\left|\omega_l^{\star 2} + (1-p)\,S_{l,l}^{\star}\right| \leq \omega_l^{\star 2} + \left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2^2. \tag{E.79}$$

- Regarding β_1, we can derive

$$\beta_1 = \frac{1}{np^2\sigma_r^{\star 4}}\left|\omega_l^{\star 2} + (1-p)\,S_{l,l}^{\star}\right|\left|\Delta_i\right|$$

$$
\begin{aligned}
&\overset{\text{(i)}}{\lesssim} \frac{1}{np^2\sigma_r^{\star 4}}\left(\omega_l^{\star 2}+\left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2^2\right)\left(\sqrt{\frac{\log^2(n+d)}{np}}\omega_{\max}^{\star 2}+\zeta_{\mathsf{1st}}\frac{\sqrt{\kappa^2\mu r^2\log(n+d)}}{d}+\sqrt{\frac{\kappa^2\mu^2 r^3\log(n+d)}{nd^2}}\sigma_1^{\star 2}\right)\\
&\asymp \underbrace{\frac{\omega_l^{\star 2}\omega_{\max}^{\star 2}}{np^2\sigma_r^{\star 4}}\sqrt{\frac{\log^2(n+d)}{np}}}_{=:\beta_{1,1}}+\underbrace{\frac{\omega_l^{\star 2}}{np^2\sigma_r^{\star 4}}\zeta_{\mathsf{1st}}\frac{\sqrt{\kappa^2\mu r^2\log(n+d)}}{d}}_{=:\beta_{1,2}}+\underbrace{\frac{\omega_l^{\star 2}}{np^2\sigma_r^{\star 2}}\sqrt{\frac{\kappa^4\mu^2 r^3\log(n+d)}{nd^2}}}_{=:\beta_{1,3}}\\
&\quad+\underbrace{\frac{\omega_{\max}^{\star 2}}{np^2\sigma_r^{\star 4}}\sqrt{\frac{\log^2(n+d)}{np}}\left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2^2}_{=:\beta_{1,4}}+\underbrace{\frac{1}{np^2\sigma_r^{\star 4}}\zeta_{\mathsf{1st}}\frac{\sqrt{\kappa^2\mu r^2\log(n+d)}}{d}\left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2^2}_{=:\beta_{1,5}}\\
&\quad+\underbrace{\frac{1}{np^2\sigma_r^{\star 2}}\sqrt{\frac{\kappa^4\mu^2 r^3\log(n+d)}{nd^2}}\left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2^2}_{=:\beta_{1,6}}\overset{\text{(ii)}}{\lesssim}\delta\lambda_{\min}\left(\boldsymbol{\Sigma}_{U,l}^{\star}\right).
\end{aligned}
$$

Here, (i) follows from (E.76) and (E.79); (ii) holds since

$$
\begin{aligned}
\beta_{1,1} &\lesssim \delta\frac{\omega_l^{\star 2}\omega_{\min}^2}{np^2\sigma_1^{\star 4}}\lesssim\delta\lambda_{\min}\left(\boldsymbol{\Sigma}_{U,l}^{\star}\right),\\
\beta_{1,2}+\beta_{1,3}+\beta_{1,4} &\lesssim \delta\frac{1}{ndp^2\kappa\sigma_1^{\star 2}}\omega_l^{\star 2}\lesssim\delta\lambda_{\min}\left(\boldsymbol{\Sigma}_{U,l}^{\star}\right),\\
\beta_{1,5}+\beta_{1,6} &\lesssim \delta\frac{1}{ndp^2\kappa\sigma_1^{\star 2}}\left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2^2\lesssim\delta\lambda_{\min}\left(\boldsymbol{\Sigma}_{U,l}^{\star}\right),
\end{aligned}
$$

provided that $np \gtrsim \delta^{-2}\kappa^6\mu^2 r^2\kappa_\omega^2\log^2(n+d)$, $\zeta_{\mathsf{1st}}/\sigma_r^{\star 2}\lesssim\delta/\sqrt{\kappa^6\mu r^2\log(n+d)}$, $n\gtrsim\delta^{-2}\kappa^8\mu^2 r^3\log(n+d)$.

- When it comes to β_2, we can see that

$$
\begin{aligned}
\beta_2 &= \frac{1}{np^2\sigma_r^{\star 4}}\left|\omega_i^{\star 2}+(1-p)\,S_{i,i}^{\star}\right|\left|\Delta_l\right|\\
&\overset{\text{(i)}}{\lesssim}\underbrace{\frac{\omega_l^{\star 2}\omega_{\max}^2}{np^2\sigma_r^{\star 4}}\left(\sqrt{\frac{\log^2(n+d)}{np}}+\theta^2\right)}_{=:\beta_{2,1}}+\underbrace{\frac{\kappa\mu r}{ndp^2\sigma_r^{\star 2}}\omega_l^{\star 2}\left(\sqrt{\frac{\log^2(n+d)}{np}}+\theta^2\right)}_{=:\beta_{2,2}}\\
&\quad+\underbrace{\frac{1}{np^2\sigma_r^{\star 4}}\omega_{\max}^2\left(\theta+\sqrt{\frac{\kappa^3 r\log(n+d)}{n}}\right)\left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2^2}_{=:\beta_{2,3}}+\underbrace{\frac{\kappa\mu r}{ndp^2\sigma_r^{\star 2}}\left(\theta+\sqrt{\frac{\kappa^3 r\log(n+d)}{n}}\right)\left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2^2}_{=:\beta_{2,4}}\\
&\quad+\underbrace{\frac{1}{np^2\sigma_r^{\star 4}}\omega_{\max}^2\omega_l^{\star}\theta\left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2}_{\beta_{2,5}}+\frac{\sqrt{\kappa}}{np^2\sigma_r^{\star 3}}\omega_{\max}^2\zeta_{\mathsf{2nd},l}\left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2+\frac{\kappa}{np^2\sigma_r^{\star 2}}\omega_{\max}^2\zeta_{\mathsf{2nd},l}^2\\
&\quad+\underbrace{\frac{\kappa\mu r}{ndp^2\sigma_r^{\star 2}}\theta\omega_l^{\star}\left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2}_{=:\beta_{2,6}}+\frac{\kappa^{3/2}\mu r}{ndp^2\sigma_r^{\star}}\zeta_{\mathsf{2nd},l}\left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2+\frac{\kappa^2\mu r}{ndp^2}\zeta_{\mathsf{2nd},l}^2\\
&\overset{\text{(ii)}}{\lesssim}\delta\lambda_{\min}\left(\boldsymbol{\Sigma}_{U,l}^{\star}\right)+\frac{\sqrt{\kappa}}{np^2\sigma_r^{\star 3}}\omega_{\max}^2\zeta_{\mathsf{2nd},l}\left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2+\frac{\kappa}{np^2\sigma_r^{\star 2}}\omega_{\max}^2\zeta_{\mathsf{2nd},l}^2+\frac{\kappa^{3/2}\mu r}{ndp^2\sigma_r^{\star}}\zeta_{\mathsf{2nd},l}\left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2+\frac{\kappa^2\mu r}{ndp^2}\zeta_{\mathsf{2nd},l}^2.
\end{aligned}
$$

Here, (i) follows from (E.77) and (E.78); (ii) holds since

$$
\beta_{2,1}\lesssim\delta\frac{\omega_l^{\star 2}\omega_{\min}^2}{np^2\sigma_1^{\star 4}}\lesssim\delta\lambda_{\min}\left(\boldsymbol{\Sigma}_{U,l}^{\star}\right),
$$

$$
\begin{aligned}
&\beta_{2,2} \lesssim \delta \frac{1}{ndp^2\kappa\sigma_1^{\star 2}}\omega_l^{\star 2} \lesssim \delta\lambda_{\min}\left(\mathbf{\Sigma}_{U,l}^{\star}\right), \\
&\beta_{2,3} \lesssim \delta \frac{1}{np^2\sigma_1^{\star 4}}\omega_{\min}^2 \left\|\boldsymbol{U}_{l,\cdot}^{\star}\mathbf{\Sigma}^{\star}\right\|_2^2 \lesssim \delta\lambda_{\min}\left(\mathbf{\Sigma}_{U,l}^{\star}\right), \\
&\beta_{2,4} \lesssim \delta \frac{1}{ndp^2\kappa\sigma_1^{\star 2}} \left\|\boldsymbol{U}_{l,\cdot}^{\star}\mathbf{\Sigma}^{\star}\right\|_2^2 \lesssim \delta\lambda_{\min}\left(\mathbf{\Sigma}_{U,l}^{\star}\right),, \\
&\beta_{2,5} \lesssim \delta \frac{1}{np^2\sigma_1^{\star 4}}\omega_{\min}^2\omega_l^{\star} \left\|\boldsymbol{U}_{l,\cdot}^{\star}\mathbf{\Sigma}^{\star}\right\|_2 \lesssim \delta \frac{1}{np^2\sigma_1^{\star 4}}\omega_{\min}^2 \left\|\boldsymbol{U}_{l,\cdot}^{\star}\mathbf{\Sigma}^{\star}\right\|_2^2 + \delta\frac{\omega_l^{\star 2}\omega_{\min}^2}{np^2\sigma_1^{\star 4}} \lesssim \delta\lambda_{\min}\left(\mathbf{\Sigma}_{U,l}^{\star}\right), \\
&\beta_{2,6} \lesssim \delta \frac{1}{ndp^2\kappa\sigma_1^{\star 2}} \left\|\boldsymbol{U}_{l,\cdot}^{\star}\mathbf{\Sigma}^{\star}\right\|_2 \omega_l^{\star} \lesssim \delta \frac{1}{ndp^2\kappa\sigma_1^{\star 2}} \left\|\boldsymbol{U}_{l,\cdot}^{\star}\mathbf{\Sigma}^{\star}\right\|_2^2 + \delta \frac{1}{ndp^2\kappa\sigma_1^{\star 2}}\omega_l^{\star 2} \lesssim \delta\lambda_{\min}\left(\mathbf{\Sigma}_{U,l}^{\star}\right),
\end{aligned}
$$

provided that $np \gtrsim \delta^{-2}\kappa^6\mu^2 r^2 \log^2(n+d)$, $\theta \lesssim \delta/(\kappa^3\mu r\kappa_\omega)$, $n \gtrsim \delta^{-2}\kappa^9\mu^2 r^3 \log(n+d)$, $n \gtrsim \delta^{-2}\kappa^7 r\kappa_\omega^2 \log(n+d)$.

- With regards to β_3, we notice that

$$
|\Delta_i| \lesssim \sqrt{\frac{\log^2(n+d)}{np}}\omega_{\max}^{\star 2} + \zeta_{\mathsf{1st}}\frac{\sqrt{\kappa^2\mu r^2\log(n+d)}}{d} + \sqrt{\frac{\kappa^2\mu^2 r^3\log(n+d)}{nd^2}}\sigma_1^{\star 2} \lesssim \omega_{\max}^2 + \frac{\mu r}{d}\sigma_1^{\star 2}
$$

provided that $np \gtrsim \log^2(n+d)$, $n \gtrsim \kappa^2 r\log(n+d)$ and $\zeta_{\mathsf{1st}}/\sigma_r^{\star 2} \lesssim 1/\sqrt{\log(n+d)}$. As a consequence, we can upper bound

$$
\beta_3 = \frac{1}{np^2\sigma_r^{\star 4}}|\Delta_i||\Delta_l| \lesssim \frac{1}{np^2\sigma_r^{\star 4}}\left|\omega_i^{\star 2} + (1-p)S_{i,i}^{\star}\right||\Delta_l|.
$$

This immediately suggests that β_3 satisfies the same upper bound we derive for β_2.

- Regarding β_4, it is seen that

$$
\begin{aligned}
\beta_4 &\lesssim \frac{1}{np^2\sigma_r^{\star 4}}\left|S_{i,l}^{\star 2} - S_{i,l}^2\right| \lesssim \frac{1}{np^2\sigma_r^{\star 4}}\left|S_{i,l}^{\star} - S_{i,l}\right|\left|S_{i,l}^{\star} + S_{i,l}\right| \\
&\lesssim \underbrace{\frac{1}{np^2\sigma_r^{\star 4}}\left|S_{i,l}^{\star} - S_{i,l}\right|S_{i,l}^{\star}}_{=:\beta_{4,1}} + \underbrace{\frac{1}{np^2\sigma_r^{\star 4}}\left|S_{i,l}^{\star} - S_{i,l}\right|^2}_{=:\beta_{4,2}}.
\end{aligned}
$$

Recall from (D.34) in Lemma 19 that

$$
\begin{aligned}
\left|S_{i,l} - S_{i,l}^{\star}\right| &\lesssim \left(\theta + \sqrt{\frac{\kappa^3 r\log(n+d)}{n}}\right)\left\|\boldsymbol{U}_{i,\cdot}^{\star}\mathbf{\Sigma}^{\star}\right\|_2\left\|\boldsymbol{U}_{l,\cdot}^{\star}\mathbf{\Sigma}^{\star}\right\|_2 + \theta\left(\omega_i^{\star}\left\|\boldsymbol{U}_{l,\cdot}^{\star}\mathbf{\Sigma}^{\star}\right\|_2 + \omega_l^{\star}\left\|\boldsymbol{U}_{i,\cdot}^{\star}\mathbf{\Sigma}^{\star}\right\|_2\right) \\
&\quad + \sigma_1^{\star}\left(\zeta_{\mathsf{2nd},i}\left\|\boldsymbol{U}_{l,\cdot}^{\star}\mathbf{\Sigma}^{\star}\right\|_2 + \zeta_{\mathsf{2nd},l}\left\|\boldsymbol{U}_{i,\cdot}^{\star}\mathbf{\Sigma}^{\star}\right\|_2\right) + \theta^2\omega_{\max}\omega_l^{\star} + \zeta_{\mathsf{2nd},i}\zeta_{\mathsf{2nd},l}\sigma_1^{\star 2} \\
&\lesssim \left(\theta + \sqrt{\frac{\kappa^3 r\log(n+d)}{n}}\right)\sqrt{\frac{\mu r}{d}}\sigma_1^{\star}\left\|\boldsymbol{U}_{l,\cdot}^{\star}\mathbf{\Sigma}^{\star}\right\|_2 + \theta^2\omega_{\max}\omega_l^{\star} + \theta\omega_l^{\star}\sqrt{\frac{\mu r}{d}}\sigma_1^{\star} \\
&\quad + \theta\omega_{\max}\left\|\boldsymbol{U}_{l,\cdot}^{\star}\mathbf{\Sigma}^{\star}\right\|_2 + \sigma_1^{\star}\left\|\boldsymbol{U}_{l,\cdot}^{\star}\mathbf{\Sigma}^{\star}\right\|_2\max_{i\in[d]}\zeta_{\mathsf{2nd},i} + \zeta_{\mathsf{2nd},l}\sqrt{\frac{\mu r}{d}}\sigma_1^{\star 2}.
\end{aligned}
\tag{E.80}
$$

provided that $\max_{i\in[d]}\zeta_{\mathsf{2nd},i}\sqrt{d} \le \sqrt{\mu r}$. The first term $\beta_{4,1}$ can be upper bounded by

$$
\begin{aligned}
\beta_{4,1} &\lesssim \frac{1}{np^2\sigma_r^{\star 4}}\left|S_{i,l}^{\star} - S_{i,l}\right|\left\|\boldsymbol{U}_{i,\cdot}^{\star}\mathbf{\Sigma}^{\star}\right\|_2\left\|\boldsymbol{U}_{l,\cdot}^{\star}\mathbf{\Sigma}^{\star}\right\|_2 \lesssim \frac{1}{np^2\sigma_r^{\star 4}}\left|S_{i,l}^{\star} - S_{i,l}\right|\sqrt{\frac{\mu r}{d}}\sigma_1^{\star}\left\|\boldsymbol{U}_{l,\cdot}^{\star}\mathbf{\Sigma}^{\star}\right\|_2 \\
&\overset{\text{(i)}}{\lesssim} \underbrace{\frac{\kappa\mu r}{ndp^2\sigma_r^{\star 2}}\left(\theta + \sqrt{\frac{\kappa^3 r\log(n+d)}{n}}\right)\left\|\boldsymbol{U}_{l,\cdot}^{\star}\mathbf{\Sigma}^{\star}\right\|_2^2}_{=:\beta_{4,1,1}} + \underbrace{\frac{1}{np^2\sigma_r^{\star 4}}\theta^2\omega_{\max}\omega_l^{\star}\sqrt{\frac{\mu r}{d}}\sigma_1^{\star}\left\|\boldsymbol{U}_{l,\cdot}^{\star}\mathbf{\Sigma}^{\star}\right\|_2}_{=:\beta_{4,1,2}}
\end{aligned}
$$

$$
\begin{aligned}
&+ \underbrace{\frac{\kappa\mu r}{ndp^2\sigma_r^{\star 2}}\theta\omega_l^\star \left\|\boldsymbol{U}_{l,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2}_{=:\beta_{4,1,3}} + \underbrace{\frac{\theta\omega_{\max}}{np^2\sigma_r^{\star 3}}\sqrt{\frac{\kappa\mu r}{d}}\left\|\boldsymbol{U}_{l,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2^2}_{=:\beta_{4,1,4}} \\
&+ \underbrace{\frac{1}{np^2\sigma_r^{\star 2}}\sqrt{\frac{\kappa^2\mu r}{d}}\left\|\boldsymbol{U}_{l,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2^2 \max_{i\in[d]}\zeta_{\mathsf{2nd},i}}_{=:\beta_{4,1,5}} + \frac{\kappa^{3/2}\mu r}{ndp^2\sigma_r^\star}\zeta_{\mathsf{2nd},l}\left\|\boldsymbol{U}_{l,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2 \\
&\overset{\text{(ii)}}{\lesssim} \delta\lambda_{\min}\left(\boldsymbol{\Sigma}_{U,l}^\star\right) + \frac{\kappa^{3/2}\mu r}{ndp^2\sigma_r^\star}\zeta_{\mathsf{2nd},l}\left\|\boldsymbol{U}_{l,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2.
\end{aligned}
$$

Here, (i) follows from (E.80), and (ii) holds since

$$
\begin{aligned}
\beta_{4,1,1} + \beta_{4,1,5} &\lesssim \delta\frac{1}{ndp^2\kappa\sigma_1^{\star 2}}\left\|\boldsymbol{U}_{l,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2^2 \lesssim \delta\lambda_{\min}\left(\boldsymbol{\Sigma}_{U,l}^\star\right), \\
\beta_{4,1,2} &\lesssim \delta\frac{1}{ndp^2\kappa\sigma_1^{\star 2}}\omega_l^{\star 2} \lesssim \delta\lambda_{\min}\left(\boldsymbol{\Sigma}_{U,l}^\star\right), \\
\beta_{4,1,3} + \beta_{4,1,4} &\lesssim \frac{\delta}{ndp^2\kappa\sigma_1^{\star 2}}\left\|\boldsymbol{U}_{l,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2\omega_l^\star \lesssim \frac{\delta}{ndp^2\kappa\sigma_1^{\star 2}}\left\|\boldsymbol{U}_{l,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2^2 + \frac{\delta}{ndp^2\kappa\sigma_1^{\star 2}}\omega_l^{\star 2} \lesssim \delta\lambda_{\min}\left(\boldsymbol{\Sigma}_{U,l}^\star\right),
\end{aligned}
$$

provided that $\theta \lesssim \delta/(\kappa^3\mu r\sqrt{\kappa_\omega})$, $n \gtrsim \delta^{-2}\kappa^9\mu^2 r^3\log(n+d)$ and $\max_{i\in[d]}\zeta_{\mathsf{2nd},i}\sqrt{d} \lesssim \delta/\sqrt{\kappa^6\mu r}$. The second term $\beta_{4,2}$ can be controlled as follows

$$
\begin{aligned}
\beta_{4,2} &\overset{\text{(i)}}{\lesssim} \underbrace{\frac{\kappa\mu r}{ndp^2\sigma_r^{\star 2}}\left(\theta + \sqrt{\frac{\kappa^3 r\log(n+d)}{n}}\right)^2\left\|\boldsymbol{U}_{l,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2^2}_{=:\beta_{4,2,1}} + \underbrace{\frac{1}{np^2\sigma_r^{\star 4}}\theta^4\omega_{\max}^2\omega_l^{\star 2}}_{=:\beta_{4,2,2}} + \underbrace{\frac{\kappa\mu r}{ndp^2\sigma_r^{\star 2}}\theta^2\omega_l^{\star 2}}_{=:\beta_{4,2,3}} \\
&+ \underbrace{\frac{1}{np^2\sigma_r^{\star 4}}\theta^2\omega_{\max}^2\left\|\boldsymbol{U}_{l,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2^2}_{=:\beta_{4,2,4}} + \underbrace{\frac{\kappa}{np^2\sigma_r^{\star 2}}\left\|\boldsymbol{U}_{l,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2^2\max_{i\in[d]}\zeta_{\mathsf{2nd},i}^2}_{=:\beta_{4,2,5}} + \frac{\kappa^2\mu r}{ndp^2}\zeta_{\mathsf{2nd},l}^2 \\
&\overset{\text{(ii)}}{\lesssim} \delta\lambda_{\min}\left(\boldsymbol{\Sigma}_{U,l}^\star\right) + \frac{\kappa^2\mu r}{ndp^2}\zeta_{\mathsf{2nd},l}^2.
\end{aligned}
$$

Here, (i) follows from (E.80), and (ii) holds since

$$
\begin{aligned}
\beta_{4,2,1} + \beta_{4,2,5} &\lesssim \delta\frac{1}{ndp^2\kappa\sigma_1^{\star 2}}\left\|\boldsymbol{U}_{l,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2^2 \lesssim \delta\lambda_{\min}\left(\boldsymbol{\Sigma}_{U,l}^\star\right), \\
\beta_{4,2,2} &\lesssim \delta\frac{\omega_l^{\star 2}\omega_{\min}^2}{np^2\sigma_1^{\star 4}} \lesssim \delta\lambda_{\min}\left(\boldsymbol{\Sigma}_{U,l}^\star\right), \\
\beta_{4,2,3} &\lesssim \delta\frac{1}{ndp^2\kappa\sigma_1^{\star 2}}\omega_l^{\star 2} \lesssim \delta\frac{1}{ndp^2\kappa\sigma_1^{\star 2}}\left\|\boldsymbol{U}_{l,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2^2 + \delta\frac{1}{4ndp^2\kappa\sigma_1^{\star 2}}\omega_l^{\star 2} \lesssim \delta\lambda_{\min}\left(\boldsymbol{\Sigma}_{U,l}^\star\right), \\
\beta_{4,2,4} &\lesssim \delta\frac{1}{np^2\sigma_1^{\star 4}}\omega_{\min}^2\left\|\boldsymbol{U}_{l,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2^2 \lesssim \delta\lambda_{\min}\left(\boldsymbol{\Sigma}_{U,l}^\star\right),
\end{aligned}
$$

provided that $\theta \lesssim \sqrt{\delta/(\kappa^3\mu r\kappa_\omega)}$, $n \gtrsim \delta^{-1}\kappa^6\mu r^2\log(n+d)$ and $\max_{i\in[d]}\zeta_{\mathsf{2nd},i}\sqrt{d} \lesssim \sqrt{\delta/\kappa^3}$. Combine the bounds on $\beta_{4,1}$ and $\beta_{4,2}$ to reach

$$
\beta_4 \le \beta_{4,1} + \beta_{4,2} \lesssim \delta\lambda_{\min}\left(\boldsymbol{\Sigma}_{U,l}^\star\right) + \frac{\kappa^{3/2}\mu r}{ndp^2\sigma_r^\star}\zeta_{\mathsf{2nd},l}\left\|\boldsymbol{U}_{l,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2 + \frac{\kappa^2\mu r}{ndp^2}\zeta_{\mathsf{2nd},l}^2.
$$

Taking together the bounds on β_1, β_2, β_3 and β_4 yields

$$
\frac{1}{\sigma_r^{\star 4}}\left|d_{l,i}^\star - d_{l,i}\right| \le \beta_1 + \beta_2 + \beta_3 + \beta_4
$$

$$
\begin{aligned}
&\lesssim \delta\lambda_{\min}\left(\boldsymbol{\Sigma}_{U,l}^{\star}\right)+\frac{\sqrt{\kappa}}{np^{2}\sigma_{r}^{\star 3}}\omega_{\max}^{2}\zeta_{\mathsf{2nd},l}\left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_{2}+\frac{\kappa}{np^{2}\sigma_{r}^{\star 2}}\omega_{\max}^{2}\zeta_{\mathsf{2nd},l}^{2}\\
&\quad+\frac{\kappa^{3/2}\mu r}{ndp^{2}\sigma_{r}^{\star}}\zeta_{\mathsf{2nd},l}\left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_{2}+\frac{\kappa^{2}\mu r}{ndp^{2}}\zeta_{\mathsf{2nd},l}^{2},
\end{aligned}
$$

with the proviso that $np \gtrsim \delta^{-2}\kappa^{6}\mu^{2}r^{2}\kappa_{\omega}^{2}\log^{2}(n+d)$, $\zeta_{\mathsf{1st}}/\sigma_{r}^{\star 2} \lesssim \delta/\sqrt{\kappa^{6}\mu r^{2}\log(n+d)}$, $\theta \lesssim \delta/(\kappa^{3}\mu r\kappa_{\omega})$, $\max_{i\in[d]}\zeta_{\mathsf{2nd},i}\sqrt{d} \lesssim \delta/\sqrt{\kappa^{6}\mu r}$, $n \gtrsim \delta^{-2}\kappa^{9}\mu^{2}r^{3}\log(n+d)$ and $n \gtrsim \kappa^{7}r\kappa_{\omega}^{2}\log(n+d)$. Given that this holds for all $i\in[d]$, it follows that

$$
\begin{aligned}
\beta &= \frac{1}{\sigma_{r}^{\star 4}}\max_{1\leq i\leq d}\left|d_{l,i}^{\star}-d_{l,i}\right|\\
&\lesssim \delta\lambda_{\min}\left(\boldsymbol{\Sigma}_{U,l}^{\star}\right)+\frac{\sqrt{\kappa}}{np^{2}\sigma_{r}^{\star 3}}\omega_{\max}^{2}\zeta_{\mathsf{2nd},l}\left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_{2}+\frac{\kappa}{np^{2}\sigma_{r}^{\star 2}}\omega_{\max}^{2}\zeta_{\mathsf{2nd},l}^{2}\\
&\quad+\frac{\kappa^{3/2}\mu r}{ndp^{2}\sigma_{r}^{\star}}\zeta_{\mathsf{2nd},l}\left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_{2}+\frac{\kappa^{2}\mu r}{ndp^{2}}\zeta_{\mathsf{2nd},l}^{2}.
\end{aligned}
$$

Step 6: putting all pieces together. From the above steps, we can demonstrate that

$$
\begin{aligned}
\left\|\boldsymbol{R}\boldsymbol{\Sigma}_{U,l}^{\star}\boldsymbol{R}^{\top}-\boldsymbol{\Sigma}_{U,l}\right\| &\leq \sum_{i=1}^{4}\left\|\boldsymbol{A}_{i}-\boldsymbol{B}_{i}\right\| \leq \sum_{i=1}^{3}\left\|\boldsymbol{A}_{i}-\boldsymbol{B}_{i}\right\|+\alpha_{4}+\beta\\
&\lesssim \delta\lambda_{\min}\left(\boldsymbol{\Sigma}_{U,l}^{\star}\right)+\underbrace{\frac{\sqrt{\kappa}}{np\sigma_{r}^{\star}}\zeta_{\mathsf{2nd},l}\left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_{2}}_{=:\gamma_{1}}+\underbrace{\frac{\kappa}{np}\zeta_{\mathsf{2nd},l}^{2}}_{=:\gamma_{2}}+\underbrace{\frac{1}{np}\zeta_{\mathsf{2nd},l}\frac{\theta}{\sqrt{\kappa}\sigma_{r}^{\star}}\omega_{l}^{\star}}_{=:\gamma_{3}}\\
&\quad+\underbrace{\frac{\kappa^{3/2}\mu r}{ndp^{2}\sigma_{r}^{\star}}\zeta_{\mathsf{2nd},l}\left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_{2}}_{=:\gamma_{4}}+\underbrace{\frac{\kappa^{2}\mu r}{ndp^{2}}\zeta_{\mathsf{2nd},l}^{2}}_{=:\gamma_{5}}\\
&\quad+\underbrace{\frac{\sqrt{\kappa}}{np^{2}\sigma_{r}^{\star 3}}\omega_{\max}^{2}\zeta_{\mathsf{2nd},l}\left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_{2}}_{=:\gamma_{6}}+\underbrace{\frac{\kappa}{np^{2}\sigma_{r}^{\star 2}}\omega_{\max}^{2}\zeta_{\mathsf{2nd},l}^{2}}_{=:\gamma_{7}}.
\end{aligned}
$$

Note that under the assumption of Lemma 17, we have shown in Appendix E.2.3 that

$$
\zeta_{\mathsf{2nd},l}\lambda_{\min}^{-1/2}\left(\boldsymbol{\Sigma}_{U,l}^{\star}\right) \lesssim \frac{1}{r^{1/4}\log^{1/2}(n+d)} \ll 1. \tag{E.81}
$$

In addition, it follows from Lemma 16 that

$$
\lambda_{\min}^{1/2}\left(\boldsymbol{\Sigma}_{U,l}^{\star}\right) \gtrsim \frac{1}{\sqrt{np}\sigma_{1}^{\star}}\left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_{2}+\frac{\omega_{l}^{\star}}{\sqrt{np}\sigma_{1}^{\star}}+\frac{1}{\sqrt{ndp^{2}\kappa}\sigma_{1}^{\star}}\left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_{2}. \tag{E.82}
$$

As a consequence, we can derive the following upper bounds

$$
\begin{aligned}
\gamma_{1} &\lesssim \frac{\sqrt{\kappa}}{np\sigma_{r}^{\star}}\lambda_{\min}^{1/2}\left(\boldsymbol{\Sigma}_{U,l}^{\star}\right)\left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_{2} \lesssim \frac{\kappa}{\sqrt{np}}\lambda_{\min}\left(\boldsymbol{\Sigma}_{U,l}^{\star}\right) \lesssim \delta\lambda_{\min}\left(\boldsymbol{\Sigma}_{U,l}^{\star}\right),\\
\gamma_{2} &\lesssim \frac{\kappa}{np}\lambda_{\min}\left(\boldsymbol{\Sigma}_{U,l}^{\star}\right) \lesssim \delta\lambda_{\min}\left(\boldsymbol{\Sigma}_{U,l}^{\star}\right),\\
\gamma_{3} &\lesssim \frac{1}{np}\lambda_{\min}^{1/2}\left(\boldsymbol{\Sigma}_{U,l}^{\star}\right)\frac{\theta}{\sqrt{\kappa}\sigma_{r}^{\star}}\omega_{l}^{\star} \lesssim \frac{\theta}{\sqrt{np}}\lambda_{\min}\left(\boldsymbol{\Sigma}_{U,l}^{\star}\right) \lesssim \delta\lambda_{\min}\left(\boldsymbol{\Sigma}_{U,l}^{\star}\right),\\
\gamma_{4} &\lesssim \frac{\kappa^{3/2}\mu r}{ndp^{2}\sigma_{r}^{\star}}\lambda_{\min}^{1/2}\left(\boldsymbol{\Sigma}_{U,l}^{\star}\right)\left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_{2} \lesssim \frac{\kappa^{5/2}\mu r}{\sqrt{ndp^{2}}}\lambda_{\min}\left(\boldsymbol{\Sigma}_{U,l}^{\star}\right) \lesssim \delta\lambda_{\min}\left(\boldsymbol{\Sigma}_{U,l}^{\star}\right),\\
\gamma_{5} &\lesssim \frac{\kappa^{2}\mu r}{ndp^{2}}\lambda_{\min}\left(\boldsymbol{\Sigma}_{U,l}^{\star}\right) \lesssim \delta\lambda_{\min}\left(\boldsymbol{\Sigma}_{U,l}^{\star}\right),
\end{aligned}
$$

provided that $\theta \lesssim 1$, $np \gtrsim \delta^{-2}\kappa^2$ and $ndp^2 \gtrsim \delta^{-2}\kappa^5\mu^2 r^2$. In addition, we also have

$$\begin{aligned}
\gamma_6 &\overset{\text{(i)}}{\lesssim} \frac{\kappa}{n^2 p^4 \sigma_r^{\star 6}\delta}\omega_{\max}^4 \left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2^2 + \delta\zeta_{\mathsf{2nd},l}^2 \lesssim \frac{\kappa^3\kappa_\omega}{np^2\sigma_r^{\star 2}\delta}\omega_{\max}^2 \cdot \frac{1}{np^2\sigma_1^{\star 4}}\omega_{\min}^2 \left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2^2 + \delta\zeta_{\mathsf{2nd},l}^2 \\
&\lesssim \frac{\kappa^3\kappa_\omega}{\delta} \cdot \frac{1}{\sqrt{ndp^2}} \cdot \frac{\omega_{\max}^2}{p\sigma_r^{\star 2}}\sqrt{\frac{d}{n}} \cdot \lambda_{\min}\left(\boldsymbol{\Sigma}_{U,l}^{\star}\right) + \delta\lambda_{\min}\left(\boldsymbol{\Sigma}_{U,l}^{\star}\right) \overset{\text{(ii)}}{\lesssim} \delta\lambda_{\min}\left(\boldsymbol{\Sigma}_{U,l}^{\star}\right)
\end{aligned}$$

as well as

$$\gamma_7 \lesssim \frac{\kappa}{np^2\sigma_r^{\star 2}}\omega_{\max}^2\lambda_{\min}\left(\boldsymbol{\Sigma}_{U,l}^{\star}\right) \lesssim \kappa\frac{1}{\sqrt{ndp^2}} \cdot \frac{\omega_{\max}^2}{p\sigma_r^{\star 2}}\sqrt{\frac{d}{n}} \cdot \lambda_{\min}\left(\boldsymbol{\Sigma}_{U,l}^{\star}\right) \lesssim \delta\lambda_{\min}\left(\boldsymbol{\Sigma}_{U,l}^{\star}\right),$$

where (i) utilizes the AM-GM inequality, while (ii) and (iii) hold when $ndp^2 \gtrsim \delta^{-2}\kappa^6\kappa_\omega^2$ and

$$\frac{\omega_{\max}^2}{p\sigma_r^{\star 2}}\sqrt{\frac{d}{n}} \lesssim \delta.$$

These allow us to conclude that

$$\left\|\boldsymbol{R}\boldsymbol{\Sigma}_{U,l}^{\star}\boldsymbol{R}^{\top} - \boldsymbol{\Sigma}_{U,l}\right\| \lesssim \delta\lambda_{\min}\left(\boldsymbol{\Sigma}_{U,l}^{\star}\right),$$

as long as the following assumptions hold: $np \gtrsim \delta^{-2}\kappa^6\mu^2 r^2\kappa_\omega^2\log^2(n+d)$, $ndp^2 \gtrsim \delta^{-2}\kappa^6\kappa_\omega^2$, $n \gtrsim \delta^{-2}\kappa^9\mu^2 r^3\log(n+d)$, $n \gtrsim \kappa^7 r\kappa_\omega^2\log(n+d)$, $d \gtrsim \kappa^3\mu r\log(n+d)$, $\zeta_{\mathsf{1st}}/\sigma_r^{\star 2} \lesssim \delta/(\kappa^3\mu r\kappa_\omega\sqrt{\log(n+d)})$, $\theta \lesssim \delta/(\kappa^3\mu r\kappa_\omega)$, $\max_{i\in[d]}\zeta_{\mathsf{2nd},i}\sqrt{d} \lesssim \delta/\sqrt{\kappa^6\mu r}$ and

$$\frac{\omega_{\max}^2}{p\sigma_r^{\star 2}}\sqrt{\frac{d}{n}} \lesssim \delta.$$

In what follows, we take a closer look at the last three assumptions.

- In view of (E.17), it is readily seen that $\zeta_{\mathsf{1st}}/\sigma_r^{\star 2} \lesssim \delta/(\kappa^3\mu r\kappa_\omega\sqrt{\log(n+d)})$ can be guaranteed by

  $$ndp^2 \gtrsim \delta^{-2}\kappa^8\mu^4 r^4\kappa_\omega^2\log^5(n+d), \qquad np \gtrsim \delta^{-2}\kappa^8\mu^3 r^3\kappa_\omega^2\log^3(n+d),$$

  and

  $$\frac{\omega_{\max}^2}{p\sigma_r^{\star 2}}\sqrt{\frac{d}{n}} \lesssim \frac{\delta}{\kappa^3\mu r\kappa_\omega\log^{3/2}(n+d)}, \qquad \frac{\omega_{\max}}{\sigma_r^{\star}}\sqrt{\frac{d}{np}} \lesssim \frac{\delta}{\kappa^{7/2}\mu r\kappa_\omega\log(n+d)}.$$

- By virtue of (E.66), it is straightforward to see that $\theta \lesssim \delta/(\kappa^3\mu r\kappa_\omega)$ can be guaranteed by $\zeta_{\mathsf{1st}}/\sigma_r^{\star 2} \lesssim \delta/(\kappa^3\mu r\kappa_\omega)$ — the latter is already guaranteed by $\zeta_{\mathsf{1st}}/\sigma_r^{\star 2} \lesssim \delta/(\kappa^3\mu r\kappa_\omega\sqrt{\log(n+d)})$.

- In addition, we know that $\max_{i\in[d]}\zeta_{\mathsf{2nd},i}\sqrt{d} \lesssim \delta/(\sqrt{\kappa^6\mu r})$ is equivalent to

  $$\frac{\zeta_{\mathsf{1st}}}{\sigma_r^{\star 2}}\frac{\sqrt{\kappa^3\mu^2 r^2\log(n+d)}}{\sqrt{d}} + \frac{\zeta_{\mathsf{1st}}^2}{\sigma_r^{\star 4}}\sqrt{\frac{\kappa^3\mu r\log(n+d)}{}} \lesssim \frac{\delta}{\sqrt{\kappa^6\mu r}},$$

  which can be guaranteed by $\zeta_{\mathsf{1st}}/\sigma_r^{\star 2} \lesssim \delta/(\kappa^3\mu r\kappa_\omega\sqrt{\log(n+d)})$ as long as $d \gtrsim \kappa^3\mu r$.

To summarize, we can conclude that the required assumptions for the above results to hold are: $n \gtrsim \delta^{-2}\kappa^9\mu^2 r^3\log(n+d)$, $n \gtrsim \kappa^7 r\kappa_\omega^2\log(n+d)$, $d \gtrsim \kappa^3\mu r\log(n+d)$,

$$ndp^2 \gtrsim \delta^{-2}\kappa^8\mu^4 r^4\kappa_\omega^2\log^5(n+d), \qquad np \gtrsim \delta^{-2}\kappa^8\mu^3 r^3\kappa_\omega^2\log^3(n+d),$$

and

$$\frac{\omega_{\max}^2}{p\sigma_r^{\star 2}}\sqrt{\frac{d}{n}} \lesssim \frac{\delta}{\kappa^3\mu r\kappa_\omega\log^{3/2}(n+d)}, \qquad \frac{\omega_{\max}}{\sigma_r^{\star}}\sqrt{\frac{d}{np}} \lesssim \frac{\delta}{\kappa^{7/2}\mu r\kappa_\omega\log(n+d)}.$$

E.3.4 Proof of Lemma 21

Recall the definition of the Euclidean ball $\mathcal{B}_{1-\alpha}$ in Algorithm 3. It is easily seen that

$$\begin{aligned}
\left(\boldsymbol{U}\boldsymbol{R}-\boldsymbol{U}^{\star}\right)_{l,\cdot}\left(\boldsymbol{\Sigma}_{U,l}^{\star}\right)^{-1/2}\in\mathcal{B}_{1-\alpha} \quad&\Longleftrightarrow\quad \left(\boldsymbol{U}\boldsymbol{R}-\boldsymbol{U}^{\star}\right)_{l,\cdot}\left(\boldsymbol{\Sigma}_{U,l}^{\star}\right)^{-1/2}\boldsymbol{R}\in\mathcal{B}_{1-\alpha}\\
&\Longleftrightarrow\quad \left(\boldsymbol{U}-\boldsymbol{U}^{\star}\boldsymbol{R}^{\top}\right)_{l,\cdot}\boldsymbol{R}\left(\boldsymbol{\Sigma}_{U,l}^{\star}\right)^{-1/2}\boldsymbol{R}^{\top}\in\mathcal{B}_{1-\alpha},
\end{aligned}$$

where the last line comes from the rotational invariance of $\mathcal{B}_{1-\alpha}$. From the definition of $\mathsf{CR}_{U,l}^{1-\alpha}$ in Algorithm 3, we also know that

$$\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{R}^{\top}\in\mathsf{CR}_{U,l}^{1-\alpha}\quad\Longleftrightarrow\quad\left(\boldsymbol{U}-\boldsymbol{U}^{\star}\boldsymbol{R}^{\top}\right)_{l,\cdot}\boldsymbol{\Sigma}_{U,l}^{-1/2}\in\mathcal{B}_{1-\alpha}.$$

Let us define

$$\boldsymbol{\Delta}\coloneqq\left(\boldsymbol{U}-\boldsymbol{U}^{\star}\boldsymbol{R}^{\top}\right)_{l,\cdot}\boldsymbol{R}\left(\boldsymbol{\Sigma}_{U,l}^{\star}\right)^{-1/2}\boldsymbol{R}^{\top}-\left(\boldsymbol{U}-\boldsymbol{U}^{\star}\boldsymbol{R}^{\top}\right)_{l,\cdot}\boldsymbol{\Sigma}_{U,l}^{-1/2},$$

then it is straightforward to check that with probability exceeding $1-O((n+d)^{-10})$

$$\begin{aligned}
\|\boldsymbol{\Delta}\|_2&\leq\left\|\left(\boldsymbol{U}-\boldsymbol{U}^{\star}\boldsymbol{R}^{\top}\right)_{l,\cdot}\right\|_2\left\|\boldsymbol{R}\left(\boldsymbol{\Sigma}_{U,l}^{\star}\right)^{-1/2}\boldsymbol{R}^{\top}-\boldsymbol{\Sigma}_{U,l}^{-1/2}\right\|\\
&=\left\|\left(\boldsymbol{U}-\boldsymbol{U}^{\star}\boldsymbol{R}^{\top}\right)_{l,\cdot}\right\|_2\left\|\boldsymbol{R}\left(\boldsymbol{\Sigma}_{U,l}^{\star}\right)^{-1/2}\boldsymbol{R}^{\top}\left(\boldsymbol{\Sigma}_{U,l}^{1/2}-\boldsymbol{R}\left(\boldsymbol{\Sigma}_{U,l}^{\star}\right)^{1/2}\boldsymbol{R}^{\top}\right)\boldsymbol{\Sigma}_{U,l}^{-1/2}\right\|\\
&\leq\left\|\left(\boldsymbol{U}-\boldsymbol{U}^{\star}\boldsymbol{R}^{\top}\right)_{l,\cdot}\right\|_2\left\|\boldsymbol{R}\left(\boldsymbol{\Sigma}_{U,l}^{\star}\right)^{-1/2}\boldsymbol{R}^{\top}\right\|\left\|\boldsymbol{R}\left(\boldsymbol{\Sigma}_{U,l}^{\star}\right)^{1/2}\boldsymbol{R}^{\top}-\boldsymbol{\Sigma}_{U,l}^{1/2}\right\|\left\|\boldsymbol{\Sigma}_{U,l}^{-1/2}\right\|\\
&\lesssim\left\|\left(\boldsymbol{U}-\boldsymbol{U}^{\star}\boldsymbol{R}^{\top}\right)_{l,\cdot}\right\|_2\lambda_{\min}^{-1}\left(\boldsymbol{\Sigma}_{U,l}^{\star}\right)\left\|\boldsymbol{R}\left(\boldsymbol{\Sigma}_{U,l}^{\star}\right)^{1/2}\boldsymbol{R}^{\top}-\boldsymbol{\Sigma}_{U,l}^{1/2}\right\|.
\end{aligned}\tag{E.83}$$

Here the last line follows from an immediate result from Lemma 20 and Weyl's inequality:

$$\lambda_{\min}\left(\boldsymbol{\Sigma}_{U,l}\right)\asymp\lambda_{\min}\left(\boldsymbol{\Sigma}_{U,l}^{\star}\right),\tag{E.84}$$

which holds as long as $\delta\ll1$. Notice that

$$\begin{aligned}
\left\|\boldsymbol{R}\left(\boldsymbol{\Sigma}_{l}^{\star}\right)^{1/2}\boldsymbol{R}^{\top}-\boldsymbol{\Sigma}_{l}^{1/2}\right\|&\overset{\text{(i)}}{\lesssim}\frac{1}{\lambda_{\min}^{1/2}\left(\boldsymbol{\Sigma}_{l}^{\star}\right)+\lambda_{\min}^{1/2}\left(\boldsymbol{\Sigma}_{l}\right)}\left\|\boldsymbol{R}\boldsymbol{\Sigma}_{l}^{\star}\boldsymbol{R}^{\top}-\boldsymbol{\Sigma}_{l}\right\|\\
&\overset{\text{(ii)}}{\lesssim}\lambda_{\min}^{-1/2}\left(\boldsymbol{\Sigma}_{l}^{\star}\right)\left\|\boldsymbol{R}\boldsymbol{\Sigma}_{l}^{\star}\boldsymbol{R}^{\top}-\boldsymbol{\Sigma}_{l}\right\|\\
&\overset{\text{(iii)}}{\lesssim}\lambda_{\min}^{1/2}\left(\boldsymbol{\Sigma}_{U,l}^{\star}\right)\delta,
\end{aligned}\tag{E.85}$$

where (i) follows from the perturbation bound of matrix square root (Schmitt, 1992, Lemma 2.1); (ii) arises from (E.84); and (iii) is a consequence of Lemma 20. We can combine (E.83) and (E.85) to achieve

$$\begin{aligned}
\|\boldsymbol{\Delta}\|_2&\lesssim\left\|\left(\boldsymbol{U}-\boldsymbol{U}^{\star}\boldsymbol{R}^{\top}\right)_{l,\cdot}\right\|_2\lambda_{\min}^{-1/2}\left(\boldsymbol{\Sigma}_{U,l}^{\star}\right)\delta\\
&\lesssim\left[\frac{\theta}{\sqrt{\kappa}\sigma_r^{\star}}\left(\left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2+\omega_l^{\star}\right)+\zeta_{\mathsf{2nd},l}\right]\lambda_{\min}^{-1/2}\left(\boldsymbol{\Sigma}_{U,l}^{\star}\right)\delta.
\end{aligned}$$

Here, $\delta\in(0,1)$ is the (unspecified) quantity appearing in Lemma 20 such that

$$\left\|\boldsymbol{R}\boldsymbol{\Sigma}_{U,l}^{\star}\boldsymbol{R}^{\top}-\boldsymbol{\Sigma}_{U,l}\right\|\lesssim\delta\lambda_{\min}\left(\boldsymbol{\Sigma}_{U,l}^{\star}\right),$$

and the last line follows from (D.32a) in Lemma 18. Let

$$\zeta\coloneqq\widetilde{C}\left[\frac{\theta}{\sqrt{\kappa}\sigma_r^{\star}}\left(\left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2+\omega_l^{\star}\right)\lambda_{\min}^{-1/2}\left(\boldsymbol{\Sigma}_{U,l}^{\star}\right)\delta+\zeta_{\mathsf{2nd},l}\lambda_{\min}^{-1/2}\left(\boldsymbol{\Sigma}_{U,l}^{\star}\right)\delta\right]$$

for some sufficiently large constant $\widetilde{C}>0$ such that $\mathbb{P}(\|\boldsymbol{\Delta}\|_2\leq\zeta)\geq1-O((n+d)^{-10})$. Recalling the definition (A.3) of $\mathcal{C}^{\varepsilon}$ for any convex set $\mathcal{C}$, we have

$$\mathbb{P}\left(\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{R}^{\top}\in\mathsf{CR}_{U,l}^{1-\alpha}\right)=\mathbb{P}\left(\left(\boldsymbol{U}-\boldsymbol{U}^{\star}\boldsymbol{R}^{\top}\right)_{l,\cdot}\boldsymbol{\Sigma}_{l}^{-1/2}\in\mathcal{B}_{1-\alpha}\right)$$

$$
\begin{aligned}
&= \mathbb{P}\left(\left(\boldsymbol{U}-\boldsymbol{U}^{\star}\boldsymbol{R}^{\top}\right)_{l,\cdot}\boldsymbol{\Sigma}_l^{-1/2}\in\mathcal{B}_{1-\alpha},\|\boldsymbol{\Delta}\|_2\leq\zeta\right)+\mathbb{P}\left(\left(\boldsymbol{U}-\boldsymbol{U}^{\star}\boldsymbol{R}^{\top}\right)_{l,\cdot}\boldsymbol{\Sigma}_l^{-1/2}\in\mathcal{B}_{1-\alpha},\|\boldsymbol{\Delta}\|_2>\zeta\right)\\
&\leq \mathbb{P}\left(\left(\boldsymbol{U}-\boldsymbol{U}^{\star}\boldsymbol{R}^{\top}\right)_{l,\cdot}\boldsymbol{R}\left(\boldsymbol{\Sigma}_l^{\star}\right)^{-1/2}\boldsymbol{R}^{\top}\in\mathcal{B}_{1-\alpha}^{\zeta}\right)+\mathbb{P}\left(\|\boldsymbol{\Delta}\|_2>\zeta\right)\\
&\overset{\text{(i)}}{=} \mathbb{P}\left(\left(\boldsymbol{U}\boldsymbol{R}-\boldsymbol{U}^{\star}\right)_{l,\cdot}\left(\boldsymbol{\Sigma}_l^{\star}\right)^{-1/2}\in\mathcal{B}_{1-\alpha}^{\zeta}\right)+O\left((n+d)^{-10}\right)\\
&\overset{\text{(ii)}}{\leq} \mathcal{N}\left(\boldsymbol{0},\boldsymbol{I}_r\right)\left\{\mathcal{B}_{1-\alpha}^{\zeta}\right\}+O\left(\log^{-1/2}(n+d)\right)\\
&\overset{\text{(iii)}}{\leq} \mathcal{N}\left(\boldsymbol{0},\boldsymbol{I}_r\right)\left\{\mathcal{B}_{1-\alpha}\right\}+\zeta\left(0.59r^{1/4}+0.21\right)+O\left(\log^{-1/2}(n+d)\right)\\
&\overset{\text{(iv)}}{\leq} 1-\alpha+O\left(\log^{-1/2}(n+d)\right).
\end{aligned}
\tag{E.86}
$$

Here (i) holds since $\mathcal{B}_{1-\alpha}^{\zeta}$ is rotational invariant; (ii) uses Lemma 17; (iii) invokes Theorem 18; and (iv) makes use of the definition of $\mathcal{B}_{1-\alpha}$ in Algorithm 3 and holds under the condition $\zeta r^{1/4}\lesssim 1/\sqrt{\log(n+d)}$. Similar to (E.86)we can show that

$$
\mathbb{P}\left(\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{R}^{\top}\in\mathsf{CR}_{U,l}^{1-\alpha}\right)\geq 1-\alpha+O\left(\log^{-1/2}(n+d)\right)
\tag{E.87}
$$

Taking (E.86) and (E.87) collectively yields

$$
\mathbb{P}\left(\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{R}^{\top}\in\mathsf{CR}_{U,l}^{1-\alpha}\right)= 1-\alpha+O\left(\log^{-1/2}(n+d)\right),
$$

provided that $\zeta r^{1/4}\lesssim 1/\sqrt{\log(n+d)}$, or equivalently,

$$
\frac{\theta}{\sqrt{\kappa}\sigma_r^{\star}}\left(\left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2+\omega_l^{\star}\right)\lambda_{\min}^{-1/2}\left(\boldsymbol{\Sigma}_{U,l}^{\star}\right)\delta+\zeta_{\mathsf{2nd},l}\lambda_{\min}^{-1/2}\left(\boldsymbol{\Sigma}_{U,l}^{\star}\right)\delta\lesssim\frac{1}{r^{1/4}\sqrt{\log(n+d)}}.
\tag{E.88}
$$

It remains to choose $\delta\in(0,1)$ to satisfy the above condition. First, we have learned from the proof of Lemma 17 (more specifically, Step 3 in Appendix E.2.3) that

$$
\zeta_{\mathsf{2nd},l}\lambda_{\min}^{-1/2}(\boldsymbol{\Sigma}_l^{\star})\lesssim 1/(r^{1/4}\log^{1/2}(n+d))
$$

holds under the conditions of Lemma 17. As a result, it is sufficient to verify that

$$
\frac{\theta}{\sqrt{\kappa}\sigma_r^{\star}}\left(\left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2+\omega_l^{\star}\right)\lambda_{\min}^{-1/2}\left(\boldsymbol{\Sigma}_{U,l}^{\star}\right)\delta\lesssim\frac{1}{r^{1/4}\sqrt{\log(n+d)}}.
\tag{E.89}
$$

To this end, recall from Lemma 16 that

$$
\lambda_{\min}^{1/2}\left(\boldsymbol{\Sigma}_{U,l}^{\star}\right)\gtrsim\left(\frac{1}{\sqrt{np}\sigma_1^{\star}}+\frac{1}{\sqrt{ndp^2\kappa}\sigma_1^{\star}}\right)\left(\left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2+\omega_l^{\star}\right)+\frac{1}{\sqrt{np^2}\sigma_1^{\star 2}}\omega_{\min}\left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2+\frac{\omega_{\min}\omega_l^{\star}}{\sqrt{np^2}\sigma_1^{\star 2}}.
$$

In view of the definition of θ (cf. (D.31)), we can show that

$$
\begin{aligned}
\frac{\theta}{\sqrt{\kappa}\sigma_r^{\star}}\left(\left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2+\omega_l^{\star}\right)&\asymp\underbrace{\frac{1}{\sigma_r^{\star}}\sqrt{\frac{r\log^2(n+d)}{np}}\left(\left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2+\omega_l^{\star}\right)}_{=:\alpha_1}+\underbrace{\frac{1}{\sigma_r^{\star}}\sqrt{\frac{\kappa\mu r^2\log^3(n+d)}{ndp^2}}\left(\left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2+\omega_l^{\star}\right)}_{=:\alpha_2}\\
&\quad+\underbrace{\frac{\omega_{\max}}{\sigma_r^{\star 2}}\sqrt{\frac{r\log^2(n+d)}{np^2}}\left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2}_{=:\alpha_3}+\underbrace{\frac{\omega_{\max}\omega_l^{\star}}{\sigma_r^{\star 2}}\sqrt{\frac{r\log^2(n+d)}{np^2}}}_{=:\alpha_4}\\
&\lesssim\sqrt{\kappa^3\mu r^2\kappa_{\omega}\log^3(n+d)}\lambda_{\min}^{1/2}\left(\boldsymbol{\Sigma}_{U,l}^{\star}\right),
\end{aligned}
$$

where the last line holds since

$$\begin{aligned}
\alpha_1 &\lesssim \sqrt{\kappa r \log^2(n+d)} \frac{1}{\sqrt{np}\sigma_1^\star} \left(\left\| \boldsymbol{U}_{l,\cdot}^\star \boldsymbol{\Sigma}^\star \right\|_2 + \omega_l^\star \right) \lesssim \sqrt{\kappa r \log^2(n+d)} \lambda_{\min}^{1/2} \left(\boldsymbol{\Sigma}_{U,l}^\star \right), \\
\alpha_2 &\lesssim \sqrt{\kappa^3 \mu r^2 \log^3(n+d)} \frac{1}{\sqrt{ndp^2 \kappa} \sigma_1^\star} \left(\left\| \boldsymbol{U}_{l,\cdot}^\star \boldsymbol{\Sigma}^\star \right\|_2 + \omega_l^\star \right) \lesssim \sqrt{\kappa^3 \mu r^2 \log^3(n+d)} \lambda_{\min}^{1/2} \left(\boldsymbol{\Sigma}_{U,l}^\star \right), \\
\alpha_3 &\lesssim \sqrt{\kappa^2 r \kappa_\omega \log^2(n+d)} \frac{1}{\sqrt{np^2 \sigma_1^{\star 2}}} \omega_{\min} \left\| \boldsymbol{U}_{l,\cdot}^\star \boldsymbol{\Sigma}^\star \right\|_2 \lesssim \sqrt{\kappa^2 r \kappa_\omega \log^2(n+d)} \lambda_{\min}^{1/2} \left(\boldsymbol{\Sigma}_{U,l}^\star \right), \\
\alpha_4 &\lesssim \sqrt{\kappa^2 r \kappa_\omega \log^2(n+d)} \frac{\omega_{\min} \omega_l^\star}{\sqrt{np^2 \sigma_1^{\star 2}}} \lesssim \sqrt{\kappa^2 r \kappa_\omega \log^2(n+d)} \lambda_{\min}^{1/2} \left(\boldsymbol{\Sigma}_{U,l}^\star \right).
\end{aligned}$$

The above bound immediately gives

$$\frac{\theta}{\sqrt{\kappa} \sigma_r^\star} \left(\left\| \boldsymbol{U}_{l,\cdot}^\star \boldsymbol{\Sigma}^\star \right\|_2 + \omega_l^\star \right) \lambda_{\min}^{-1/2} \left(\boldsymbol{\Sigma}_{U,l}^\star \right) \delta \lesssim \sqrt{\kappa^3 \mu r^2 \kappa_\omega \log^3(n+d)} \delta.$$

Therefore, the desired condition (E.89) — and hence $\zeta r^{1/4} \lesssim 1/\sqrt{\log(n+d)}$ — can be guaranteed by taking

$$\delta = \frac{1}{\kappa^{3/2} \mu^{1/2} r^{5/4} \kappa_\omega^{1/2} \log^2(n+d)}.$$

After δ is specified, we can easily check that under our choice of δ, the assumptions in Lemma 20 read $n \gtrsim \kappa^{12} \mu^3 r^{11/2} \kappa_\omega \log^5(n+d)$, $d \gtrsim \kappa^3 \mu r \log(n+d)$,

$$ndp^2 \gtrsim \kappa^{11} \mu^5 r^{13/2} \kappa_\omega^3 \log^9(n+d), \qquad np \gtrsim \kappa^{11} \mu^4 r^{11/2} \kappa_\omega^3 \log^7(n+d),$$

and

$$\frac{\omega_{\max}^2}{p \sigma_r^{\star 2}} \sqrt{\frac{d}{n}} \lesssim \frac{1}{\kappa^{9/2} \mu^{3/2} r^{9/4} \kappa_\omega^{3/2} \log^{7/2}(n+d)}, \qquad \frac{\omega_{\max}}{\sigma_r^\star} \sqrt{\frac{d}{np}} \lesssim \frac{1}{\kappa^5 \mu^{3/2} r^{9/4} \kappa_\omega^{3/2} \log^3(n+d)}.$$

This concludes the proof.

E.4 Auxiliary lemmas for Theorem 13

E.4.1 Proof of Lemma 22

Throughout this section, all the probabilistic arguments are conditional on $\boldsymbol{F}$, and we shall always assume that the $\sigma(\boldsymbol{F})$-measurable high-probability event $\mathcal{E}_{\mathsf{good}}$ occurs. Following the same analysis as in Appendix E.2.1 (Proof of Lemma 15), we can obtain

$$\boldsymbol{U} \boldsymbol{R}_{\boldsymbol{U}} - \boldsymbol{U}^\natural = \boldsymbol{Z} + \boldsymbol{\Psi}_{\boldsymbol{U}}$$

with $\boldsymbol{R}_{\boldsymbol{U}} = \arg\min_{\boldsymbol{O} \in \mathcal{O}^{r \times r}} \|\boldsymbol{U}\boldsymbol{O} - \boldsymbol{U}^\natural\|_{\mathrm{F}}^2$, where

$$\boldsymbol{Z} = \left[\boldsymbol{E} \boldsymbol{M}^{\natural \top} + \mathcal{P}_{\mathsf{off\text{-}diag}} \left(\boldsymbol{E} \boldsymbol{E}^\top \right) \right] \boldsymbol{U}^\natural \left(\boldsymbol{\Sigma}^\natural \right)^{-2}$$

and $\boldsymbol{\Psi}_{\boldsymbol{U}}$ is a residual matrix obeying

$$\mathbb{P}\left(\left\| \boldsymbol{e}_l^\top \boldsymbol{\Psi}_{\boldsymbol{U}} \right\|_2 \mathbb{1}_{\mathcal{E}_{\mathsf{good}}} \lesssim \zeta_{\mathsf{2nd},l} \,\middle|\, \boldsymbol{F} \right) \geq 1 - O\left((n+d)^{-10} \right) \tag{E.90}$$

for all $l \in [d]$. Here, $\zeta_{\mathsf{2nd},l}$ is a quantity defined in Lemma 15. From the proof of Lemma 18, we know that

$$\boldsymbol{R}_{\boldsymbol{U}}^\top \boldsymbol{\Sigma}^2 \boldsymbol{R}_{\boldsymbol{U}} = \boldsymbol{\Sigma}^{\natural 2} + \boldsymbol{\Psi}_{\boldsymbol{\Sigma}}$$

holds for some matrix $\boldsymbol{\Psi}_{\boldsymbol{\Sigma}}$ satisfying

$$\mathbb{P}\left(\left\| \boldsymbol{\Psi}_{\boldsymbol{\Sigma}} \right\| \mathbb{1}_{\mathcal{E}_{\mathsf{good}}} \lesssim \sqrt{\frac{\kappa^3 \mu r \log(n+d)}{d}} \zeta_{\mathsf{1st}} + \kappa \frac{\zeta_{\mathsf{1st}}^2}{\sigma_r^{\star 2}} \,\middle|\, \boldsymbol{F} \right) \geq 1 - O\left((n+d)^{-10} \right). \tag{E.91}$$

Armed with the above facts, we can write

$$
\begin{aligned}
\boldsymbol{S} - \boldsymbol{M}^{\natural}\boldsymbol{M}^{\natural\top} &= \boldsymbol{U}\boldsymbol{\Sigma}^{2}\boldsymbol{U}^{\top} - \boldsymbol{U}^{\natural}\boldsymbol{\Sigma}^{\natural 2}\boldsymbol{U}^{\natural\top} \\
&= \boldsymbol{U}\boldsymbol{R}_{\boldsymbol{U}}\boldsymbol{R}_{\boldsymbol{U}}^{\top}\boldsymbol{\Sigma}^{2}\boldsymbol{R}_{\boldsymbol{U}}\boldsymbol{R}_{\boldsymbol{U}}^{\top}\boldsymbol{U}^{\top} - \boldsymbol{U}^{\natural}\boldsymbol{\Sigma}^{\natural 2}\boldsymbol{U}^{\natural\top} \\
&= \boldsymbol{U}\boldsymbol{R}_{\boldsymbol{U}}\boldsymbol{\Sigma}^{\natural 2}\boldsymbol{R}_{\boldsymbol{U}}^{\top}\boldsymbol{U}^{\top} + \boldsymbol{U}\boldsymbol{R}_{\boldsymbol{U}}\boldsymbol{\Psi}_{\boldsymbol{\Sigma}}\boldsymbol{R}_{\boldsymbol{U}}^{\top}\boldsymbol{U}^{\top} - \boldsymbol{U}^{\natural}\boldsymbol{\Sigma}^{\natural 2}\boldsymbol{U}^{\natural\top} \\
&= \left(\boldsymbol{U}^{\natural} + \boldsymbol{Z} + \boldsymbol{\Psi}_{\boldsymbol{U}}\right)\boldsymbol{\Sigma}^{\natural 2}\left(\boldsymbol{U}^{\natural} + \boldsymbol{Z} + \boldsymbol{\Psi}_{\boldsymbol{U}}\right)^{\top} + \boldsymbol{U}\boldsymbol{R}_{\boldsymbol{U}}\boldsymbol{\Psi}_{\boldsymbol{\Sigma}}\boldsymbol{R}_{\boldsymbol{U}}^{\top}\boldsymbol{U}^{\top} - \boldsymbol{U}^{\natural}\boldsymbol{\Sigma}^{\natural 2}\boldsymbol{U}^{\natural\top} \\
&= \boldsymbol{X} + \boldsymbol{\Phi},
\end{aligned}
$$

where

$$
\begin{aligned}
\boldsymbol{X} &= \boldsymbol{U}^{\natural}\boldsymbol{\Sigma}^{\natural 2}\boldsymbol{Z}^{\top} + \boldsymbol{Z}\boldsymbol{\Sigma}^{\natural 2}\boldsymbol{U}^{\natural} \\
&= \boldsymbol{E}\boldsymbol{M}^{\natural\top} + \boldsymbol{M}^{\natural}\boldsymbol{E}^{\top} + \mathcal{P}_{\textsf{off-diag}}\left(\boldsymbol{E}\boldsymbol{E}^{\top}\right)\boldsymbol{U}^{\natural}\boldsymbol{U}^{\natural\top} + \boldsymbol{U}^{\natural}\boldsymbol{U}^{\natural\top}\mathcal{P}_{\textsf{off-diag}}\left(\boldsymbol{E}\boldsymbol{E}^{\top}\right)
\end{aligned}
$$

and

$$
\boldsymbol{\Phi} = \boldsymbol{U}^{\natural}\boldsymbol{\Sigma}^{\natural 2}\boldsymbol{\Psi}_{\boldsymbol{U}}^{\top} + \boldsymbol{Z}\boldsymbol{\Sigma}^{\natural 2}\left(\boldsymbol{Z} + \boldsymbol{\Psi}_{\boldsymbol{U}}\right)^{\top} + \boldsymbol{\Psi}_{\boldsymbol{U}}\boldsymbol{\Sigma}^{\natural 2}\left(\boldsymbol{U}\boldsymbol{R}_{\boldsymbol{U}}\right)^{\top} + \boldsymbol{U}\boldsymbol{R}_{\boldsymbol{U}}\boldsymbol{\Psi}_{\boldsymbol{\Sigma}}\left(\boldsymbol{U}\boldsymbol{R}_{\boldsymbol{U}}\right)^{\top}.
$$

It has already been shown in Lemma 18 that for each $l \in [d]$,

$$
\left\|\boldsymbol{Z}_{l,\cdot}\right\|_{2} \lesssim \frac{\theta}{\sqrt{\kappa}\sigma_{r}^{\star}}\left(\left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_{2} + \omega_{l}^{\star}\right) \tag{E.92}
$$

and

$$
\left\|\left(\boldsymbol{U}\boldsymbol{R} - \boldsymbol{U}^{\star}\right)_{l,\cdot}\right\|_{2} \lesssim \frac{\theta}{\sqrt{\kappa}\sigma_{r}^{\star}}\left(\left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_{2} + \omega_{l}^{\star}\right) + \zeta_{\textsf{2nd},l}
$$

hold with probability at least $1 - O\big((n+d)^{-10}\big)$. As an immediate consequence,

$$
\left\|\boldsymbol{U}_{l,\cdot}\right\|_{2} \leq \left\|\boldsymbol{U}_{l,\cdot}^{\star}\right\|_{2} + \left\|\left(\boldsymbol{U}\boldsymbol{R} - \boldsymbol{U}^{\star}\right)_{l,\cdot}\right\|_{2} \lesssim \left\|\boldsymbol{U}_{l,\cdot}^{\star}\right\|_{2} + \theta\frac{\omega_{l}^{\star}}{\sigma_{1}^{\star}} + \zeta_{\textsf{2nd},l}, \tag{E.93}
$$

provided that $\theta \lesssim 1$. Then with probability exceeding $1 - O((n+d)^{-10})$, we have

$$
\begin{aligned}
\left|\Phi_{i,j}\right| &\overset{\text{(i)}}{\lesssim} \sigma_{1}^{\star 2}\left(\left\|\boldsymbol{U}_{i,\cdot}^{\star}\right\|_{2}\left\|\boldsymbol{e}_{j}^{\top}\boldsymbol{\Psi}_{\boldsymbol{U}}\right\|_{2} + \left\|\boldsymbol{Z}_{i,\cdot}\right\|_{2}\left\|\boldsymbol{Z}_{j,\cdot}\right\|_{2} + \left\|\boldsymbol{Z}_{i,\cdot}\right\|_{2}\left\|\boldsymbol{e}_{j}^{\top}\boldsymbol{\Psi}_{\boldsymbol{U}}\right\|_{2} + \left\|\boldsymbol{e}_{i}^{\top}\boldsymbol{\Psi}_{\boldsymbol{U}}\right\|_{2}\left\|\boldsymbol{U}_{j,\cdot}\right\|_{2}\right) \\
&\quad + \left\|\boldsymbol{\Psi}_{\boldsymbol{\Sigma}}\right\|\left\|\boldsymbol{U}_{i,\cdot}\right\|_{2}\left\|\boldsymbol{U}_{j,\cdot}\right\|_{2} \\
&\overset{\text{(ii)}}{\lesssim} \zeta_{\textsf{2nd},j}\sigma_{1}^{\star 2}\left(\left\|\boldsymbol{U}_{i,\cdot}^{\star}\right\|_{2} + \theta\frac{\omega_{i}^{\star}}{\sigma_{1}^{\star}}\right) + \sigma_{1}^{\star 2}\zeta_{\textsf{2nd},i}\left(\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_{2} + \theta\frac{\omega_{j}^{\star}}{\sigma_{1}^{\star}} + \zeta_{\textsf{2nd},j}\right) \\
&\quad + \theta^{2}\left(\left\|\boldsymbol{U}_{i,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_{2} + \omega_{i}^{\star}\right)\left(\left\|\boldsymbol{U}_{j,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_{2} + \omega_{j}^{\star}\right) \\
&\quad + \left(\sqrt{\frac{\kappa^{3}\mu r\log(n+d)}{d}}\zeta_{\textsf{1st}} + \kappa\frac{\zeta_{\textsf{1st}}^{2}}{\sigma_{r}^{\star 2}}\right)\left(\left\|\boldsymbol{U}_{i,\cdot}^{\star}\right\|_{2} + \theta\frac{\omega_{i}^{\star}}{\sigma_{1}^{\star}} + \zeta_{\textsf{2nd},i}\right)\left(\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_{2} + \theta\frac{\omega_{j}^{\star}}{\sigma_{1}^{\star}} + \zeta_{\textsf{2nd},j}\right) \\
&\overset{\text{(iii)}}{\lesssim} \theta^{2}\left(\left\|\boldsymbol{U}_{i,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_{2} + \omega_{i}^{\star}\right)\left(\left\|\boldsymbol{U}_{j,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_{2} + \omega_{j}^{\star}\right) + \sigma_{1}^{\star 2}\zeta_{\textsf{2nd},i}\zeta_{\textsf{2nd},j} \\
&\quad + \sigma_{1}^{\star 2}\zeta_{\textsf{2nd},i}\left(\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_{2} + \theta\frac{\omega_{j}^{\star}}{\sigma_{1}^{\star}}\right) + \zeta_{\textsf{2nd},j}\sigma_{1}^{\star 2}\left(\left\|\boldsymbol{U}_{i,\cdot}^{\star}\right\|_{2} + \theta\frac{\omega_{i}^{\star}}{\sigma_{1}^{\star}}\right).
\end{aligned}
$$

for each $i, j \in [d]$. Here, (i) makes use of (D.13) and the fact that $\boldsymbol{U}^{\natural} = \boldsymbol{U}^{\star}\boldsymbol{Q}$ for some orthonormal matrix $\boldsymbol{Q}$ (cf. (D.5)); (ii) follows from (E.90), (E.91), (E.92) and (E.93), and holds provided that $\theta \lesssim 1$; (iii) holds provided that $\zeta_{\textsf{1st}}/\sigma_{r}^{\star 2} \lesssim 1$ and $d \gtrsim \kappa\mu r\log(n+d)$.

E.4.2 Proof of Lemma 23

In this subsection, we shall focus on establishing the claimed result for the case when $i \neq j$; the case when $i = j$ can be proved in a similar (in fact, easier) manner. In view of the expression (D.40), we can write

$$\mathsf{var}\left(X_{i,j}|\boldsymbol{F}\right) = \underbrace{\sum_{l=1}^{n} M_{j,l}^{\natural 2}\sigma_{i,l}^{2}}_{=:\alpha_1} + \underbrace{\sum_{l=1}^{n} M_{i,l}^{\natural 2}\sigma_{j,l}^{2}}_{=:\alpha_2} + \underbrace{\sum_{l=1}^{n}\sum_{k:k\neq i}\sigma_{i,l}^{2}\sigma_{k,l}^{2}\left(\boldsymbol{U}_{k,\cdot}^{\star}\boldsymbol{U}_{j,\cdot}^{\star\top}\right)^{2}}_{=:\alpha_3} + \underbrace{\sum_{l=1}^{n}\sum_{k:k\neq j}\sigma_{j,l}^{2}\sigma_{k,l}^{2}\left(\boldsymbol{U}_{k,\cdot}^{\star}\boldsymbol{U}_{i,\cdot}^{\star\top}\right)^{2}}_{=:\alpha_4},$$

thus motivating us to study the behavior of α_1, α_2, α_3 and α_4 respectively.

- Let us begin with the term α_1. By virtue of (D.3) and (D.9), we have

$$\begin{aligned}
\alpha_1 &= \sum_{l=1}^{n} M_{j,l}^{\natural 2}\sigma_{i,l}^{2} = \sum_{l=1}^{n}\left(\frac{1}{\sqrt{n}}\boldsymbol{U}_{j,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\boldsymbol{f}_l\right)^{2}\left[\frac{1-p}{np}\left(\boldsymbol{U}_{i,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\boldsymbol{f}_l\right)^{2} + \frac{\omega_i^{\star 2}}{np}\right] \\
&= \frac{1-p}{n^2 p}\sum_{l=1}^{n}\left(\boldsymbol{U}_{j,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\boldsymbol{f}_l\right)^{2}\left(\boldsymbol{U}_{i,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\boldsymbol{f}_l\right)^{2} + \frac{\omega_i^{\star 2}}{n^2 p}\sum_{l=1}^{n}\left(\boldsymbol{U}_{j,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\boldsymbol{f}_l\right)^{2}.
\end{aligned}$$

  On the event $\mathcal{E}_{\mathsf{good}}$ (see Lemma 14), we know from the basic facts in (D.21b) and (D.21c) that

$$\begin{aligned}
\left|\frac{1}{n}\sum_{l=1}^{n}\left(\boldsymbol{U}_{j,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\boldsymbol{f}_l\right)^{2}\left(\boldsymbol{U}_{i,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\boldsymbol{f}_l\right)^{2} - S_{i,i}^{\star}S_{j,j}^{\star} - 2S_{i,j}^{\star 2}\right| &\lesssim \sqrt{\frac{\log^3(n+d)}{n}}S_{i,i}^{\star}S_{j,j}^{\star}; \\
\left|\frac{1}{n}\sum_{l=1}^{n}\left(\boldsymbol{U}_{j,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\boldsymbol{f}_l\right)^{2} - S_{j,j}^{\star}\right| &\lesssim \sqrt{\frac{\log(n+d)}{n}}S_{j,j}^{\star}.
\end{aligned}$$

  As a result, we can express α_1 as

$$\alpha_1 = \underbrace{\frac{1-p}{np}\left(S_{i,i}^{\star}S_{j,j}^{\star} + 2S_{i,j}^{\star 2}\right) + \frac{\omega_i^{\star 2}}{np}S_{j,j}^{\star}}_{=:\alpha_1^{\star}} + r_1$$

  for some residual term r_1 satisfying

$$|r_1| \lesssim \sqrt{\frac{\log^3(n+d)}{n}}\alpha_1^{\star}.$$

- Akin to our analysis for α_1, we can also demonstrate that

$$\alpha_2 = \underbrace{\frac{1-p}{np}\left[S_{i,i}^{\star}S_{j,j}^{\star} + 2S_{i,j}^{\star 2}\right] + \frac{\omega_j^{\star 2}}{np}S_{i,i}^{\star}}_{=:\alpha_2^{\star}} + r_2$$

  for some residual term r_2 obeying

$$|r_2| \lesssim \sqrt{\frac{\log^3(n+d)}{n}}\alpha_2^{\star}.$$

- When it comes to α_3, we make the observation that

$$\alpha_3 = \sum_{l=1}^{n}\sum_{k:k\neq i}\left[\frac{1-p}{np}\left(\boldsymbol{U}_{i,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\boldsymbol{f}_l\right)^{2} + \frac{\omega_i^{\star 2}}{np}\right]\left[\frac{1-p}{np}\left(\boldsymbol{U}_{k,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\boldsymbol{f}_l\right)^{2} + \frac{\omega_k^{\star 2}}{np}\right]\left(\boldsymbol{U}_{k,\cdot}^{\star}\boldsymbol{U}_{j,\cdot}^{\star\top}\right)^{2}$$

$$
\begin{aligned}
&= \left(\frac{1-p}{np}\right)^2 \sum_{k:k\neq i} \left[\sum_{l=1}^{n} \left(\boldsymbol{U}_{i,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\boldsymbol{f}_l\right)^2 \left(\boldsymbol{U}_{k,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\boldsymbol{f}_l\right)^2\right] \left(\boldsymbol{U}_{k,\cdot}^{\star}\boldsymbol{U}_{j,\cdot}^{\star\top}\right)^2 \\
&\quad + \frac{1-p}{n^2p^2} \sum_{k:k\neq i} \sum_{l=1}^{n} \left(\boldsymbol{U}_{i,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\boldsymbol{f}_l\right)^2 \omega_k^{\star 2} \left(\boldsymbol{U}_{k,\cdot}^{\star}\boldsymbol{U}_{j,\cdot}^{\star\top}\right)^2 \\
&\quad + \frac{(1-p)\,\omega_i^{\star 2}}{n^2p^2} \sum_{k:k\neq i} \sum_{l=1}^{n} \left(\boldsymbol{U}_{k,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\boldsymbol{f}_l\right)^2 \left(\boldsymbol{U}_{k,\cdot}^{\star}\boldsymbol{U}_{j,\cdot}^{\star\top}\right)^2 + \frac{\omega_i^{\star 2}}{np^2} \sum_{k:k\neq i} \omega_k^{\star 2} \left(\boldsymbol{U}_{k,\cdot}^{\star}\boldsymbol{U}_{j,\cdot}^{\star\top}\right)^2.
\end{aligned}
$$

On the event $\mathcal{E}_{\mathsf{good}}$, it is seen from (D.21b) and (D.21c) that

$$
\begin{aligned}
\left|\frac{1}{n}\sum_{l=1}^{n} \left(\boldsymbol{U}_{i,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\boldsymbol{f}_l\right)^2 \left(\boldsymbol{U}_{k,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\boldsymbol{f}_l\right)^2 - S_{k,k}^{\star}S_{i,i}^{\star} - 2S_{i,k}^{\star 2}\right| &\lesssim \sqrt{\frac{\log^3(n+d)}{n}} S_{k,k}^{\star}S_{i,i}^{\star} \\
\left|\frac{1}{n}\sum_{l=1}^{n} \left(\boldsymbol{U}_{i,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\boldsymbol{f}_l\right)^2 - S_{i,i}^{\star}\right| &\lesssim \sqrt{\frac{\log(n+d)}{n}} S_{i,i}^{\star} \\
\left|\frac{1}{n}\sum_{l=1}^{n} \left(\boldsymbol{U}_{k,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\boldsymbol{f}_l\right)^2 - S_{k,k}^{\star}\right| &\lesssim \sqrt{\frac{\log(n+d)}{n}} S_{k,k}^{\star}
\end{aligned}
$$

for each $k \in [d]\backslash\{i\}$. As a consequence, one can express

$$
\begin{aligned}
\alpha_3 &= \frac{\omega_i^{\star 2}}{np^2} \sum_{k:k\neq i} \omega_k^{\star 2} \left(\boldsymbol{U}_{k,\cdot}^{\star}\boldsymbol{U}_{j,\cdot}^{\star\top}\right)^2 + \frac{1-p}{np^2} S_{i,i}^{\star} \sum_{k:k\neq i} \omega_k^{\star 2} \left(\boldsymbol{U}_{k,\cdot}^{\star}\boldsymbol{U}_{j,\cdot}^{\star\top}\right)^2 \\
&\quad + \frac{(1-p)\,\omega_i^{\star 2}}{np^2} \sum_{k:k\neq i} S_{k,k}^{\star} \left(\boldsymbol{U}_{k,\cdot}^{\star}\boldsymbol{U}_{j,\cdot}^{\star\top}\right)^2 + \frac{(1-p)^2}{np^2} \sum_{k:k\neq i} \left(S_{k,k}^{\star}S_{i,i}^{\star} + 2S_{i,k}^{\star 2}\right) \left(\boldsymbol{U}_{k,\cdot}^{\star}\boldsymbol{U}_{j,\cdot}^{\star\top}\right)^2 + \widetilde{r}_3
\end{aligned}
$$

for some residual term $\widetilde{r}_3$ satisfying

$$
|\widetilde{r}_3| \lesssim \sqrt{\frac{\log^3(n+d)}{n}}\, |\alpha_3 - \widetilde{r}_3| \asymp \sqrt{\frac{\log^3(n+d)}{n}}\, |\alpha_3|,
$$

where the last relation holds provided that $n \gg \log^3(n+d)$. This allows one to decompose α_3 as follows

$$
\begin{aligned}
\alpha_3 &= \frac{1}{np^2} \sum_{k:k\neq i} \left\{\left[\omega_i^{\star 2} + (1-p)\,S_{i,i}^{\star}\right]\left[\omega_k^{\star 2} + (1-p)\,S_{k,k}^{\star}\right] + 2\,(1-p)^2\, S_{i,k}^{\star 2}\right\} \left(\boldsymbol{U}_{k,\cdot}^{\star}\boldsymbol{U}_{j,\cdot}^{\star\top}\right)^2 + \widetilde{r}_3 \\
&= \underbrace{\frac{1}{np^2} \sum_{k=1}^{d} \left\{\left[\omega_i^{\star 2} + (1-p)\,S_{i,i}^{\star}\right]\left[\omega_k^{\star 2} + (1-p)\,S_{k,k}^{\star}\right] + 2\,(1-p)^2\, S_{i,k}^{\star 2}\right\} \left(\boldsymbol{U}_{k,\cdot}^{\star}\boldsymbol{U}_{j,\cdot}^{\star\top}\right)^2}_{=:\alpha_3^{\star}} + r_3,
\end{aligned}
$$

where we define

$$
r_3 = \widetilde{r}_3 - \underbrace{\frac{1}{np^2} \left\{\left[\omega_i^{\star 2} + (1-p)\,S_{i,i}^{\star}\right]^2 + 2\,(1-p)^2\, S_{i,i}^{\star 2}\right\} \left(\boldsymbol{U}_{i,\cdot}^{\star}\boldsymbol{U}_{j,\cdot}^{\star\top}\right)^2}_{=:\delta}.
$$

Recalling from Claim 1 that

$$
\sum_{k=1}^{d} S_{k,k}^{\star} \boldsymbol{U}_{k,\cdot}^{\star\top}\boldsymbol{U}_{k,\cdot}^{\star} = \sum_{k=1}^{d} \left\|\boldsymbol{U}_{k,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2^2 \boldsymbol{U}_{k,\cdot}^{\star\top}\boldsymbol{U}_{k,\cdot}^{\star} \succeq \frac{\sigma_r^{\star 2}}{4d}\boldsymbol{I}_r, \tag{E.94}
$$

we can reach

$$
\alpha_3^{\star} \geq \frac{1}{np^2} \sum_{k=1}^{d} \left\{\left[\omega_i^{\star 2} + (1-p)\,S_{i,i}^{\star}\right]\left[\omega_k^{\star 2} + (1-p)\,S_{k,k}^{\star}\right]\right\} \left(\boldsymbol{U}_{k,\cdot}^{\star}\boldsymbol{U}_{j,\cdot}^{\star\top}\right)^2
$$

$$
\begin{aligned}
&\gtrsim \frac{1}{np^2}\left[\omega_i^{\star 2}+(1-p)\,S_{i,i}^{\star}\right]\boldsymbol{U}_{j,\cdot}^{\star}\left[\sum_{k=1}^{d}\omega_k^{\star 2}\boldsymbol{U}_{k,\cdot}^{\star\top}\boldsymbol{U}_{k,\cdot}^{\star}+\sum_{k=1}^{d}S_{k,k}^{\star}\boldsymbol{U}_{k,\cdot}^{\star\top}\boldsymbol{U}_{k,\cdot}^{\star}\right]\boldsymbol{U}_{j,\cdot}^{\star\top}\\
&\gtrsim \frac{1}{np^2}\left[\omega_i^{\star 2}+(1-p)\,S_{i,i}^{\star}\right]\left[\omega_{\min}^2\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2^2+\frac{\sigma_r^{\star 2}}{d}\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2^2\right] &\text{(E.95)}\\
&\gtrsim \frac{1}{np^2}\left[\omega_i^{\star 2}+(1-p)\,S_{i,i}^{\star}\right]\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2^2\left(\omega_{\min}^2+\frac{\sigma_r^{\star 2}}{d}\right), &\text{(E.96)}
\end{aligned}
$$

where the second line uses the assumption that p is strictly bounded away from 1 (so that $1-p\asymp 1$), and the penultimate line makes use of (E.94). This immediately leads to

$$
\begin{aligned}
\delta &\lesssim \frac{1}{np^2}\left[\omega_i^{\star 2}+(1-p)\,S_{i,i}^{\star}\right]^2\left\|\boldsymbol{U}_{i,\cdot}^{\star}\right\|_2^2\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2^2\\
&\lesssim \frac{\mu r}{ndp^2}\left[\omega_i^{\star 2}+(1-p)\,S_{i,i}^{\star}\right]\left(\omega_{\max}^2+\frac{\mu r}{d}\sigma_1^{\star 2}\right)\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2^2\\
&\lesssim \frac{\kappa\mu^2 r^2+\kappa_\omega \mu r}{d}\alpha_3^{\star},
\end{aligned}
$$

where the penultimate line comes from the assumption $1-p\asymp 1$ and the fact that $S_{i,i}^{\star}=\|\boldsymbol{U}_{i,\cdot}^{\star}\|_2^2\|\boldsymbol{\Sigma}^{\star}\|^2\leq \frac{\mu r}{d}\sigma_1^{\star 2}$, and the last inequality results from (E.96). As a consequence, we arrive at

$$
\begin{aligned}
|r_3| \leq |\widetilde{r}_3|+|\delta| &\lesssim \sqrt{\frac{\log^3(n+d)}{n}}\,|\alpha_3|+\frac{\kappa\mu^2 r^2+\kappa_\omega \mu r}{d}\alpha_3^{\star}\\
&\lesssim \sqrt{\frac{\log^3(n+d)}{n}}\left(\alpha_3^{\star}+|r_3|\right)+\frac{\kappa\mu^2 r^2+\kappa_\omega \mu r}{d}\alpha_3^{\star}\\
&\lesssim \left(\sqrt{\frac{\log^3(n+d)}{n}}+\frac{\kappa\mu^2 r^2+\kappa_\omega \mu r}{d}\right)\alpha_3^{\star},
\end{aligned}
$$

where the last relation holds true as long as $n\gg \log^3(n+d)$.

- Repeating our analysis for α_3 allows one to show that

$$
\alpha_4=\underbrace{\frac{1}{np^2}\sum_{k=1}^{d}\left\{\left[\omega_j^{\star 2}+(1-p)\,S_{j,j}^{\star}\right]\left[\omega_k^{\star 2}+(1-p)\,S_{k,k}^{\star}\right]+2\,(1-p)^2\,S_{j,k}^{\star 2}\right\}\left(\boldsymbol{U}_{k,\cdot}^{\star}\boldsymbol{U}_{i,\cdot}^{\star\top}\right)^2}_{=:\alpha_4^{\star}}+r_4,
$$

where the residual term r_4 obeys

$$
|r_4|\lesssim \left(\sqrt{\frac{\log^3(n+d)}{n}}+\frac{\kappa\mu^2 r^2+\kappa_\omega \mu r}{d}\right)\alpha_4^{\star}.
$$

Putting the above bounds together, we can conclude that

$$
\mathsf{var}\left(X_{i,j}|\boldsymbol{F}\right)=\alpha_1+\alpha_2+\alpha_3+\alpha_4=\widetilde{v}_{i,j}+r_{i,j},
$$

where

$$
\begin{aligned}
\widetilde{v}_{i,j} &\coloneqq \alpha_1^{\star}+\alpha_2^{\star}+\alpha_3^{\star}+\alpha_4^{\star}\\
&=\frac{2\,(1-p)}{np}\left(S_{i,i}^{\star}S_{j,j}^{\star}+2S_{i,j}^{\star 2}\right)+\frac{1}{np}\left(\omega_i^{\star 2}S_{j,j}^{\star}+\omega_j^{\star 2}S_{i,i}^{\star}\right)\\
&\quad+\frac{1}{np^2}\sum_{k=1}^{d}\left\{\left[\omega_i^{\star 2}+(1-p)\,S_{i,i}^{\star}\right]\left[\omega_k^{\star 2}+(1-p)\,S_{k,k}^{\star}\right]+2\,(1-p)^2\,S_{i,k}^{\star 2}\right\}\left(\boldsymbol{U}_{k,\cdot}^{\star}\boldsymbol{U}_{j,\cdot}^{\star\top}\right)^2
\end{aligned}
$$

$$+\frac{1}{np^2}\sum_{k=1}^{d}\left\{\left[\omega_j^{\star 2}+(1-p)S_{j,j}^{\star}\right]\left[\omega_k^{\star 2}+(1-p)S_{k,k}^{\star}\right]+2(1-p)^2S_{j,k}^{\star 2}\right\}\left(\boldsymbol{U}_{k,\cdot}^{\star}\boldsymbol{U}_{i,\cdot}^{\star\top}\right)^2,$$

and the residual term $r_{i,j}$ is bounded in magnitude by

$$\begin{aligned}
|r_{i,j}| &\le |r_1|+|r_2|+|r_3|+|r_4| \\
&\lesssim \sqrt{\frac{\log^3(n+d)}{n}}\left(\alpha_1^{\star}+\alpha_2^{\star}\right)+\left(\sqrt{\frac{\log^3(n+d)}{n}}+\frac{\kappa\mu^2r^2+\kappa_\omega\mu r}{d}\right)\left(\alpha_3^{\star}+\alpha_4^{\star}\right) \\
&\lesssim \left(\sqrt{\frac{\log^3(n+d)}{n}}+\frac{\kappa\mu^2r^2+\kappa_\omega\mu r}{d}\right)\widetilde{v}_{i,j}.
\end{aligned}$$

To finish up, it remains to develop a lower bound on $\widetilde{v}_{i,j}$. Towards this end, we first observe that

$$\begin{aligned}
\alpha_1^{\star}+\alpha_2^{\star} &= \frac{2(1-p)}{np}\left(S_{i,i}^{\star}S_{j,j}^{\star}+2S_{i,j}^{\star 2}\right)+\frac{1}{np}\left(\omega_i^{\star 2}S_{j,j}^{\star}+\omega_j^{\star 2}S_{i,i}^{\star}\right) \\
&\gtrsim \frac{1}{np}\left\|\boldsymbol{U}_{i,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2^2\left\|\boldsymbol{U}_{j,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2^2+\frac{\sigma_r^{\star 2}}{np}\left(\omega_j^{\star 2}\left\|\boldsymbol{U}_{i,\cdot}^{\star}\right\|_2^2+\omega_i^{\star 2}\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2^2\right),
\end{aligned}$$

where we have made use of the assumption $1-p\asymp 1$ and the elementary inequality $S_{i,i}^{\star}=\left\|\boldsymbol{U}_{i,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2^2\ge\sigma_r^{\star 2}\left\|\boldsymbol{U}_{i,\cdot}^{\star}\right\|_2^2$. In view of (E.95), the assumption $1-p\asymp 1$ as well as the bound $\left\|\boldsymbol{U}_{j,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2^2\le\sigma_1^{\star 2}\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2^2$, we can further lower bound

$$\begin{aligned}
\alpha_3^{\star} &\gtrsim \frac{1}{np^2}\left[\omega_i^{\star 2}+(1-p)S_{i,i}^{\star}\right]\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2^2\left(\omega_{\min}^2+\frac{\sigma_r^{\star 2}}{d}\right) \\
&\ge \frac{\omega_i^{\star 2}\omega_{\min}^2}{np^2}\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2^2+\frac{\omega_i^{\star 2}}{np^2}\frac{\sigma_r^{\star 2}}{d}\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2^2+\frac{(1-p)S_{i,i}^{\star}}{np^2}\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2^2\frac{\sigma_r^{\star 2}}{d} \\
&\gtrsim \frac{\omega_i^{\star 2}\omega_{\min}^2}{np^2}\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2^2+\frac{\omega_i^{\star 2}\sigma_r^{\star 2}}{ndp^2}\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2^2+\frac{1}{ndp^2\kappa}\left\|\boldsymbol{U}_{i,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2^2\left\|\boldsymbol{U}_{j,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2^2,
\end{aligned}$$

and similarly,

$$\alpha_4^{\star}\gtrsim\frac{\omega_j^{\star 2}\omega_{\min}^2}{np^2}\left\|\boldsymbol{U}_{i,\cdot}^{\star}\right\|_2^2+\frac{\omega_j^{\star 2}\sigma_r^{\star 2}}{ndp^2}\left\|\boldsymbol{U}_{i,\cdot}^{\star}\right\|_2^2+\frac{1}{ndp^2\kappa}\left\|\boldsymbol{U}_{i,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2^2\left\|\boldsymbol{U}_{j,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2^2.$$

The above bounds taken collectively yield

$$\begin{aligned}
\widetilde{v}_{i,j} &= \alpha_1^{\star}+\alpha_2^{\star}+\alpha_3^{\star}+\alpha_4^{\star} \\
&\gtrsim \frac{1}{\min\left\{ndp^2\kappa,np\right\}}\left\|\boldsymbol{U}_{i,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2^2\left\|\boldsymbol{U}_{j,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2^2+\left(\frac{\sigma_r^{\star 2}}{\min\left\{ndp^2,np\right\}}+\frac{\omega_{\min}^2}{np^2}\right)\left(\omega_j^{\star 2}\left\|\boldsymbol{U}_{i,\cdot}^{\star}\right\|_2^2+\omega_i^{\star 2}\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2^2\right)
\end{aligned}$$

as claimed.

E.4.3 Proof of Lemma 24

As before, all the probabilistic arguments in this subsection are conditional on $\boldsymbol{F}$, and it is assumed, unless otherwise noted, that the $\sigma(\boldsymbol{F})$-measurable high-probability event $\mathcal{E}_{\mathsf{good}}$ occurs.

Step 1: Gaussian approximation of $X_{i,j}$ using the Berry-Esseen Theorem. Recalling the definition (D.39) of $X_{i,j}$, let us denote

$$X_{i,j}=\sum_{l=1}^{n}\underbrace{\left\{M_{j,l}^{\natural}E_{i,l}+M_{i,l}^{\natural}E_{j,l}+E_{i,l}\left[\mathcal{P}_{-i,\cdot}\left(\boldsymbol{E}_{\cdot,l}\right)\right]^{\top}\boldsymbol{U}^{\star}\boldsymbol{U}_{j,\cdot}^{\star\top}+E_{j,l}\left[\mathcal{P}_{-j,\cdot}\left(\boldsymbol{E}_{\cdot,l}\right)\right]^{\top}\boldsymbol{U}^{\star}\boldsymbol{U}_{i,\cdot}^{\star\top}\right\}}_{=:Y_l},$$

where the Y_l's are statistically independent. Apply the Berry-Esseen Theorem (cf. Theorem 15) to reach

$$\sup_{t\in\mathbb{R}}\left|\mathbb{P}\left(X_{i,j}/\sqrt{\mathsf{var}\left(X_{i,j}|\boldsymbol{F}\right)}\leq t\,\Big|\,\boldsymbol{F}\right)-\Phi\left(t\right)\right|\lesssim\gamma\left(\boldsymbol{F}\right),$$

where

$$\gamma\left(\boldsymbol{F}\right)\coloneqq\frac{1}{\mathsf{var}^{3/2}\left(X_{i,j}|\boldsymbol{F}\right)}\sum_{l=1}^{n}\mathbb{E}\left[\left|Y_l\right|^3\left|\boldsymbol{F}\right.\right].$$

It thus boils down to controlling the quantity $\gamma(\boldsymbol{F})$, which forms the main content of the remaining proof.

- We first develop a high-probability bound on each $|Y_l|$. For any $l\in[n]$, observe that

  $$\left[\mathcal{P}_{-i,\cdot}\left(\boldsymbol{E}_{\cdot,l}\right)\right]^{\top}\boldsymbol{U}^{\star}\boldsymbol{U}_{j,\cdot}^{\star\top}=\sum_{k:k\neq i}E_{k,l}\left(\boldsymbol{U}^{\star}\boldsymbol{U}^{\star\top}\right)_{k,j}.$$

  It is straightforward to calculate that

  $$L\coloneqq\max_{k:k\neq i}\left|E_{k,l}\left(\boldsymbol{U}^{\star}\boldsymbol{U}^{\star\top}\right)_{k,j}\right|\leq B\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2\left\|\boldsymbol{U}^{\star}\right\|_{2,\infty}\lesssim B\sqrt{\frac{\mu r}{d}}\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2,$$
  $$V\coloneqq\sum_{k:k\neq i}\mathsf{var}\left(E_{k,l}\left(\boldsymbol{U}^{\star}\boldsymbol{U}^{\star\top}\right)_{k,j}\right)\leq\sum_{k=1}^{d}\sigma_{k,l}^2\left(\boldsymbol{U}^{\star}\boldsymbol{U}^{\star\top}\right)_{k,j}^2\leq\sigma_{\mathsf{ub}}^2\left\|\boldsymbol{U}^{\star}\boldsymbol{U}_{j,\cdot}^{\star\top}\right\|_2^2.$$

  In view of the Bernstein inequality (Vershynin, 2018, Theorem 2.8.4), with probability exceeding $1-O((n+d)^{-101})$ we have

  $$\begin{aligned}\left|\left[\mathcal{P}_{-i,\cdot}\left(\boldsymbol{E}_{\cdot,l}\right)\right]^{\top}\boldsymbol{U}^{\star}\boldsymbol{U}_{j,\cdot}^{\star\top}\right|&\lesssim\sqrt{V\log\left(n+d\right)}+L\log\left(n+d\right)\\&\lesssim\sigma_{\mathsf{ub}}\left\|\boldsymbol{U}^{\star}\boldsymbol{U}_{j,\cdot}^{\star\top}\right\|_2\sqrt{\log\left(n+d\right)}+B\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2\sqrt{\frac{\mu r}{d}}\log\left(n+d\right)\\&\lesssim\sigma_{\mathsf{ub}}\sqrt{\log\left(n+d\right)}\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2+B\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2\sqrt{\frac{\mu r}{d}}\log\left(n+d\right).\end{aligned}$$

  By defining two quantities B_i and B_j such that

  $$\max_{l\in[n]}\left|E_{i,l}\right|\leq B_i\qquad\text{and}\qquad\max_{l\in[n]}\left|E_{j,l}\right|\leq B_j,$$

  we can immediately obtain from the preceding inequality that

  $$\left|E_{i,l}\left[\mathcal{P}_{-i,\cdot}\left(\boldsymbol{E}_{\cdot,l}\right)\right]^{\top}\boldsymbol{U}^{\star}\boldsymbol{U}_{j,\cdot}^{\star\top}\right|\lesssim\sigma_{\mathsf{ub}}B_i\sqrt{\log\left(n+d\right)}\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2+BB_i\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2\sqrt{\frac{\mu r}{d}}\log\left(n+d\right).\tag{E.97}$$

  Similarly, we can also show that with probability exceeding $1-O((n+d)^{-101})$,

  $$\left|E_{j,l}\left[\mathcal{P}_{-j,\cdot}\left(\boldsymbol{E}_{\cdot,l}\right)\right]^{\top}\boldsymbol{U}^{\star}\boldsymbol{U}_{i,\cdot}^{\star\top}\right|\lesssim\sigma_{\mathsf{ub}}B_j\sqrt{\log\left(n+d\right)}\left\|\boldsymbol{U}_{i,\cdot}^{\star}\right\|_2+BB_j\left\|\boldsymbol{U}_{i,\cdot}\right\|_2\sqrt{\frac{\mu r}{d_1}}\log\left(n+d\right).$$

  In addition, it is easily seen from the inequality $|M_{i,l}^{\natural}|\leq\|\boldsymbol{U}_{i,\cdot}^{\natural}\|_2\|\boldsymbol{\Sigma}^{\natural}\|\|\boldsymbol{V}^{\natural}\|_{2,\infty}$ that

  $$\begin{aligned}\left|M_{j,l}^{\natural}E_{i,l}+M_{i,l}^{\natural}E_{j,l}\right|&\leq\sigma_1^{\natural}\left\|\boldsymbol{U}_{j,\cdot}^{\natural}\right\|_2\left\|\boldsymbol{V}^{\natural}\right\|_{2,\infty}B_i+\sigma_1^{\natural}\left\|\boldsymbol{U}_{i,\cdot}^{\natural}\right\|_2\left\|\boldsymbol{V}^{\natural}\right\|_{2,\infty}B_j\\&=\sigma_1^{\natural}\left\|\boldsymbol{V}^{\natural}\right\|_{2,\infty}\left(B_i\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2+B_j\left\|\boldsymbol{U}_{i,\cdot}\right\|_2\right).\end{aligned}$$

  Therefore we know that with probability exceeding $1-O((n+d)^{-101})$

  $$\left|Y_l\right|\lesssim\left[\sigma_{\mathsf{ub}}\sqrt{\log\left(n+d\right)}+B\sqrt{\frac{\mu r}{d}}\log\left(n+d\right)+\sigma_1^{\natural}\left\|\boldsymbol{V}^{\natural}\right\|_{2,\infty}\right]\left(B_i\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2+B_j\left\|\boldsymbol{U}_{i,\cdot}\right\|_2\right).$$

Let

$$C_{\mathsf{prob}} \coloneqq \widetilde{C}_1 \left[\sigma_{\mathsf{ub}}\sqrt{\log(n+d)} + B\sqrt{\frac{\mu r}{d}}\log(n+d) + \sigma_1^{\natural}\left\|\boldsymbol{V}^{\natural}\right\|_{2,\infty}\right]\left(B_i\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2 + B_j\left\|\boldsymbol{U}_{i,\cdot}\right\|_2\right)$$

for some sufficiently large constant $\tilde{C} > 0$ such that with probability exceeding $1 - O((n+d)^{-101})$,

$$\max_{l\in[n]} |Y_l| \le C_{\mathsf{prob}}. \tag{E.98}$$

- In addition, we are also in need of a deterministic upper bound on each $|Y_l|$. Observe that

$$\left|\left[\mathcal{P}_{-i,\cdot}\left(\boldsymbol{E}_{\cdot,l}\right)\right]^{\top}\boldsymbol{U}^{\star}\boldsymbol{U}_{j,\cdot}^{\star\top}\right| \le \left\|\mathcal{P}_{-i,\cdot}\left(\boldsymbol{E}_{\cdot,l}\right)\right\|_2\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2 \le \left\|\boldsymbol{E}_{\cdot,l}\right\|_2\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2 \le \sqrt{d}B\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2,$$

and similarly,

$$\left|\left[\mathcal{P}_{-j,\cdot}\left(\boldsymbol{E}_{\cdot,l}\right)\right]^{\top}\boldsymbol{U}^{\star}\boldsymbol{U}_{i,\cdot}^{\star\top}\right| \le \sqrt{d}B\left\|\boldsymbol{U}_{i,\cdot}^{\star}\right\|_2. \tag{E.99}$$

As a result, we can derive

$$\begin{aligned}|Y_l| &\le \sigma_1^{\natural 2}\left\|\boldsymbol{V}^{\natural}\right\|_{2,\infty}\left(B_i\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2 + B_j\left\|\boldsymbol{U}_{i,\cdot}\right\|_2\right) + B_i\sqrt{d}B\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2 + B_j\sqrt{d}B\left\|\boldsymbol{U}_{i,\cdot}^{\star}\right\|_2 \\ &\le \left(\sigma_1^{\natural 2}\left\|\boldsymbol{V}^{\natural}\right\|_{2,\infty} + B\sqrt{d}\right)\left(B_i\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2 + B_j\left\|\boldsymbol{U}_{i,\cdot}\right\|_2\right) =: C_{\mathsf{det}} \end{aligned} \tag{E.100}$$

With the above probabilistic and deterministic bounds in place (see (E.98) and (E.100)), we can decompose $\mathbb{E}\left[|Y_l|^3|\boldsymbol{F}\right]$ for each $l \in [n]$ as follows

$$\begin{aligned}\mathbb{E}\left[|Y_l|^3|\boldsymbol{F}\right] &= \mathbb{E}\left[|Y_l|^3\mathbb{1}_{|Y_l|\le C_{\mathsf{prob}}}|\boldsymbol{F}\right] + \mathbb{E}\left[|Y_l|^3\mathbb{1}_{|Y_l|>C_{\mathsf{prob}}}|\boldsymbol{F}\right] \\ &\le C_{\mathsf{prob}}\mathbb{E}\left[Y_l^2|\boldsymbol{F}\right] + C_{\mathsf{det}}^3\mathbb{P}\left(|Y_l| > C_{\mathsf{prob}}\right) \\ &\lesssim C_{\mathsf{prob}}\mathbb{E}\left[Y_l^2|\boldsymbol{F}\right] + C_{\mathsf{det}}^3(n+d)^{-101}. \end{aligned}$$

Recognizing that $\sum_{l=1}^n \mathbb{E}[Y_l^2|\boldsymbol{F}] = \mathsf{var}(X_{i,j}|\boldsymbol{F})$, we obtain

$$\gamma(\boldsymbol{F}) \le \frac{1}{\mathsf{var}^{3/2}\left(X_{i,j}|\boldsymbol{F}\right)}\sum_{l=1}^n\mathbb{E}\left[|Y_l|^3|\boldsymbol{F}\right] \lesssim \underbrace{\frac{C_{\mathsf{prob}}}{\mathsf{var}^{1/2}\left(X_{i,j}|\boldsymbol{F}\right)}}_{=:\alpha} + \underbrace{\frac{C_{\mathsf{det}}^3(n+d)^{-100}}{\mathsf{var}^{3/2}\left(X_{i,j}|\boldsymbol{F}\right)}}_{\beta}. \tag{E.101}$$

It remains to control the terms $\mathsf{var}\left(X_{i,j}|\boldsymbol{F}\right)$, C_{prob} and C_{det}. On the event $\mathcal{E}_{\mathsf{good}}$, we have learned from Lemma 23 that

$$\begin{aligned}\mathsf{var}^{1/2}\left(X_{i,j}|\boldsymbol{F}\right) \asymp \widetilde{v}_{i,j}^{1/2} &\gtrsim \frac{1}{\sqrt{ndp^2\kappa \wedge np}}\left\|\boldsymbol{U}_{i,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2\left\|\boldsymbol{U}_{j,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2 + \frac{\omega_{\min}}{\sqrt{np}}\left(\omega_j^{\star}\left\|\boldsymbol{U}_{i,\cdot}^{\star}\right\|_2 + \omega_i^{\star}\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2\right) \\ &\quad + \frac{\sigma_r^{\star}}{\sqrt{ndp^2 \wedge np}}\left(\omega_j^{\star}\left\|\boldsymbol{U}_{i,\cdot}^{\star}\right\|_2 + \omega_i^{\star}\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2\right), \end{aligned} \tag{E.102}$$

provided that $n \gg \log^3(n+d)$ and $d \gg \kappa\mu^2 r^2 + \kappa_\omega \mu r$. Moreover, it is seen from (D.18) that

$$B_i \lesssim \frac{1}{p}\sqrt{\frac{\log(n+d)}{n}}\left(\left\|\boldsymbol{U}_{i,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2 + \omega_i^{\star}\right) \qquad \text{and} \qquad B_j \lesssim \frac{1}{p}\sqrt{\frac{\log(n+d)}{n}}\left(\left\|\boldsymbol{U}_{j,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2 + \omega_j^{\star}\right), \tag{E.103}$$

which immediately lead to

$$B_i\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2 + B_j\left\|\boldsymbol{U}_{i,\cdot}\right\|_2 \lesssim \frac{1}{p}\sqrt{\frac{\log(n+d)}{n}}\left[\frac{1}{\sigma_r^{\star}}\left\|\boldsymbol{U}_{i,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2\left\|\boldsymbol{U}_{j,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2 + \left(\omega_j^{\star}\left\|\boldsymbol{U}_{i,\cdot}^{\star}\right\|_2 + \omega_i^{\star}\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2\right)\right].$$

As a result, we can bound

$$C_{\mathsf{prob}} \lesssim \left[\frac{\sigma_{\mathsf{ub}}}{p}\sqrt{\frac{\log^2(n+d)}{n}} + B\sqrt{\frac{\mu r\log(n+d)}{ndp^2}}\log(n+d) + \sigma_1^\star\frac{\log(n+d)}{np}\right]$$
$$\cdot\left[\frac{1}{\sigma_r^\star}\left\|\boldsymbol{U}_{i,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2\left\|\boldsymbol{U}_{j,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2 + \left(\omega_j^\star\left\|\boldsymbol{U}_{i,\cdot}^\star\right\|_2 + \omega_i^\star\left\|\boldsymbol{U}_{j,\cdot}^\star\right\|_2\right)\right]$$

as well as

$$C_{\mathsf{det}} \lesssim \left(\sigma_1^\star\frac{\log(n+d)}{np} + \frac{B}{p}\sqrt{\frac{d\log(n+d)}{n}}\right)$$
$$\cdot\left[\frac{1}{\sigma_r^\star}\left\|\boldsymbol{U}_{i,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2\left\|\boldsymbol{U}_{j,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2 + \left(\omega_j^\star\left\|\boldsymbol{U}_{i,\cdot}^\star\right\|_2 + \omega_i^\star\left\|\boldsymbol{U}_{j,\cdot}^\star\right\|_2\right)\right],$$

where we have utilized (D.13) and (D.16). In what follows, we bound the terms α and β in (E.101) separately.

- Regarding α, one first observes that

$$\begin{aligned}
\alpha\widetilde{v}_{i,j}^{1/2} &\lesssim \underbrace{\frac{\sigma_{\mathsf{ub}}}{p\sigma_r^\star}\sqrt{\frac{\log^2(n+d)}{n}}\left\|\boldsymbol{U}_{i,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2\left\|\boldsymbol{U}_{j,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2}_{=:\alpha_1} + \underbrace{\frac{\sigma_{\mathsf{ub}}}{p}\sqrt{\frac{\log^2(n+d)}{n}}\left(\omega_j^\star\left\|\boldsymbol{U}_{i,\cdot}^\star\right\|_2 + \omega_i^\star\left\|\boldsymbol{U}_{j,\cdot}^\star\right\|_2\right)}_{=:\alpha_2} \\
&\quad + \underbrace{B\sqrt{\frac{\mu r\log^3(n+d)}{ndp^2}}\frac{1}{\sigma_r^\star}\left\|\boldsymbol{U}_{i,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2\left\|\boldsymbol{U}_{j,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2}_{=:\alpha_3} + \underbrace{B\sqrt{\frac{\mu r\log^3(n+d)}{ndp^2}}\left(\omega_j^\star\left\|\boldsymbol{U}_{i,\cdot}^\star\right\|_2 + \omega_i^\star\left\|\boldsymbol{U}_{j,\cdot}^\star\right\|_2\right)}_{=:\alpha_4} \\
&\quad + \underbrace{\frac{\sqrt{\kappa}\log(n+d)}{np}\left\|\boldsymbol{U}_{i,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2\left\|\boldsymbol{U}_{j,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2}_{=:\alpha_5} + \underbrace{\sigma_1^\star\frac{\log(n+d)}{np}\left(\omega_j^\star\left\|\boldsymbol{U}_{i,\cdot}^\star\right\|_2 + \omega_i^\star\left\|\boldsymbol{U}_{j,\cdot}^\star\right\|_2\right)}_{=:\alpha_6} \\
&\lesssim \frac{1}{\sqrt{\log(n+d)}}\widetilde{v}_{i,j}^{1/2}.
\end{aligned}$$

  To see why the last step holds, we note that due to (D.17), (D.18), the following inequalities hold:

$$\begin{aligned}
\alpha_1 &\asymp \frac{1}{p\sigma_r^\star}\left(\sqrt{\frac{\mu r\log(n+d)}{ndp}}\sigma_1^\star + \frac{\omega_{\max}}{\sqrt{np}}\right)\sqrt{\frac{\log^2(n+d)}{n}}\left\|\boldsymbol{U}_{i,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2\left\|\boldsymbol{U}_{j,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2 \\
&\lesssim \frac{1}{\sqrt{\log(n+d)}}\left(\frac{1}{\sqrt{ndp^2\kappa}} + \frac{1}{\sqrt{np}}\right)\left\|\boldsymbol{U}_{i,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2\left\|\boldsymbol{U}_{j,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2 \lesssim \frac{1}{\sqrt{\log(n+d)}}\widetilde{v}_{i,j}^{1/2}, \\
\alpha_2 &\asymp \frac{1}{p}\left(\sqrt{\frac{\mu r\log(n+d)}{ndp}}\sigma_1^\star + \frac{\omega_{\max}}{\sqrt{np}}\right)\sqrt{\frac{\log^2(n+d)}{n}}\left(\omega_j^\star\left\|\boldsymbol{U}_{i,\cdot}^\star\right\|_2 + \omega_i^\star\left\|\boldsymbol{U}_{j,\cdot}^\star\right\|_2\right) \\
&\lesssim \frac{1}{\sqrt{\log(n+d)}}\left(\frac{1}{\sqrt{ndp^2}} + \frac{1}{\sqrt{np}}\right)\sigma_r^\star\left(\omega_j^\star\left\|\boldsymbol{U}_{i,\cdot}^\star\right\|_2 + \omega_i^\star\left\|\boldsymbol{U}_{j,\cdot}^\star\right\|_2\right) \lesssim \frac{1}{\sqrt{\log(n+d)}}\widetilde{v}_{i,j}^{1/2}, \\
\alpha_3 &\asymp \left(\frac{1}{p}\sqrt{\frac{\kappa\mu^2r^2\log^4(n+d)}{n^2d^2p^2}} + \frac{\omega_{\max}}{p\sigma_r^\star}\sqrt{\frac{\mu r\log^4(n+d)}{n^2dp^2}}\right)\left\|\boldsymbol{U}_{i,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2\left\|\boldsymbol{U}_{j,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2 \\
&\lesssim \frac{1}{\sqrt{\log(n+d)}}\frac{1}{\sqrt{ndp^2\kappa}}\left\|\boldsymbol{U}_{i,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2\left\|\boldsymbol{U}_{j,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2 \lesssim \frac{1}{\sqrt{\log(n+d)}}\widetilde{v}_{i,j}^{1/2}, \\
\alpha_4 &\asymp \left(\frac{\mu r\log^2(n+d)}{ndp^2}\sigma_1^\star + \frac{\omega_{\max}}{p}\sqrt{\frac{\mu r\log^4(n+d)}{n^2dp^2}}\right)\left(\omega_j^\star\left\|\boldsymbol{U}_{i,\cdot}^\star\right\|_2 + \omega_i^\star\left\|\boldsymbol{U}_{j,\cdot}^\star\right\|_2\right)
\end{aligned}$$

$$\lesssim \frac{1}{\sqrt{\log(n+d)}} \frac{\sigma_r^\star}{\sqrt{ndp^2}} \left(\omega_j^\star \left\|\boldsymbol{U}_{i,\cdot}^\star\right\|_2 + \omega_i^\star \left\|\boldsymbol{U}_{j,\cdot}^\star\right\|_2\right) \lesssim \frac{1}{\sqrt{\log(n+d)}} \widetilde{v}_{i,j}^{1/2},$$

$$\alpha_5 \lesssim \frac{1}{\sqrt{\log(n+d)}} \frac{1}{\sqrt{np}} \left\|\boldsymbol{U}_{i,\cdot}^\star \boldsymbol{\Sigma}^\star\right\|_2 \left\|\boldsymbol{U}_{j,\cdot}^\star \boldsymbol{\Sigma}^\star\right\|_2 \lesssim \frac{1}{\sqrt{\log(n+d)}} \widetilde{v}_{i,j}^{1/2},$$

$$\alpha_6 \lesssim \frac{1}{\sqrt{\log(n+d)}} \frac{\sigma_r^\star}{\sqrt{np}} \left(\omega_j^\star \left\|\boldsymbol{U}_{i,\cdot}^\star\right\|_2 + \omega_i^\star \left\|\boldsymbol{U}_{j,\cdot}^\star\right\|_2\right) \lesssim \frac{1}{\sqrt{\log(n+d)}} \widetilde{v}_{i,j}^{1/2},$$

provided that $np \gtrsim \kappa^2 \mu r \log^4(n+d)$, $ndp^2 \gtrsim \kappa^2 \mu^2 r^2 \log^5(n+d)$,

$$\omega_{\max} \sqrt{\frac{d}{np}} \lesssim \frac{1}{\sqrt{\kappa \log^3(n+d)}}, \qquad \text{and} \qquad \frac{\omega_{\max}}{\sigma_r^\star} \sqrt{\frac{1}{np^2}} \lesssim \frac{1}{\sqrt{\kappa \mu r \log^5(n+d)}}. \tag{E.104}$$

In view of (E.65), we know that the second condition in (E.104) can be guaranteed by $ndp^2 \gtrsim \kappa \mu r \log^5(n+d)$ and

$$\frac{\omega_{\max}^2}{p\sigma_r^{\star 2}} \sqrt{\frac{d}{n}} \lesssim \frac{1}{\sqrt{\kappa \mu r \log^5(n+d)}}.$$

As a result, we conclude that

$$\alpha \lesssim \frac{1}{\sqrt{\log(n+d)}}$$

holds as long as $np \gtrsim \kappa^2 \mu r \log^4(n+d)$, $ndp^2 \gtrsim \kappa^2 \mu^2 r^2 \log^5(n+d)$,

$$\omega_{\max} \sqrt{\frac{d}{np}} \lesssim \frac{1}{\sqrt{\kappa \log^3(n+d)}}, \qquad \text{and} \qquad \frac{\omega_{\max}^2}{p\sigma_r^{\star 2}} \sqrt{\frac{d}{n}} \lesssim \frac{1}{\sqrt{\kappa \mu r \log^5(n+d)}}.$$

- Turning to the term β, we make the observation that

$$\begin{aligned}
C_{\mathsf{det}} \widetilde{v}_{i,j}^{-1/2} &\overset{\text{(i)}}{\lesssim} \left(\sigma_1^\star \frac{\log(n+d)}{np} + \frac{B}{p}\sqrt{\frac{d\log(n+d)}{n}}\right) \left[\frac{1}{\sigma_r^\star} \left\|\boldsymbol{U}_{i,\cdot}^\star \boldsymbol{\Sigma}^\star\right\|_2 \left\|\boldsymbol{U}_{j,\cdot}^\star \boldsymbol{\Sigma}^\star\right\|_2 + \left(\omega_j^\star \left\|\boldsymbol{U}_{i,\cdot}^\star\right\|_2 + \omega_i^\star \left\|\boldsymbol{U}_{j,\cdot}^\star\right\|_2\right)\right] \\
&\quad \cdot \left\{\frac{1}{\sqrt{np}} \left[\left\|\boldsymbol{U}_{i,\cdot}^\star \boldsymbol{\Sigma}^\star\right\|_2 \left\|\boldsymbol{U}_{j,\cdot}^\star \boldsymbol{\Sigma}^\star\right\|_2 + \sigma_r^\star \left(\omega_j^\star \left\|\boldsymbol{U}_{i,\cdot}^\star\right\|_2 + \omega_i^\star \left\|\boldsymbol{U}_{j,\cdot}^\star\right\|_2\right)\right]\right\}^{-1} \\
&\lesssim \left(\sigma_1^\star \frac{\log(n+d)}{np} + \frac{B}{p}\sqrt{\frac{d\log(n+d)}{n}}\right) \frac{\sqrt{np}}{\sigma_r^\star} \overset{\text{(ii)}}{\lesssim} \left(\frac{\log(n+d)}{np} + \frac{1}{p}\sqrt{\frac{d\log(n+d)}{n}}\right)\sqrt{\kappa np} \\
&\lesssim \frac{\sqrt{\kappa}\log(n+d)}{\sqrt{np}} + \sqrt{\frac{\kappa d\log(n+d)}{p}} \overset{\text{(iii)}}{\lesssim} \sqrt{nd}.
\end{aligned}$$

Here (i) follows from (E.102); (ii) holds as long as $B \lesssim \sigma_r^\star$, which can be guaranteed by $ndp^2 \gtrsim \kappa \mu r \log(n+d)$ and

$$\frac{\omega_{\max}}{p\sigma_r^\star} \sqrt{\frac{1}{n}} \lesssim \frac{\omega_{\max}^2}{p\sigma_r^{\star 2}} \sqrt{\frac{d}{n}} + \frac{1}{\sqrt{ndp}} \lesssim \frac{1}{\sqrt{\log(n+d)}};$$

and (iii) holds as long as $np \gtrsim \kappa \log(n+d)$. This immediately results in

$$\beta \lesssim \frac{C_{\mathsf{det}}^3 (n+d)^{-100}}{\widetilde{v}_{i,j}^{3/2}} \lesssim (nd)^{3/2} (n+d)^{-100} \lesssim (n+d)^{-50}.$$

Putting the above pieces together, we have shown that

$$\sup_{t\in\mathbb{R}} \left|\mathbb{P}\left(X_{i,j}/\sqrt{\mathsf{var}(X_{i,j}|\boldsymbol{F})} \le t \,\middle|\, \boldsymbol{F}\right) - \Phi(t)\right| \lesssim \gamma(\boldsymbol{F}) \lesssim \alpha + \beta \lesssim \frac{1}{\sqrt{\log(n+d)}}, \tag{E.105}$$

provided that $np \gtrsim \kappa^2 \mu r \log^4(n+d)$, $ndp^2 \gtrsim \kappa^2 \mu^2 r^2 \log^5(n+d)$,

$$\omega_{\max}\sqrt{\frac{d}{np}} \lesssim \frac{1}{\sqrt{\kappa \log^3(n+d)}}, \qquad \text{and} \qquad \frac{\omega_{\max}^2}{p\sigma_r^{\star 2}}\sqrt{\frac{d}{n}} \lesssim \frac{1}{\sqrt{\kappa\mu r \log^5(n+d)}}.$$

Step 2: derandomizing the conditional variance. In this step, we intend to replace $\mathsf{var}(X_{i,j}|\boldsymbol{F})$ in (E.105) with $\widetilde{v}_{i,j}$. Towards this, it is first observed that

$$\begin{aligned}\mathbb{P}\left(X_{i,j}/\sqrt{\widetilde{v}_{i,j}} \le t \,\middle|\, \boldsymbol{F}\right) - \Phi(t) &= \underbrace{\mathbb{P}\left(X_{i,j}/\sqrt{\widetilde{v}_{i,j}} \le t \,\middle|\, \boldsymbol{F}\right) - \mathbb{P}\left(X_{i,j}/\sqrt{\mathsf{var}(X_{i,j}|\boldsymbol{F})} \le t \,\middle|\, \boldsymbol{F}\right)}_{=:\beta_1} \\ &\quad + \underbrace{\mathbb{P}\left(X_{i,j}/\sqrt{\mathsf{var}(X_{i,j}|\boldsymbol{F})} \le t \,\middle|\, \boldsymbol{F}\right) - \Phi(t)}_{=:\beta_2}.\end{aligned}$$

Regarding β_2, it follows from (E.105) that when $\mathcal{E}_{\mathsf{good}}$ occurs, one has

$$|\beta_2| \lesssim \frac{1}{\sqrt{\log(n+d)}}.$$

We then turn attention to bounding β_1. In view of Lemma 23, we know that when $\mathcal{E}_{\mathsf{good}}$ happens, one has

$$\mathsf{var}(X_{i,j}|\boldsymbol{F}) = \widetilde{v}_{i,j} + O\left(\sqrt{\frac{\log^3(n+d)}{n}} + \frac{\kappa\mu^2 r^2}{d}\right)\widetilde{v}_{i,j}. \tag{E.106}$$

An immediate consequence is that

$$\frac{1}{2}\widetilde{v}_{i,j} \le \mathsf{var}(X_{i,j}|\boldsymbol{F}) \le 2\widetilde{v}_{i,j}, \tag{E.107}$$

with the proviso that $n \gg \log^3(n+d)$ and $d \gg \kappa\mu^2 r^2$. Taking the above two equations collectively yields

$$\left|\sqrt{\mathsf{var}(X_{i,j}|\boldsymbol{F})} - \sqrt{\widetilde{v}_{i,j}}\right| = \frac{|\mathsf{var}(X_{i,j}|\boldsymbol{F}) - \widetilde{v}_{i,j}|}{\sqrt{\mathsf{var}(X_{i,j}|\boldsymbol{F})} + \sqrt{\widetilde{v}_{i,j}}} \le \underbrace{\tilde{C}\left(\sqrt{\frac{\log(n+d)^3}{n}} + \frac{\kappa\mu^2 r^2}{d}\right)}_{=:\delta}\sqrt{\widetilde{v}_{i,j}} \tag{E.108}$$

for some sufficiently large constant $\tilde{C} > 0$. Consequently, we arrive at

$$\begin{aligned}\mathbb{P}\left(X_{i,j}/\sqrt{\widetilde{v}_{i,j}} \le t \,\middle|\, \boldsymbol{F}\right) &= \mathbb{P}\left(X_{i,j} \le t\sqrt{\widetilde{v}_{i,j}} \,\middle|\, \boldsymbol{F}\right) \overset{\text{(i)}}{\le} \mathbb{P}\left(X_{i,j} \le t\sqrt{\mathsf{var}(X_{i,j}|\boldsymbol{F})} + t\delta\sqrt{\widetilde{v}_{i,j}} \,\middle|\, \boldsymbol{F}\right) \\ &\overset{\text{(ii)}}{\le} \mathbb{P}\left(X_{i,j}/\sqrt{\mathsf{var}(X_{i,j}|\boldsymbol{F})} \le t + \sqrt{2}t\delta \,\middle|\, \boldsymbol{F}\right) \\ &\overset{\text{(iii)}}{\le} \Phi\left(t + \sqrt{2}t\delta\right) + O\left(\frac{1}{\sqrt{\log(n+d)}}\right) \overset{\text{(iv)}}{\le} \Phi(t) + O\left(\frac{1}{\sqrt{\log(n+d)}}\right) \\ &\overset{\text{(v)}}{\le} \mathbb{P}\left(X_{i,j}/\sqrt{\mathsf{var}(X_{i,j}|\boldsymbol{F})} \le t \,\middle|\, \boldsymbol{F}\right) + O\left(\frac{1}{\sqrt{\log(n+d)}}\right).\end{aligned}$$

Here, (i) follows from (E.108); (ii) is a consequence of (E.107); (iii) and (v) comes from (E.105); and (iv) arises from

$$\Phi\left(t + \sqrt{2}t\delta\right) - \Phi(t) = \int_t^{t+\sqrt{2}t\delta} \phi(s)\,\mathrm{d}s \le \sqrt{2}t\phi(t)\,\delta \le \sqrt{2}\cdot\sqrt{\frac{1}{2\pi e}}\delta \lesssim \frac{1}{\sqrt{\log(n+d)}}$$

provided that $n \gtrsim \log^4(n+d)$ and $d \gtrsim \kappa\mu^2 r^2\sqrt{\log(n+d)}$, where we use the fact that $\sup_{t\in\mathbb{R}} t\phi(t) = \phi(1) = \sqrt{1/(2\pi e)}$. Similarly, we can show that

$$\mathbb{P}\left(X_{i,j}/\sqrt{\widetilde{v}_{i,j}} \leq t \,\middle|\, \boldsymbol{F}\right) \geq \left[\mathbb{P}\left(X_{i,j}/\sqrt{\mathsf{var}\left(X_{i,j}|\boldsymbol{F}\right)} \leq t \,\middle|\, \boldsymbol{F}\right) - O\left(\frac{1}{\sqrt{\log\left(n+d\right)}}\right)\right].$$

As a consequence, we arrive at

$$|\beta_1| \lesssim \frac{1}{\sqrt{\log\left(n+d\right)}}.$$

Combine the preceding bounds on β_1 and β_2 to reach

$$\left|\mathbb{P}\left(X_{i,j}/\sqrt{\widetilde{v}_{i,j}} \leq t \,\middle|\, \boldsymbol{F}\right) - \Phi\left(t\right)\right| \leq \left(|\beta_1| + |\beta_2|\right) \lesssim \frac{1}{\sqrt{\log\left(n+d\right)}}$$

for all $t \in \mathbb{R}$.

Step 3: taking higher-order errors into account. By following the same analysis as Step 3 in Appendix E.2.3 (proof of Lemma 17), we know that if one can show

$$\mathbb{P}\left(\widetilde{v}_{i,j}^{-1/2}\left|\Phi_{i,j}\right| \lesssim \log^{-1/2}\left(n+d\right) \,\middle|\, \boldsymbol{F}\right) \geq 1 - O\left(\left(n+d\right)^{-10}\right), \tag{E.109}$$

then it holds that

$$\sup_{t\in\mathbb{R}}\left|\mathbb{P}\left(\widetilde{v}_{i,j}^{-1/2}\left(\boldsymbol{S} - \boldsymbol{M}^{\natural}\boldsymbol{M}^{\natural\top}\right)_{i,j} \leq t \,\middle|\, \boldsymbol{F}\right) - \Phi\left(t\right)\right| \lesssim \frac{1}{\sqrt{\log\left(n+d\right)}}.$$

As a result, we shall focus on proving (E.109) from now on. Recall from Lemma 22 that with probability exceeding $1 - O((n+d)^{-10})$,

$$\begin{aligned}
\left|\Phi_{i,j}\right| \lesssim \zeta_{i,j} \asymp\ & \underbrace{\theta^2\left[\sigma_1^\star\left(\omega_i^\star\left\|\boldsymbol{U}_{j,\cdot}^\star\right\|_2 + \omega_j^\star\left\|\boldsymbol{U}_{i,\cdot}^\star\right\|_2\right) + \left\|\boldsymbol{U}_{i,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2\left\|\boldsymbol{U}_{j,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2\right]}_{=:\gamma_1} \\
& + \underbrace{\sigma_1^{\star 2}\left(\left\|\boldsymbol{U}_{i,\cdot}^\star\right\|_2\zeta_{\mathsf{2nd},j} + \left\|\boldsymbol{U}_{j,\cdot}^\star\right\|_2\zeta_{\mathsf{2nd},i}\right)}_{=:\gamma_2} + \underbrace{\theta^2\omega_i^\star\omega_j^\star}_{=:\gamma_3} \\
& + \underbrace{\sigma_1^{\star 2}\zeta_{\mathsf{2nd},i}\zeta_{\mathsf{2nd},j}}_{=:\gamma_4} + \underbrace{\theta\sigma_1^\star\left(\omega_i^\star\zeta_{\mathsf{2nd},j} + \omega_j^\star\zeta_{\mathsf{2nd},i}\right)}_{=:\gamma_5},
\end{aligned} \tag{E.110}$$

and from Lemma 23 that

$$\begin{aligned}
\widetilde{v}_{i,j}^{1/2} \gtrsim\ & \frac{1}{\sqrt{\min\left\{ndp^2\kappa, np\right\}}}\left\|\boldsymbol{U}_{i,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2\left\|\boldsymbol{U}_{j,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2 + \frac{\omega_{\min}}{\sqrt{np}}\left(\omega_j^\star\left\|\boldsymbol{U}_{i,\cdot}^\star\right\|_2 + \omega_i^\star\left\|\boldsymbol{U}_{j,\cdot}^\star\right\|_2\right) \\
& + \frac{\sigma_r^\star}{\sqrt{ndp^2 \wedge np}}\left(\omega_j^\star\left\|\boldsymbol{U}_{i,\cdot}^\star\right\|_2 + \omega_i^\star\left\|\boldsymbol{U}_{j,\cdot}^\star\right\|_2\right).
\end{aligned}$$

Armed with these bounds, we seek to derive sufficient conditions that guarantee $\zeta_{i,j} \lesssim \delta\widetilde{v}_{i,j}^{1/2}$.

- Regarding the quantity γ_1 defined in (E.110), we note that

$$\begin{aligned}
\gamma_1 \lesssim\ & \left(\frac{\kappa r\log^2\left(n+d\right)}{np} + \frac{\kappa^2\mu r^2\log^3\left(n+d\right)}{ndp^2} + \frac{\omega_{\max}^2}{\sigma_r^{\star 2}}\frac{\kappa r\log^2\left(n+d\right)}{np^2}\right) \\
& \cdot\left[\sigma_1^\star\left(\omega_i^\star\left\|\boldsymbol{U}_{j,\cdot}^\star\right\|_2 + \omega_j^\star\left\|\boldsymbol{U}_{i,\cdot}^\star\right\|_2\right) + \left\|\boldsymbol{U}_{i,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2\left\|\boldsymbol{U}_{j,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2\right] \\
\lesssim\ & \delta\frac{1}{\sqrt{\min\left\{ndp^2\kappa, np\right\}}}\left\|\boldsymbol{U}_{i,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2\left\|\boldsymbol{U}_{j,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2 + \delta\frac{\sigma_r^\star}{\sqrt{\min\left\{ndp^2, np\right\}}}\left(\omega_j^\star\left\|\boldsymbol{U}_{i,\cdot}^\star\right\|_2 + \omega_i^\star\left\|\boldsymbol{U}_{j,\cdot}^\star\right\|_2\right)
\end{aligned}$$

$$\lesssim \delta \widetilde{v}_{i,j}^{1/2},$$

provided that $np \gtrsim \delta^{-2}\kappa^3 r^2 \log^4(n+d)$, $ndp^2 \gtrsim \delta^{-2}\kappa^5\mu^2 r^4 \log^6(n+d)$, and

$$\frac{\omega_{\max}^2}{p\sigma_r^{\star 2}}\sqrt{\frac{d}{n}} \lesssim \frac{\delta}{\kappa^{3/2} r \log^2(n+d)}.$$

- Regarding the term γ_2 defined in (E.110), we recall from the proof of Lemma 17 (Step 3 in Appendix E.2.3) that with probability exceeding $1-O((n+d)^{-10})$,

$$\zeta_{\mathsf{2nd},i} \lesssim \frac{\delta}{\sqrt{\kappa}} \frac{\left\|\boldsymbol{U}_{i,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2 + \omega_i^{\star}}{\sqrt{\min\{ndp^2\kappa, np\}}\sigma_1^{\star}} + \frac{\delta}{\sqrt{\kappa}} \frac{\omega_{\min}\omega_i^{\star}}{\sqrt{np^2}\sigma_1^{\star 2}}, \tag{E.111a}$$

$$\zeta_{\mathsf{2nd},j} \lesssim \frac{\delta}{\sqrt{\kappa}} \frac{\left\|\boldsymbol{U}_{j,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2 + \omega_j^{\star}}{\sqrt{\min\{ndp^2\kappa, np\}}\sigma_1^{\star}} + \frac{\delta}{\sqrt{\kappa}} \frac{\omega_{\min}\omega_j^{\star}}{\sqrt{np^2}\sigma_1^{\star 2}}, \tag{E.111b}$$

provided that $d \gtrsim \delta^{-2}\kappa^8\mu^3 r^3 \kappa_\omega^2 \log^4(n+d)$,

$$ndp^2 \gtrsim \delta^{-2}\kappa^{10}\mu^4 r^4 \kappa_\omega^2 \log^8(n+d), \qquad np \gtrsim \delta^{-2}\kappa^{10}\mu^3 r^3 \kappa_\omega^2 \log^6(n+d),$$

$$\text{and} \qquad \frac{\omega_{\max}^2}{p\sigma_r^{\star 2}}\sqrt{\frac{d}{n}} \lesssim \frac{\delta}{\sqrt{\kappa^7\mu^2 r^2 \kappa_\omega \log^6(n+d)}}, \qquad \frac{\omega_{\max}}{\sigma_r^{\star}}\sqrt{\frac{d}{np}} \lesssim \frac{\delta}{\sqrt{\kappa^8\mu^2 r^2 \kappa_\omega \log^5(n+d)}}.$$

Consequently, we obtain

$$\begin{aligned}
\gamma_2 &\lesssim \delta \frac{\left\|\boldsymbol{U}_{i,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2 \left\|\boldsymbol{U}_{j,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2}{\sqrt{\min\{ndp^2\kappa, np\}}} + \delta\sigma_r^{\star} \frac{\omega_j^{\star}\left\|\boldsymbol{U}_{i,\cdot}^{\star}\right\|_2 + \omega_i^{\star}\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2}{\sqrt{\min\{ndp^2\kappa, np\}}} \\
&\quad + \frac{\delta}{\sqrt{\kappa}} \frac{\omega_{\min}}{\sqrt{np^2}} \left(\omega_j^{\star}\left\|\boldsymbol{U}_{i,\cdot}^{\star}\right\|_2 + \omega_i^{\star}\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2\right) \\
&\lesssim \delta \widetilde{v}_{i,j}^{1/2}.
\end{aligned}$$

- We now move on to the term γ_3 defined in (E.110), which obeys

$$\gamma_3 \asymp \underbrace{\frac{\kappa r \log^2(n+d)}{np}\omega_i^{\star}\omega_j^{\star}}_{\gamma_{3,1}} + \underbrace{\frac{\kappa^2\mu r^2 \log^3(n+d)}{ndp^2}\omega_i^{\star}\omega_j^{\star}}_{=:\gamma_{3,2}} + \underbrace{\frac{\omega_{\max}^2\omega_i^{\star}\omega_j^{\star}}{\sigma_r^{\star 2}} \frac{\kappa r \log^2(n+d)}{np^2}}_{=:\gamma_{3,3}} \lesssim \delta \widetilde{v}_{i,j}^{1/2}.$$

Here, the last relation holds since

$$\gamma_{3,1} \lesssim \delta \frac{\sigma_r^{\star}}{\sqrt{np}} \left(\omega_j^{\star}\left\|\boldsymbol{U}_{i,\cdot}^{\star}\right\|_2 + \omega_i^{\star}\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2\right) \lesssim \delta \widetilde{v}_{i,j}^{1/2},$$

$$\gamma_{3,2} + \gamma_{3,3} \lesssim \delta \frac{\omega_{\min}}{\sqrt{np}} \left(\omega_j^{\star}\left\|\boldsymbol{U}_{i,\cdot}^{\star}\right\|_2 + \omega_i^{\star}\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2\right) \lesssim \delta \widetilde{v}_{i,j}^{1/2},$$

with the proviso that

$$\left\|\boldsymbol{U}_{i,\cdot}^{\star}\right\|_2 + \left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2 \gtrsim \delta^{-1}\kappa r \log^2(n+d) \left[\frac{\omega_{\max}}{\sigma_r^{\star}}\sqrt{\frac{d}{np}} + \kappa_\omega \frac{\omega_{\max}^2}{p\sigma_r^{\star 2}}\sqrt{\frac{d}{n}} + \frac{\kappa\mu r\kappa_\omega \log(n+d)}{\sqrt{ndp^2}}\right]\sqrt{\frac{1}{d}}. \tag{E.112}$$

- With regards to the term γ_4 defined in (E.110), we can see from (E.111) that

$$\gamma_4 \lesssim \delta^2\sigma_r^{\star 2} \left(\frac{\left\|\boldsymbol{U}_{i,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2 + \omega_i^{\star}}{\sqrt{\min\{ndp^2\kappa, np\}}\sigma_1^{\star}} + \frac{\omega_{\min}\omega_i^{\star}}{\sqrt{np^2}\sigma_1^{\star 2}}\right)\left(\frac{\left\|\boldsymbol{U}_{j,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2 + \omega_j^{\star}}{\sqrt{\min\{ndp^2\kappa, np\}}\sigma_1^{\star}} + \frac{\omega_{\min}\omega_j^{\star}}{\sqrt{np^2}\sigma_1^{\star 2}}\right)$$

$$
\begin{aligned}
&\lesssim \underbrace{\delta^{2} \frac{\left\|\boldsymbol{U}_{i,\cdot}^{\star} \boldsymbol{\Sigma}^{\star}\right\|_{2}\left\|\boldsymbol{U}_{j,\cdot}^{\star} \boldsymbol{\Sigma}^{\star}\right\|_{2}}{\min \left\{n d p^{2} \kappa, n p\right\} \kappa}}_{=: \gamma_{4,1}}+\underbrace{\delta^{2} \sigma_{1}^{\star} \frac{\omega_{i}^{\star}\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_{2}+\omega_{j}^{\star}\left\|\boldsymbol{U}_{i,\cdot}^{\star}\right\|_{2}}{\min \left\{n d p^{2} \kappa, n p\right\} \kappa}}_{=: \gamma_{4,2}}+\underbrace{\delta^{2} \frac{\omega_{i}^{\star} \omega_{j}^{\star}}{\min \left\{n d p^{2} \kappa, n p\right\} \kappa}}_{=: \gamma_{4,3}} \\
&\quad+\underbrace{\delta^{2} \frac{\omega_{\min }\left(\omega_{i}^{\star}\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_{2}+\omega_{j}^{\star}\left\|\boldsymbol{U}_{i,\cdot}^{\star}\right\|_{2}\right)}{\sqrt{n p^{2}} \sqrt{\min \left\{n d p^{2} \kappa, n p\right\}} \kappa}}_{=: \gamma_{4,4}}+\underbrace{\delta^{2} \frac{\omega_{\min } \omega_{i}^{\star} \omega_{j}^{\star}}{\sqrt{n p^{2}} \sqrt{\min \left\{n d p^{2} \kappa, n p\right\}} \sigma_{1}^{\star} \kappa}}_{=: \gamma_{4,5}}+\underbrace{\delta^{2} \sigma_{r}^{\star 2} \frac{\omega_{\min }^{2} \omega_{i}^{\star} \omega_{j}^{\star}}{n p^{2} \sigma_{1}^{\star 4}}}_{=: \gamma_{4,6}} \\
&\lesssim \delta \widetilde{v}_{i,j}^{1/2}.
\end{aligned}
$$

Here the last inequality follows from

$$
\begin{aligned}
\gamma_{4,1} &\lesssim \frac{\delta}{\sqrt{\min \left\{n d p^{2} \kappa, n p\right\}}}\left\|\boldsymbol{U}_{i,\cdot}^{\star} \boldsymbol{\Sigma}^{\star}\right\|_{2}\left\|\boldsymbol{U}_{j,\cdot}^{\star} \boldsymbol{\Sigma}^{\star}\right\|_{2} \lesssim \delta \widetilde{v}_{i,j}^{1/2}, \\
\gamma_{4,2} &\lesssim \frac{\delta \sigma_{r}^{\star}}{\sqrt{n d p^{2} \wedge n p}}\left(\omega_{j}^{\star}\left\|\boldsymbol{U}_{i,\cdot}^{\star}\right\|_{2}+\omega_{i}^{\star}\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_{2}\right) \lesssim \delta \widetilde{v}_{i,j}^{1/2}, \\
\gamma_{4,3} &\lesssim \gamma_{3} \lesssim \delta \widetilde{v}_{i,j}^{1/2}, \\
\gamma_{4,4} &\lesssim \delta \frac{\omega_{\min }}{\sqrt{n p}}\left(\omega_{j}^{\star}\left\|\boldsymbol{U}_{i,\cdot}^{\star}\right\|_{2}+\omega_{i}^{\star}\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_{2}\right) \lesssim \delta \widetilde{v}_{i,j}^{1/2}, \\
\gamma_{4,5} &\lesssim \frac{\omega_{\min }}{\sqrt{n p^{2}} \sigma_{1}^{\star}} \gamma_{4,3} \lesssim\left(\frac{\omega_{\min }^{2}}{p \sigma_{1}^{\star}} \sqrt{\frac{d}{n}}+\frac{1}{\sqrt{n d p^{2}}}\right) \gamma_{4,3} \lesssim \delta \widetilde{v}_{i,j}^{1/2}, \\
\gamma_{4,6} &\lesssim \delta \frac{\omega_{\min }}{\sqrt{n p}}\left(\omega_{j}^{\star}\left\|\boldsymbol{U}_{i,\cdot}^{\star}\right\|_{2}+\omega_{i}^{\star}\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_{2}\right) \lesssim \delta \widetilde{v}_{i,j}^{1/2},
\end{aligned}
$$

provided that $\delta \lesssim 1$, $ndp^2 \gtrsim 1$, $np \gtrsim 1$,

$$
\frac{\omega_{\max }^{2}}{p \sigma_{r}^{\star}} \sqrt{\frac{d}{n}} \lesssim 1
$$

and

$$
\left\|\boldsymbol{U}_{i,\cdot}^{\star}\right\|_{2}+\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_{2} \gtrsim \frac{\omega_{\max }^{2}}{p \sigma_{r}^{\star 2}} \sqrt{\frac{d}{n}} \cdot \sqrt{\frac{1}{d}},
$$

which is guaranteed by (E.112).

- We are left with γ_5, where we can utilize (E.111) to achieve

$$
\begin{aligned}
\gamma_{5} &\lesssim \theta \sigma_{1}^{\star}\left(\omega_{i}^{\star} \zeta_{\mathsf{2nd}, j}+\omega_{j}^{\star} \zeta_{\mathsf{2nd}, i}\right) \\
&\lesssim \underbrace{\delta \theta \sigma_{r}^{\star} \frac{\omega_{i}^{\star}\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_{2}+\omega_{j}^{\star}\left\|\boldsymbol{U}_{i,\cdot}^{\star}\right\|_{2}}{\sqrt{\min \left\{n d p^{2} \kappa, n p\right\}}}}_{=: \gamma_{5,1}}+\underbrace{\delta \theta \frac{\omega_{i}^{\star} \omega_{j}^{\star}}{\sqrt{\min \left\{n d p^{2} \kappa, n p\right\} \kappa}}}_{=: \gamma_{5,2}}+\underbrace{\delta \theta \frac{\omega_{\min } \omega_{i}^{\star} \omega_{j}^{\star}}{\sqrt{n p^{2} \kappa} \sigma_{1}^{\star}}}_{=: \gamma_{5,3}} \\
&\lesssim \delta \widetilde{v}_{i,j}^{1/2}.
\end{aligned}
$$

Here the last relation holds since

$$
\begin{aligned}
\gamma_{5,1}+\gamma_{5,2} &\lesssim \delta \frac{\sigma_{r}^{\star}}{\sqrt{n d p^{2} \wedge n p}}\left(\omega_{j}^{\star}\left\|\boldsymbol{U}_{i,\cdot}^{\star}\right\|_{2}+\omega_{i}^{\star}\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_{2}\right) \lesssim \delta \widetilde{v}_{i,j}^{1/2}, \\
\gamma_{5,3} &\lesssim \delta \frac{\omega_{\min }}{\sqrt{n p}}\left(\omega_{j}^{\star}\left\|\boldsymbol{U}_{i,\cdot}^{\star}\right\|_{2}+\omega_{i}^{\star}\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_{2}\right) \lesssim \delta \widetilde{v}_{i,j}^{1/2},
\end{aligned}
$$

provided that $\theta \lesssim 1$ and

$$
\left\|\boldsymbol{U}_{i,\cdot}^{\star}\right\|_{2}+\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_{2} \gtrsim \theta \frac{\omega_{\max }}{\sigma_{1}^{\star}} \asymp\left(\frac{\omega_{\max }}{\sigma_{1}^{\star}} \sqrt{\frac{d \kappa}{n p}}+\frac{\omega_{\max }}{\sigma_{1}^{\star}} \sqrt{\frac{\kappa^{2} \mu r \log (n+d)}{n p^{2}}}+\frac{\omega_{\max }^{2}}{p \sigma_{r}^{\star 2}} \sqrt{\frac{d}{n}}\right) \sqrt{\frac{r \log ^{2}(n+d)}{d}},
$$

which can be guaranteed by

$$\left\|\boldsymbol{U}_{i,\cdot}^{\star}\right\|_{2}+\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_{2}\gtrsim\left(\frac{\omega_{\max}}{\sigma_{1}^{\star}}\sqrt{\frac{d\kappa}{np}}+\frac{\omega_{\max}^{2}}{p\sigma_{r}^{\star2}}\sqrt{\frac{d}{n}}+\frac{\kappa^{2}\mu r\log\left(n+d\right)}{\sqrt{ndp^{2}}}\right)\sqrt{\frac{r\log^{2}\left(n+d\right)}{d}}$$

where we have used (E.64) and a result from the AM-GM inequality

$$\frac{\omega_{\max}}{\sigma_{r}^{\star}}\sqrt{\frac{\kappa^{2}\mu r\log\left(n+d\right)}{np^{2}}}\lesssim\frac{\omega_{\max}^{2}}{\sigma_{r}^{\star2}}\sqrt{\frac{d}{np^{2}}}+\frac{\kappa^{2}\mu r\log\left(n+d\right)}{\sqrt{ndp^{2}}}.$$

As a consequence, we have demonstrated that with probability exceeding $1-O((n+d)^{-10})$,

$$\left(\widetilde{v}_{i,j}\right)^{-1/2}\zeta_{i,j}\lesssim\left(\widetilde{v}_{i,j}\right)^{-1/2}\left(\gamma_{1}+\gamma_{2}+\gamma_{3}+\gamma_{4}+\gamma_{5}\right)\lesssim\delta$$

holds as long as $d\gtrsim\delta^{-2}\kappa^{8}\mu^{3}r^{3}\kappa_{\omega}^{2}\log^{4}(n+d)$,

$$ndp^{2}\gtrsim\delta^{-2}\kappa^{10}\mu^{4}r^{4}\kappa_{\omega}^{2}\log^{8}\left(n+d\right),\qquad np\gtrsim\delta^{-2}\kappa^{10}\mu^{3}r^{3}\kappa_{\omega}^{2}\log^{6}\left(n+d\right),$$

$$\frac{\omega_{\max}^{2}}{p\sigma_{r}^{\star2}}\sqrt{\frac{d}{n}}\lesssim\frac{\delta}{\sqrt{\kappa^{7}\mu^{2}r^{2}\kappa_{\omega}\log^{6}\left(n+d\right)}},\qquad\frac{\omega_{\max}}{\sigma_{r}^{\star}}\sqrt{\frac{d}{np}}\lesssim\frac{\delta}{\sqrt{\kappa^{8}\mu^{2}r^{2}\kappa_{\omega}\log^{5}\left(n+d\right)}},$$

and

$$\left\|\boldsymbol{U}_{i,\cdot}^{\star}\right\|_{2}+\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_{2}\gtrsim\delta^{-1}\kappa r\log^{2}\left(n+d\right)\left[\frac{\omega_{\max}}{\sigma_{r}^{\star}}\sqrt{\frac{d}{np}}+\kappa_{\omega}\frac{\omega_{\max}^{2}}{p\sigma_{r}^{\star2}}\sqrt{\frac{d}{n}}+\frac{\kappa\mu r\kappa_{\omega}\log\left(n+d\right)}{\sqrt{ndp^{2}}}\right]\sqrt{\frac{1}{d}}.$$

Taking $\delta\asymp\log^{-1/2}(n+d)$ in the above bounds directly establishes the advertised result.

E.4.4 Proof of Lemma 25

Define

$$a_{l}=n^{-1}(\boldsymbol{U}_{i,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\boldsymbol{f}_{l})(\boldsymbol{U}_{j,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\boldsymbol{f}_{l})$$

for each $l=1,\dots,n$. In view of the expression (D.44), we can write

$$A_{i,j}=\sum_{l=1}^{n}\left(a_{l}-\mathbb{E}\left[a_{l}\right]\right).$$

Apply Theorem 15 to show that

$$\sup_{z\in\mathbb{R}}\left|\mathbb{P}\left(\overline{v}_{i,j}^{-1/2}A_{i,j}\leq z\right)-\Phi\left(z\right)\right|\lesssim\gamma,$$

where $\overline{v}_{i,j}$ is defined in (D.45) and

$$\gamma\coloneqq\overline{v}_{i,j}^{-3/2}\sum_{l=1}^{n}\mathbb{E}\left[\left|a_{l}^{3}\right|\right].$$

It remains to control the term γ. Given that $\boldsymbol{U}_{i,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\boldsymbol{f}_{l}\sim\mathcal{N}(0,\|\boldsymbol{U}_{i,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\|_{2}^{2})$ and $\boldsymbol{U}_{j,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\boldsymbol{f}_{l}\sim\mathcal{N}(0,\|\boldsymbol{U}_{j,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\|_{2}^{2})$, it is straightforward to check that for each $l\in[n]$,

$$\mathbb{E}\left[\left|a_{l}^{3}\right|\right]\leq\frac{1}{n^{3}}\mathbb{E}\left[\left(\boldsymbol{U}_{i,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\boldsymbol{f}_{l}\right)^{6}\right]^{1/2}\mathbb{E}\left[\left(\boldsymbol{U}_{j,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\boldsymbol{f}_{l}\right)^{6}\right]^{1/2}\lesssim\frac{1}{n^{3}}\left\|\boldsymbol{U}_{i,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_{2}^{3}\left\|\boldsymbol{U}_{j,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_{2}^{3}.$$

Recognizing that

$$\overline{v}_{i,j}=\frac{1}{n}\left(S_{i,i}^{\star}S_{j,j}^{\star}+S_{i,j}^{\star2}\right)\geq\frac{1}{n}S_{i,i}^{\star}S_{j,j}^{\star}=\frac{1}{n}\left\|\boldsymbol{U}_{i,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_{2}^{2}\left\|\boldsymbol{U}_{j,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_{2}^{2},$$

we can combine the above bounds to arrive at

$$\gamma=\overline{v}_{i,j}^{-3/2}\sum_{l=1}^{n}\mathbb{E}\left[\left|a_{l}^{3}\right|\right]\lesssim\frac{n^{-2}\left\|\boldsymbol{U}_{i,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_{2}^{3}\left\|\boldsymbol{U}_{j,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_{2}^{3}}{n^{-3/2}\left\|\boldsymbol{U}_{i,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_{2}^{3}\left\|\boldsymbol{U}_{j,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_{2}^{3}}\lesssim\frac{1}{\sqrt{n}}$$

as claimed.

E.4.5 Proof of Lemma 26

For any $z \in \mathbb{R}$, we can decompose

$$\begin{aligned}\mathbb{P}\left(\left(S_{i,j}-S_{i,j}^{\star}\right)/\sqrt{v_{i,j}^{\star}}\leq z\right) &= \mathbb{P}\left(A_{i,j}+W_{i,j}\leq\sqrt{v_{i,j}^{\star}}z\right)=\mathbb{E}\left[\mathbb{E}\left[\mathbb{1}_{A_{i,j}+W_{i,j}\leq\sqrt{v_{i,j}^{\star}}z}\,\Big|\,\boldsymbol{F}\right]\right]\\ &=\underbrace{\mathbb{E}\left[\mathbb{E}\left[\mathbb{1}_{A_{i,j}+W_{i,j}\leq\sqrt{v_{i,j}^{\star}}z}\mathbb{1}_{\mathcal{E}_{\mathsf{good}}}\,\Big|\,\boldsymbol{F}\right]\right]}_{=:\alpha_1}+\underbrace{\mathbb{E}\left[\mathbb{E}\left[\mathbb{1}_{A_{i,j}+W_{i,j}\leq\sqrt{v_{i,j}^{\star}}z}\mathbb{1}_{\mathcal{E}_{\mathsf{good}}^{\mathsf{c}}}\,\Big|\,\boldsymbol{F}\right]\right]}_{=:\alpha_2},\end{aligned}$$

leaving us with two terms to control.

- Regarding the first term α_1, we note that $\mathcal{E}_{\mathsf{good}}$ is $\sigma(\boldsymbol{F})$-measurable, and consequently,

$$\begin{aligned}\mathbb{E}\left[\mathbb{1}_{A_{i,j}+W_{i,j}\leq\sqrt{v_{i,j}^{\star}}z}\mathbb{1}_{\mathcal{E}_{\mathsf{good}}}\,\Big|\,\boldsymbol{F}\right] &= \mathbb{1}_{\mathcal{E}_{\mathsf{good}}}\mathbb{E}\left[\mathbb{1}_{A_{i,j}+W_{i,j}\leq\sqrt{v_{i,j}^{\star}}z}\,\Big|\,\boldsymbol{F}\right]\\ &= \mathbb{1}_{\mathcal{E}_{\mathsf{good}}}\mathbb{P}\left(W_{i,j}\leq\sqrt{v_{i,j}^{\star}}z-A_{i,j}\,|\,\boldsymbol{F}\right).\end{aligned}$$

  In view of Lemma 24, we can see that on the event $\mathcal{E}_{\mathsf{good}}$,

$$\sup_{t\in\mathbb{R}}\left|\mathbb{P}\left(\left(\widetilde{v}_{i,j}\right)^{-1/2}W_{i,j}\leq t\,|\,\boldsymbol{F}\right)-\Phi\left(t\right)\right|\lesssim\frac{1}{\sqrt{\log\left(n+d\right)}}.$$

  Since $A_{i,j}$ is $\sigma(\boldsymbol{F})$-measurable, by choosing $t=\left(\widetilde{v}_{i,j}\right)^{-1/2}(\sqrt{v_{i,j}^{\star}}z-A_{i,j})$ we have

$$\left|\mathbb{P}\left(W_{i,j}\leq\sqrt{v_{i,j}^{\star}}z-A_{i,j}\,|\,\boldsymbol{F}\right)-\Phi\left(\frac{\sqrt{v_{i,j}^{\star}}z-A_{i,j}}{\widetilde{v}_{i,j}^{1/2}}\right)\right|\mathbb{1}_{\mathcal{E}_{\mathsf{good}}}\lesssim\frac{1}{\sqrt{\log\left(n+d\right)}}.$$

  This in turn leads to

$$\begin{aligned}\alpha_1 &= \mathbb{E}\left[\Phi\left(\frac{\sqrt{v_{i,j}^{\star}}z-A_{i,j}}{\widetilde{v}_{i,j}^{1/2}}\right)\mathbb{1}_{\mathcal{E}_{\mathsf{good}}}\right]+O\left(\frac{1}{\sqrt{\log\left(n+d\right)}}\right)\\ &= \mathbb{E}\left[\Phi\left(\frac{\sqrt{v_{i,j}^{\star}}z-A_{i,j}}{\widetilde{v}_{i,j}^{1/2}}\right)\right]-\mathbb{E}\left[\Phi\left(\frac{\sqrt{v_{i,j}^{\star}}z-A_{i,j}}{\widetilde{v}_{i,j}^{1/2}}\right)\mathbb{1}_{\mathcal{E}_{\mathsf{good}}^{\mathsf{c}}}\right]+O\left(\frac{1}{\sqrt{\log\left(n+d\right)}}\right)\\ &= \mathbb{E}\left[\Phi\left(\frac{\sqrt{v_{i,j}^{\star}}z-A_{i,j}}{\widetilde{v}_{i,j}^{1/2}}\right)\right]+O\left(\frac{1}{\sqrt{\log\left(n+d\right)}}\right),\end{aligned}$$

  where the last identity is valid since

$$\mathbb{E}\left[\Phi\left(\frac{\sqrt{v_{i,j}^{\star}}z-A_{i,j}}{\widetilde{v}_{i,j}^{1/2}}\right)\mathbb{1}_{\mathcal{E}_{\mathsf{good}}^{\mathsf{c}}}\right]\leq\mathbb{P}\left(\mathcal{E}_{\mathsf{good}}^{\mathsf{c}}\right)\lesssim\left(n+d\right)^{-10}\lesssim\frac{1}{\sqrt{\log\left(n+d\right)}}.$$

  Let $\phi(\cdot)$ denote the probability density of $\mathcal{N}(0,1)$, then it is readily seen that

$$\begin{aligned}\mathbb{E}\left[\Phi\left(\frac{\sqrt{v_{i,j}^{\star}}z-A_{i,j}}{\widetilde{v}_{i,j}^{1/2}}\right)\right] &= \mathbb{E}\left[\int_{-\infty}^{+\infty}\phi\left(t\right)\mathbb{1}_{t\leq\left(\widetilde{v}_{i,j}\right)^{-1/2}\left(\sqrt{v_{i,j}^{\star}}z-A_{i,j}\right)}\,\mathrm{d}t\right]\\ &\overset{\text{(i)}}{=}\int_{-\infty}^{+\infty}\mathbb{E}\left[\phi\left(t\right)\mathbb{1}_{t\leq\left(\widetilde{v}_{i,j}\right)^{-1/2}\left(\sqrt{v_{i,j}^{\star}}z-A_{i,j}\right)}\right]\mathrm{d}t\\ &= \int_{-\infty}^{+\infty}\phi\left(t\right)\mathbb{P}\left(t\leq\left(\widetilde{v}_{i,j}\right)^{-1/2}\left(\sqrt{v_{i,j}^{\star}}z-A_{i,j}\right)\right)\mathrm{d}t\\ &= \int_{-\infty}^{+\infty}\phi\left(t\right)\mathbb{P}\left(A_{i,j}\leq\sqrt{v_{i,j}^{\star}}z-t\widetilde{v}_{i,j}^{1/2}\right)\mathrm{d}t\end{aligned}$$

$$\overset{\text{(ii)}}{=} \int_{-\infty}^{+\infty} \phi\left(t\right) \Phi\left[\left(\overline{v}_{i,j}\right)^{-1/2}\left(\sqrt{v_{i,j}^{\star}}z - t\widetilde{v}_{i,j}^{1/2}\right)\right] \mathrm{d}t + O\left(\frac{1}{\sqrt{n}}\right).$$

Here, (i) invokes Fubini's Theorem for nonnegative measurable functions, whereas (ii) follows from Lemma 25. Finally, letting U and V be two independent $\mathcal{N}(0,1)$ random variables, we can readily calculate

$$\begin{aligned}\int_{-\infty}^{+\infty} \phi\left(t\right) \Phi\left(\frac{z\sqrt{v_{i,j}^{\star}} - t\sqrt{\widetilde{v}_{i,j}}}{\sqrt{\overline{v}_{i,j}}}\right) \mathrm{d}t &= \mathbb{P}\left(U \leq \frac{z\sqrt{v_{i,j}^{\star}} - V\sqrt{\widetilde{v}_{i,j}}}{\sqrt{\overline{v}_{i,j}}}\right) \\ &= \mathbb{P}\left(U\sqrt{\overline{v}_{i,j}} + V\sqrt{\widetilde{v}_{i,j}} \leq z\sqrt{v_{i,j}^{\star}}\right) \\ &= \Phi\left(z\right),\end{aligned}$$

where the last relation follows from the fact that $U\sqrt{\overline{v}_{i,j}} + V\sqrt{\widetilde{v}_{i,j}} \sim \mathcal{N}(0, v_{i,j}^{\star})$. This allows us to conclude that

$$\alpha_1 = \Phi\left(z\right) + O\left(\frac{1}{\sqrt{\log\left(n+d\right)}} + \frac{1}{\sqrt{n}}\right).$$

- We then move on to the other term α_2. Towards this, it is straightforward to derive that

$$\alpha_2 \leq \mathbb{P}\left(\mathcal{E}_{\mathsf{good}}^{\mathsf{c}}\right) \lesssim \left(n+d\right)^{-10}.$$

Note that the above analysis holds for all $z \in \mathbb{R}$. Therefore, for any $z \in \mathbb{R}$, taking the above calculation together yields

$$\left|\mathbb{P}\left(\left(S_{i,j} - S_{i,j}^{\star}\right)/\sqrt{v_{i,j}^{\star}} \leq z\right) - \Phi\left(z\right)\right| \leq \left|\alpha_1 - \Phi\left(z\right)\right| + \alpha_2 \lesssim \frac{1}{\sqrt{\log\left(n+d\right)}} + \frac{1}{\sqrt{n}} \asymp \frac{1}{\sqrt{\log\left(n+d\right)}}$$

as claimed, provided that $n \gtrsim \log(n+d)$.

E.5 Auxiliary lemmas for Theorem 14

E.5.1 Proof of Lemma 27

In the sequel, we shall only consider the case when $i \neq j$; the analysis for $i = j$ is similar and in fact simpler, and hence we omit it here for brevity. Let us denote

$$\begin{aligned}v_{i,j}^{\star} = {} & \underbrace{\frac{2-p}{np}S_{i,i}^{\star}S_{j,j}^{\star}}_{=:\alpha_1} + \underbrace{\frac{4-3p}{np}S_{i,j}^{\star 2}}_{=:\alpha_2} + \underbrace{\frac{1}{np}\left(\omega_i^{\star 2}S_{j,j}^{\star} + \omega_j^{\star 2}S_{i,i}^{\star}\right)}_{=:\alpha_3} \\ & + \underbrace{\frac{1}{np^2}\sum_{k=1}^{d}\left[\omega_i^{\star 2} + \left(1-p\right)S_{i,i}^{\star}\right]\left[\omega_k^{\star 2} + \left(1-p\right)S_{k,k}^{\star}\right]\left(\boldsymbol{U}_{k,\cdot}^{\star}\boldsymbol{U}_{j,\cdot}^{\star\top}\right)^2}_{=:\alpha_4} + \underbrace{\frac{2\left(1-p\right)^2}{np^2}\sum_{k=1}^{d}S_{i,k}^{\star 2}\left(\boldsymbol{U}_{k,\cdot}^{\star}\boldsymbol{U}_{j,\cdot}^{\star\top}\right)^2}_{=:\alpha_5} \\ & + \underbrace{\frac{1}{np^2}\sum_{k=1}^{d}\left[\omega_j^{\star 2} + \left(1-p\right)S_{j,j}^{\star}\right]\left[\omega_k^{\star 2} + \left(1-p\right)S_{k,k}^{\star}\right]\left(\boldsymbol{U}_{k,\cdot}^{\star}\boldsymbol{U}_{i,\cdot}^{\star\top}\right)^2}_{=:\alpha_6} + \underbrace{\frac{2\left(1-p\right)^2}{np^2}\sum_{k=1}^{d}S_{j,k}^{\star 2}\left(\boldsymbol{U}_{k,\cdot}^{\star}\boldsymbol{U}_{i,\cdot}^{\star\top}\right)^2}_{=:\alpha_7}\end{aligned}$$

and

$$\begin{aligned}v_{i,j} = {} & \underbrace{\frac{2-p}{np}S_{i,i}S_{j,j}}_{=:\beta_1} + \underbrace{\frac{4-3p}{np}S_{i,j}^{2}}_{=:\beta_2} + \underbrace{\frac{1}{np}\left(\omega_i^{2}S_{j,j} + \omega_j^{2}S_{i,i}\right)}_{=:\beta_3} \\ & + \underbrace{\frac{1}{np^2}\sum_{k=1}^{d}\left[\omega_i^{2} + \left(1-p\right)S_{i,i}\right]\left[\omega_k^{2} + \left(1-p\right)S_{k,k}\right]\left(\boldsymbol{U}_{k,\cdot}\boldsymbol{U}_{j,\cdot}^{\top}\right)^2}_{=:\beta_4} + \underbrace{\frac{2\left(1-p\right)^2}{np^2}\sum_{k=1}^{d}S_{i,k}^{2}\left(\boldsymbol{U}_{k,\cdot}\boldsymbol{U}_{j,\cdot}^{\top}\right)^2}_{=:\beta_5}\end{aligned}$$

$$+\underbrace{\frac{1}{np^2}\sum_{k=1}^{d}\left[\omega_j^2+(1-p)\,S_{j,j}\right]\left[\omega_k^2+(1-p)\,S_{k,k}\right]\left(\boldsymbol{U}_{k,\cdot}\boldsymbol{U}_{i,\cdot}^{\top}\right)^2}_{=:\beta_6}+\underbrace{\frac{2\,(1-p)^2}{np^2}\sum_{k=1}^{d}S_{j,k}^2\left(\boldsymbol{U}_{k,\cdot}\boldsymbol{U}_{i,\cdot}^{\top}\right)^2}_{=:\beta_7}.$$

It follows from Lemma 23 that

$$v_{i,j}^{\star}\gtrsim\frac{1}{\min\left\{ndp^2\kappa,np\right\}}\left\|\boldsymbol{U}_{i,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2^2\left\|\boldsymbol{U}_{j,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2^2+\left(\frac{\sigma_r^{\star 2}}{\min\left\{ndp^2,np\right\}}+\frac{\omega_{\min}^2}{np^2}\right)\left(\omega_j^{\star 2}\left\|\boldsymbol{U}_{i,\cdot}^{\star}\right\|_2^2+\omega_i^{\star 2}\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2^2\right).$$

In view of the bounds on γ_2, γ_3, γ_4 and γ_5 in Step 3 of Appendix E.4.3 as well as (E.111), we know that for any $\varepsilon\in(0,1)$, one has

$$\sigma_1^{\star 2}\left(\left\|\boldsymbol{U}_{i,\cdot}^{\star}\right\|_2\zeta_{\mathsf{2nd},j}+\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2\zeta_{\mathsf{2nd},i}\right)\lesssim\varepsilon\widetilde{v}_{i,j}^{1/2}\lesssim\varepsilon\left(v_{i,j}^{\star}\right)^{1/2},\tag{E.113}$$

and

$$\zeta_{\mathsf{2nd},i}\lesssim\frac{\varepsilon}{\sqrt{\kappa}}\frac{\left\|\boldsymbol{U}_{i,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2+\omega_i^{\star}}{\sqrt{\min\left\{ndp^2\kappa,np\right\}}\sigma_1^{\star}}+\frac{\varepsilon}{\sqrt{\kappa}}\frac{\omega_{\min}\omega_i^{\star}}{\sqrt{np^2}\sigma_1^{\star 2}},\tag{E.114a}$$

$$\zeta_{\mathsf{2nd},j}\lesssim\frac{\varepsilon}{\sqrt{\kappa}}\frac{\left\|\boldsymbol{U}_{j,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2+\omega_j^{\star}}{\sqrt{\min\left\{ndp^2\kappa,np\right\}}\sigma_1^{\star}}+\frac{\varepsilon}{\sqrt{\kappa}}\frac{\omega_{\min}\omega_j^{\star}}{\sqrt{np^2}\sigma_1^{\star 2}},\tag{E.114b}$$

and

$$\theta^2\omega_i^{\star}\omega_j^{\star}+\sigma_1^{\star 2}\zeta_{\mathsf{2nd},i}\zeta_{\mathsf{2nd},j}+\theta\sigma_1^{\star}\left(\omega_i^{\star}\zeta_{\mathsf{2nd},j}+\omega_j^{\star}\zeta_{\mathsf{2nd},i}\right)\lesssim\varepsilon\widetilde{v}_{i,j}^{1/2}\lesssim\varepsilon\left(v_{i,j}^{\star}\right)^{1/2},\tag{E.115}$$

provided that the following conditions hold: $d\gtrsim\varepsilon^{-2}\kappa^8\mu^3r^3\kappa_\omega^2\log^4(n+d)$,

$$ndp^2\gtrsim\varepsilon^{-2}\kappa^{10}\mu^4r^4\kappa_\omega^2\log^8(n+d),\qquad np\gtrsim\varepsilon^{-2}\kappa^{10}\mu^3r^3\kappa_\omega^2\log^6(n+d),$$

$$\frac{\omega_{\max}^2}{p\sigma_r^{\star 2}}\sqrt{\frac{d}{n}}\lesssim\frac{\varepsilon}{\sqrt{\kappa^7\mu^2r^2\kappa_\omega\log^6(n+d)}},\qquad\frac{\omega_{\max}}{\sigma_r^{\star}}\sqrt{\frac{d}{np}}\lesssim\frac{\varepsilon}{\sqrt{\kappa^8\mu^2r^2\kappa_\omega\log^5(n+d)}},$$

and

$$\left\|\boldsymbol{U}_{i,\cdot}^{\star}\right\|_2+\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2\gtrsim\varepsilon^{-1}\kappa r\log^2(n+d)\left[\frac{\omega_{\max}}{\sigma_r^{\star}}\sqrt{\frac{d}{np}}+\kappa_\omega\frac{\omega_{\max}^2}{p\sigma_r^{\star 2}}\sqrt{\frac{d}{n}}+\frac{\kappa\mu r\kappa_\omega\log(n+d)}{\sqrt{ndp^2}}\right]\sqrt{\frac{1}{d}}.$$

These basic facts will be very useful for us to control the difference between $v_{i,j}^{\star}$ and $v_{i,j}$, towards which we shall bound $\alpha_i-\beta_i$, $1\leq i\leq 7$, separately.

Step 1: bounding $|\alpha_1-\beta_1|$. Recall from Lemma 19 that

$$\left|S_{i,i}-S_{i,i}^{\star}\right|\lesssim\left(\theta+\sqrt{\frac{\kappa^3r\log(n+d)}{n}}\right)\left\|\boldsymbol{U}_{i,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2^2+\theta^2\omega_i^{\star 2}+\left(\theta\omega_i^{\star}+\zeta_{\mathsf{2nd},i}\sigma_1^{\star}\right)\left\|\boldsymbol{U}_{i,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2+\zeta_{\mathsf{2nd},i}^2\sigma_1^{\star 2},\tag{E.116a}$$

$$\left|S_{j,j}-S_{j,j}^{\star}\right|\lesssim\left(\theta+\sqrt{\frac{\kappa^3r\log(n+d)}{n}}\right)\left\|\boldsymbol{U}_{j,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2^2+\theta^2\omega_j^{\star 2}+\left(\theta\omega_j^{\star}+\zeta_{\mathsf{2nd},j}\sigma_1^{\star}\right)\left\|\boldsymbol{U}_{j,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2+\zeta_{\mathsf{2nd},j}^2\sigma_1^{\star 2}.\tag{E.116b}$$

We proceed with the following elementary inequality:

$$\begin{aligned}\left|\alpha_1-\beta_1\right|&\lesssim\frac{1}{np}\left|S_{i,i}S_{j,j}-S_{i,i}^{\star}S_{j,j}^{\star}\right|\lesssim\frac{1}{np}S_{i.i}\left|S_{j,j}-S_{j,j}^{\star}\right|+\frac{1}{np}S_{j,j}^{\star}\left|S_{i,i}-S_{i,i}^{\star}\right|\\&\lesssim\underbrace{\frac{1}{np}S_{i.i}^{\star}\left|S_{j,j}-S_{j,j}^{\star}\right|+\frac{1}{np}S_{j,j}^{\star}\left|S_{i,i}-S_{i,i}^{\star}\right|}_{=:\gamma_{1,1}}+\underbrace{\frac{1}{np}\left|S_{i,i}-S_{i,i}^{\star}\right|\left|S_{j,j}-S_{j,j}^{\star}\right|}_{=:\gamma_{1,2}}.\end{aligned}$$

- Regarding $\gamma_{1,1}$, it is seen that

$$\gamma_{1,1} \lesssim \underbrace{\frac{1}{np}\left(\theta+\sqrt{\frac{\kappa^3 r\log(n+d)}{n}}\right)\left\|\boldsymbol{U}_{i,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2^2\left\|\boldsymbol{U}_{j,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2^2}_{=:\gamma_{1,1,1}}+\underbrace{\frac{1}{np}\theta^2\left(\omega_j^{\star 2}\left\|\boldsymbol{U}_{i,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2^2+\omega_i^{\star 2}\left\|\boldsymbol{U}_{j,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2^2\right)}_{=:\gamma_{1,1,2}}$$
$$+\underbrace{\frac{1}{np}\theta\left(\omega_i^{\star}\left\|\boldsymbol{U}_{i,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2\left\|\boldsymbol{U}_{j,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2^2+\omega_j^{\star}\left\|\boldsymbol{U}_{j,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2\left\|\boldsymbol{U}_{i,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2^2\right)}_{=:\gamma_{1,1,3}}$$
$$+\underbrace{\frac{1}{np}\sigma_1^{\star 2}\left(\left\|\boldsymbol{U}_{i,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2^2\zeta_{\mathsf{2nd},j}^2+\left\|\boldsymbol{U}_{j,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2^2\zeta_{\mathsf{2nd},i}^2\right)}_{=:\gamma_{1,1,4}}$$
$$+\underbrace{\frac{1}{np}\sigma_1^{\star}\left\|\boldsymbol{U}_{i,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2\left\|\boldsymbol{U}_{j,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2\left(\left\|\boldsymbol{U}_{i,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2\zeta_{\mathsf{2nd},j}+\left\|\boldsymbol{U}_{j,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2\zeta_{\mathsf{2nd},i}\right)}_{=:\gamma_{1,1,5}}\lesssim\delta v_{i,j}^{\star},$$

where the last relation holds true since

$$\begin{aligned}
\gamma_{1,1,1} &\lesssim \frac{\delta}{np}\left\|\boldsymbol{U}_{i,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2^2\left\|\boldsymbol{U}_{j,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2^2\lesssim\delta v_{i,j}^{\star},\\
\gamma_{1,1,2} &\lesssim \delta\frac{\sigma_r^{\star 2}}{np}\left(\omega_j^{\star 2}\left\|\boldsymbol{U}_{i,\cdot}^{\star}\right\|_2^2+\omega_i^{\star 2}\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2^2\right)\lesssim\delta v_{i,j}^{\star},\\
\gamma_{1,1,3} &\lesssim \frac{\delta}{\sqrt{\kappa}}\frac{1}{np}\left(\omega_i^{\star}\left\|\boldsymbol{U}_{i,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2\left\|\boldsymbol{U}_{j,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2^2+\omega_j^{\star}\left\|\boldsymbol{U}_{j,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2\left\|\boldsymbol{U}_{i,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2^2\right)\\
&\overset{\text{(i)}}{\lesssim}\delta\frac{1}{np}\left\|\boldsymbol{U}_{i,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2^2\left\|\boldsymbol{U}_{j,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2^2+\delta\frac{\sigma_r^{\star 2}}{np}\left(\omega_j^{\star 2}\left\|\boldsymbol{U}_{i,\cdot}^{\star}\right\|_2^2+\omega_i^{\star 2}\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2^2\right)\lesssim\delta v_{i,j}^{\star},\\
\gamma_{1,1,4} &\lesssim \frac{1}{np}\sigma_1^{\star 4}\left(\left\|\boldsymbol{U}_{i,\cdot}^{\star}\right\|_2^2\zeta_{\mathsf{2nd},j}^2+\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2^2\zeta_{\mathsf{2nd},i}^2\right)\overset{\text{(ii)}}{\lesssim}\frac{\varepsilon^2}{np}v_{i,j}^{\star}\lesssim\delta v_{i,j}^{\star},\\
\gamma_{1,1,5} &\lesssim \frac{1}{np}\left\|\boldsymbol{U}_{i,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2\left\|\boldsymbol{U}_{j,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2\cdot\sigma_1^{\star 2}\left(\left\|\boldsymbol{U}_{i,\cdot}^{\star}\right\|_2\zeta_{\mathsf{2nd},j}+\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2\zeta_{\mathsf{2nd},i}\right)\overset{\text{(iii)}}{\lesssim}\frac{\varepsilon}{\sqrt{np}}v_{i,j}^{\star}\lesssim\delta v_{i,j}^{\star},
\end{aligned}$$

provided that $\theta\lesssim\delta/\sqrt{\kappa}$, $n\gtrsim\delta^{-2}\kappa^3 r\log(n+d)$, $np\gtrsim\delta^{-2}$ and $\varepsilon\lesssim 1$. Here, the relation (i) invokes the AM-GM inequality, while (ii) and (iii) make use of (E.113).

- Regarding $\gamma_{1,2}$, we make the observation that

$$\gamma_{1,2}\lesssim\left(\theta+\sqrt{\frac{\kappa^3 r\log(n+d)}{n}}\right)\gamma_{1,1}+\frac{1}{np}\left[\theta^2\omega_i^{\star 2}+\left(\theta\omega_i^{\star}+\zeta_{\mathsf{2nd},i}\sigma_1^{\star}\right)\left\|\boldsymbol{U}_{i,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2+\zeta_{\mathsf{2nd},i}^2\sigma_1^{\star 2}\right]$$
$$\cdot\left[\theta^2\omega_j^{\star 2}+\left(\theta\omega_j^{\star}+\zeta_{\mathsf{2nd},j}\sigma_1^{\star}\right)\left\|\boldsymbol{U}_{j,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2+\zeta_{\mathsf{2nd},j}^2\sigma_1^{\star 2}\right]$$
$$\lesssim\delta v_{i,j}^{\star}+\underbrace{\frac{1}{np}\theta^4\omega_i^{\star 2}\omega_j^{\star 2}}_{=:\gamma_{1,2,1}}+\underbrace{\frac{1}{np}\left[\theta^2\omega_i^{\star}\omega_j^{\star}+\theta\sigma_1^{\star}\left(\omega_i^{\star}\zeta_{\mathsf{2nd},j}+\omega_j^{\star}\zeta_{\mathsf{2nd},i}\right)+\sigma_1^{\star 2}\zeta_{\mathsf{2nd},i}\zeta_{\mathsf{2nd},j}\right]\left\|\boldsymbol{U}_{i,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2\left\|\boldsymbol{U}_{j,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2}_{=:\gamma_{1,2,2}}$$
$$+\underbrace{\frac{1}{np}\zeta_{\mathsf{2nd},i}^2\zeta_{\mathsf{2nd},j}^2\sigma_1^{\star 4}}_{=:\gamma_{1,2,3}}+\underbrace{\frac{1}{np}\theta^3\omega_i^{\star}\omega_j^{\star}\left(\omega_j^{\star}\left\|\boldsymbol{U}_{i,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2+\omega_i^{\star}\left\|\boldsymbol{U}_{j,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2\right)}_{=:\gamma_{1,2,4}}$$
$$+\underbrace{\frac{1}{np}\zeta_{\mathsf{2nd},i}\zeta_{\mathsf{2nd},j}\sigma_1^{\star 4}\left(\left\|\boldsymbol{U}_{i,\cdot}^{\star}\right\|_2\zeta_{\mathsf{2nd},j}+\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2\zeta_{\mathsf{2nd},i}\right)}_{=:\gamma_{1,2,5}}$$
$$+\underbrace{\frac{1}{np}\theta^2\sigma_1^{\star}\left(\omega_j^{\star 2}\zeta_{\mathsf{2nd},i}\left\|\boldsymbol{U}_{i,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2+\omega_i^{\star 2}\zeta_{\mathsf{2nd},j}\left\|\boldsymbol{U}_{j,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2\right)}_{=:\gamma_{1,2,6}}+\underbrace{\frac{1}{np}\theta^2\sigma_1^{\star 2}\left(\omega_i^{\star 2}\zeta_{\mathsf{2nd},j}^2+\omega_j^{\star 2}\zeta_{\mathsf{2nd},i}^2\right)}_{=:\gamma_{1,2,7}}$$

$$+\underbrace{\frac{1}{np}\theta\sigma_1^{\star 2}\left(\omega_i^\star\left\|\boldsymbol{U}_{i,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2\zeta_{\mathsf{2nd},j}^2+\omega_j^\star\left\|\boldsymbol{U}_{j,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2\zeta_{\mathsf{2nd},i}^2\right)}_{=:\gamma_{1,2,8}}\lesssim\delta v_{i,j}^\star.$$

Here, the penultimate relation follows from the previous bound on $\gamma_{1,1}$ as well as the assumptions that $\theta\lesssim 1$ and $n\gtrsim\kappa^3 r\log(n+d)$; the last line is valid since

$$\begin{aligned}
\gamma_{1,2,1}+\gamma_{1,2,3}&\overset{\text{(i)}}{\lesssim}\frac{\varepsilon^2}{np}v_{i,j}^\star\leq\delta v_{i,j}^\star,\\
\gamma_{1,2,2}&\overset{\text{(ii)}}{\lesssim}\frac{1}{np}\left\|\boldsymbol{U}_{i,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2\left\|\boldsymbol{U}_{j,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2\varepsilon\left(v_{i,j}^\star\right)^{1/2}\lesssim\frac{\varepsilon}{\sqrt{np}}v_{i,j}^\star\lesssim\delta v_{i,j}^\star,\\
\gamma_{1,2,4}&\lesssim\sqrt{\frac{\kappa}{np}}\theta\cdot\theta^2\omega_i^\star\omega_j^\star\cdot\frac{\sigma_r^\star}{\sqrt{np}}\left(\omega_j^\star\left\|\boldsymbol{U}_{i,\cdot}^\star\right\|_2+\omega_i^\star\left\|\boldsymbol{U}_{j,\cdot}^\star\right\|_2\right)\overset{\text{(iii)}}{\lesssim}\sqrt{\frac{\kappa}{np}}\theta\varepsilon v_{i,j}^\star\lesssim\delta v_{i,j}^\star,\\
\gamma_{1,2,5}&\lesssim\frac{1}{np}\cdot\zeta_{\mathsf{2nd},i}\zeta_{\mathsf{2nd},j}\sigma_1^{\star 2}\cdot\sigma_1^{\star 2}\left(\left\|\boldsymbol{U}_{i,\cdot}^\star\right\|_2\zeta_{\mathsf{2nd},j}+\left\|\boldsymbol{U}_{j,\cdot}^\star\right\|_2\zeta_{\mathsf{2nd},i}\right)\overset{\text{(iv)}}{\lesssim}\frac{\varepsilon^2}{np}v_{i,j}^\star\lesssim\delta v_{i,j}^\star,\\
\gamma_{1,2,6}&\lesssim\sqrt{\frac{\kappa}{np}}\theta\cdot\theta\sigma_1^\star\left(\omega_i^\star\zeta_{\mathsf{2nd},j}+\omega_j^\star\zeta_{\mathsf{2nd},i}\right)\cdot\frac{\sigma_r^\star}{\sqrt{np}}\left(\omega_j^\star\left\|\boldsymbol{U}_{i,\cdot}^\star\right\|_2+\omega_i^\star\left\|\boldsymbol{U}_{j,\cdot}^\star\right\|_2\right)\overset{\text{(v)}}{\lesssim}\sqrt{\frac{\kappa}{np}}\theta\varepsilon v_{i,j}^\star\lesssim\delta v_{i,j}^\star,\\
\gamma_{1,2,7}&\lesssim\frac{1}{np}\left[\sigma_1^\star\left(\omega_i^\star\zeta_{\mathsf{2nd},j}+\omega_j^\star\zeta_{\mathsf{2nd},i}\right)\right]^2\overset{\text{(vi)}}{\lesssim}\frac{\varepsilon}{np}v_{i,j}^\star\lesssim\delta v_{i,j}^\star,\\
\gamma_{1,2,8}&\lesssim\frac{1}{np}\cdot\theta\sigma_1^\star\left(\omega_i^\star\zeta_{\mathsf{2nd},j}+\omega_j^\star\zeta_{\mathsf{2nd},i}\right)\cdot\sigma_1^{\star 2}\left(\left\|\boldsymbol{U}_{i,\cdot}^\star\right\|_2\zeta_{\mathsf{2nd},j}+\left\|\boldsymbol{U}_{j,\cdot}^\star\right\|_2\zeta_{\mathsf{2nd},i}\right)\overset{\text{(vii)}}{\lesssim}\frac{\varepsilon^2}{np}v_{i,j}^\star\lesssim\delta v_{i,j}^\star,
\end{aligned}$$

provided that $np\gtrsim\delta^{-2}$, $\varepsilon\lesssim 1$ and $\theta\lesssim\delta/\sqrt{\kappa}$. Here, the inequalities (i)-(vii) rely on (E.113) and (E.115).

Taking the above bounds on $\gamma_{1,1}$ and $\gamma_{1,2}$ collectively, we arrive at

$$|\alpha_1-\beta_1|\lesssim|\gamma_{1,1}|+|\gamma_{1,2}|\lesssim\delta v_{i,j}^\star,$$

provided that $\theta\lesssim\delta/\sqrt{\kappa}$, $n\gtrsim\delta^{-2}\kappa^3 r\log(n+d)$, $\varepsilon\lesssim 1$, and $np\gtrsim\delta^{-2}$.

Step 2: bounding $|\alpha_2-\beta_2|$. It comes from Lemma 19 that

$$\begin{aligned}
\left|S_{i,j}-S_{i,j}^\star\right|&\lesssim\left(\theta+\sqrt{\frac{\kappa^3 r\log(n+d)}{n}}\right)\left\|\boldsymbol{U}_{i,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2\left\|\boldsymbol{U}_{j,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2+\theta\left(\omega_i^\star\left\|\boldsymbol{U}_{j,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2+\omega_j^\star\left\|\boldsymbol{U}_{i,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2\right)\\
&\quad+\sigma_1^\star\left(\zeta_{\mathsf{2nd},i}\left\|\boldsymbol{U}_{j,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2+\zeta_{\mathsf{2nd},j}\left\|\boldsymbol{U}_{i,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2\right)+\theta^2\omega_i^\star\omega_j^\star+\zeta_{\mathsf{2nd},i}\zeta_{\mathsf{2nd},j}\sigma_1^{\star 2}.
\end{aligned}$$

With this in place, we shall proceed to decompose $|\alpha_2-\beta_2|$ as follows:

$$|\alpha_2-\beta_2|\lesssim\frac{1}{np}\left|S_{i,j}^{\star 2}-S_{i,j}^2\right|\lesssim\underbrace{\frac{1}{np}S_{i,j}^\star\left|S_{i,j}-S_{i,j}^\star\right|}_{=:\gamma_{2,1}}+\underbrace{\frac{1}{np}\left|S_{i,j}-S_{i,j}^\star\right|^2}_{=:\gamma_{2,2}}.$$

- With regards to $\gamma_{2,1}$, it is observed that

$$\begin{aligned}
\gamma_{2,1}&\lesssim\underbrace{\frac{1}{np}\left(\theta+\sqrt{\frac{\kappa^3 r\log(n+d)}{n}}\right)\left\|\boldsymbol{U}_{i,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2^2\left\|\boldsymbol{U}_{j,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2^2}_{=:\gamma_{2,1,1}}+\underbrace{\frac{1}{np}\theta^2\omega_i^\star\omega_j^\star\left\|\boldsymbol{U}_{i,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2\left\|\boldsymbol{U}_{j,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2}_{=:\gamma_{2,1,2}}\\
&\quad+\underbrace{\frac{1}{np}\theta\left\|\boldsymbol{U}_{i,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2\left\|\boldsymbol{U}_{j,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2\left(\omega_i^\star\left\|\boldsymbol{U}_{j,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2+\omega_j^\star\left\|\boldsymbol{U}_{i,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2\right)}_{=:\gamma_{2,1,3}}
\end{aligned}$$

$$
+ \underbrace{\frac{1}{np}\sigma_1^\star \left\|\boldsymbol{U}_{i,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2 \left\|\boldsymbol{U}_{j,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2 \left(\zeta_{\mathsf{2nd},i}\left\|\boldsymbol{U}_{j,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2 + \zeta_{\mathsf{2nd},j}\left\|\boldsymbol{U}_{i,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2\right)}_{=:\gamma_{2,1,4}}
$$

$$
+ \underbrace{\frac{1}{np}\zeta_{\mathsf{2nd},i}\zeta_{\mathsf{2nd},j}\sigma_1^{\star 2}\left\|\boldsymbol{U}_{i,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2 \left\|\boldsymbol{U}_{j,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2}_{=:\gamma_{2,1,5}} \lesssim \delta v_{i,j}^\star.
$$

Here, the last relation holds due to the following bounds

$$
\begin{aligned}
\gamma_{2,1,1} &\lesssim \frac{\delta}{np}\left\|\boldsymbol{U}_{i,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2^2 \left\|\boldsymbol{U}_{j,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2^2 \lesssim \delta v_{i,j}^\star, \\
\gamma_{2,1,2} &\lesssim \frac{1}{\sqrt{np}}\cdot\theta^2\omega_i^\star\omega_j^\star\cdot\frac{1}{\sqrt{np}}\left\|\boldsymbol{U}_{i,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2 \left\|\boldsymbol{U}_{j,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2 \overset{\text{(i)}}{\lesssim} \frac{\varepsilon}{\sqrt{np}} v_{i,j}^\star \lesssim \delta v_{i,j}^\star, \\
\gamma_{2,1,3} &\lesssim \sqrt{\kappa}\theta\cdot\frac{1}{\sqrt{np}}\left\|\boldsymbol{U}_{i,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2 \left\|\boldsymbol{U}_{j,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2\cdot\frac{\sigma_r^\star}{\sqrt{np}}\left(\omega_i^\star\left\|\boldsymbol{U}_{j,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2 + \omega_j^\star\left\|\boldsymbol{U}_{i,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2\right) \\
&\lesssim \sqrt{\kappa}\theta v_{i,j}^\star \lesssim \delta v_{i,j}^\star, \\
\gamma_{2,1,4} &\lesssim \frac{1}{\sqrt{np}}\cdot\frac{1}{\sqrt{np}}\left\|\boldsymbol{U}_{i,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2 \left\|\boldsymbol{U}_{j,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2\cdot\sigma_1^{\star 2}\left(\left\|\boldsymbol{U}_{i,\cdot}^\star\right\|_2\zeta_{\mathsf{2nd},j} + \left\|\boldsymbol{U}_{j,\cdot}^\star\right\|_2\zeta_{\mathsf{2nd},i}\right) \\
&\overset{\text{(ii)}}{\lesssim} \frac{\varepsilon}{\sqrt{np}} v_{i,j}^\star \lesssim \delta v_{i,j}^\star, \\
\gamma_{2,1,5} &\lesssim \frac{1}{np}\left\|\boldsymbol{U}_{i,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2 \left\|\boldsymbol{U}_{j,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2\cdot\sigma_1^{\star 2}\zeta_{\mathsf{2nd},i}\zeta_{\mathsf{2nd},j} \overset{\text{(iii)}}{\lesssim} \frac{\varepsilon}{\sqrt{np}} v_{i,j}^\star \lesssim \delta v_{i,j}^\star,
\end{aligned}
$$

with the proviso that $\theta \lesssim \delta/\sqrt{\kappa}$, $n \gtrsim \delta^{-2}\kappa^3 r\log(n+d)$, $\varepsilon \lesssim 1$ and $np \gtrsim \delta^{-2}$. Note that (i)-(iii) follow from (E.113) and (E.115).

- Regarding $\gamma_{2,2}$, we obtain

$$
\begin{aligned}
\gamma_{2,2} \lesssim\ & \underbrace{\frac{1}{np}\left(\theta^2 + \frac{\kappa^3 r\log(n+d)}{n}\right)\left\|\boldsymbol{U}_{i,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2^2 \left\|\boldsymbol{U}_{j,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2^2}_{=:\gamma_{2,2,1}} + \underbrace{\frac{1}{np}\theta^4\omega_i^{\star 2}\omega_j^{\star 2}}_{=:\gamma_{2,2,2}} + \underbrace{\frac{1}{np}\zeta_{\mathsf{2nd},i}^2\zeta_{\mathsf{2nd},j}^2\sigma_1^{\star 4}}_{=:\gamma_{2,2,3}} \\
& + \underbrace{\frac{1}{np}\theta^2\left(\omega_i^\star\left\|\boldsymbol{U}_{j,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2 + \omega_j^\star\left\|\boldsymbol{U}_{i,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2\right)^2}_{=:\gamma_{2,2,4}} + \underbrace{\frac{1}{np}\sigma_1^{\star 2}\left(\zeta_{\mathsf{2nd},i}\left\|\boldsymbol{U}_{j,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2 + \zeta_{\mathsf{2nd},j}\left\|\boldsymbol{U}_{i,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2\right)^2}_{=:\gamma_{2,2,5}} \\
\lesssim\ & \delta v_{i,j}^\star,
\end{aligned}
$$

where the last inequality holds true since

$$
\begin{aligned}
\gamma_{2,2,1} &\lesssim \left(\theta + \sqrt{\frac{\kappa^3 r\log(n+d)}{n}}\right)\gamma_{2,1,1} \lesssim \delta v_{i,j}^\star, \\
\gamma_{2,2,2} + \gamma_{2,2,3} &\overset{\text{(i)}}{\lesssim} \frac{\varepsilon^2}{np} v_{i,j}^\star \lesssim \delta v_{i,j}^\star, \\
\gamma_{2,2,4} &\lesssim \theta^2\kappa\cdot\frac{\sigma_r^{\star 2}}{np}\left(\omega_j^{\star 2}\left\|\boldsymbol{U}_{i,\cdot}^\star\right\|_2^2 + \omega_i^{\star 2}\left\|\boldsymbol{U}_{j,\cdot}^\star\right\|_2^2\right) \lesssim \theta^2\kappa v_{i,j}^\star \lesssim \delta v_{i,j}^\star, \\
\gamma_{2,2,5} &\lesssim \frac{1}{np}\sigma_1^{\star 4}\left(\left\|\boldsymbol{U}_{i,\cdot}^\star\right\|_2\zeta_{\mathsf{2nd},j} + \left\|\boldsymbol{U}_{j,\cdot}^\star\right\|_2\zeta_{\mathsf{2nd},i}\right)^2 \overset{\text{(ii)}}{\lesssim} \frac{\varepsilon^2}{np} v_{i,j}^\star \lesssim \delta v_{i,j}^\star,
\end{aligned}
$$

provided that $\theta \lesssim \delta/\sqrt{\kappa}$, $n \gtrsim \kappa^3 r\log(n+d)$, $\varepsilon \lesssim 1$ and $np \geq \delta^{-1}$. Here, (i) follows from (E.115), whereas (ii) follows from (E.113).

Taking the above bounds on $\gamma_{2,1}$ and $\gamma_{2,2}$ together yields

$$\left|\alpha_2-\beta_2\right| \lesssim \left|\gamma_{2,1}\right|+\left|\gamma_{2,2}\right| \lesssim \delta v_{i,j}^{\star},$$

as long as $\theta \lesssim \delta/\sqrt{\kappa}$, $n \gtrsim \delta^{-2}\kappa^3 r\log(n+d)$, $\varepsilon \lesssim 1$ and $np \gtrsim \delta^{-2}$.

Step 3: bounding $|\alpha_3-\beta_3|$. For each $l \in [d]$, let us first define UB_l as follows

$$\mathsf{UB}_l \coloneqq \left(\theta+\sqrt{\frac{\kappa^3 r\log(n+d)}{n}}\right)\left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2^2+\theta^2\omega_l^{\star 2}+\left(\theta\omega_l^{\star}+\zeta_{\mathsf{2nd},l}\sigma_1^{\star}\right)\left\|\boldsymbol{U}_{l,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2+\zeta_{\mathsf{2nd},l}^2\sigma_1^{\star 2}.$$

According to Lemma 19, we can obtain

$$\left|S_{l,l}-S_{l,l}^{\star}\right| \lesssim \mathsf{UB}_l \qquad \forall\, l \in [d]\,.$$

Lemma 19 also tells us that

$$\left|\omega_i^2-\omega_i^{\star 2}\right| \lesssim \sqrt{\frac{\log^2(n+d)}{np}}\omega_i^{\star 2}+\mathsf{UB}_i.$$

With these basic bounds in mind, we proceed with the following decomposition:

$$\begin{aligned}
\frac{1}{np}\left|\omega_i^2 S_{j,j}-\omega_i^{\star 2}S_{j,j}^{\star}\right| &\lesssim \frac{1}{np}S_{j,j}\left|\omega_i^2-\omega_i^{\star 2}\right|+\frac{1}{np}\omega_i^{\star 2}\left|S_{j,j}-S_{j,j}^{\star}\right| \\
&\lesssim \underbrace{\frac{1}{np}S_{j,j}^{\star}\left|\omega_i^2-\omega_i^{\star 2}\right|}_{=:\gamma_{3,1}}+\underbrace{\frac{1}{np}\omega_i^{\star 2}\left|S_{j,j}-S_{j,j}^{\star}\right|}_{=:\gamma_{3,2}}+\underbrace{\frac{1}{np}\left|S_{j,j}-S_{j,j}^{\star}\right|\left|\omega_i^2-\omega_i^{\star 2}\right|}_{=:\gamma_{3,3}},
\end{aligned}$$

leaving us with three terms to cope with.

- Regarding $\gamma_{3,1}$, we can upper bound

$$\gamma_{3,1} \lesssim \underbrace{\frac{1}{np}\left\|\boldsymbol{U}_{j,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2^2\sqrt{\frac{\log^2(n+d)}{np}}\omega_i^{\star 2}}_{=:\gamma_{3,1,1}}+\underbrace{\frac{1}{np}S_{j,j}^{\star}\mathsf{UB}_i}_{=:\gamma_{3,1,2}} \lesssim \delta v_{i,j}^{\star}.$$

  Here, the last relation holds since (i) the first term

$$\gamma_{3,1,1} \lesssim \sqrt{\frac{\kappa\log^2(n+d)}{n^3p^3}}\omega_i^{\star 2}\sigma_r^{\star 2}\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2^2 \lesssim \sqrt{\frac{\kappa\log^2(n+d)}{np}}v_{i,j}^{\star} \lesssim \delta v_{i,j}^{\star}$$

  provided that $np \gtrsim \delta^{-2}\kappa\log^2(n+d)$; and (ii) we can also demonstrate that the second term obeys $\gamma_{3,1,2} \lesssim \delta v_{i,j}^{\star}$, since it is easy seen that $\gamma_{3,1,2}$ admits the same upper bound as for $\gamma_{1,1}$.

- Regarding $\gamma_{3,2}$, it can be seen that

$$\begin{aligned}
\gamma_{3,2} &\lesssim \underbrace{\frac{1}{np}\omega_i^{\star 2}\left(\theta+\sqrt{\frac{\kappa^3 r\log(n+d)}{n}}\right)\left\|\boldsymbol{U}_{j,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2^2}_{=:\gamma_{3,2,1}}+\underbrace{\frac{1}{np}\theta^2\omega_i^{\star 2}\omega_j^{\star 2}}_{=:\gamma_{3,2,2}} \\
&\quad+\underbrace{\frac{1}{np}\theta\omega_i^{\star 2}\omega_j^{\star}\left\|\boldsymbol{U}_{j,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2}_{=:\gamma_{3,2,3}}+\underbrace{\frac{1}{np}\omega_i^{\star 2}\zeta_{\mathsf{2nd},j}\sigma_1^{\star}\left\|\boldsymbol{U}_{j,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2}_{=:\gamma_{3,2,4}}+\underbrace{\frac{1}{np}\omega_i^{\star 2}\zeta_{\mathsf{2nd},j}^2\sigma_1^{\star 2}}_{=:\gamma_{3,2,5}} \\
&\lesssim \delta v_{i,j}^{\star},
\end{aligned}$$

where the last line holds true since

$$
\begin{aligned}
\gamma_{3,2,1} &\lesssim \frac{\delta}{np}\omega_i^{\star 2}\left\|\boldsymbol{U}_{j,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2^2 \lesssim \delta v_{i,j}^{\star}, \\
\gamma_{3,2,2}+\gamma_{3,2,3} &\lesssim \delta\frac{\sigma_r^{\star 2}}{np}\left(\omega_j^{\star 2}\left\|\boldsymbol{U}_{i,\cdot}^{\star}\right\|_2^2+\omega_i^{\star 2}\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2^2\right) \lesssim \delta v_{i,j}^{\star}, \\
\gamma_{3,2,4} &\lesssim \frac{\sigma_r^{\star}}{\sqrt{np}}\omega_i^{\star}\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2 \cdot \sqrt{\frac{\kappa}{np}}\omega_i^{\star}\zeta_{\mathsf{2nd},j}\sigma_1^{\star} \lesssim v_{i,j}^{\star 1/2}\cdot\sqrt{\frac{\kappa}{np}}\omega_i^{\star}\zeta_{\mathsf{2nd},j}\sigma_1^{\star} \\
&\overset{\text{(i)}}{\lesssim} v_{i,j}^{\star 1/2}\cdot\left(\sqrt{\frac{1}{np}}\varepsilon\frac{\omega_i^{\star}\left\|\boldsymbol{U}_{j,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2+\omega_i^{\star}\omega_j^{\star}}{\sqrt{\min\left\{ndp^2\kappa, np\right\}}}+\sqrt{\frac{1}{np}}\varepsilon\frac{\omega_{\min}\omega_i^{\star}\omega_j^{\star}}{\sqrt{np^2}\sigma_1^{\star}}\right) \\
&\lesssim v_{i,j}^{\star 1/2}\cdot\delta\left(\frac{\sigma_r^{\star}}{\sqrt{np}}+\frac{\omega_{\min}}{\sqrt{np^2}}\right)\left(\omega_j^{\star}\left\|\boldsymbol{U}_{i,\cdot}^{\star}\right\|_2+\omega_i^{\star}\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2\right) \lesssim \delta v_{i,j}^{\star}, \\
\gamma_{3,2,5} &\overset{\text{(ii)}}{\lesssim} \left(\sqrt{\frac{1}{np}}\varepsilon\frac{\omega_i^{\star}\left\|\boldsymbol{U}_{j,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2+\omega_i^{\star}\omega_j^{\star}}{\sqrt{\min\left\{ndp^2\kappa, np\right\}}}+\sqrt{\frac{1}{np}}\varepsilon\frac{\omega_{\min}\omega_i^{\star}\omega_j^{\star}}{\sqrt{np^2}\sigma_1^{\star}}\right)^2 \\
&\lesssim \delta\left(\frac{\sigma_r^{\star 2}}{np}+\frac{\omega_{\min}^2}{np^2}\right)\left(\omega_j^{\star 2}\left\|\boldsymbol{U}_{i,\cdot}^{\star}\right\|_2^2+\omega_i^{\star 2}\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2^2\right) \lesssim \delta v_{i,j}^{\star},
\end{aligned}
$$

provided that $\theta \lesssim \delta$, $n \gtrsim \delta^{-2}\kappa^3 r\log(n+d)$, $\varepsilon \lesssim 1$, $np \gtrsim \delta^{-2}\kappa$, $ndp^2 \gtrsim \delta^{-2}$ and

$$
\left\|\boldsymbol{U}_{i,\cdot}^{\star}\right\|_2+\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2 \gtrsim \delta^{-1}\kappa^{1/2}\theta\frac{\omega_{\max}}{\sigma_r^{\star}}+\delta^{-1}\frac{\omega_{\max}}{\sigma_r^{\star}}\sqrt{\frac{1}{ndp^2\kappa}}+\delta^{-1}\frac{\omega_{\max}}{\sigma_r^{\star}}\sqrt{\frac{1}{np}}.
$$

Here, (i) and (ii) arise from (E.114).

- When it comes to $\gamma_{3,3}$, we obtain

$$
\gamma_{3,3} \lesssim \underbrace{\frac{1}{np}\left|S_{j,j}-S_{j,j}^{\star}\right|\sqrt{\frac{\log^2(n+d)}{np}}\omega_i^{\star 2}}_{=:\gamma_{3,3,1}}+\underbrace{\frac{1}{np}\left|S_{j,j}-S_{j,j}^{\star}\right|\mathsf{UB}_i}_{=:\gamma_{3,3,2}} \lesssim \delta v_{i,j}^{\star}.
$$

  The last relation holds since (i) the first term

$$
\gamma_{3,3,1} \lesssim \sqrt{\frac{\log^2(n+d)}{np}}\gamma_{3,2} \lesssim \delta v_{i,j}^{\star}
$$

  provided that $np \gtrsim \log^2(n+d)$; and (ii) one can easily check that the second term $\gamma_{3,3,2}$ admits the same upper bound as for $\gamma_{1,2}$.

With the above results in hand, one can conclude that

$$
\frac{1}{np}\left|\omega_i^2 S_{j,j}-\omega_i^{\star 2}S_{j,j}^{\star}\right| \lesssim \gamma_{3,1}+\gamma_{3,2}+\gamma_{3,3} \lesssim \delta v_{i,j}^{\star},
$$

and similarly,

$$
\frac{1}{np}\left|\omega_j^2 S_{i,i}-\omega_j^{\star 2}S_{i,i}^{\star}\right| \lesssim \delta v_{i,j}^{\star}.
$$

These allow us to reach

$$
|\alpha_3-\beta_3| \lesssim \frac{1}{np}\left|\omega_i^2 S_{j,j}-\omega_i^{\star 2}S_{j,j}^{\star}\right|+\frac{1}{np}\left|\omega_j^2 S_{i,i}-\omega_j^{\star 2}S_{i,i}^{\star}\right| \lesssim \delta v_{i,j}^{\star},
$$

provided that $np \gtrsim \delta^{-2}\kappa\log^2(n+d)$, $\theta \lesssim \delta$, $n \gtrsim \delta^{-2}\kappa^3 r\log(n+d)$, $\varepsilon \lesssim 1$, and

$$
\left\|\boldsymbol{U}_{i,\cdot}^{\star}\right\|_2+\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2 \gtrsim \delta^{-1}\kappa^{1/2}\theta\frac{\omega_{\max}}{\sigma_r^{\star}}. \tag{E.117}
$$

In view of (E.64), we know that

$$
\begin{aligned}
\theta\frac{\omega_{\max}}{\sigma_r^\star} &\asymp \left(\frac{\omega_{\max}}{\sigma_r^\star}\sqrt{\frac{d\kappa\log^2(n+d)}{np}}+\frac{\omega_{\max}}{\sigma_r^\star}\sqrt{\frac{\kappa^2\mu r\log^3(n+d)}{np^2}}+\frac{\omega_{\max}^2}{p\sigma_r^{\star 2}}\sqrt{\frac{d\kappa\log^2(n+d)}{n}}\right)\cdot\sqrt{\frac{r}{d}}\\
&\asymp \sqrt{\kappa\log^2(n+d)}\left(\frac{\omega_{\max}}{\sigma_r^\star}\sqrt{\frac{d}{np}}+\frac{\omega_{\max}^2}{p\sigma_r^{\star 2}}\sqrt{\frac{d}{n}}+\frac{\omega_{\max}}{\sigma_r^\star}\sqrt{\frac{\kappa\mu r\log(n+d)}{np^2}}\right)\cdot\sqrt{\frac{r}{d}}\\
&\asymp \sqrt{\kappa\log^2(n+d)}\left[\frac{\omega_{\max}}{\sigma_r^\star}\sqrt{\frac{d}{np}}+\frac{\omega_{\max}^2}{p\sigma_r^{\star 2}}\sqrt{\frac{d}{n}}+\frac{\kappa\mu r\log(n+d)}{\sqrt{ndp^2}}\right]\cdot\sqrt{\frac{r}{d}},
\end{aligned}
\tag{E.118}
$$

where we have used the AM-GM inequality in the last line. As a result, (E.117) is guaranteed by

$$
\left\|\boldsymbol{U}_{i,\cdot}^\star\right\|_2+\left\|\boldsymbol{U}_{j,\cdot}^\star\right\|_2 \gtrsim \delta^{-1}\kappa\log(n+d)\left[\frac{\omega_{\max}}{\sigma_r^\star}\sqrt{\frac{d}{np}}+\frac{\omega_{\max}^2}{p\sigma_r^{\star 2}}\sqrt{\frac{d}{n}}+\frac{\kappa\mu r\log(n+d)}{\sqrt{ndp^2}}\right]\cdot\sqrt{\frac{r}{d}}.
$$

Step 4: bounding $|\alpha_4-\beta_4|$ and $|\alpha_6-\beta_6|$. For each $l\in[d]$, let us denote

$$
\Delta_l := \left|\omega_l^{\star 2}+(1-p)S_{l,l}^{\star 2}-\omega_l^2-(1-p)S_{l,l}^2\right|.
$$

Lemma 19 tells us that, for each $l\in[d]$,

$$
\Delta_l \lesssim \sqrt{\frac{\log^2(n+d)}{np}}\omega_{\max}^2+\zeta_{\mathsf{1st}}\frac{\sqrt{\kappa^2\mu r^2\log(n+d)}}{d}+\sqrt{\frac{\kappa^2\mu^2 r^3\log(n+d)}{nd^2}}\sigma_1^{\star 2}.
\tag{E.119}
$$

We also know that

$$
\Delta_i \lesssim \sqrt{\frac{\log^2(n+d)}{np}}\omega_i^{\star 2}+\mathsf{UB}_i \qquad\text{and}\qquad \Delta_j \lesssim \sqrt{\frac{\log^2(n+d)}{np}}\omega_j^{\star 2}+\mathsf{UB}_j.
\tag{E.120}
$$

In addition, for each $l\in[d]$, it holds that

$$
\omega_l^{\star 2}+(1-p)\,S_{l,l}^\star \le \omega_{\max}^2+\left\|\boldsymbol{U}_{l,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2^2 \lesssim \omega_{\max}^2+\frac{\mu r}{d}\sigma_1^{\star 2}.
\tag{E.121}
$$

We also have the following bound

$$
\begin{aligned}
\left\|\boldsymbol{U}_{j,\cdot}\right\|_2 &\le \left\|\boldsymbol{U}_{j,\cdot}^\star\right\|_2+\left\|\left(\boldsymbol{U}\boldsymbol{R}-\boldsymbol{U}^\star\right)_{j,\cdot}\right\|_2\\
&\overset{\text{(i)}}{\lesssim} \left\|\boldsymbol{U}_{j,\cdot}^\star\right\|_2+\frac{\theta}{\sqrt{\kappa}\sigma_r^\star}\left(\left\|\boldsymbol{U}_{j,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2+\omega_j^\star\right)+\zeta_{\mathsf{2nd},j}\\
&\overset{\text{(ii)}}{\lesssim} \left\|\boldsymbol{U}_{j,\cdot}^\star\right\|_2+\frac{\theta}{\sqrt{\kappa}\sigma_r^\star}\omega_j^\star+\zeta_{\mathsf{2nd},j}\\
&\overset{\text{(iii)}}{\lesssim} \left\|\boldsymbol{U}_{j,\cdot}^\star\right\|_2+\frac{\theta}{\sqrt{\kappa}\sigma_r^\star}\omega_j^\star.
\end{aligned}
\tag{E.122}
$$

Here (i) follows from Lemma 18, (ii) holds provided that $\theta\lesssim 1$, (iii) utilizes (E.114) and (E.64) and holds provided that $\varepsilon\lesssim 1$. We can thus decompose

$$
\begin{aligned}
|\alpha_4-\beta_4| \lesssim\; &\underbrace{\frac{1}{np^2}\left[\omega_i^{\star 2}+(1-p)\,S_{i,i}^\star\right]\sum_{k=1}^d\Delta_k\left(\boldsymbol{U}_{k,\cdot}\boldsymbol{U}_{j,\cdot}^\top\right)^2}_{=:\gamma_{4,1}}+\underbrace{\frac{\Delta_i}{np^2}\sum_{k=1}^d\left[\omega_k^2+(1-p)\,S_{k,k}\right]\left(\boldsymbol{U}_{k,\cdot}\boldsymbol{U}_{j,\cdot}^\top\right)^2}_{=:\gamma_{4,2}}\\
&+\underbrace{\frac{1}{np^2}\left[\omega_i^{\star 2}+(1-p)\,S_{i,i}^\star\right]\sum_{k=1}^d\left[\omega_k^{\star 2}+(1-p)\,S_{k,k}^\star\right]\left|\left(\boldsymbol{U}_{k,\cdot}^\star\boldsymbol{U}_{j,\cdot}^{\star\top}\right)^2-\left(\boldsymbol{U}_{k,\cdot}\boldsymbol{U}_{j,\cdot}^\top\right)^2\right|}_{=:\gamma_{4,3}}.
\end{aligned}
$$

Step 4.1: bounding $\gamma_{4,1}$. We have learned from (E.122) that

$$
\begin{aligned}
\gamma_{4,1} &\lesssim \frac{1}{np^2}\left[\omega_i^{\star 2} + (1-p)\, S_{i,i}^{\star}\right] \left\|\boldsymbol{U}_{j,\cdot}\right\|_2^2 \max_{1\le k\le d} \Delta_k \\
&\lesssim \underbrace{\frac{1}{np^2}\left[\omega_i^{\star 2} + (1-p)\, S_{i,i}^{\star}\right] \left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2^2 \max_{1\le k\le d} \Delta_k}_{=:\gamma_{4,1,1}} + \underbrace{\frac{\theta^2 \omega_j^{\star 2}}{np^2 \sigma_1^{\star 2}}\left[\omega_i^{\star 2} + (1-p)\, S_{i,i}^{\star}\right] \max_{1\le k\le d} \Delta_k}_{=:\gamma_{4,1,2}}.
\end{aligned}
$$

Regarding $\gamma_{4,1,1}$, we can derive

$$
\begin{aligned}
\gamma_{4,1,1} \lesssim\;& \underbrace{\frac{1}{np^2}\omega_{\max}^2 \omega_i^{\star 2} \left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2^2 \sqrt{\frac{\log^2(n+d)}{np}}}_{=:\gamma_{4,1,1,1}} + \underbrace{\frac{\sqrt{\kappa^2 \mu r^2 \log(n+d)}}{ndp^2}\omega_i^{\star 2}\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2^2 \zeta_{\mathsf{1st}}}_{=:\gamma_{4,1,1,2}} \\
&+ \underbrace{\frac{1}{np^2}\omega_i^{\star 2}\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2^2 \sqrt{\frac{\kappa^2 \mu^2 r^3 \log(n+d)}{nd^2}}\sigma_1^{\star 2}}_{=:\gamma_{4,1,1,3}} + \underbrace{\frac{1}{np^2}\left\|\boldsymbol{U}_{i,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2^2 \left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2^2 \sqrt{\frac{\log^2(n+d)}{np}}\omega_{\max}^2}_{=:\gamma_{4,1,1,4}} \\
&+ \underbrace{\frac{\sqrt{\kappa^2 \mu r^2 \log(n+d)}}{ndp^2}\left\|\boldsymbol{U}_{i,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2^2 \left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2^2 \zeta_{\mathsf{1st}}}_{=:\gamma_{4,1,1,5}} + \underbrace{\frac{1}{np^2}\left\|\boldsymbol{U}_{i,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2^2 \left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2^2 \sqrt{\frac{\kappa^2 \mu^2 r^3 \log(n+d)}{nd^2}}\sigma_1^{\star 2}}_{=:\gamma_{4,1,1,6}} \\
\lesssim\;& \delta v_{i,j}^{\star}.
\end{aligned}
$$

Here, the last line holds since

$$
\begin{aligned}
\gamma_{4,1,1,1} &\lesssim \delta \frac{\omega_{\min}^2}{np^2}\left(\omega_j^{\star 2}\left\|\boldsymbol{U}_{i,\cdot}^{\star}\right\|_2^2 + \omega_i^{\star 2}\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2^2\right) \lesssim \delta v_{i,j}^{\star}, \\
\gamma_{4,1,1,2} + \gamma_{4,1,1,3} + \gamma_{4,1,1,4} &\lesssim \delta \frac{\sigma_r^{\star 2}}{ndp^2}\left(\omega_j^{\star 2}\left\|\boldsymbol{U}_{i,\cdot}^{\star}\right\|_2^2 + \omega_i^{\star 2}\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2^2\right) \lesssim \delta v_{i,j}^{\star}, \\
\gamma_{4,1,1,5} + \gamma_{4,1,1,6} &\lesssim \delta \frac{1}{ndp^2\kappa}\left\|\boldsymbol{U}_{i,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2^2 \left\|\boldsymbol{U}_{j,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2^2 \lesssim \delta v_{i,j}^{\star},
\end{aligned}
$$

provided that $np \gtrsim \delta^{-2}\kappa^2\mu^2 r^2 \log^2(n+d)$, $n \gtrsim \delta^{-2}\kappa^4\mu^2 r^3 \log(n+d)$, $np \gtrsim \delta^{-2}\kappa_\omega^2 \log^2(n+d)$and

$$
\frac{\zeta_{\mathsf{1st}}}{\sigma_r^{\star 2}} \lesssim \frac{\delta}{\sqrt{\kappa^2 \mu r^2 \log(n+d)}}.
$$

Regarding $\gamma_{4,1,2}$, we can upper bound

$$
\begin{aligned}
\gamma_{4,1,2} \lesssim\;& \underbrace{\frac{\theta^2}{np^2}\frac{\omega_i^{\star 2}\omega_j^{\star 2}\omega_{\max}^2}{\sigma_1^{\star 2}}\sqrt{\frac{\log^2(n+d)}{np}}}_{=:\gamma_{4,1,2,1}} + \underbrace{\frac{\sqrt{\kappa^2 \mu r^2 \log(n+d)}}{ndp^2}\frac{\omega_i^{\star 2}\omega_j^{\star 2}}{\sigma_1^{\star 2}}\theta^2 \zeta_{\mathsf{1st}}}_{=:\gamma_{4,1,2,2}} \\
&+ \underbrace{\frac{1}{np^2}\omega_i^{\star 2}\omega_j^{\star 2}\theta^2 \sqrt{\frac{\kappa^2 \mu^2 r^3 \log(n+d)}{nd^2}}}_{=:\gamma_{4,1,2,3}} + \underbrace{\frac{\theta^2}{np^2}\left\|\boldsymbol{U}_{i,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2^2 \frac{\omega_{\max}^2 \omega_j^{\star 2}}{\sigma_1^{\star 2}}\sqrt{\frac{\log^2(n+d)}{np}}}_{=:\gamma_{4,1,2,4}} \\
&+ \underbrace{\frac{\sqrt{\kappa^2 \mu r^2 \log(n+d)}}{ndp^2}\left\|\boldsymbol{U}_{i,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2^2 \frac{\omega_j^{\star 2}}{\sigma_1^{\star 2}}\theta^2 \zeta_{\mathsf{1st}}}_{=:\gamma_{4,1,2,5}} + \underbrace{\frac{1}{np^2}\left\|\boldsymbol{U}_{i,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2^2 \theta^2 \omega_j^{\star 2}\sqrt{\frac{\kappa^2 \mu^2 r^3 \log(n+d)}{nd^2}}}_{=:\gamma_{4,1,2,6}}.
\end{aligned}
$$

There are six terms on the right-hand side of the above inequality, which we shall control separately.

- With regards to $\gamma_{4,1,2,1}$, we observe that

$$\gamma_{4,1,2,1} \asymp \underbrace{\frac{\omega_i^{\star 2}\omega_j^{\star 2}\omega_{\max}^2}{\sigma_r^{\star 2}}\frac{r\log^3(n+d)}{n^{2.5}p^{3.5}}}_{=:\gamma_{4,1,2,1,1}} + \underbrace{\frac{\omega_i^{\star 2}\omega_j^{\star 2}\omega_{\max}^2}{\sigma_r^{\star 2}}\frac{\kappa\mu r^2\log^4(n+d)}{n^{2.5}dp^{4.5}}}_{=:\gamma_{4,1,2,1,2}} + \underbrace{\frac{\omega_i^{\star 2}\omega_j^{\star 2}\omega_{\max}^4}{\sigma_r^{\star 4}}\frac{r\log^3(n+d)}{n^{2.5}p^{4.5}}}_{=:\gamma_{4,1,2,1,3}} \lesssim \delta v_{i,j}^{\star}.$$

  Here, the last relation holds since

$$\gamma_{4,1,2,1,1} + \gamma_{4,1,2,1,3} \lesssim \delta\frac{\omega_{\min}^2}{np^2}\left(\omega_j^{\star 2}\left\|\boldsymbol{U}_{i,\cdot}^{\star}\right\|_2^2 + \omega_i^{\star 2}\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2^2\right) \lesssim \delta v_{i,j}^{\star},$$
$$\gamma_{4,1,2,1,2} \lesssim \delta\frac{\sigma_r^{\star 2}}{ndp^2}\left(\omega_j^{\star 2}\left\|\boldsymbol{U}_{i,\cdot}^{\star}\right\|_2^2 + \omega_i^{\star 2}\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2^2\right) \lesssim \delta v_{i,j}^{\star},$$

  provided that

$$\left\|\boldsymbol{U}_{i,\cdot}^{\star}\right\|_2 + \left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2 \gtrsim \delta^{-1/2}\frac{\sqrt{\kappa_\omega\log^3(n+d)}}{n^{1/4}p^{1/4}}\left[\frac{\omega_{\max}}{\sigma_r^{\star}}\sqrt{\frac{d}{np}} + \frac{\omega_{\max}^2}{p\sigma_r^{\star 2}}\sqrt{\frac{d}{n}}\right]\sqrt{\frac{r}{d}}.$$

- When it comes to $\gamma_{4,1,2,2}$, we obtain

$$\gamma_{4,1,2,2} \asymp \underbrace{\frac{\omega_i^{\star 2}\omega_j^{\star 2}}{\sigma_r^{\star 2}}\frac{\kappa\mu^{1/2}r^2\log^{5/2}(n+d)}{n^2dp^3}\zeta_{\mathsf{1st}}}_{=:\gamma_{4,1,2,2,1}} + \underbrace{\frac{\omega_i^{\star 2}\omega_j^{\star 2}}{\sigma_r^{\star 2}}\frac{\kappa^2\mu^{3/2}r^3\log^{7/2}(n+d)}{n^2d^2p^4}\zeta_{\mathsf{1st}}}_{=:\gamma_{4,1,2,2,2}}$$
$$+ \underbrace{\frac{\omega_i^{\star 2}\omega_j^{\star 2}\omega_{\max}^2}{\sigma_r^{\star 4}}\frac{\kappa\mu^{1/2}r^2\log^{5/2}(n+d)}{n^2dp^4}\zeta_{\mathsf{1st}}}_{=:\gamma_{4,1,2,2,3}} \lesssim \delta v_{i,j}^{\star}.$$

  Here, the last relation holds since

$$\gamma_{4,1,2,2,1} + \gamma_{4,1,2,2,3} \lesssim \delta\frac{\sigma_r^{\star 2}}{ndp^2}\left(\omega_j^{\star 2}\left\|\boldsymbol{U}_{i,\cdot}^{\star}\right\|_2^2 + \omega_i^{\star 2}\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2^2\right) \lesssim \delta v_{i,j}^{\star},$$
$$\gamma_{4,1,2,2,2} \lesssim \delta\frac{\omega_{\min}^2}{np^2}\left(\omega_j^{\star 2}\left\|\boldsymbol{U}_{i,\cdot}^{\star}\right\|_2^2 + \omega_i^{\star 2}\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2^2\right) \lesssim \delta v_{i,j}^{\star},$$

  provided that

$$\left\|\boldsymbol{U}_{i,\cdot}^{\star}\right\|_2 + \left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2 \gtrsim \delta^{-1/2}\sqrt{\frac{\zeta_{\mathsf{1st}}}{\sigma_r^{\star 2}}}\left[\frac{\omega_{\max}}{\sigma_r^{\star}}\sqrt{\frac{d\kappa\mu^{1/2}r\log^{5/2}(n+d)}{np}} + \sqrt{\frac{\kappa^2\mu^{3/2}r^2\kappa_\omega\log^{7/2}(n+d)}{ndp^2}}\right.$$
$$\left. + \frac{\omega_{\max}^2}{p\sigma_r^{\star 2}}\sqrt{\frac{d\kappa\mu^{1/2}r\log^{5/2}(n+d)}{n}}\right]\sqrt{\frac{r}{d}}.$$

- Regarding $\gamma_{4,1,2,3}$, we can obtain

$$\gamma_{4,1,2,3} \asymp \underbrace{\frac{\omega_i^{\star 2}\omega_j^{\star 2}}{n^{2.5}dp^3}\kappa^2\mu r^{5/2}\log^{5/2}(n+d)}_{=:\gamma_{4,1,2,3,1}} + \underbrace{\frac{\omega_i^{\star 2}\omega_j^{\star 2}}{n^{2.5}d^2p^4}\kappa^3\mu^2 r^{7/2}\log^{7/2}(n+d)}_{=:\gamma_{4,1,2,3,2}}$$
$$+ \underbrace{\frac{\omega_{\max}^2\omega_i^{\star 2}\omega_j^{\star 2}}{\sigma_r^{\star 2}}\frac{\kappa^2\mu r^{5/2}\log^{5/2}(n+d)}{n^{5/2}dp^4}}_{=:\gamma_{4,1,2,3,3}} \lesssim \delta v_{i,j}^{\star}.$$

Here, the last relation holds since

$$\gamma_{4,1,2,3,1}+\gamma_{4,1,2,3,3}\lesssim\delta\frac{\sigma_r^{\star2}}{ndp^2}\left(\omega_j^{\star2}\left\|\boldsymbol{U}_{i,\cdot}^{\star}\right\|_2^2+\omega_i^{\star2}\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2^2\right)\lesssim\delta v_{i,j}^{\star},$$

$$\gamma_{4,1,2,3,2}\lesssim\delta\frac{\omega_{\min}^2}{np^2}\left(\omega_j^{\star2}\left\|\boldsymbol{U}_{i,\cdot}^{\star}\right\|_2^2+\omega_i^{\star2}\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2^2\right)\lesssim\delta v_{i,j}^{\star},$$

provided that

$$\begin{aligned}\left\|\boldsymbol{U}_{i,\cdot}^{\star}\right\|_2+\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2\gtrsim\delta^{-1/2}\frac{\kappa^{1/2}\mu^{1/4}r^{1/4}}{n^{1/4}}&\left[\frac{\omega_{\max}}{\sigma_r^{\star}}\sqrt{\frac{d\kappa\mu^{1/2}r\log^{5/2}(n+d)}{np}}+\sqrt{\frac{\kappa^2\mu^{3/2}r^2\kappa_\omega\log^{7/2}(n+d)}{ndp^2}}\right.\\&\left.+\frac{\omega_{\max}^2}{p\sigma_r^{\star2}}\sqrt{\frac{d\kappa\mu^{1/2}r\log^{5/2}(n+d)}{n}}\right]\sqrt{\frac{r}{d}}.\end{aligned}$$

- For $\gamma_{4,1,2,4}$, $\gamma_{4,1,2,5}$ and $\gamma_{4,1,2,6}$, it is straightforward to check that

$$\gamma_{4,1,2,4}\lesssim\delta\frac{\omega_{\min}^2}{np^2}\left(\omega_j^{\star2}\left\|\boldsymbol{U}_{i,\cdot}^{\star}\right\|_2^2+\omega_i^{\star2}\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2^2\right)\lesssim\delta v_{i,j}^{\star},$$

$$\gamma_{4,1,2,5}+\gamma_{4,1,2,6}\lesssim\delta\frac{\sigma_r^{\star2}}{ndp^2}\left(\omega_j^{\star2}\left\|\boldsymbol{U}_{i,\cdot}^{\star}\right\|_2^2+\omega_i^{\star2}\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2^2\right)\lesssim\delta v_{i,j}^{\star},$$

provided that $\theta\lesssim1$, $np\gtrsim\delta^{-2}\kappa_\omega^2\log^2(n+d)$, $n\gtrsim\delta^{-2}\kappa^4\mu^2r^3\log(n+d)$ and

$$\frac{\zeta_{\mathsf{1st}}}{\sigma_r^{\star2}}\lesssim\frac{\delta}{\sqrt{\kappa^2\mu r^2\log(n+d)}}.$$

Taking the bounds on $\gamma_{4,1,2,1}$ to $\gamma_{4,1,2,6}$ collectively yields

$$\gamma_{4,1,2}\lesssim\gamma_{4,1,2,1}+\gamma_{4,1,2,2}+\gamma_{4,1,2,3}+\gamma_{4,1,2,4}+\gamma_{4,1,2,5}+\gamma_{4,1,2,6}\lesssim\delta v_{i,j}^{\star}.$$

Then we can combine the bounds on $\gamma_{4,1,1}$ and $\gamma_{4,1,2}$ to arrive at

$$\gamma_{4,1}\lesssim\gamma_{4,1,1}+\gamma_{4,1,2}\lesssim\delta v_{i,j}^{\star},$$

provided that $\theta\lesssim1$, $np\gtrsim\delta^{-2}\kappa^2\mu^2r^2\kappa_\omega^2\log^2(n+d)$, $n\gtrsim\delta^{-2}\kappa^4\mu^2r^3\log(n+d)$ and

$$\frac{\zeta_{\mathsf{1st}}}{\sigma_r^{\star2}}\lesssim\frac{\delta}{\sqrt{\kappa^2\mu r^2\log(n+d)}},$$

$$\left\|\boldsymbol{U}_{i,\cdot}^{\star}\right\|_2+\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2\gtrsim\delta^{-1/2}\frac{\sqrt{\kappa_\omega\log^3(n+d)}}{n^{1/4}p^{1/4}}\left[\frac{\omega_{\max}}{\sigma_r^{\star}}\sqrt{\frac{d}{np}}+\frac{\omega_{\max}^2}{p\sigma_r^{\star2}}\sqrt{\frac{d}{n}}\right]\sqrt{\frac{r}{d}},$$

and

$$\begin{aligned}\left\|\boldsymbol{U}_{i,\cdot}^{\star}\right\|_2+\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2\gtrsim\delta^{-1/2}\left(\sqrt{\frac{\zeta_{\mathsf{1st}}}{\sigma_r^{\star2}}}+\frac{\kappa^{1/2}\mu^{1/4}r^{1/4}}{n^{1/4}}\right)&\left[\frac{\omega_{\max}}{\sigma_r^{\star}}\sqrt{\frac{d\kappa\mu^{1/2}r\log^{5/2}(n+d)}{np}}\right.\\&\left.+\sqrt{\frac{\kappa^2\mu^{3/2}r^2\kappa_\omega\log^{7/2}(n+d)}{ndp^2}}+\frac{\omega_{\max}^2}{p\sigma_r^{\star2}}\sqrt{\frac{d\kappa\mu^{1/2}r\log^{5/2}(n+d)}{n}}\right]\sqrt{\frac{r}{d}}.\end{aligned}$$

Step 4.2: bounding $\gamma_{4,2}$. In view of (E.120), (E.121) and (E.122), we can develop the following upper bound:

$$
\begin{aligned}
\gamma_{4,2} &\lesssim \frac{\Delta_i}{np^2}\max_{1\le k\le d}\left[\omega_k^2+(1-p)\,S_{k,k}\right]\|\boldsymbol{U}_{j,\cdot}\|_2^2 \\
&\lesssim \frac{1}{np^2}\left(\sqrt{\frac{\log^2(n+d)}{np}}\omega_i^{\star 2}+\mathsf{UB}_i\right)\left(\omega_{\max}^2+\frac{\mu r}{d}\sigma_1^{\star 2}\right)\left(\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2^2+\frac{\theta^2}{\kappa\sigma_r^{\star 2}}\omega_j^{\star 2}\right) \\
&\lesssim \delta v_{i,j}^{\star}+\frac{1}{np^2}\mathsf{UB}_i\left(\omega_{\max}^2+\frac{\mu r}{d}\sigma_1^{\star 2}\right)\left(\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2^2+\frac{\theta^2}{\kappa\sigma_r^{\star 2}}\omega_j^{\star 2}\right) \\
&\quad+\frac{1}{np^2}\sqrt{\frac{\log^2(n+d)}{np}}\omega_i^{\star 2}\frac{\mu r}{d}\sigma_1^{\star 2}\left(\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2^2+\frac{\theta^2}{\kappa\sigma_r^{\star 2}}\omega_j^{\star 2}\right) \\
&\lesssim \delta v_{i,j}^{\star}+\underbrace{\frac{1}{np^2}\mathsf{UB}_i\omega_{\max}^2\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2^2}_{=:\gamma_{4,2,1}}+\underbrace{\frac{1}{np^2}\mathsf{UB}_i\omega_{\max}^2\omega_j^{\star 2}\frac{\theta^2}{\kappa\sigma_r^{\star 2}}}_{=:\gamma_{4,2,2}}+\underbrace{\frac{\mu r}{ndp^2}\mathsf{UB}_i\sigma_1^{\star 2}\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2^2}_{=:\gamma_{4,2,3}} \\
&\quad+\underbrace{\frac{\mu r}{ndp^2}\mathsf{UB}_i\theta^2\omega_j^{\star 2}}_{=:\gamma_{4,2,4}}+\underbrace{\frac{\mu r}{ndp^2}\sqrt{\frac{\log^2(n+d)}{np}}\omega_i^{\star 2}\sigma_1^{\star 2}\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2^2}_{=:\gamma_{4,2,5}}+\underbrace{\frac{\mu r}{ndp^2}\sqrt{\frac{\log^2(n+d)}{np}}\omega_i^{\star 2}\omega_j^{\star 2}\theta^2}_{=:\gamma_{4,2,6}},
\end{aligned}
$$

where the penultimate relation holds since

$$
\frac{1}{np^2}\sqrt{\frac{\log^2(n+d)}{np}}\omega_i^{\star 2}\omega_{\max}^2\left(\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2^2+\frac{\theta^2}{\kappa\sigma_r^{\star 2}}\omega_j^{\star 2}\right)\lesssim\gamma_{4,1}\lesssim\delta v_{i,j}^{\star}.
$$

In what follows, we shall bound the six terms from $\gamma_{4,2,1}$ to $\gamma_{4,2,6}$ separately.

- Regarding $\gamma_{4,2,1}$, we can derive

$$
\begin{aligned}
\gamma_{4,2,1} &\asymp \underbrace{\frac{1}{np^2}\left(\theta+\sqrt{\frac{\kappa^3 r\log(n+d)}{n}}\right)\left\|\boldsymbol{U}_{i,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2^2\omega_{\max}^2\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2^2}_{=:\gamma_{4,2,1,1}}+\underbrace{\frac{1}{np^2}\theta^2\omega_i^{\star 2}\omega_{\max}^2\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2^2}_{=:\gamma_{4,2,1,2}} \\
&\quad+\underbrace{\frac{1}{np^2}\theta\omega_i^{\star}\left\|\boldsymbol{U}_{i,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2\omega_{\max}^2\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2^2}_{=:\gamma_{4,2,1,3}}+\underbrace{\frac{1}{np^2}\zeta_{\mathsf{2nd},i}\sigma_1^{\star}\left\|\boldsymbol{U}_{i,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2\omega_{\max}^2\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2^2}_{=:\gamma_{4,2,1,4}} \\
&\quad+\underbrace{\frac{1}{np^2}\zeta_{\mathsf{2nd},i}^2\sigma_1^{\star 2}\omega_{\max}^2\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2^2}_{=:\gamma_{4,2,1,5}}\lesssim\delta v_{i,j}^{\star}.
\end{aligned}
$$

  Here, the last relation holds since

$$
\begin{aligned}
\gamma_{4,2,1,1}+\gamma_{4,2,1,3}+\gamma_{4,2,1,4}+\gamma_{4,2,1,5} &\lesssim \delta\frac{\sigma_r^{\star 2}}{ndp^2}\left(\omega_j^{\star 2}\left\|\boldsymbol{U}_{i,\cdot}^{\star}\right\|_2^2+\omega_i^{\star 2}\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2^2\right)\lesssim\delta v_{i,j}^{\star}, \\
\gamma_{4,2,1,2} &\lesssim \delta\frac{\omega_{\min}^2}{np^2}\left(\omega_j^{\star 2}\left\|\boldsymbol{U}_{i,\cdot}^{\star}\right\|_2^2+\omega_i^{\star 2}\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2^2\right)\lesssim\delta v_{i,j}^{\star},
\end{aligned}
$$

  provided that $\theta\lesssim\delta/(\kappa\mu r\kappa_\omega)$, $n\gtrsim\delta^{-2}\kappa^5\mu^2r^3\kappa_\omega^2\log(n+d)$, $\zeta_{\mathsf{2nd},i}\sqrt{d}\lesssim\delta/\sqrt{\kappa^2\mu r\kappa_\omega^2}$,

$$
\theta\frac{\omega_{\max}}{\sigma_r^{\star}}\sqrt{d}\lesssim\frac{\delta}{\sqrt{\kappa\mu r\kappa_\omega}}. \tag{E.123}
$$

  Note that in view of (E.64) and (E.65), (E.123) is guaranteed by $ndp^2\gtrsim\delta^{-2}\kappa^4\mu^3r^4\kappa_\omega\log^4(n+d)$,

$$
\frac{\omega_{\max}}{\sigma_r^{\star}}\sqrt{\frac{d}{np}}\lesssim\frac{\delta}{\sqrt{\kappa^2\mu r^2\kappa_\omega\log^2(n+d)}},\qquad\frac{\omega_{\max}^2}{p\sigma_r^{\star 2}}\sqrt{\frac{d}{n}}\lesssim\frac{\delta}{\sqrt{\kappa^2\mu r^2\kappa_\omega\log^2(n+d)}}.
$$

- Regarding $\gamma_{4,2,2}$, we have

$$\gamma_{4,2,2} \asymp \underbrace{\frac{1}{np^2}\left(\theta+\sqrt{\frac{\kappa^3 r\log(n+d)}{n}}\right)\left\|\boldsymbol{U}_{i,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2^2\omega_{\max}^2\omega_j^{\star 2}\frac{\theta^2}{\kappa\sigma_r^{\star 2}}}_{=:\gamma_{4,2,2,1}}+\underbrace{\frac{1}{np^2}\theta^2\omega_{\max}^2\omega_i^{\star 2}\omega_j^{\star 2}\frac{\theta^2}{\kappa\sigma_r^{\star 2}}}_{=:\gamma_{4,2,2,2}}$$
$$+\underbrace{\frac{1}{np^2}\theta\omega_{\max}^2\omega_i^{\star}\omega_j^{\star 2}\left\|\boldsymbol{U}_{i,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2\frac{\theta^2}{\kappa\sigma_r^{\star 2}}}_{=:\gamma_{4,2,2,3}}+\underbrace{\frac{1}{np^2}\zeta_{\mathsf{2nd},i}\omega_{\max}^2\omega_j^{\star 2}\left\|\boldsymbol{U}_{i,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2\frac{\theta^2}{\sqrt{\kappa}\sigma_r^{\star}}}_{=:\gamma_{4,2,2,4}}$$
$$+\underbrace{\frac{1}{np^2}\zeta_{\mathsf{2nd},i}^2\omega_{\max}^2\omega_j^{\star 2}\theta^2}_{=:\gamma_{4,2,2,5}}\lesssim \delta v_{i,j}^{\star}.$$

Here, the last relation holds since

$$\gamma_{4,2,2,1}\lesssim \delta\frac{\omega_{\min}^2}{np^2}\left(\omega_j^{\star 2}\left\|\boldsymbol{U}_{i,\cdot}^{\star}\right\|_2^2+\omega_i^{\star 2}\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2^2\right)\lesssim \delta v_{i,j}^{\star},$$
$$\gamma_{4,2,2,2}\lesssim \frac{\omega_{\max}^2}{\sigma_r^{\star 2}}\frac{1}{np^2\kappa}\cdot\theta^4\omega_i^{\star 2}\omega_j^{\star 2}\overset{\text{(i)}}{\lesssim}\frac{\omega_{\max}^2}{\sigma_r^{\star 2}}\frac{1}{np^2\kappa}\varepsilon^2 v_{i,j}^{\star}\lesssim \delta v_{i,j}^{\star},$$
$$\gamma_{4,2,2,3}\lesssim \theta\frac{\omega_{\max}}{\sqrt{\kappa}\sigma_r^{\star}}\sqrt{\frac{\kappa_\omega}{np^2}}\cdot\frac{\omega_{\min}\omega_j^{\star}}{\sqrt{np^2}}\left\|\boldsymbol{U}_{i,\cdot}^{\star}\right\|_2\cdot\theta^2\omega_i^{\star}\omega_j^{\star}\overset{\text{(ii)}}{\lesssim}\varepsilon\theta\frac{\omega_{\max}}{\sqrt{\kappa}\sigma_r^{\star}}\sqrt{\frac{\kappa_\omega}{np^2}}v_{i,j}^{\star}\lesssim \delta v_{i,j}^{\star},$$
$$\gamma_{4,2,2,4}\lesssim \frac{\kappa_\omega}{\sqrt{ndp^2}}\zeta_{\mathsf{2nd},i}\sqrt{d}\cdot\theta^2\omega_i^{\star}\omega_j^{\star}\cdot\frac{\omega_{\min}\omega_j^{\star}}{\sqrt{np^2}}\left\|\boldsymbol{U}_{i,\cdot}^{\star}\right\|_2\overset{\text{(iii)}}{\lesssim}\frac{\varepsilon\kappa_\omega}{\sqrt{ndp^2}}\zeta_{\mathsf{2nd},i}\sqrt{d}v_{i,j}^{\star}\lesssim \delta v_{i,j}^{\star},$$
$$\gamma_{4,2,2,5}\lesssim \frac{\omega_{\max}^2}{\sigma_1^{\star 2}}\frac{1}{np^2}\cdot\zeta_{\mathsf{2nd},i}^2\sigma_1^{\star 2}\cdot\theta^2\omega_{\max}^2\lesssim\frac{\omega_{\max}^2}{\sigma_1^{\star 2}}\frac{\kappa_\omega}{np^2}\varepsilon^2 v_{i,j}^{\star}\lesssim \delta v_{i,j}^{\star},$$

provided that $\theta\lesssim\delta/\kappa_\omega$, $n\gtrsim\kappa^3 r\log(n+d)$, $\varepsilon\lesssim 1$, $\zeta_{\mathsf{2nd},i}\sqrt{d}\lesssim\delta$, $ndp^2\gtrsim\kappa_\omega^2$,

$$\frac{\omega_{\max}^2}{\sigma_r^{\star 2}}\frac{1}{np^2\kappa}\lesssim\frac{\delta}{\kappa_\omega}.$$

Here, (i)-(iii) rely on (E.115). In view of the fact that

$$\frac{1}{\sqrt{ndp^2}}\cdot\frac{\omega_{\max}^2}{p\sigma_r^{\star 2}}\sqrt{\frac{d}{n}}=\frac{\omega_{\max}^2}{\sigma_r^{\star 2}}\frac{1}{np^2},$$

the last condition above is guaranteed by $ndp^2\gtrsim\kappa_\omega^2$ and

$$\frac{\omega_{\max}^2}{p\sigma_r^{\star 2}}\sqrt{\frac{d}{n}}\lesssim\delta.$$

- Regarding $\gamma_{4,2,3}$, we have

$$\gamma_{4,2,3}\asymp\underbrace{\frac{\mu r}{ndp^2}\sigma_1^{\star 2}\left(\theta+\sqrt{\frac{\kappa^3 r\log(n+d)}{n}}\right)\left\|\boldsymbol{U}_{i,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2^2\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2^2}_{=:\gamma_{4,2,3,1}}+\underbrace{\frac{\mu r}{ndp^2}\sigma_1^{\star 2}\theta^2\omega_i^{\star 2}\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2^2}_{=:\gamma_{4,2,3,2}}$$
$$+\underbrace{\frac{\mu r}{ndp^2}\sigma_1^{\star 2}\theta\omega_i^{\star}\left\|\boldsymbol{U}_{i,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2^2}_{=:\gamma_{4,2,3,3}}+\underbrace{\frac{\mu r}{ndp^2}\sigma_1^{\star 3}\zeta_{\mathsf{2nd},i}\left\|\boldsymbol{U}_{i,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2^2}_{=:\gamma_{4,2,3,4}}$$
$$+\underbrace{\frac{\mu r}{ndp^2}\sigma_1^{\star 2}\zeta_{\mathsf{2nd},i}^2\sigma_1^{\star 2}\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2^2}_{=:\gamma_{4,2,3,5}}\lesssim\delta v_{i,j}^{\star}.$$

Here, the last relation holds since

$$
\begin{aligned}
\gamma_{4,2,3,1} &\lesssim \frac{\delta}{ndp^2\kappa}\left\|\boldsymbol{U}_{i,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2^2\left\|\boldsymbol{U}_{j,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2^2 \lesssim \delta v_{i,j}^{\star},\\
\gamma_{4,2,3,2} &\lesssim \delta\frac{\sigma_r^{\star 2}}{ndp^2}\left(\omega_j^{\star 2}\left\|\boldsymbol{U}_{i,\cdot}^{\star}\right\|_2^2+\omega_i^{\star 2}\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2^2\right) \lesssim \delta v_{i,j}^{\star},\\
\gamma_{4,2,3,3} &\overset{\text{(i)}}{\lesssim} \delta\frac{1}{ndp^2\kappa}\left\|\boldsymbol{U}_{i,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2^2\left\|\boldsymbol{U}_{j,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2^2+\delta\frac{\sigma_r^{\star 2}}{ndp^2}\omega_i^{\star 2}\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2^2 \lesssim \delta v_{i,j}^{\star},\\
\gamma_{4,2,3,4} &\lesssim \frac{\kappa\mu r}{\sqrt{ndp^2}}\cdot\frac{1}{\sqrt{ndp^2\kappa}}\left\|\boldsymbol{U}_{i,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2\left\|\boldsymbol{U}_{j,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2\cdot\sigma_1^{\star 2}\left(\left\|\boldsymbol{U}_{i,\cdot}^{\star}\right\|_2\zeta_{\mathsf{2nd},j}+\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2\zeta_{\mathsf{2nd},i}\right)\\
&\overset{\text{(ii)}}{\lesssim} \frac{\kappa\mu r}{\sqrt{ndp^2}}\cdot\frac{1}{\sqrt{ndp^2\kappa}}\left\|\boldsymbol{U}_{i,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2\left\|\boldsymbol{U}_{j,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2\cdot\varepsilon(v_{i,j}^{\star})^{1/2} \lesssim \frac{\kappa\mu r}{\sqrt{ndp^2}}\varepsilon v_{i,j}^{\star} \lesssim \delta v_{i,j}^{\star},\\
\gamma_{4,2,3,5} &\lesssim \frac{\mu r}{ndp^2}\left[\sigma_1^{\star 2}\left(\left\|\boldsymbol{U}_{i,\cdot}^{\star}\right\|_2\zeta_{\mathsf{2nd},j}+\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2\zeta_{\mathsf{2nd},i}\right)\right]^2 \overset{\text{(iii)}}{\lesssim} \frac{\mu r}{ndp^2}\varepsilon v_{i,j}^{\star} \lesssim \delta v_{i,j}^{\star},
\end{aligned}
$$

provided that $\theta \lesssim \delta/(\kappa^2\mu r)$, $n \gtrsim \delta^{-2}\kappa^7\mu^2 r^3\log(n+d)$, $\varepsilon \lesssim 1$ and $ndp^2 \gtrsim \delta^{-2}\kappa^2\mu^2 r^2$. Here, (i) utilizes the AM-GM inequality, whereas (ii) and (iii) utilize (E.114).

- Regarding $\gamma_{4,2,4}$, it follows that

$$
\begin{aligned}
\gamma_{4,2,4} \asymp\, & \underbrace{\frac{\mu r}{ndp^2}\theta^2\omega_j^{\star 2}\left(\theta+\sqrt{\frac{\kappa^3 r\log(n+d)}{n}}\right)\left\|\boldsymbol{U}_{i,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2^2}_{=:\gamma_{4,2,4,1}}+\underbrace{\frac{\mu r}{ndp^2}\theta^4\omega_i^{\star 2}\omega_j^{\star 2}}_{=:\gamma_{4,2,4,2}}\\
&+\underbrace{\frac{\mu r}{ndp^2}\theta^3\omega_i^{\star}\omega_j^{\star 2}\left\|\boldsymbol{U}_{i,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2}_{=:\gamma_{4,2,4,3}}+\underbrace{\frac{\mu r}{ndp^2}\theta^2\omega_j^{\star 2}\sigma_1^{\star}\zeta_{\mathsf{2nd},i}\left\|\boldsymbol{U}_{i,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2}_{=:\gamma_{4,2,4,4}}\\
&+\underbrace{\frac{\mu r}{ndp^2}\theta^2\omega_j^{\star 2}\zeta_{\mathsf{2nd},i}^2\sigma_1^{\star 2}}_{=:\gamma_{4,2,4,5}} \lesssim \delta v_{i,j}^{\star},
\end{aligned}
$$

where the last relation holds since

$$
\begin{aligned}
\gamma_{4,2,4,1} &\lesssim \delta\frac{\sigma_r^{\star 2}}{ndp^2}\left(\omega_j^{\star 2}\left\|\boldsymbol{U}_{i,\cdot}^{\star}\right\|_2^2+\omega_i^{\star 2}\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2^2\right) \lesssim \delta v_{i,j}^{\star},\\
\gamma_{4,2,4,2} &\overset{\text{(i)}}{\lesssim} \frac{\mu r}{ndp^2}\varepsilon^2 v_{i,j}^{\star} \lesssim \delta v_{i,j}^{\star},\\
\gamma_{4,2,4,3} &\overset{\text{(ii)}}{\lesssim} \frac{\sqrt{\kappa}\mu r}{\sqrt{ndp^2}}\theta\cdot\theta^2\omega_i^{\star}\omega_j^{\star}\cdot\frac{\sigma_r^{\star}}{\sqrt{ndp^2}}\omega_j^{\star}\left\|\boldsymbol{U}_{i,\cdot}^{\star}\right\|_2 \lesssim \frac{\sqrt{\kappa}\mu r}{\sqrt{ndp^2}}\varepsilon\theta v_{i,j}^{\star} \lesssim \delta v_{i,j}^{\star},\\
\gamma_{4,2,4,4} &\lesssim \frac{\mu r\kappa_\omega^{1/2}}{ndp^2}\cdot\theta^2\omega_i^{\star}\omega_j^{\star}\cdot\sigma_1^{\star 2}\left(\left\|\boldsymbol{U}_{i,\cdot}^{\star}\right\|_2\zeta_{\mathsf{2nd},j}+\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2\zeta_{\mathsf{2nd},i}\right) \overset{\text{(iii)}}{\lesssim} \frac{\mu r\kappa_\omega^{1/2}\varepsilon^2}{ndp^2}v_{i,j}^{\star} \lesssim \delta v_{i,j}^{\star},\\
\gamma_{4,2,4,5} &\lesssim \frac{\mu r}{ndp^2}\cdot\left(\theta\sigma_1^{\star}\omega_j^{\star}\zeta_{\mathsf{2nd},i}\right)^2 \overset{\text{(iv)}}{\lesssim} \frac{\mu r}{ndp^2}\varepsilon^2 v_{i,j}^{\star} \lesssim \delta v_{i,j}^{\star},
\end{aligned}
$$

as long as $\theta \lesssim \delta/(\kappa\mu r)$, $n \gtrsim \kappa^3 r\log(n+d)$, $ndp^2 \gtrsim \delta^{-2}\mu r$, $\varepsilon \lesssim 1$. Here, (i), (ii) and (iv) utilize (E.115); (iii) utilizes (E.113).

- Regarding $\gamma_{4,2,5}$, we can derive

$$
\gamma_{4,2,5} \lesssim \delta\frac{\sigma_r^{\star 2}}{ndp^2}\left(\omega_j^{\star 2}\left\|\boldsymbol{U}_{i,\cdot}^{\star}\right\|_2^2+\omega_i^{\star 2}\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2^2\right) \lesssim \delta v_{i,j}^{\star}
$$

provided that $np \gtrsim \delta^{-2}\kappa^2\mu^2 r^2\log^2(n+d)$.

- Finally, when it comes to $\gamma_{4,2,6}$, we can bound

$$\gamma_{4,2,6} \overset{\text{(i)}}{\lesssim} \theta^2 \sqrt{\frac{1}{np}\omega_i^{\star 2}\omega_j^{\star 2}\theta^2} \overset{\text{(ii)}}{\lesssim} \sqrt{\frac{1}{np}} v_{i,j}^{\star} \overset{\text{(iii)}}{\lesssim} \delta v_{i,j}^{\star}.$$

Here (i) follows from (E.64), (ii) comes from (E.115), whereas (iii) holds provided that $np \gtrsim \delta^{-2}$.

Take the bounds on the terms (from $\gamma_{4,2,1}$ to $\gamma_{4,2,6}$) together to arrive at

$$\gamma_{4,2} \lesssim \gamma_{4,2,1} + \gamma_{4,2,2} + \gamma_{4,2,3} + \gamma_{4,2,4} + \gamma_{4,2,5} + \gamma_{4,2,6} \lesssim \delta v_{i,j}^{\star},$$

provided that $\theta \lesssim \delta/(\kappa^2\mu r\kappa_\omega)$, $\varepsilon \lesssim 1$, $\zeta_{\mathsf{2nd},i}\sqrt{d} \lesssim \delta/\sqrt{\kappa^2\mu r\kappa_\omega^2}$, $n \gtrsim \delta^{-2}\kappa^7\mu^2 r^3\kappa_\omega^2 \log(n+d)$, $ndp^2 \gtrsim \delta^{-2}\kappa^4\mu^3 r^4\kappa_\omega^2 \log^4(n+d)$, $np \gtrsim \delta^{-2}\kappa^2\mu^2 r^2 \log^2(n+d)$

$$\frac{\omega_{\max}}{\sigma_r^{\star}}\sqrt{\frac{d}{np}} \lesssim \frac{\delta}{\sqrt{\kappa^2\mu r^2\kappa_\omega \log^2(n+d)}} \qquad \text{and} \qquad \frac{\omega_{\max}^2}{p\sigma_r^{\star 2}}\sqrt{\frac{d}{n}} \lesssim \frac{\delta}{\sqrt{\kappa^2\mu r^2\kappa_\omega \log^2(n+d)}}.$$

Step 4.3: bounding $\gamma_{4,3}$. In view of (E.121), we can upper bound

$$\begin{aligned}
\gamma_{4,3} &\lesssim \frac{1}{np^2}\left(\omega_i^{\star 2} + \left\|\boldsymbol{U}_{i,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2^2\right)\left(\omega_{\max}^2 + \frac{\mu r}{d}\sigma_1^{\star 2}\right)\sum_{k=1}^{d}\left|\left(\boldsymbol{U}_{k,\cdot}^{\star}\boldsymbol{U}_{j,\cdot}^{\star\top}\right)^2 - \left(\boldsymbol{U}_{k,\cdot}\boldsymbol{U}_{j,\cdot}^{\top}\right)^2\right| \\
&\lesssim \underbrace{\frac{\omega_i^{\star 2}\omega_{\max}^2}{np^2}\sum_{k=1}^{d}\left|\left(\boldsymbol{U}_{k,\cdot}^{\star}\boldsymbol{U}_{j,\cdot}^{\star\top}\right)^2 - \left(\boldsymbol{U}_{k,\cdot}\boldsymbol{U}_{j,\cdot}^{\top}\right)^2\right|}_{=:\gamma_{4,3,1}} + \underbrace{\frac{\mu r}{ndp^2}\sigma_1^{\star 2}\omega_i^{\star 2}\sum_{k=1}^{d}\left|\left(\boldsymbol{U}_{k,\cdot}^{\star}\boldsymbol{U}_{j,\cdot}^{\star\top}\right)^2 - \left(\boldsymbol{U}_{k,\cdot}\boldsymbol{U}_{j,\cdot}^{\top}\right)^2\right|}_{=:\gamma_{4,3,2}} \\
&\quad + \underbrace{\frac{\mu r}{ndp^2}\sigma_1^{\star 2}\left\|\boldsymbol{U}_{i,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2^2\sum_{k=1}^{d}\left|\left(\boldsymbol{U}_{k,\cdot}^{\star}\boldsymbol{U}_{j,\cdot}^{\star\top}\right)^2 - \left(\boldsymbol{U}_{k,\cdot}\boldsymbol{U}_{j,\cdot}^{\top}\right)^2\right|}_{=:\gamma_{4,3,3}} \\
&\quad + \underbrace{\frac{1}{np^2}\left\|\boldsymbol{U}_{i,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2^2\omega_{\max}^2\sum_{k=1}^{d}\left|\left(\boldsymbol{U}_{k,\cdot}^{\star}\boldsymbol{U}_{j,\cdot}^{\star\top}\right)^2 - \left(\boldsymbol{U}_{k,\cdot}\boldsymbol{U}_{j,\cdot}^{\top}\right)^2\right|}_{=:\gamma_{4,3,4}}.
\end{aligned}$$

Note that for each $k \in [d]$,

$$\begin{aligned}
\left|\boldsymbol{U}_{k,\cdot}^{\star}\boldsymbol{U}_{j,\cdot}^{\star\top} - \boldsymbol{U}_{k,\cdot}\boldsymbol{U}_{j,\cdot}^{\top}\right| &= \left|\boldsymbol{U}_{k,\cdot}^{\star}\boldsymbol{U}_{j,\cdot}^{\star\top} - \boldsymbol{U}_{k,\cdot}\boldsymbol{R}\left(\boldsymbol{U}_{j,\cdot}\boldsymbol{R}\right)^{\top}\right| \\
&\leq \left|\left(\boldsymbol{U}\boldsymbol{R} - \boldsymbol{U}^{\star}\right)_{k,\cdot}\boldsymbol{U}_{j,\cdot}^{\star}\right| + \left|\left(\boldsymbol{U}\boldsymbol{R}\right)_{k,\cdot}\left(\boldsymbol{U}\boldsymbol{R} - \boldsymbol{U}^{\star}\right)_{j,\cdot}^{\top}\right| \\
&\leq \left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2\left\|\boldsymbol{U}\boldsymbol{R} - \boldsymbol{U}^{\star}\right\|_{2,\infty} + \left\|\boldsymbol{U}_{k,\cdot}\right\|_2\left\|\left(\boldsymbol{U}\boldsymbol{R} - \boldsymbol{U}^{\star}\right)_{j,\cdot}\right\|_2 \\
&\overset{\text{(i)}}{\lesssim} \left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2\frac{\zeta_{\mathsf{1st}}}{\sigma_r^{\star 2}}\sqrt{\frac{r\log(n+d)}{d}} + \left\|\boldsymbol{U}_{k,\cdot}\right\|_2\left[\frac{\theta}{\sqrt{\kappa}\sigma_r^{\star}}\left(\left\|\boldsymbol{U}_{j,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2 + \omega_j^{\star}\right)\right] \\
&\overset{\text{(ii)}}{\lesssim} \left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2\frac{\zeta_{\mathsf{1st}}}{\sigma_r^{\star 2}}\sqrt{\frac{r\log(n+d)}{d}} + \left\|\boldsymbol{U}_{k,\cdot}\right\|_2\frac{\theta}{\sqrt{\kappa}\sigma_r^{\star}}\omega_j^{\star}.
\end{aligned}$$

Here, (i) follows from Lemma 18 as well as a direct consequence of Lemma 18, (E.114) and (E.64):

$$\begin{aligned}
\left\|\left(\boldsymbol{U}\boldsymbol{R} - \boldsymbol{U}^{\star}\right)_{j,\cdot}\right\|_2 &\lesssim \frac{\theta}{\sqrt{\kappa}\sigma_r^{\star}}\left(\left\|\boldsymbol{U}_{j,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2 + \omega_j^{\star}\right) + \zeta_{\mathsf{2nd},j} \\
&\lesssim \frac{\theta}{\sqrt{\kappa}\sigma_r^{\star}}\left(\left\|\boldsymbol{U}_{j,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2 + \omega_j^{\star}\right) + \frac{\varepsilon}{\sqrt{\kappa}}\frac{\left\|\boldsymbol{U}_{j,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2 + \omega_j^{\star}}{\sqrt{\min\{ndp^2\kappa, np\}}\sigma_1^{\star}} + \frac{\varepsilon}{\sqrt{\kappa}}\frac{\omega_{\min}\omega_j^{\star}}{\sqrt{np^2}\sigma_1^{\star 2}}
\end{aligned}$$

$$\lesssim \frac{\theta}{\sqrt{\kappa}\sigma_r^\star}\left(\left\|\boldsymbol{U}_{j,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2+\omega_j^\star\right);$$

and (ii) utilizes (E.66). Therefore, one can derive

$$\begin{aligned}
\sum_{k=1}^d\left|\left(\boldsymbol{U}_{k,\cdot}^\star\boldsymbol{U}_{j,\cdot}^{\star\top}\right)^2-\left(\boldsymbol{U}_{k,\cdot}\boldsymbol{U}_{j,\cdot}^{\top}\right)^2\right| &\lesssim \sum_{k=1}^d\left|\boldsymbol{U}_{k,\cdot}^\star\boldsymbol{U}_{j,\cdot}^{\star\top}-\boldsymbol{U}_{k,\cdot}\boldsymbol{U}_{j,\cdot}^{\top}\right|\left\|\boldsymbol{U}_{j,\cdot}^\star\right\|_2\left\|\boldsymbol{U}_{k,\cdot}^\star\right\|_2+\sum_{k=1}^d\left|\boldsymbol{U}_{k,\cdot}^\star\boldsymbol{U}_{j,\cdot}^{\star\top}-\boldsymbol{U}_{k,\cdot}\boldsymbol{U}_{j,\cdot}^{\top}\right|^2\\
&\lesssim \sum_{k=1}^d\left\|\boldsymbol{U}_{j,\cdot}^\star\right\|_2^2\frac{\zeta_{\mathsf{1st}}}{\sigma_r^{\star2}}\sqrt{\frac{r\log(n+d)}{d}}\left\|\boldsymbol{U}_{k,\cdot}^\star\right\|_2+\sum_{k=1}^d\left\|\boldsymbol{U}_{k,\cdot}\right\|_2\left\|\boldsymbol{U}_{k,\cdot}^\star\right\|_2\frac{\theta}{\sqrt{\kappa}\sigma_r^\star}\omega_j^\star\left\|\boldsymbol{U}_{j,\cdot}^\star\right\|_2\\
&\quad+\sum_{k=1}^d\left\|\boldsymbol{U}_{j,\cdot}^\star\right\|_2^2\frac{\zeta_{\mathsf{1st}}^2}{\sigma_r^{\star4}}\frac{r\log(n+d)}{d}+\sum_{k=1}^d\left\|\boldsymbol{U}_{k,\cdot}\right\|_2^2\frac{\theta^2}{\kappa\sigma_r^{\star2}}\omega_j^{\star2}\\
&\overset{\text{(i)}}{\lesssim}\left\|\boldsymbol{U}_{j,\cdot}^\star\right\|_2^2\frac{\zeta_{\mathsf{1st}}}{\sigma_r^{\star2}}\sqrt{r^2\log(n+d)}+\frac{\theta r}{\sqrt{\kappa}\sigma_r^\star}\omega_j^\star\left\|\boldsymbol{U}_{j,\cdot}^\star\right\|_2\\
&\quad+\left\|\boldsymbol{U}_{j,\cdot}^\star\right\|_2^2\frac{\zeta_{\mathsf{1st}}^2}{\sigma_r^{\star4}}r^2\log(n+d)+\frac{\theta^2 r}{\kappa\sigma_r^{\star2}}\omega_j^{\star2}\\
&\overset{\text{(ii)}}{\lesssim}\left\|\boldsymbol{U}_{j,\cdot}^\star\right\|_2^2\frac{\zeta_{\mathsf{1st}}}{\sigma_r^{\star2}}\sqrt{r^2\log(n+d)}+\frac{\theta r}{\sqrt{\kappa}\sigma_r^\star}\omega_j^\star\left\|\boldsymbol{U}_{j,\cdot}^\star\right\|_2+\frac{\theta^2 r}{\kappa\sigma_r^{\star2}}\omega_j^{\star2}. \qquad \text{(E.124)}
\end{aligned}$$

Here, (i) follows from the Cauchy-Schwarz inequality

$$\sum_{k=1}^d\left\|\boldsymbol{U}_{k,\cdot}^\star\right\|_2\le\sqrt{d\sum_{k=1}^d\left\|\boldsymbol{U}_{k,\cdot}^\star\right\|_2^2}\le\sqrt{d\left\|\boldsymbol{U}^\star\right\|_{\mathrm{F}}^2}\le\sqrt{dr},$$

as well as the following bound

$$\begin{aligned}
\sum_{k=1}^d\left\|\boldsymbol{U}_{k,\cdot}\right\|_2\left\|\boldsymbol{U}_{k,\cdot}^\star\right\|_2 &\le\sum_{k=1}^d\left\|\boldsymbol{U}_{k,\cdot}^\star\right\|_2^2+\left\|\boldsymbol{U}\boldsymbol{R}-\boldsymbol{U}^\star\right\|_{2,\infty}\sum_{k=1}^d\left\|\boldsymbol{U}_{k,\cdot}^\star\right\|_2\\
&\lesssim\left\|\boldsymbol{U}^\star\right\|_{\mathrm{F}}^2+\frac{\zeta_{\mathsf{1st}}}{\sigma_r^{\star2}}\sqrt{\frac{r\log(n+d)}{d}}\cdot\sqrt{dr}\lesssim r,
\end{aligned}$$

which utilizes Lemma 18 and holds provided that $\zeta_{\mathsf{1st}}/\sigma_r^{\star2}\lesssim 1/\sqrt{\log(n+d)}$; (ii) holds provided that $\zeta_{\mathsf{1st}}/\sigma_r^{\star2}\lesssim 1/\sqrt{\log(n+d)}$. Then we shall bound the terms $\gamma_{4,3,1}$, $\gamma_{4,3,2}$, $\gamma_{4,3,3}$ and $\gamma_{4,3,4}$ respectively.

- Regarding $\gamma_{4,3,1}$, we have

$$\begin{aligned}
\gamma_{4,3,1} &\lesssim \underbrace{\frac{\sqrt{r^2\log(n+d)}}{np^2}\omega_i^{\star2}\omega_{\max}^2\left\|\boldsymbol{U}_{j,\cdot}^\star\right\|_2^2\frac{\zeta_{\mathsf{1st}}}{\sigma_r^{\star2}}}_{=:\gamma_{4,3,1,1}}+\underbrace{\frac{r}{np^2}\omega_i^{\star2}\omega_j^\star\omega_{\max}^2\frac{\theta}{\sqrt{\kappa}\sigma_r^\star}\left\|\boldsymbol{U}_{j,\cdot}^\star\right\|_2}_{=:\gamma_{4,3,1,2}}\\
&\quad+\underbrace{\frac{r}{np^2}\omega_i^{\star2}\omega_j^{\star2}\omega_{\max}^2\frac{\theta^2}{\kappa\sigma_r^{\star2}}}_{=:\gamma_{4,3,1,3}}\lesssim\delta v_{i,j}^\star,
\end{aligned}$$

  where the last relation holds since

$$\begin{aligned}
\gamma_{4,3,1,1}+\gamma_{4,3,1,2} &\lesssim\delta\frac{\omega_{\min}^2}{np^2}\left(\omega_j^{\star2}\left\|\boldsymbol{U}_{i,\cdot}^\star\right\|_2^2+\omega_i^{\star2}\left\|\boldsymbol{U}_{j,\cdot}^\star\right\|_2^2\right)\lesssim\delta v_{i,j}^\star,\\
\gamma_{4,3,1,3} &\lesssim\theta^2\omega_i^\star\omega_j^\star\cdot\frac{r}{np^2}\frac{\omega_{\max}^2\omega_i^\star\omega_j^\star}{\kappa\sigma_r^{\star2}}\overset{\text{(i)}}{\lesssim}\varepsilon(v_{i,j}^\star)^{1/2}\cdot\frac{r}{np^2}\frac{\omega_{\max}^2\omega_i^\star\omega_j^\star}{\kappa\sigma_r^{\star2}}\\
&\lesssim\varepsilon(v_{i,j}^\star)^{1/2}\cdot\delta\frac{\omega_{\min}}{\sqrt{np^2}}\left(\omega_j^\star\left\|\boldsymbol{U}_{i,\cdot}^\star\right\|_2+\omega_i^\star\left\|\boldsymbol{U}_{j,\cdot}^\star\right\|_2\right)\lesssim\delta v_{i,j}^\star,
\end{aligned}$$

provided that $\zeta_{\mathsf{1st}}/\sigma_r^{\star 2} \lesssim \delta/\sqrt{r^2\kappa_\omega^2 \log(n+d)}$, $\varepsilon \lesssim 1$,

$$\left\|\boldsymbol{U}_{i,\cdot}^{\star}\right\|_2 + \left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2 \gtrsim \delta^{-1} r\kappa_\omega \frac{\omega_{\max}}{\sigma_1^\star}\theta + \delta^{-1} r\kappa_\omega^{1/2}\frac{\omega_{\max}^2}{\sigma_1^{\star 2}}\sqrt{\frac{1}{np^2}}.$$

Here (i) utilizes (E.115).

- Regarding $\gamma_{4,3,2}$, we have

$$\gamma_{4,3,2} \lesssim \underbrace{\frac{\mu r^2\sqrt{\log(n+d)}}{ndp^2}\sigma_1^{\star 2}\omega_i^{\star 2}\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2^2\frac{\zeta_{\mathsf{1st}}}{\sigma_r^{\star 2}}}_{=:\gamma_{4,3,2,1}} + \underbrace{\frac{\mu r^2}{ndp^2}\sigma_1^\star\omega_i^{\star 2}\omega_j^\star\theta\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2}_{=:\gamma_{4,3,2,2}}$$
$$+ \underbrace{\frac{\mu r^2}{ndp^2}\omega_i^{\star 2}\omega_j^{\star 2}\theta^2}_{=:\gamma_{4,3,2,3}} \lesssim \delta v_{i,j}^\star.$$

Here, the last relation holds since

$$\gamma_{4,3,2,1} + \gamma_{4,3,2,2} \lesssim \delta\frac{\sigma_r^{\star 2}}{ndp^2}\left(\omega_j^{\star 2}\left\|\boldsymbol{U}_{i,\cdot}^{\star}\right\|_2^2 + \omega_i^{\star 2}\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2^2\right) \lesssim \delta v_{i,j}^\star,$$
$$\gamma_{4,3,2,3} \lesssim \frac{\mu r^2}{ndp^2}\omega_i^\star\omega_j^\star\cdot\theta^2\omega_i^\star\omega_j^\star \overset{\text{(i)}}{\lesssim} \frac{\mu r^2}{ndp^2}\omega_i^\star\omega_j^\star\cdot\varepsilon(v_{i,j}^\star)^{1/2},$$
$$\lesssim \delta\frac{\omega_{\min}}{\sqrt{np^2}}\left(\omega_j^\star\left\|\boldsymbol{U}_{i,\cdot}^{\star}\right\|_2 + \omega_i^\star\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2\right)\cdot\varepsilon(v_{i,j}^\star)^{1/2} \lesssim \delta v_{i,j}^\star,$$

provided that $\zeta_{\mathsf{1st}}/\sigma_r^{\star 2} \lesssim \delta/\sqrt{\kappa^2\mu^2 r^4\log(n+d)}$ and

$$\left\|\boldsymbol{U}_{i,\cdot}^{\star}\right\|_2 + \left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2 \gtrsim \delta^{-1}\sqrt{\kappa}\mu r^2\frac{\omega_{\max}}{\sigma_r^\star}\theta + \delta^{-1}\kappa_\omega^{1/2}\frac{\mu r^2}{\sqrt{ndp^2}}\cdot\sqrt{\frac{1}{d}}.$$

Here (i) follows from (E.115).

- Regarding $\gamma_{4,3,3}$, we have

$$\gamma_{4,3,3} \lesssim \underbrace{\frac{\mu r^2\sqrt{\log(n+d)}}{ndp^2}\sigma_1^{\star 2}\left\|\boldsymbol{U}_{i,\cdot}^{\star}\boldsymbol{\Sigma}^\star\right\|_2^2\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2^2\frac{\zeta_{\mathsf{1st}}}{\sigma_r^{\star 2}}}_{=:\gamma_{4,3,3,1}} + \underbrace{\frac{\mu r^2}{ndp^2}\sigma_1^\star\omega_j^\star\theta\left\|\boldsymbol{U}_{i,\cdot}^{\star}\boldsymbol{\Sigma}^\star\right\|_2^2\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2}_{=:\gamma_{4,3,3,2}}$$
$$+ \underbrace{\frac{\mu r^2}{ndp^2}\omega_j^{\star 2}\theta^2\left\|\boldsymbol{U}_{i,\cdot}^{\star}\boldsymbol{\Sigma}^\star\right\|_2^2}_{=:\gamma_{4,3,3,3}} \lesssim \delta v_{i,j}^\star,$$

where the last relation holds since

$$\gamma_{4,3,3,1} \lesssim \frac{\delta}{ndp^2\kappa}\left\|\boldsymbol{U}_{i,\cdot}^{\star}\boldsymbol{\Sigma}^\star\right\|_2^2\left\|\boldsymbol{U}_{j,\cdot}^{\star}\boldsymbol{\Sigma}^\star\right\|_2^2 \lesssim \delta v_{i,j}^\star,$$
$$\gamma_{4,3,3,2} \overset{\text{(i)}}{\lesssim} \frac{\delta}{ndp^2\kappa}\left\|\boldsymbol{U}_{i,\cdot}^{\star}\boldsymbol{\Sigma}^\star\right\|_2^2\left\|\boldsymbol{U}_{j,\cdot}^{\star}\boldsymbol{\Sigma}^\star\right\|_2^2 + \delta\frac{\sigma_r^{\star 2}}{ndp^2}\omega_j^{\star 2}\left\|\boldsymbol{U}_{i,\cdot}^{\star}\right\|_2^2 \lesssim \delta v_{i,j}^\star,$$
$$\gamma_{4,3,3,3} \lesssim \delta\frac{\sigma_r^{\star 2}}{ndp^2}\omega_j^{\star 2}\left\|\boldsymbol{U}_{i,\cdot}^{\star}\right\|_2^2 \lesssim \delta v_{i,j}^\star,$$

provided that $\zeta_{\mathsf{1st}}/\sigma_r^{\star 2} \lesssim \delta/(\kappa^2\mu r^2\sqrt{\log(n+d)})$ and $\theta \lesssim \delta/(\kappa^{3/2}\mu r^2)$. Here (i) invokes the AM-GM inequality.

- Regarding $\gamma_{4,3,4}$, we have

$$\gamma_{4,3,4} \lesssim \underbrace{\frac{1}{np^2}\left\|\boldsymbol{U}_{i,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2^2\omega_{\max}^2\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2^2\frac{\zeta_{\mathsf{1st}}}{\sigma_r^{\star 2}}\sqrt{r^2\log(n+d)}}_{=:\gamma_{4,3,4,1}} + \underbrace{\frac{1}{np^2}\left\|\boldsymbol{U}_{i,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2^2\omega_{\max}^2\frac{\theta r}{\sigma_1^{\star}}\omega_j^{\star}\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2}_{=:\gamma_{4,3,4,2}}$$
$$+ \underbrace{\frac{1}{np^2}\left\|\boldsymbol{U}_{i,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2^2\omega_{\max}^2\frac{\theta^2 r}{\kappa\sigma_r^{\star 2}}\omega_j^{\star 2}}_{=:\gamma_{4,3,4,3}} \lesssim \delta v_{i,j}^{\star},$$

where the last relation holds since

$$\gamma_{4,3,2,1}+\gamma_{4,3,2,2} \lesssim \delta\frac{\sigma_r^{\star 2}}{ndp^2}\left(\omega_j^{\star 2}\left\|\boldsymbol{U}_{i,\cdot}^{\star}\right\|_2^2+\omega_i^{\star 2}\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2^2\right) \lesssim \delta v_{i,j}^{\star},$$
$$\gamma_{4,3,2,3} \lesssim \delta\frac{\omega_{\min}^2}{np^2}\left(\omega_j^{\star 2}\left\|\boldsymbol{U}_{i,\cdot}^{\star}\right\|_2^2+\omega_i^{\star 2}\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2^2\right) \lesssim \delta v_{i,j}^{\star},$$

provided that $\zeta_{\mathsf{1st}}/\sigma_r^{\star 2} \lesssim \delta/\sqrt{\kappa^2\mu^2 r^4\kappa_\omega^2\log(n+d)}$, $\theta \lesssim \sqrt{\delta/(\kappa_\omega r)}$, and

$$\theta\frac{\omega_{\max}}{\sigma_r^{\star}}\sqrt{d} \lesssim \frac{\delta}{\sqrt{\kappa_\omega\kappa\mu r^3}}.$$

Here, (i) utilizes (E.113) and (ii) follows from (E.115).

Taking the bounds on $\gamma_{4,3,1}$, $\gamma_{4,3,2}$ and $\gamma_{4,3,3}$ collectively yields

$$\gamma_{4,3} \lesssim \gamma_{4,3,1}+\gamma_{4,3,2}+\gamma_{4,3,3} \lesssim \delta v_{i,j}^{\star},$$

provided that $\zeta_{\mathsf{1st}}/\sigma_r^{\star 2} \lesssim \delta/(\kappa^2\mu r^2\kappa_\omega\sqrt{\log(n+d)})$, $\theta \lesssim \delta/(\kappa^{3/2}\mu r^2\kappa_\omega^{1/2})$, $\varepsilon \lesssim 1$,

$$\theta\frac{\omega_{\max}}{\sigma_r^{\star}}\sqrt{d} \lesssim \frac{\delta}{\sqrt{\kappa_\omega\kappa\mu r^3}}$$

and

$$\left\|\boldsymbol{U}_{i,\cdot}^{\star}\right\|_2+\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2 \gtrsim \delta^{-1}\sqrt{\kappa}\mu r^2\kappa_\omega\frac{\omega_{\max}}{\sigma_r^{\star}}\theta+\delta^{-1}r\kappa_\omega^{1/2}\frac{\omega_{\max}^2}{\sigma_1^{\star 2}}\sqrt{\frac{1}{np^2}}+\delta^{-1}\frac{\mu r^2\kappa_\omega^{1/2}}{\sqrt{ndp^2}}\cdot\sqrt{\frac{1}{d}}.$$

In view of (E.118), we know that the last two conditions above can be guaranteed by $ndp^2 \gtrsim \delta^{-2}\kappa^4\mu^3 r^6\kappa_\omega\log^4(n+d)$,

$$\frac{\omega_{\max}}{\sigma_r^{\star}}\sqrt{\frac{d}{np}} \lesssim \frac{\delta}{\kappa\mu^{1/2}r^2\kappa_\omega^{1/2}\log(n+d)},\qquad \frac{\omega_{\max}^2}{p\sigma_r^{\star 2}}\sqrt{\frac{d}{n}} \lesssim \frac{\delta}{\kappa\mu^{1/2}r^2\kappa_\omega^{1/2}\log(n+d)},$$

and

$$\left\|\boldsymbol{U}_{i,\cdot}^{\star}\right\|_2+\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2 \gtrsim \delta^{-1}\kappa\mu r^2\kappa_\omega\log(n+d)\left[\frac{\omega_{\max}}{\sigma_r^{\star}}\sqrt{\frac{d}{np}}+\frac{\omega_{\max}^2}{p\sigma_r^{\star 2}}\sqrt{\frac{d}{n}}+\frac{\kappa\mu r\log(n+d)}{\sqrt{ndp^2}}\right]\cdot\sqrt{\frac{r}{d}}.$$

Step 4.4: putting the bounds on $\gamma_{4,1}$, $\gamma_{4,2}$ and $\gamma_{4,3}$ together. In view of (E.66), we can take the bounds on $\gamma_{4,1}$, $\gamma_{4,2}$ and $\gamma_{4,3}$ together to reach

$$|\alpha_4-\beta_4| \lesssim \gamma_{4,1}+\gamma_{4,2}+\gamma_{4,3} \lesssim \delta v_{i,j}^{\star},$$

provided that $\varepsilon \lesssim 1$, $n \gtrsim \delta^{-2}\kappa^7\mu^2 r^3\kappa_\omega^2\log(n+d)$, $ndp^2 \gtrsim \delta^{-2}\kappa^4\mu^3 r^6\kappa_\omega^2\log^4(n+d)$, $np \gtrsim \delta^{-2}\kappa^2\mu^2 r^2\kappa_\omega^2\log^2(n+d)$,

$$\frac{\zeta_{\mathsf{1st}}}{\sigma_r^{\star 2}} \lesssim \frac{\delta}{\kappa^2\mu r^2\kappa_\omega\sqrt{\log(n+d)}},\qquad \zeta_{\mathsf{2nd},i}\sqrt{d} \lesssim \frac{\delta}{\sqrt{\kappa^2\mu r\kappa_\omega^2}},\tag{E.125}$$

$$\frac{\omega_{\max}}{\sigma_r^\star}\sqrt{\frac{d}{np}} \lesssim \frac{\delta}{\sqrt{\kappa^2\mu r^2\kappa_\omega \log^2(n+d)}}, \qquad \frac{\omega_{\max}^2}{p\sigma_r^{\star 2}}\sqrt{\frac{d}{n}} \lesssim \frac{\delta}{\sqrt{\kappa^2\mu r^2\kappa_\omega \log^2(n+d)}},$$

and

$$\left\|\boldsymbol{U}_{i,\cdot}^\star\right\|_2 + \left\|\boldsymbol{U}_{j,\cdot}^\star\right\|_2 \gtrsim \delta^{-1}\kappa\mu r^2\kappa_\omega \log(n+d)\left[\frac{\kappa\mu r\log(n+d)}{\sqrt{ndp}} + \frac{\omega_{\max}^2}{p\sigma_r^{\star 2}}\sqrt{\frac{d}{n}} + \frac{\omega_{\max}}{\sigma_r^\star}\sqrt{\frac{d}{np}}\right]\sqrt{\frac{r}{d}}.$$

Let us now take a closer look at (E.125). In view of (E.65) and the definition of $\zeta_{\mathsf{2nd},i}$, we know that (E.125) is equivalent to

$$ndp^2 \gtrsim \delta^{-2}\kappa^6\mu^4 r^6\kappa_\omega^2 \log^5(n+d), \qquad np \gtrsim \delta^{-2}\kappa^6\mu^3 r^5\kappa_\omega^2\log^3(n+d),$$
$$\frac{\omega_{\max}^2}{p\sigma_r^{\star 2}}\sqrt{\frac{d}{n}} \lesssim \frac{\delta}{\kappa^2\mu r^2\kappa_\omega \log^{3/2}(n+d)}, \qquad \frac{\omega_{\max}}{\sigma_r^\star}\sqrt{\frac{d}{np}} \lesssim \frac{\delta}{\kappa^{5/2}\mu r^2\kappa_\omega\log(n+d)},$$

provided that $d \gtrsim \kappa^2\mu\log(n+d)$. This concludes our bound on $|\alpha_4 - \beta_4|$ and the required conditions.

Similarly, we can also prove that (which we omit here for the sake of brevity)

$$|\alpha_6 - \beta_6| \lesssim \delta v_{i,j}^\star$$

under the above conditions.

Step 5: bounding $|\alpha_5 - \beta_5|$ and $|\alpha_7 - \beta_7|$. Regarding $|\alpha_5 - \beta_5|$, we first make the observation that

$$\begin{aligned}
|\alpha_5 - \beta_5| &\lesssim \frac{1}{np^2}\left|\sum_{k=1}^d S_{i,k}^{\star 2}\left(\boldsymbol{U}_{k,\cdot}^\star\boldsymbol{U}_{j,\cdot}^{\star\top}\right)^2 - \sum_{k=1}^d S_{i,k}^2\left(\boldsymbol{U}_{k,\cdot}\boldsymbol{U}_{j,\cdot}^\top\right)^2\right| \\
&\lesssim \frac{1}{np^2}\left|\sum_{k=1}^d \left(S_{i,k}^2 - S_{i,k}^{\star 2}\right)\left(\boldsymbol{U}_{k,\cdot}\boldsymbol{U}_{j,\cdot}^\top\right)^2\right| + \frac{1}{np^2}\left|\sum_{k=1}^d S_{i,k}^{\star 2}\left[\left(\boldsymbol{U}_{k,\cdot}^\star\boldsymbol{U}_{j,\cdot}^{\star\top}\right)^2 - \left(\boldsymbol{U}_{k,\cdot}\boldsymbol{U}_{j,\cdot}^\top\right)^2\right]\right| \\
&\lesssim \frac{1}{np^2}\left\|\boldsymbol{U}_{j,\cdot}\right\|_2^2\max_{k\in[d]}\left|S_{i,k}^2 - S_{i,k}^{\star 2}\right| + \frac{1}{np^2}\left\|\boldsymbol{U}_{i,\cdot}\boldsymbol{\Sigma}^\star\right\|_2^2\left\|\boldsymbol{U}^\star\boldsymbol{\Sigma}^\star\right\|_{2,\infty}^2\sum_{k=1}^d\left|\left(\boldsymbol{U}_{k,\cdot}^\star\boldsymbol{U}_{j,\cdot}^{\star\top}\right)^2 - \left(\boldsymbol{U}_{k,\cdot}\boldsymbol{U}_{j,\cdot}^\top\right)^2\right| \\
&\overset{\text{(i)}}{\lesssim} \frac{1}{np^2}\left(\left\|\boldsymbol{U}_{j,\cdot}^\star\right\|_2^2 + \frac{\theta^2}{\kappa\sigma_r^{\star 2}}\omega_j^{\star 2}\right)\max_{k\in[d]}\left|S_{i,k}^2 - S_{i,k}^{\star 2}\right| + \frac{\mu r}{ndp^2}\sigma_1^{\star 2}\left\|\boldsymbol{U}_{i,\cdot}\boldsymbol{\Sigma}^\star\right\|_2^2\sum_{k=1}^d\left|\left(\boldsymbol{U}_{k,\cdot}^\star\boldsymbol{U}_{j,\cdot}^{\star\top}\right)^2 - \left(\boldsymbol{U}_{k,\cdot}\boldsymbol{U}_{j,\cdot}^\top\right)^2\right| \\
&\overset{\text{(ii)}}{\lesssim} \frac{1}{np^2}\left\|\boldsymbol{U}_{j,\cdot}^\star\right\|_2^2\max_{k\in[d]}\left|S_{i,k}^2 - S_{i,k}^{\star 2}\right| + \frac{1}{np^2}\frac{\theta^2}{\kappa\sigma_r^{\star 2}}\omega_j^{\star 2}\max_{k\in[d]}\left|S_{i,k}^2 - S_{i,k}^{\star 2}\right| + \delta v_{i,j}^\star \\
&\overset{\text{(iii)}}{\lesssim} \underbrace{\frac{1}{np^2}\sqrt{\frac{\mu r}{d}}\sigma_1^\star\left\|\boldsymbol{U}_{i,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2\left\|\boldsymbol{U}_{j,\cdot}^\star\right\|_2^2\max_{k\in[d]}\left|S_{i,k} - S_{i,k}^\star\right|}_{=:\gamma_{5,1}} + \underbrace{\frac{1}{np^2}\sqrt{\frac{\mu r}{d}}\left\|\boldsymbol{U}_{i,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2\frac{\theta^2}{\sqrt{\kappa}\sigma_r^\star}\omega_j^{\star 2}\max_{k\in[d]}\left|S_{i,k} - S_{i,k}^\star\right|}_{=:\gamma_{5,2}} \\
&\quad + \underbrace{\frac{1}{np^2}\left\|\boldsymbol{U}_{j,\cdot}^\star\right\|_2^2\max_{k\in[d]}\left|S_{i,k} - S_{i,k}^\star\right|^2}_{=:\gamma_{5,3}} + \underbrace{\frac{1}{np^2}\frac{\theta^2}{\kappa\sigma_r^{\star 2}}\omega_j^{\star 2}\max_{k\in[d]}\left|S_{i,k} - S_{i,k}^\star\right|^2}_{=:\gamma_{5,4}} + \delta v_{i,j}^\star.
\end{aligned}$$

Here, (i) follows from (E.122); (ii) holds since

$$\frac{\mu r}{ndp^2}\sigma_1^{\star 2}\left\|\boldsymbol{U}_{i,\cdot}\boldsymbol{\Sigma}^\star\right\|_2^2\sum_{k=1}^d\left|\left(\boldsymbol{U}_{k,\cdot}^\star\boldsymbol{U}_{j,\cdot}^{\star\top}\right)^2 - \left(\boldsymbol{U}_{k,\cdot}\boldsymbol{U}_{j,\cdot}^\top\right)^2\right| = \gamma_{4,3,3} \lesssim \delta v_{i,j}^\star;$$

and (iii) follows from the fact that

$$\max_{k\in[d]}\left|S_{i,k}^2 - S_{i,k}^{\star 2}\right| \lesssim \max_{k\in[d]}\left|S_{i,k} - S_{i,k}^\star\right|S_{i,k}^\star + \max_{k\in[d]}\left|S_{i,k} - S_{i,k}^\star\right|^2$$

$$\lesssim \left\| \boldsymbol{U}_{i,\cdot}^{\star} \boldsymbol{\Sigma}^{\star} \right\|_2 \left\| \boldsymbol{U}^{\star} \boldsymbol{\Sigma}^{\star} \right\|_{2,\infty} \max_{k \in [d]} \left| S_{i,k} - S_{i,k}^{\star} \right| + \max_{k \in [d]} \left| S_{i,k} - S_{i,k}^{\star} \right|^2$$
$$\lesssim \sqrt{\frac{\mu r}{d}} \sigma_1^{\star} \left\| \boldsymbol{U}_{i,\cdot}^{\star} \boldsymbol{\Sigma}^{\star} \right\|_2 \max_{k \in [d]} \left| S_{i,k} - S_{i,k}^{\star} \right| + \max_{k \in [d]} \left| S_{i,k} - S_{i,k}^{\star} \right|^2.$$

In view of Lemma 19, we know that for each $k \in [d]$,

$$\begin{aligned}
\left| S_{i,k} - S_{i,k}^{\star} \right| &\lesssim \left(\theta + \sqrt{\frac{\kappa^3 r \log(n+d)}{n}} \right) \left\| \boldsymbol{U}_{i,\cdot}^{\star} \boldsymbol{\Sigma}^{\star} \right\|_2 \sqrt{\frac{\mu r}{d}} \sigma_1^{\star} + \theta^2 \omega_i^{\star} \omega_{\max} + \theta \omega_i^{\star} \sqrt{\frac{\mu r}{d}} \sigma_1^{\star} \\
&\quad + \theta \omega_{\max} \left\| \boldsymbol{U}_{i,\cdot}^{\star} \boldsymbol{\Sigma}^{\star} \right\|_2 + \zeta_{\mathsf{2nd},i} \sqrt{\frac{\mu r}{d}} \sigma_1^{\star 2} + \zeta_{\mathsf{2nd},k} \left\| \boldsymbol{U}_{i,\cdot}^{\star} \boldsymbol{\Sigma}^{\star} \right\|_2 \sigma_1^{\star} + \zeta_{\mathsf{2nd},i} \zeta_{\mathsf{2nd},k} \sigma_1^{\star 2} \\
&\lesssim \left(\theta + \sqrt{\frac{\kappa^3 r \log(n+d)}{n}} \right) \left\| \boldsymbol{U}_{i,\cdot}^{\star} \boldsymbol{\Sigma}^{\star} \right\|_2 \sqrt{\frac{\mu r}{d}} \sigma_1^{\star} + \theta \omega_i^{\star} \sqrt{\frac{\mu r}{d}} \sigma_1^{\star} \\
&\quad + \zeta_{\mathsf{2nd},i} \sqrt{\frac{\mu r}{d}} \sigma_1^{\star 2} + \theta \omega_{\max} \left\| \boldsymbol{U}_{i,\cdot}^{\star} \boldsymbol{\Sigma}^{\star} \right\|_2,
\end{aligned}$$

where the last relation holds due to (E.123) and (E.114), provided that $\varepsilon \lesssim 1$, $np \gtrsim 1$ and $ndp^2 \gtrsim 1$. With these preparations in place, we now proceed to bound the terms $\gamma_{5,1}$, $\gamma_{5,2}$, $\gamma_{5,3}$ and $\gamma_{5,4}$ separately.

- Regarding $\gamma_{5,1}$, we have

$$\begin{aligned}
\gamma_{5,1} &\lesssim \underbrace{\frac{\mu r}{ndp^2} \sigma_1^{\star 2} \left\| \boldsymbol{U}_{i,\cdot}^{\star} \boldsymbol{\Sigma}^{\star} \right\|_2^2 \left\| \boldsymbol{U}_{j,\cdot}^{\star} \right\|_2^2 \left(\theta + \sqrt{\frac{\kappa^3 r \log(n+d)}{n}} \right)}_{=:\gamma_{5,1,1}} \\
&\quad + \underbrace{\frac{\mu r}{ndp^2} \sigma_1^{\star 2} \left\| \boldsymbol{U}_{i,\cdot}^{\star} \boldsymbol{\Sigma}^{\star} \right\|_2 \left\| \boldsymbol{U}_{j,\cdot}^{\star} \right\|_2^2 \theta \omega_i^{\star}}_{=:\gamma_{5,1,2}} + \underbrace{\frac{\mu r}{ndp^2} \sigma_1^{\star 2} \left\| \boldsymbol{U}_{i,\cdot}^{\star} \boldsymbol{\Sigma}^{\star} \right\|_2 \left\| \boldsymbol{U}_{j,\cdot}^{\star} \right\|_2^2 \zeta_{\mathsf{2nd},i} \sigma_1^{\star}}_{=:\gamma_{5,1,3}} \\
&\quad + \underbrace{\frac{1}{np^2} \sqrt{\frac{\mu r}{d}} \sigma_1^{\star} \left\| \boldsymbol{U}_{i,\cdot}^{\star} \boldsymbol{\Sigma}^{\star} \right\|_2^2 \left\| \boldsymbol{U}_{j,\cdot}^{\star} \right\|_2^2 \theta \omega_{\max}}_{=:\gamma_{5,1,4}} \lesssim \delta v_{i,j}^{\star},
\end{aligned}$$

  where the last relation holds since

$$\begin{aligned}
\gamma_{5,1,1} &\lesssim \frac{\delta}{ndp^2 \kappa} \left\| \boldsymbol{U}_{i,\cdot}^{\star} \boldsymbol{\Sigma}^{\star} \right\|_2^2 \left\| \boldsymbol{U}_{j,\cdot}^{\star} \boldsymbol{\Sigma}^{\star} \right\|_2^2 \lesssim \delta v_{i,j}^{\star}, \\
\gamma_{5,1,2} &\overset{\text{(i)}}{\lesssim} \frac{\delta}{ndp^2 \kappa} \left\| \boldsymbol{U}_{i,\cdot}^{\star} \boldsymbol{\Sigma}^{\star} \right\|_2^2 \left\| \boldsymbol{U}_{j,\cdot}^{\star} \boldsymbol{\Sigma}^{\star} \right\|_2^2 + \delta \frac{\sigma_r^{\star 2}}{ndp^2} \omega_i^{\star 2} \left\| \boldsymbol{U}_{j,\cdot}^{\star} \right\|_2^2 \lesssim \delta v_{i,j}^{\star}, \\
\gamma_{5,1,3} &\lesssim \frac{\kappa \mu r}{\sqrt{ndp^2}} \cdot \sqrt{\frac{1}{ndp^2 \kappa}} \left\| \boldsymbol{U}_{i,\cdot}^{\star} \boldsymbol{\Sigma}^{\star} \right\|_2 \left\| \boldsymbol{U}_{j,\cdot}^{\star} \boldsymbol{\Sigma}^{\star} \right\|_2 \cdot \sigma_1^{\star 2} \left(\left\| \boldsymbol{U}_{i,\cdot}^{\star} \right\|_2 \zeta_{\mathsf{2nd},j} + \left\| \boldsymbol{U}_{j,\cdot}^{\star} \right\|_2 \zeta_{\mathsf{2nd},i} \right) \\
&\overset{\text{(ii)}}{\lesssim} \frac{\kappa \mu r}{\sqrt{ndp^2}} \cdot \sqrt{\frac{1}{ndp^2 \kappa}} \left\| \boldsymbol{U}_{i,\cdot}^{\star} \boldsymbol{\Sigma}^{\star} \right\|_2 \left\| \boldsymbol{U}_{j,\cdot}^{\star} \boldsymbol{\Sigma}^{\star} \right\|_2 \cdot \varepsilon (v_{i,j}^{\star})^{1/2} \lesssim \frac{\kappa \mu r}{\sqrt{ndp^2}} \varepsilon v_{i,j}^{\star} \lesssim \delta v_{i,j}^{\star}, \\
\gamma_{5,1,4} &\lesssim \delta \frac{1}{ndp^2 \kappa} \left\| \boldsymbol{U}_{i,\cdot}^{\star} \boldsymbol{\Sigma}^{\star} \right\|_2^2 \left\| \boldsymbol{U}_{j,\cdot}^{\star} \boldsymbol{\Sigma}^{\star} \right\|_2^2 \lesssim \delta v_{i,j}^{\star},
\end{aligned}$$

  provided that $\theta \lesssim \delta/(\kappa^2 \mu r)$, $n \gtrsim \delta^{-2} \kappa^7 \mu^2 r^3 \log(n+d)$, $\varepsilon \lesssim 1$, $ndp^2 \gtrsim \delta^{-2} \kappa^2 \mu^2 r^2$, and

$$\theta \frac{\omega_{\max}}{\sigma_r^{\star}} \sqrt{d} \lesssim \frac{\delta}{\sqrt{\kappa^3 \mu r}}.$$

  In the above relations, (i) uses the AM-GM inequality, where (ii) utilizes (E.113).

- Regarding $\gamma_{5,2}$, we can derive

$$\gamma_{5,2} \lesssim \underbrace{\frac{\mu r}{ndp^2}\sigma_1^\star \left\|\boldsymbol{U}_{i,\cdot}^\star \boldsymbol{\Sigma}^\star\right\|_2^2 \frac{\theta^2}{\sqrt{\kappa}\sigma_r^\star}\omega_j^{\star 2}\left(\theta+\sqrt{\frac{\kappa^3 r\log\left(n+d\right)}{n}}\right)}_{=:\gamma_{5,2,1}}$$
$$+\underbrace{\frac{\mu r}{ndp^2}\left\|\boldsymbol{U}_{i,\cdot}^\star \boldsymbol{\Sigma}^\star\right\|_2 \theta^3\omega_j^{\star 2}\omega_i^\star}_{=:\gamma_{5,2,2}} + \underbrace{\frac{\mu r}{ndp^2}\sigma_1^\star \left\|\boldsymbol{U}_{i,\cdot}^\star \boldsymbol{\Sigma}^\star\right\|_2 \theta^2\omega_j^{\star 2}\zeta_{\mathsf{2nd},i}}_{=:\gamma_{5,2,3}}$$
$$+\underbrace{\frac{1}{np^2}\sqrt{\frac{\mu r}{d}}\left\|\boldsymbol{U}_{i,\cdot}^\star \boldsymbol{\Sigma}^\star\right\|_2^2 \frac{\theta^3}{\sqrt{\kappa}\sigma_r^\star}\omega_j^{\star 2}\omega_{\max}}_{=:\gamma_{5,2,4}} \lesssim \delta v_{i,j}^\star,$$

$$\frac{\mu r}{ndp^2}\sigma_1^\star \left\|\boldsymbol{U}_{i,\cdot}^\star \boldsymbol{\Sigma}^\star\right\|_2 \theta^2\omega_j^{\star 2}\zeta_{\mathsf{2nd},i} \lesssim \theta\frac{\mu r}{ndp^2}\left\|\boldsymbol{U}_{i,\cdot}^\star \boldsymbol{\Sigma}^\star\right\|_2 \omega_j^\star \cdot \theta\sigma_1^\star\left(\omega_i^\star\zeta_{\mathsf{2nd},j}+\omega_j^\star\zeta_{\mathsf{2nd},i}\right)$$

where the last relation holds since

$$\gamma_{5,2,1} \lesssim \delta\frac{\sigma_r^{\star 2}}{ndp^2}\left(\omega_j^{\star 2}\left\|\boldsymbol{U}_{i,\cdot}^\star\right\|_2^2+\omega_i^{\star 2}\left\|\boldsymbol{U}_{j,\cdot}^\star\right\|_2^2\right) \lesssim \delta v_{i,j}^\star,$$
$$\gamma_{5,2,2} \lesssim \frac{\mu r}{ndp^2}\left\|\boldsymbol{U}_{i,\cdot}^\star \boldsymbol{\Sigma}^\star\right\|_2 \theta^3\omega_j^{\star 2}\omega_i^\star \lesssim \theta^2\gamma_{4,3,2,2} \lesssim \delta v_{i,j}^\star,$$
$$\gamma_{5,2,3} \lesssim \theta\frac{\sqrt{\kappa}\mu r}{\sqrt{ndp^2}}\cdot\frac{\sigma_r^\star}{\sqrt{ndp^2}}\omega_j^\star\left\|\boldsymbol{U}_{i,\cdot}^\star\right\|_2\cdot\theta\sigma_1^\star\omega_j^\star\zeta_{\mathsf{2nd},i} \overset{\text{(i)}}{\lesssim} \theta\frac{\sqrt{\kappa}\mu r}{\sqrt{ndp^2}}\varepsilon v_{i,j}^\star \lesssim \delta v_{i,j}^\star,$$
$$\gamma_{5,2,4} \lesssim \delta\frac{\sigma_r^{\star 2}}{ndp^2}\left(\omega_j^{\star 2}\left\|\boldsymbol{U}_{i,\cdot}^\star\right\|_2^2+\omega_i^{\star 2}\left\|\boldsymbol{U}_{j,\cdot}^\star\right\|_2^2\right) \lesssim \delta v_{i,j}^\star,$$

provided that $\theta \lesssim \delta/(\kappa^{1/2}\mu r)$, $n \gtrsim \kappa^3 r\log(n+d)$, $\varepsilon \lesssim 1$, $ndp^2 \gtrsim 1$ and

$$\theta^3\frac{\omega_{\max}}{\sigma_r^\star}\sqrt{d} \lesssim \frac{\delta}{\sqrt{\kappa\mu r}}.$$

Here, the relation (i) in the above inequality arises from (E.115).

- With regards to $\gamma_{5,3}$, we make the observation that

$$\gamma_{5,3} \lesssim \underbrace{\frac{\mu r}{ndp^2}\sigma_1^{\star 2}\left\|\boldsymbol{U}_{i,\cdot}^\star \boldsymbol{\Sigma}^\star\right\|_2^2\left\|\boldsymbol{U}_{j,\cdot}^\star\right\|_2^2\left(\theta^2+\frac{\kappa^3 r\log\left(n+d\right)}{n}\right)}_{=:\gamma_{5,3,1}} + \underbrace{\frac{\mu r}{ndp^2}\sigma_1^{\star 2}\left\|\boldsymbol{U}_{j,\cdot}^\star\right\|_2^2\theta^2\omega_i^{\star 2}}_{=:\gamma_{5,3,2}}$$
$$+\underbrace{\frac{\mu r}{ndp^2}\left\|\boldsymbol{U}_{j,\cdot}^\star\right\|_2^2\zeta_{\mathsf{2nd},i}^2\sigma_1^{\star 4}}_{=:\gamma_{5,3,3}} + \underbrace{\frac{1}{np^2}\left\|\boldsymbol{U}_{j,\cdot}^\star\right\|_2^2\theta^2\omega_{\max}^2\left\|\boldsymbol{U}_{i,\cdot}^\star \boldsymbol{\Sigma}^\star\right\|_2^2}_{=:\gamma_{5,3,4}} \lesssim \delta v_{i,j}^\star,$$

where the last relation holds since

$$\gamma_{5,3,1} \lesssim \left(\theta+\sqrt{\frac{\kappa^3 r\log\left(n+d\right)}{n}}\right)\gamma_{5,1,1} \lesssim \delta v_{i,j}^\star,$$
$$\gamma_{5,3,2} \lesssim \delta\frac{\sigma_r^{\star 2}}{ndp^2}\left(\omega_j^{\star 2}\left\|\boldsymbol{U}_{i,\cdot}^\star\right\|_2^2+\omega_i^{\star 2}\left\|\boldsymbol{U}_{j,\cdot}^\star\right\|_2^2\right) \lesssim \delta v_{i,j}^\star,$$
$$\gamma_{5,3,3} \lesssim \frac{\mu r}{ndp^2}\left[\sigma_1^{\star 2}\left(\left\|\boldsymbol{U}_{i,\cdot}^\star\right\|_2\zeta_{\mathsf{2nd},j}+\left\|\boldsymbol{U}_{j,\cdot}^\star\right\|_2\zeta_{\mathsf{2nd},i}\right)\right]^2 \overset{\text{(i)}}{\lesssim} \frac{\mu r}{ndp^2}\varepsilon^2 v_{i,j}^\star \lesssim \delta v_{i,j}^\star,$$
$$\gamma_{5,3,4} \lesssim \delta\frac{1}{ndp^2\kappa}\left\|\boldsymbol{U}_{i,\cdot}^\star \boldsymbol{\Sigma}^\star\right\|_2^2\left\|\boldsymbol{U}_{j,\cdot}^\star \boldsymbol{\Sigma}^\star\right\|_2^2 \lesssim \delta v_{i,j}^\star,$$

provided that $\theta \lesssim \sqrt{\delta/(\kappa\mu r)}$, $n \gtrsim \kappa^3 r \log(n+d)$, $\varepsilon \lesssim 1$, $ndp^2 \gtrsim \delta^{-1}\mu r$ and

$$\theta \frac{\omega_{\max}}{\sigma_r^\star}\sqrt{d} \lesssim \frac{\delta}{\sqrt{\kappa}}.$$

Note that the inequality (i) in the above relation results from (E.113).

- When it comes to $\gamma_{5,4}$, we have the following upper bound

$$\begin{aligned}
\gamma_{5,4} &\lesssim \underbrace{\frac{\mu r}{ndp^2}\left\|\boldsymbol{U}_{i,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2^2\theta^2\omega_j^{\star 2}\left(\theta^2+\frac{\kappa^3 r\log(n+d)}{n}\right)}_{=:\gamma_{5,4,1}} + \underbrace{\frac{\mu r}{ndp^2}\theta^4\omega_j^{\star 2}\omega_{\max}^2}_{=:\gamma_{5,4,2}} \\
&\quad + \underbrace{\frac{\mu r}{ndp^2}\frac{\theta^2}{\kappa\sigma_r^{\star 2}}\omega_j^{\star 2}\zeta_{\mathsf{2nd},i}^2\sigma_1^{\star 4}}_{=:\gamma_{5,4,3}} + \underbrace{\frac{1}{np^2}\frac{\theta^4}{\kappa\sigma_r^{\star 2}}\omega_j^{\star 2}\omega_{\max}^2\left\|\boldsymbol{U}_{i,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2^2}_{=:\gamma_{5,4,4}} \lesssim \delta v_{i,j}^\star,
\end{aligned}$$

where the last inequality holds true since

$$\begin{aligned}
\gamma_{5,4,1} &\lesssim \delta\frac{\sigma_r^{\star 2}}{ndp^2}\left(\omega_j^{\star 2}\left\|\boldsymbol{U}_{i,\cdot}^\star\right\|_2^2+\omega_i^{\star 2}\left\|\boldsymbol{U}_{j,\cdot}^\star\right\|_2^2\right) \lesssim \delta v_{i,j}^\star, \\
\gamma_{5,4,2} &\overset{\text{(i)}}{\lesssim} \frac{\mu r\kappa_\omega}{ndp^2}\varepsilon^2 v_{i,j}^\star \lesssim \delta v_{i,j}^\star, \\
\gamma_{5,4,3} &\overset{\text{(ii)}}{\lesssim} \frac{\mu r\kappa_\omega^{1/2}}{ndp^2\kappa}\varepsilon^2 v_{i,j}^\star \lesssim \delta v_{i,j}^\star, \\
\gamma_{5,4,4} &\lesssim \delta\frac{\omega_{\min}^2}{np^2}\left(\omega_j^{\star 2}\left\|\boldsymbol{U}_{i,\cdot}^\star\right\|_2^2+\omega_i^{\star 2}\left\|\boldsymbol{U}_{j,\cdot}^\star\right\|_2^2\right)
\end{aligned}$$

provided that $\theta \lesssim \delta/(\kappa\mu r\kappa_\omega)$, $n \gtrsim \kappa^3 r\log(n+d)$ and $ndp^2 \gtrsim \delta^{-1}\mu r\kappa_\omega$. Here, the inequalities (i) and (ii) follow from (E.115).

Combining the above bounds on $\gamma_{5,1}$, $\gamma_{5,2}$, $\gamma_{5,3}$ and $\gamma_{5,4}$ yields

$$|\alpha_5-\beta_5| \lesssim \gamma_{5,1}+\gamma_{5,2}+\gamma_{5,3}+\gamma_{5,4} \lesssim \delta v_{i,j}^\star,$$

provided that $\theta \lesssim \delta/(\kappa^2\mu r)$, $n \gtrsim \delta^{-2}\kappa^7\mu^2 r^3\log(n+d)$, $\varepsilon \lesssim 1$, $ndp^2 \gtrsim \delta^{-2}\kappa^2\mu^2 r^2$ and

$$\theta\frac{\omega_{\max}}{\sigma_r^\star}\sqrt{d} \lesssim \frac{\delta}{\sqrt{\kappa^3\mu r}}. \tag{E.126}$$

Note that in view of (E.64) and (E.65), (E.126) is guaranteed by $ndp^2 \gtrsim \delta^{-2}\kappa^6\mu^3 r^4\log^4(n+d)$,

$$\frac{\omega_{\max}}{\sigma_r^\star}\sqrt{\frac{d}{np}} \lesssim \frac{\delta}{\sqrt{\kappa^4\mu r^2\log^2(n+d)}}, \qquad \frac{\omega_{\max}^2}{p\sigma_r^{\star 2}}\sqrt{\frac{d}{n}} \lesssim \frac{\delta}{\sqrt{\kappa^4\mu r^2\log^2(n+d)}}.$$

Similarly we can also show that (which we omit here for brevity)

$$|\alpha_7-\beta_7| \lesssim \delta v_{i,j}^\star$$

holds true under these conditions.

Step 6: putting everything together. We are now ready to combine the above bounds on $|\alpha_k-\beta_k|$, $k=1,\ldots,7$, to conclude that

$$\left|v_{i,j}-v_{i,j}^\star\right| \le \sum_{k=1}^{7}|\alpha_k-\beta_k| \lesssim \delta v_{i,j}^\star,$$

provided that $\varepsilon \lesssim 1$, $n \gtrsim \delta^{-2}\kappa^7\mu^2r^3\kappa_\omega^2\log(n+d)$, $d \gtrsim \kappa^2\mu\log(n+d)$,

$$ndp^2 \gtrsim \delta^{-2}\kappa^6\mu^4r^6\kappa_\omega^2\log^5(n+d), \qquad np \gtrsim \delta^{-2}\kappa^6\mu^3r^5\kappa_\omega^2\log^3(n+d),$$
$$\frac{\omega_{\max}^2}{p\sigma_r^{\star 2}}\sqrt{\frac{d}{n}} \lesssim \frac{\delta}{\kappa^2\mu r^2\kappa_\omega\log^{3/2}(n+d)}, \qquad \frac{\omega_{\max}}{\sigma_r^\star}\sqrt{\frac{d}{np}} \lesssim \frac{\delta}{\kappa^{5/2}\mu r^2\kappa_\omega\log(n+d)},$$

and

$$\left\|\boldsymbol{U}_{i,\cdot}^\star\right\|_2 + \left\|\boldsymbol{U}_{j,\cdot}^\star\right\|_2 \gtrsim \delta^{-1}\kappa\mu r^2\kappa_\omega\log(n+d)\left[\frac{\kappa\mu r\log(n+d)}{\sqrt{ndp}} + \frac{\omega_{\max}^2}{p\sigma_r^{\star 2}}\sqrt{\frac{d}{n}} + \frac{\omega_{\max}}{\sigma_r^\star}\sqrt{\frac{d}{np}}\right]\sqrt{\frac{r}{d}}.$$

To finish up, we shall take $\varepsilon \asymp 1$, and note that (E.113), (E.114) as well as E.115 are guaranteed by the conditions of Lemma 24.

E.5.2 Proof of Lemma 28

To begin with, we know that

$$\begin{aligned} S_{i,j}^\star \in \mathsf{CI}_{i,j}^{1-\alpha} \quad &\Longleftrightarrow \quad S_{i,j}^\star \in \left[S_{i,j} \pm \Phi^{-1}(1-\alpha/2)\sqrt{v_{i,j}}\right] \\ &\Longleftrightarrow \quad \frac{S_{i,j}-S_{i,j}^\star}{\sqrt{v_{i,j}}} \in \left[-\Phi^{-1}(1-\alpha/2), \Phi^{-1}(1-\alpha/2)\right]. \end{aligned} \tag{E.127}$$

Note that we have learned from Lemma 27 that with probability exceeding $1-O((n+d)^{-10})$,

$$\left|v_{i,j}^\star - v_{i,j}\right| \lesssim \delta v_{i,j}^\star,$$

where δ is the (unspecified) quantity that has appeared in Lemma 27. When $\delta \ll 1$, an immediate result is that $v_{i,j} \asymp v_{i,j}^\star$, thus indicating that

$$\Delta \coloneqq \left|\frac{S_{i,j}-S_{i,j}^\star}{\sqrt{v_{i,j}^\star}} - \frac{S_{i,j}-S_{i,j}^\star}{\sqrt{v_{i,j}}}\right| = \left|S_{i,j}-S_{i,j}^\star\right|\left|\frac{v_{i,j}^\star - v_{i,j}}{\sqrt{v_{i,j}^\star v_{i,j}}\left(\sqrt{v_{i,j}^\star}+\sqrt{v_{i,j}}\right)}\right| \lesssim \delta\left|S_{i,j}-S_{i,j}^\star\right|/\sqrt{v_{i,j}^\star}.$$

Suppose for the moment that

$$\Delta \lesssim \frac{1}{\sqrt{\log(n+d)}} \tag{E.128}$$

holds with probability exceeding $1-O((n+d)^{-10})$. From Lemma 26 we know that for any $t\in\mathbb{R}$,

$$\mathbb{P}\left(\frac{S_{i,j}-S_{i,j}^\star}{\sqrt{v_{i,j}^\star}} \le t\right) = \Phi(t) + O\left(\frac{1}{\sqrt{\log(n+d)}}\right). \tag{E.129}$$

Then we know that for any $t\in\mathbb{R}$,

$$\begin{aligned} \mathbb{P}\left(\frac{S_{i,j}-S_{i,j}^\star}{\sqrt{v_{i,j}}} \le t\right) &\le \mathbb{P}\left(\frac{S_{i,j}-S_{i,j}^\star}{\sqrt{v_{i,j}}} \le t, \Delta \lesssim \frac{1}{\sqrt{\log(n+d)}}\right) + O\left((n+d)^{-10}\right) \\ &\le \mathbb{P}\left(\frac{S_{i,j}-S_{i,j}^\star}{\sqrt{v_{i,j}^\star}} \le t + \frac{1}{\sqrt{\log(n+d)}}\right) + O\left((n+d)^{-10}\right) \\ &\overset{\text{(i)}}{\le} \Phi\left(t+\frac{1}{\sqrt{\log(n+d)}}\right) + O\left(\frac{1}{\sqrt{\log(n+d)}}\right) \\ &\overset{\text{(ii)}}{\le} \Phi(t) + O\left(\frac{1}{\sqrt{\log(n+d)}}\right). \end{aligned}$$

Here (i) follows from (E.129); (ii) holds since $\Phi(\cdot)$ is a $1/\sqrt{2\pi}$-Lipschitz continuous function. Similarly we can show that

$$\mathbb{P}\left(\frac{S_{i,j}-S_{i,j}^{\star}}{\sqrt{v_{i,j}}}\leq t\right)\geq\Phi\left(t\right)+O\left(\frac{1}{\sqrt{\log\left(n+d\right)}}\right).$$

Therefore we have

$$\mathbb{P}\left(\frac{S_{i,j}-S_{i,j}^{\star}}{\sqrt{v_{i,j}}}\leq t\right)=\Phi\left(t\right)+O\left(\frac{1}{\sqrt{\log\left(n+d\right)}}\right). \tag{E.130}$$

By taking $t=\Phi^{-1}(1-\alpha/2)$ in (E.130), we have

$$\mathbb{P}\left(\frac{S_{i,j}-S_{i,j}^{\star}}{\sqrt{v_{i,j}}}\leq\Phi^{-1}\left(1-\alpha/2\right)\right)=1-\frac{\alpha}{2}+O\left(\frac{1}{\sqrt{\log\left(n+d\right)}}\right). \tag{E.131}$$

In addition, for all $\varepsilon\in(0,1/\sqrt{\log(n+d)}]$, by taking $t=-\Phi^{-1}(1-\alpha/2)-\varepsilon$ in (E.130), we have

$$\begin{aligned}\mathbb{P}\left(\frac{S_{i,j}-S_{i,j}^{\star}}{\sqrt{v_{i,j}}}\leq-\Phi^{-1}\left(1-\alpha/2\right)-\varepsilon\right)&=\Phi\left(-\Phi^{-1}\left(1-\alpha/2\right)-\varepsilon\right)+O\left(\frac{1}{\sqrt{\log\left(n+d\right)}}\right)\\&=\frac{\alpha}{2}+O\left(\frac{1}{\sqrt{\log\left(n+d\right)}}\right),\end{aligned}$$

where the constant hide in $O(\cdot)$ is independent of ε. Here the last relation holds since $\Phi(\cdot)$ is a $1/\sqrt{2\pi}$-Lipschitz continuous function and $\varepsilon\lesssim1/\sqrt{\log(n+d)}$. As a result,

$$\begin{aligned}\mathbb{P}\left(\frac{S_{i,j}-S_{i,j}^{\star}}{\sqrt{v_{i,j}}}<-\Phi^{-1}\left(1-\alpha/2\right)\right)&=\lim_{\varepsilon\searrow0}\mathbb{P}\left(\frac{S_{i,j}-S_{i,j}^{\star}}{\sqrt{v_{i,j}}}\leq-\Phi^{-1}\left(1-\alpha/2\right)-\varepsilon\right)\\&=\frac{\alpha}{2}+O\left(\frac{1}{\sqrt{\log\left(n+d\right)}}\right).\end{aligned} \tag{E.132}$$

Taking (E.127), (E.130) and (E.132) collectively yields

$$\begin{aligned}\mathbb{P}\left(S_{i,j}^{\star}\in\mathsf{CI}_{i,j}^{1-\alpha}\right)&=\mathbb{P}\left(\frac{S_{i,j}-S_{i,j}^{\star}}{\sqrt{v_{i,j}}}\in\left[-\Phi^{-1}\left(1-\alpha/2\right),\Phi^{-1}\left(1-\alpha/2\right)\right]\right)\\&=\mathbb{P}\left(\frac{S_{i,j}-S_{i,j}^{\star}}{\sqrt{v_{i,j}}}\leq\Phi^{-1}\left(1-\alpha/2\right)\right)-\mathbb{P}\left(\frac{S_{i,j}-S_{i,j}^{\star}}{\sqrt{v_{i,j}}}<-\Phi^{-1}\left(1-\alpha/2\right)\right)\\&=1-\alpha+O\left(\frac{1}{\sqrt{\log\left(n+d\right)}}\right).\end{aligned}$$

as long as (E.128) holds with probability exceeding $1-O((n+d)^{-10})$. With the above calculations in mind, everything boils down to bounding the quantity Δ to the desired level. Towards this, we are in need of the following lemma.

Claim 2. Suppose that the conditions of Lemma 22 hold. Suppose that $np\gtrsim\log^4(n+d)$ and $ndp^2\gtrsim\mu r\log^5(n+d)$. Then with probability exceeding $1-O((n+d)^{-10})$, we have

$$\begin{aligned}\left|S_{i,j}-S_{i,j}^{\star}\right|&\lesssim\zeta_{i,j}+\left\|\boldsymbol{U}_{j,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_{2}\left\|\boldsymbol{U}_{j,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_{2}\left(\sqrt{\frac{\kappa\log\left(n+d\right)}{np}}+\sqrt{\frac{\kappa\mu r\log^{2}\left(n+d\right)}{ndp^{2}}}+\frac{\omega_{\max}}{\sigma_{r}^{\star}}\sqrt{\frac{\log\left(n+d\right)}{np^{2}}}\right).\\&\quad+\left(\sqrt{\frac{\kappa}{np}}+\sqrt{\frac{\kappa\mu r\log^{2}\left(n+d\right)}{ndp^{2}}}+\frac{\omega_{\max}}{\sigma_{r}^{\star}}\sqrt{\frac{\log^{2}\left(n+d\right)}{np^{2}}}\right)\sigma_{r}^{\star}\left(\omega_{j}^{\star}\left\|\boldsymbol{U}_{i,\cdot}^{\star}\right\|_{2}+\omega_{i}^{\star}\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_{2}\right).\end{aligned}$$

Recall from the proof of Lemma 24 (more specifically, Step 3 in Appendix E.4.3) that

$$(v_{i,j}^{\star})^{-1/2}\zeta_{i,j} \le \widetilde{v}_{i,j}^{-1/2}\zeta_{i,j} \lesssim \frac{1}{\sqrt{\log(n+d)}} \tag{E.133}$$

holds under the assumptions of Lemma 24. In addition, recall from Lemma 23 that

$$\sqrt{v_{i,j}^{\star}} \gtrsim \frac{1}{\sqrt{\min\{ndp^2\kappa, np\}}}\left\|\boldsymbol{U}_{i,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2\left\|\boldsymbol{U}_{j,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2 + \left(\frac{\sigma_r^{\star}}{\sqrt{\min\{ndp^2, np\}}} + \frac{\omega_{\min}}{\sqrt{np^2}}\right)\left(\omega_j^{\star}\left\|\boldsymbol{U}_{i,\cdot}^{\star}\right\|_2 + \omega_i^{\star}\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2\right). \tag{E.134}$$

Therefore we can take (E.133), (E.134) and Claim 2 collectively to show that

$$\Delta \lesssim \delta\sqrt{\kappa^2\mu r\kappa_\omega \log^2(n+d)} + \frac{1}{\sqrt{\log(n+d)}} \lesssim \frac{1}{\sqrt{\log(n+d)}}$$

with probability exceeding $1 - O((n+d)^{-10})$ as long as we take

$$\delta \asymp \frac{1}{\kappa\mu^{1/2}r^{1/2}\kappa_\omega^{1/2}\log^{3/2}(n+d)},$$

where we use the AM-GM inequality that

$$\begin{aligned}
\left\|\boldsymbol{U}_{j,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2\left\|\boldsymbol{U}_{j,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2\frac{\omega_{\max}}{\sigma_r^{\star}}\sqrt{\frac{\log(n+d)}{np^2}} &\lesssim \sqrt{\kappa_\omega\log(n+d)}\left\|\boldsymbol{U}_{j,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2\left\|\boldsymbol{U}_{j,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2\frac{\omega_{\min}}{\sigma_r^{\star}}\sqrt{\frac{1}{np^2}} \\
&\lesssim \sqrt{\kappa\kappa_\omega\log(n+d)}\left[\frac{1}{\sqrt{np^2}}\left\|\boldsymbol{U}_{i,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2\left\|\boldsymbol{U}_{j,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2\sqrt{\frac{\mu r}{d}} + \frac{\omega_{\min}^2}{\sqrt{np^2}}\left(\left\|\boldsymbol{U}_{i,\cdot}^{\star}\right\|_2 + \left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2\right)\right] \\
&\lesssim \sqrt{\kappa^2\mu r\kappa_\omega\log(n+d)}\left[\frac{1}{\sqrt{ndp^2\kappa}}\left\|\boldsymbol{U}_{i,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2\left\|\boldsymbol{U}_{j,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2 + \frac{\omega_{\min}}{\sqrt{np^2}}\left(\omega_j^{\star}\left\|\boldsymbol{U}_{i,\cdot}^{\star}\right\|_2 + \omega_i^{\star}\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2\right)\right] \\
&\lesssim \sqrt{\kappa^2\mu r\kappa_\omega\log(n+d)}\sqrt{v_{i,j}^{\star}}.
\end{aligned}$$

This in turn confirms (E.128).

It remains to specify the conditions of Lemma 27 when we choose $\delta \asymp \kappa^{-1}\mu^{-1/2}r^{-1/2}\kappa_\omega^{-1/2}\log^{-3/2}(n+d)$. In this case, these conditions read $n \gtrsim \kappa^9\mu^3r^4\kappa_\omega^3\log^4(n+d)$, $d \gtrsim \kappa^2\mu\log(n+d)$,

$$ndp^2 \gtrsim \kappa^8\mu^5r^7\kappa_\omega^3\log^8(n+d), \qquad np \gtrsim \kappa^8\mu^4r^6\kappa_\omega^3\log^6(n+d),$$

$$\frac{\omega_{\max}^2}{p\sigma_r^{\star 2}}\sqrt{\frac{d}{n}} \lesssim \frac{1}{\kappa^3\mu^{3/2}r^{5/2}\kappa_\omega^{3/2}\log^3(n+d)}, \qquad \frac{\omega_{\max}}{\sigma_r^{\star}}\sqrt{\frac{d}{np}} \lesssim \frac{1}{\kappa^{7/2}\mu^{3/2}r^{5/2}\kappa_\omega^{3/2}\log^{5/2}(n+d)},$$

and

$$\left\|\boldsymbol{U}_{i,\cdot}^{\star}\right\|_2 + \left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2 \gtrsim \kappa^2\mu^{3/2}r^{5/2}\kappa_\omega^{3/2}\log^{5/2}(n+d)\left[\frac{\kappa\mu r\log(n+d)}{\sqrt{ndp}} + \frac{\omega_{\max}^2}{p\sigma_r^{\star 2}}\sqrt{\frac{d}{n}} + \frac{\omega_{\max}}{\sigma_r^{\star}}\sqrt{\frac{d}{np}}\right]\sqrt{\frac{r}{d}},$$

in addition to all the assumptions of Lemma 24. This concludes the proof of this lemma, as long as Claim 2 can be established.

Proof of Claim 2. Let us start by decomposing

$$X_{i,j} = \sum_{l=1}^{n}\Bigg[\underbrace{M_{j,l}^{\natural}E_{i,l}}_{=:a_l} + \underbrace{M_{i,l}^{\natural}E_{j,l}}_{=:b_l} + \underbrace{\sum_{k:k\neq i}E_{i,l}E_{k,l}\left(\boldsymbol{U}_{k,\cdot}^{\star}(\boldsymbol{U}_{j,\cdot}^{\star})^{\top}\right)}_{=:c_l} + \underbrace{\sum_{k:k\neq j}E_{j,l}E_{k,l}\left(\boldsymbol{U}_{k,\cdot}^{\star}(\boldsymbol{U}_{i,\cdot}^{\star})^{\top}\right)}_{=:d_l}\Bigg].$$

We shall then bound $\sum_{l=1}^{n}a_l$, $\sum_{l=1}^{n}b_l$, $\sum_{l=1}^{n}c_l$, and $\sum_{l=1}^{n}d_l$ separately.

- We first bound $\sum_{l=1}^{n} a_l$. It is straightforward to calculate

$$L_a := \max_{1\le l\le n} |a_l| \le B_i \left\| \boldsymbol{M}_{j,\cdot}^{\natural} \right\|_{\infty} \le B_i \left\| \boldsymbol{U}_{j,\cdot}^{\natural} \right\|_2 \sigma_1^{\natural} \left\| \boldsymbol{V}^{\natural} \right\|_{2,\infty},$$

$$V_a := \sum_{l=1}^{n} \mathsf{var}\left(M_{j,l}^{\natural} E_{i,l} | \boldsymbol{F} \right) = \sum_{l=1}^{n} M_{j,l}^{\natural 2} \sigma_{i,l}^2 = \sum_{l=1}^{n} M_{j,l}^{\natural 2} \sigma_{i,l}^2 = \sum_{l=1}^{n} M_{j,l}^{\natural 2} \left[\frac{1-p}{np} \left(\boldsymbol{U}_{i,\cdot}^{\star} \boldsymbol{\Sigma}^{\star} \boldsymbol{f}_l \right)^2 + \frac{\omega_i^{\star 2}}{np} \right],$$

whereB_i and B_j are two (random) quantities such that

$$\max_{l\in[n]} |E_{i,l}| \le B_i \qquad \text{and} \qquad \max_{l\in[n]} |E_{j,l}| \le B_j.$$

On the event $\mathcal{E}_{\mathsf{good}}$, we know that

$$\begin{aligned} L_a &\overset{\text{(i)}}{\lesssim} \frac{1}{p} \sqrt{\frac{\log (n+d)}{n}} \left(\left\| \boldsymbol{U}_{i,\cdot}^{\star} \boldsymbol{\Sigma}^{\star} \right\|_2 + \omega_i^{\star} \right) \left\| \boldsymbol{U}_{j,\cdot}^{\star} \right\|_2 \sigma_1^{\star} \sqrt{\frac{\log (n+d)}{n}}, \\ &\lesssim \frac{\log (n+d)}{np} \left(\left\| \boldsymbol{U}_{i,\cdot}^{\star} \boldsymbol{\Sigma}^{\star} \right\|_2 + \omega_i^{\star} \right) \left\| \boldsymbol{U}_{j,\cdot}^{\star} \right\|_2 \sigma_1^{\star}, \end{aligned}$$

and

$$\begin{aligned} V_a &\overset{\text{(ii)}}{\lesssim} \sum_{l=1}^{n} M_{j,l}^{\natural 2} \left[\frac{1-p}{np} \left\| \boldsymbol{U}_{i,\cdot}^{\star} \boldsymbol{\Sigma}^{\star} \right\|_2^2 \log (n+d) + \frac{\omega_i^{\star 2}}{np} \right] \\ &= \left\| \boldsymbol{M}_{j,\cdot}^{\natural} \right\|_2^2 \left[\frac{1-p}{np} \left\| \boldsymbol{U}_{i,\cdot}^{\star} \boldsymbol{\Sigma}^{\star} \right\|_2^2 \log (n+d) + \frac{\omega_i^{\star 2}}{np} \right] \\ &\lesssim \left\| \boldsymbol{U}_{j,\cdot}^{\natural} \right\|_2^2 \sigma_1^{\natural 2} \left[\frac{1-p}{np} \left\| \boldsymbol{U}_{i,\cdot}^{\star} \boldsymbol{\Sigma}^{\star} \right\|_2^2 \log (n+d) + \frac{\omega_i^{\star 2}}{np} \right] \\ &\overset{\text{(iii)}}{\lesssim} \left[\frac{1-p}{np} \left\| \boldsymbol{U}_{i,\cdot}^{\star} \boldsymbol{\Sigma}^{\star} \right\|_2^2 \log (n+d) + \frac{\omega_i^{\star 2}}{np} \right] \left\| \boldsymbol{U}_{j,\cdot}^{\star} \right\|_2^2 \sigma_1^{\star 2}. \end{aligned}$$

Here (i) follows from (D.13), (D.16) and (D.18); (ii) follows from (D.20); (iii) follows from (D.13). Therefore in view of the Bernstein inequality (Vershynin, 2018, Theorem 2.8.4), conditional on $\boldsymbol{F}$, with probability exceeding $1 - O((n+d)^{-10})$,

$$\begin{aligned} \sum_{l=1}^{n} a_l &\lesssim \sqrt{V_a \log (n+d)} + L_a \log (n+d) \\ &\lesssim \left[\sqrt{\frac{\log (n+d)}{np}} \left\| \boldsymbol{U}_{i,\cdot}^{\star} \boldsymbol{\Sigma}^{\star} \right\|_2 + \frac{\omega_i^{\star}}{\sqrt{np}} \right] \left\| \boldsymbol{U}_{j,\cdot}^{\star} \right\|_2 \sigma_1^{\star} \frac{\log^2 (n+d)}{np} \left(\left\| \boldsymbol{U}_{i,\cdot}^{\star} \boldsymbol{\Sigma}^{\star} \right\|_2 + \omega_j^{\star} \right) \left\| \boldsymbol{U}_{j,\cdot}^{\star} \right\|_2 \sigma_1^{\star} \\ &\lesssim \left[\sqrt{\frac{\log (n+d)}{np}} \left\| \boldsymbol{U}_{i,\cdot}^{\star} \boldsymbol{\Sigma}^{\star} \right\|_2 + \frac{\omega_i^{\star}}{\sqrt{np}} \right] \left\| \boldsymbol{U}_{j,\cdot}^{\star} \right\|_2 \sigma_1^{\star}, \end{aligned}$$

provided that $np \gtrsim \log^4(n+d)$. Similarly we can show that with probability exceeding $1 - O((n+d)^{-10})$,

$$\sum_{l=1}^{n} b_l \lesssim \left[\sqrt{\frac{\log (n+d)}{np}} \left\| \boldsymbol{U}_{j,\cdot}^{\star} \boldsymbol{\Sigma}^{\star} \right\|_2 + \frac{\omega_j^{\star}}{\sqrt{np}} \right] \left\| \boldsymbol{U}_{i,\cdot}^{\star} \right\|_2 \sigma_1^{\star}.$$

- We then move on to bound $\sum_{l=1}^{n} c_l$. It has been shown in (E.97) that with probability exceeding $1 - O((n+d)^{-101})$,

$$\max_{1\le l\le n} |c_l| \le \underbrace{\widetilde{C} \sigma_{\mathsf{ub}} B_i \sqrt{\log (n+d)} \left\| \boldsymbol{U}_{j,\cdot}^{\star} \right\|_2 + \widetilde{C} B B_i \left\| \boldsymbol{U}_{j,\cdot}^{\star} \right\|_2 \sqrt{\frac{\mu r}{d}} \log (n+d)}_{=: C_{\mathsf{prob}}},$$

where $\widetilde{C} > 0$ is some sufficiently large constant. We also know from (E.99) that $\max_{1\leq l\leq n}|c_l|$ satisfies the following deterministic bound:

$$\max_{1\leq l\leq n}|c_l| \leq \underbrace{B_i\sqrt{d}B\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2}_{=:C_{\mathsf{det}}}.$$

Then with probability exceeding $1-O((n+d)^{-101})$, one has

$$\sum_{l=1}^{n} c_l = \sum_{l=1}^{n} c_l\,\mathbb{1}_{|c_l|\leq C_{\mathsf{prob}}}.$$

It is then straightforward to calculate that

$$\begin{aligned}
L_c &\coloneqq \max_{1\leq l\leq n}\left|c_l\,\mathbb{1}_{|c_l|\leq C_{\mathsf{prob}}}\right| \leq C_{\mathsf{prob}},\\
V_c &\coloneqq \sum_{l=1}^{n}\mathsf{var}\left(c_l\,\mathbb{1}_{|c_l|\leq C_{\mathsf{prob}}}\right) \leq \sum_{l=1}^{n}\mathbb{E}\left[c_l^2\,\mathbb{1}_{|c_l|\leq C_{\mathsf{prob}}}\right] \leq \sum_{l=1}^{n}\mathbb{E}\left[c_l^2\right] = \sum_{l=1}^{n}\mathsf{var}\left(c_l\right)\\
&= \sum_{l=1}^{n}\mathsf{var}\left[\sum_{k:k\neq i}E_{i,l}E_{k,l}\left(\boldsymbol{U}_{k,\cdot}^{\star}\boldsymbol{U}_{j,\cdot}^{\star\top}\right)\right] \leq \sum_{l=1}^{n}\sum_{k=1}^{d}\sigma_{i,l}^2\sigma_{k,l}^2\left(\boldsymbol{U}_{k,\cdot}^{\star}\boldsymbol{U}_{j,\cdot}^{\star\top}\right)^2,\\
M_c &\coloneqq \sum_{l=1}^{n}\mathbb{E}\left[c_l\,\mathbb{1}_{|c_l|> C_{\mathsf{prob}}}\right] \leq C_{\mathsf{det}}\sum_{l=1}^{n}\mathbb{P}\left(|c_l|>C_{\mathsf{prob}}\right) \lesssim C_{\mathsf{det}}\left(n+d\right)^{-100}.
\end{aligned}$$

On the event $\mathcal{E}_{\mathsf{good}}$, it is seen that

$$\begin{aligned}
L_c &\lesssim \sigma_{\mathsf{ub}}B_i\sqrt{\log\left(n+d\right)}\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2 + BB_i\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2\sqrt{\frac{\mu r}{d}}\log\left(n+d\right),\\
&\overset{\text{(i)}}{\lesssim} \sigma_{\mathsf{ub}}B_i\sqrt{\log\left(n+d\right)}\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2 + \sigma_{\mathsf{ub}}\sqrt{\frac{\log\left(n+d\right)}{p}}B_i\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2\sqrt{\frac{\mu r}{d}}\log\left(n+d\right),\\
&\lesssim \sigma_{\mathsf{ub}}B_i\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2\sqrt{\log\left(n+d\right)}\left(1+\sqrt{\frac{\mu r}{dp}}\log\left(n+d\right)\right),
\end{aligned}$$

$$\begin{aligned}
V_c &\leq \sigma_{\mathsf{ub}}^2\sum_{l=1}^{n}\sigma_{i,l}^2\sum_{k=1}^{d}\left(\boldsymbol{U}_{k,\cdot}^{\star}\boldsymbol{U}_{j,\cdot}^{\star\top}\right)^2 = n\sigma_{\mathsf{ub}}^2\left(\frac{1-p}{np}\left(\boldsymbol{U}_{i,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\boldsymbol{f}_l\right)^2+\frac{\omega_i^{\star 2}}{np}\right)\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2^2,\\
&\overset{\text{(ii)}}{\lesssim} n\sigma_{\mathsf{ub}}^2\left(\frac{1-p}{np}\left\|\boldsymbol{U}_{i,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2^2\log\left(n+d\right)+\frac{\omega_i^{\star 2}}{np}\right)\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2^2,
\end{aligned}$$

and

$$\begin{aligned}
M_c &\lesssim B_i\sqrt{d}B\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2\left(n+d\right)^{-100} \overset{\text{(iii)}}{\lesssim} \sigma_{\mathsf{ub}}B_i\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2\sqrt{\frac{d\log\left(n+d\right)}{p}}\left(n+d\right)^{-100}\\
&\overset{\text{(iv)}}{\lesssim} \sigma_{\mathsf{ub}}B_i\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2\sqrt{nd}\left(n+d\right)^{-100} \lesssim \sigma_{\mathsf{ub}}B_i\left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2\left(n+d\right)^{-98}.
\end{aligned}$$

Here, (i) and (iii) follow from (D.18) and (D.17); (ii) arises from (D.20); and (iv) holds true provided that $np \gtrsim \log(n+d)$. Therefore, by virtue of the Bernstein inequality (Vershynin, 2018, Theorem 2.8.4), we see that conditional on $\boldsymbol{F}$,

$$\sum_{l=1}^{n} c_l = \sum_{l=1}^{n} c_l\,\mathbb{1}_{|c_l|\leq C_{\mathsf{prob}}} \lesssim M_c + \sqrt{V_c\log\left(n+d\right)} + L_c\log\left(n+d\right)$$

$$
\begin{aligned}
&\lesssim \sigma_{\mathsf{ub}} B_i \left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2 (n+d)^{-98} + \sigma_{\mathsf{ub}} \left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2 \sqrt{\frac{\log(n+d)}{p}} \left(\left\|\boldsymbol{U}_{i,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2 \sqrt{\log(n+d)} + \omega_i^{\star}\right) \\
&\quad + \sigma_{\mathsf{ub}} B_i \left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2 \log^{3/2}(n+d) \left(1 + \sqrt{\frac{\mu r}{dp}} \log(n+d)\right) \\
&\overset{\text{(i)}}{\lesssim} \sigma_{\mathsf{ub}} \left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2 \left(\left\|\boldsymbol{U}_{i,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2 + \omega_i^{\star}\right) \sqrt{\frac{\log(n+d)}{p}} \\
&\lesssim \left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2 \left(\left\|\boldsymbol{U}_{i,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2 + \omega_i^{\star}\right) \left(\sqrt{\frac{\mu r \log^2(n+d)}{ndp^2}} \sigma_1^{\star} + \omega_{\max} \sqrt{\frac{\log(n+d)}{np^2}}\right)
\end{aligned}
$$

holds with probability exceeding $1-O((n+d)^{-10})$. Here, (i) holds true provided that $np \gtrsim \log^3(n+d)$ and $ndp^2 \gtrsim \mu r \log^5(n+d)$. Similarly we can demonstrate that with probability exceeding $1-O((n+d)^{-10})$,

$$
\sum_{l=1}^{n} d_l \lesssim \left\|\boldsymbol{U}_{i,\cdot}^{\star}\right\|_2 \left(\left\|\boldsymbol{U}_{j,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2 + \omega_j^{\star}\right) \left(\sqrt{\frac{\mu r \log^2(n+d)}{ndp^2}} \sigma_1^{\star} + \omega_{\max} \sqrt{\frac{\log(n+d)}{np^2}}\right).
$$

Therefore, the above two bounds taken together lead to

$$
\begin{aligned}
\sum_{l=1}^{n} (c_l + d_l) &\lesssim \left[\frac{1}{\sigma_r^{\star}} \left\|\boldsymbol{U}_{j,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2 \left\|\boldsymbol{U}_{j,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2 + \left(\omega_j^{\star} \left\|\boldsymbol{U}_{i,\cdot}^{\star}\right\|_2 + \omega_i^{\star} \left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2\right)\right] \\
&\qquad \cdot \left(\sqrt{\frac{\mu r \log^2(n+d)}{ndp^2}} \sigma_1^{\star} + \omega_{\max} \sqrt{\frac{\log(n+d)}{np^2}}\right).
\end{aligned}
$$

$$
\sum_{l=1}^{n} (a_l + b_l) \lesssim \sqrt{\frac{\kappa \log(n+d)}{np}} \left\|\boldsymbol{U}_{j,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2 \left\|\boldsymbol{U}_{j,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2 + \frac{\sigma_1^{\star}}{\sqrt{np}} \left(\omega_j^{\star} \left\|\boldsymbol{U}_{i,\cdot}^{\star}\right\|_2 + \omega_i^{\star} \left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2\right),
$$

Therefore, putting the above bounds together, one can conclude that

$$
\begin{aligned}
\left|X_{i,j}\right| &\le \sum_{l=1}^{n} (a_l + b_l + c_l + d_l) \\
&\lesssim \left\|\boldsymbol{U}_{j,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2 \left\|\boldsymbol{U}_{j,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2 \left(\sqrt{\frac{\kappa \log(n+d)}{np}} + \sqrt{\frac{\kappa \mu r \log^2(n+d)}{ndp^2}}\right) \\
&\quad + \sigma_r^{\star} \left(\omega_j^{\star} \left\|\boldsymbol{U}_{i,\cdot}^{\star}\right\|_2 + \omega_i^{\star} \left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2\right) \left(\sqrt{\frac{\kappa}{np}} + \sqrt{\frac{\kappa \mu r \log^2(n+d)}{ndp^2}}\right) \\
&\quad + \left\|\boldsymbol{U}_{j,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2 \left\|\boldsymbol{U}_{j,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2 \frac{\omega_{\max}}{\sigma_r^{\star}} \sqrt{\frac{\log(n+d)}{np^2}} + \left(\omega_j^{\star} \left\|\boldsymbol{U}_{i,\cdot}^{\star}\right\|_2 + \omega_i^{\star} \left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2\right) \omega_{\max} \sqrt{\frac{\log(n+d)}{np^2}} \\
&\lesssim \left\|\boldsymbol{U}_{j,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2 \left\|\boldsymbol{U}_{j,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2 \left(\sqrt{\frac{\kappa \log(n+d)}{np}} + \sqrt{\frac{\kappa \mu r \log^2(n+d)}{ndp^2}} + \frac{\omega_{\max}}{\sigma_r^{\star}} \sqrt{\frac{\log(n+d)}{np^2}}\right) \\
&\quad + \left[\sqrt{\frac{\kappa}{np}} + \sqrt{\frac{\kappa \mu r \log^2(n+d)}{ndp^2}} + \frac{\omega_{\max}}{\sigma_r^{\star}} \sqrt{\frac{\log^2(n+d)}{np^2}}\right] \sigma_r^{\star} \left(\omega_j^{\star} \left\|\boldsymbol{U}_{i,\cdot}^{\star}\right\|_2 + \omega_i^{\star} \left\|\boldsymbol{U}_{j,\cdot}^{\star}\right\|_2\right),
\end{aligned}
$$

Combine the above result with Lemma 22 to arrive at

$$
\left|S_{i,j} - S_{i,j}^{\star}\right| \lesssim \zeta_{i,j} + \left\|\boldsymbol{U}_{j,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2 \left\|\boldsymbol{U}_{j,\cdot}^{\star}\boldsymbol{\Sigma}^{\star}\right\|_2 \left(\sqrt{\frac{\kappa \log(n+d)}{np}} + \sqrt{\frac{\kappa \mu r \log^2(n+d)}{ndp^2}} + \frac{\omega_{\max}}{\sigma_r^{\star}} \sqrt{\frac{\log(n+d)}{np^2}}\right).
$$

$$+\left(\sqrt{\frac{\kappa}{np}}+\sqrt{\frac{\kappa\mu r\log^2(n+d)}{ndp^2}}+\frac{\omega_{\max}}{\sigma_r^\star}\sqrt{\frac{\log^2(n+d)}{np^2}}\right)\sigma_r^\star\left(\omega_j^\star\left\|\boldsymbol{U}_{i,\cdot}^\star\right\|_2+\omega_i^\star\left\|\boldsymbol{U}_{j,\cdot}^\star\right\|_2\right).$$

□

F Other useful lemmas

In this section, we gather a few useful results from prior literature that prove useful for our analysis. The first theorem is a non-asymptotic version of the Berry-Esseen Theorem, which has been established using Stein's method; see Chen et al. (2010, Theorem 3.7).

Theorem 15. *Let $\xi_1,\dots,\xi_n$ be independent random variables with zero means, satisfying $\sum_{i=1}^n \mathsf{var}(\xi_i)=1$. Then $W=\sum_{i=1}^n \xi_i$ satisfies*

$$\sup_{z\in\mathbb{R}}\left|\mathbb{P}\left(W\le z\right)-\Phi\left(z\right)\right|\le 10\gamma,\qquad \text{where } \gamma=\sum_{i=1}^n\mathbb{E}\left[\left|\xi_i\right|^3\right].$$

The next theorem, which we borrow from Raič (2019, Theorem 1.1), generalizes Theorem 15 to the vector case.

Theorem 16. *Let $\boldsymbol{\xi}_1,\dots,\boldsymbol{\xi}_n$ be independent, $\mathbb{R}^d$-valued random vectors with zero means. Let $\boldsymbol{W}=\sum_{i=1}^n\boldsymbol{\xi}_i$ and assume $\boldsymbol{\Sigma}=\mathsf{cov}(\boldsymbol{W})$ is invertible. Let $\boldsymbol{Z}$ be a d-dimensional Gaussian random vector with zero mean and covariance matrix $\boldsymbol{\Sigma}$. Then we have*

$$\sup_{\mathcal{C}\in\mathscr{C}^d}\left|\mathbb{P}\left(\boldsymbol{W}\in\mathcal{C}\right)-\mathbb{P}\left(\boldsymbol{Z}\in\mathcal{C}\right)\right|\le\left(42d^{1/4}+16\right)\gamma,\qquad \text{where } \gamma=\sum_{i=1}^n\mathbb{E}\left[\left\|\boldsymbol{\Sigma}^{-1/2}\boldsymbol{\xi}_i\right\|_2^3\right].$$

Here, $\mathscr{C}^d$ represents the set of all convex sets in $\mathbb{R}^d$.

Moreover, deriving our distributional theory involves some basic results about the total-variation distance between two Gaussian distributions, as stated below. This result can be found in Devroye et al. (2018, Theorem 1.1).

Theorem 17. *Let $\boldsymbol{\mu}\in\mathbb{R}^d$, $\boldsymbol{\Sigma}_1$ and $\boldsymbol{\Sigma}_2$ be positive definite $d\times d$ matrices. Then the total-variation distance between $\mathcal{N}\left(\boldsymbol{\mu},\boldsymbol{\Sigma}_1\right)$ and $\mathcal{N}\left(\boldsymbol{\mu},\boldsymbol{\Sigma}_2\right)$ — denoted by $\mathsf{TV}\left(\mathcal{N}\left(\boldsymbol{\mu},\boldsymbol{\Sigma}_1\right),\mathcal{N}\left(\boldsymbol{\mu},\boldsymbol{\Sigma}_2\right)\right)$ — satisfies*

$$\frac{1}{100}\le\frac{\mathsf{TV}\left(\mathcal{N}\left(\boldsymbol{\mu},\boldsymbol{\Sigma}_1\right),\mathcal{N}\left(\boldsymbol{\mu},\boldsymbol{\Sigma}_2\right)\right)}{\min\left\{1,\left\|\boldsymbol{\Sigma}_1^{-1/2}\boldsymbol{\Sigma}_2\boldsymbol{\Sigma}_1^{-1/2}-\boldsymbol{I}_d\right\|_{\mathrm{F}}\right\}}\le\frac{3}{2}.$$

Recall the definition of $\mathcal{C}^\varepsilon$ in (A.3). Then for any convex set $\mathcal{C}\in\mathscr{C}^d$, let us define the following quantity related to Gaussian distributions

$$\gamma\left(\mathcal{C}\right)\coloneqq\sup_{\varepsilon>0}\max\left\{\frac{1}{\varepsilon}\mathcal{N}\left(\boldsymbol{0},\boldsymbol{I}_d\right)\left\{\mathcal{C}^\varepsilon\setminus\mathcal{C}\right\},\frac{1}{\varepsilon}\mathcal{N}\left(\boldsymbol{0},\boldsymbol{I}_d\right)\left\{\mathcal{C}\setminus\mathcal{C}^{-\varepsilon}\right\}\right\},\tag{F.1}$$

and, in addition,

$$\gamma_d\coloneqq\sup_{\mathcal{C}\in\mathscr{C}^d}\gamma\left(\mathcal{C}\right).\tag{F.2}$$

The following theorem from Raič (2019, Theorem 1.2) delivers an upper bound on the quantity γ_d.

Theorem 18. *For all $d\in\mathbb{N}$, we have*

$$\gamma_d<0.59d^{1/4}+0.21.$$

Finally, we are in need of the following basic lemma in order to translate results derived for bounded random variables to the ones concerned with sub-Gaussian random variables.

Lemma 31. *There exist two universal constants $C_\delta, C_\sigma > 0$ such that: for any sub-Gaussian random variable X with $\mathbb{E}[X] = 0$, $\mathsf{var}(X) = \sigma^2$, $\|X\|_{\psi_2} \lesssim \sigma$, and any $\delta \in (0, C_\delta \sigma)$, one can construct a random variable $\widetilde{X}$ satisfying the following properties:*

1. *$\widetilde{X}$ is equal to X with probability at least $1 - \delta$;*
2. *$\mathbb{E}[\widetilde{X}] = 0$;*
3. *$\widetilde{X}$ is a bounded random variable: $|\widetilde{X}| \le C_\sigma \sigma \sqrt{\log(\delta^{-1})}$;*
4. *The variance of $\widetilde{X}$ obeys: $\mathsf{var}(\widetilde{X}) = (1 + O(\sqrt{\delta}))\sigma^2$;*
5. *$\widetilde{X}$ is a sub-Gaussian random variable obeying $\|\widetilde{X}\|_{\psi_2} \lesssim \sigma$.*

F.1 Proof of Lemma 31

Step 1: lower bounding $\mathbb{E}|X|$. For any $t > 0$, it is easily seen that

$$\mathbb{E}\left[X^2 \mathbb{1}_{|X|>t}\right] \overset{\text{(i)}}{\le} \left(\mathbb{E}\left[X^4\right]\right)^{\frac{1}{2}} \left(\mathbb{P}\left(|X| > t\right)\right)^{\frac{1}{2}} \overset{\text{(ii)}}{\lesssim} \sigma^2 \exp\left(-\frac{t^2}{C\sigma^2}\right),$$

where $C > 0$ is some absolute constant. Here, (i) results from the Cauchy-Schwarz inequality, whereas (ii) follows from standard properties of sub-Gaussian random variables (Vershynin, 2018, Proposition 2.5.2). By taking $t = c_{\mathsf{lb}}^{-1}\sigma/4$ for some sufficiently small constant $c_{\mathsf{lb}} > 0$, we can guarantee that $\mathbb{E}[X^2 \mathbb{1}_{|X|>t}] \le \sigma^2/2$, which in turn results in

$$\mathbb{E}\left[X^2 \mathbb{1}_{|X|\le t}\right] = \mathbb{E}\left[X^2\right] - \mathbb{E}\left[X^2 \mathbb{1}_{|X|>t}\right] \ge \sigma^2 - \frac{1}{2}\sigma^2 \ge \frac{1}{2}\sigma^2.$$

As a consequence, we obtain (with the above choice $t = c_{\mathsf{lb}}^{-1}\sigma/4$)

$$\mathbb{E}\left[|X|\right] \ge \mathbb{E}\left[|X| \mathbb{1}_{|X|\le t}\right] \ge \frac{\mathbb{E}\left[X^2 \mathbb{1}_{|X|\le t}\right]}{t} \ge 2c_{\mathsf{lb}}\sigma. \tag{F.3}$$

Step 2: constructing $\widetilde{X}$ by truncating X randomly. For notational simplicity, let us define $X^+ = X \vee 0$ and $X^- = (-X) \vee 0$. Given that $\mathbb{E}[X] = 0$, we have

$$\mathbb{E}\left[X^+\right] = \mathbb{E}\left[X^-\right] = \frac{1}{2}\mathbb{E}\left[|X|\right] \ge c_{\mathsf{lb}}\sigma, \tag{F.4}$$

where the last inequality comes from (F.3). Define a function $f : \mathbb{R}^+ \mapsto \mathbb{R}^+$ as follows

$$f\left(x\right) \coloneqq \mathbb{E}\left[X \mathbb{1}_{X\ge x}\right].$$

It is straightforward to check that $f(x)$ is monotonically non-increasing within the domain $x \in [0, \infty)$. In view of the monotone convergence theorem, we know that $f(x)$ is a left continuous function with $\lim_{x\searrow 0} f(x) = \mathbb{E}[X^+]$ and $\lim_{x\to+\infty} f(x) = 0$. In addition, for any $x \in \mathbb{R}^+$, one has

$$f\left(x\right) = \mathbb{E}\left[X \mathbb{1}_{X\ge x}\right] \overset{\text{(i)}}{\le} \left(\mathbb{E}\left[X^2\right]\right)^{\frac{1}{2}} \left(\mathbb{P}\left(X \ge x\right)\right)^{\frac{1}{2}} \overset{\text{(ii)}}{\le} \sigma \exp\left(-\frac{x^2}{C\sigma^2}\right), \tag{F.5}$$

where $C > 0$ is some absolute constant. Here, (i) comes from the Cauchy-Schwarz inequality, while (ii) follows from standard properties of sub-Gaussian random variables (Vershynin, 2018, Proposition 2.5.2).

For any given $\varepsilon \in (0, c_{\mathsf{lb}}\sigma/2)$, we take

$$x_\varepsilon \coloneqq \sup\left\{x \in \mathbb{R}^+ : f\left(x\right) \ge \varepsilon\right\}.$$

Since f is known to be left continuous, we know that

$$\lim_{x\nearrow x_\varepsilon} f\left(x\right) = f\left(x_\varepsilon\right) \ge \varepsilon \ge \lim_{x\searrow x_\varepsilon} f\left(x\right).$$

Taking the definition of x_ε and (F.5) collectively yields

$$\varepsilon \le f(x_\varepsilon) \le \sigma \exp\left(-\frac{x_\varepsilon^2}{C\sigma^2}\right),$$

which further gives an upper bound on x_ε as follows

$$x_\varepsilon \le \sigma\sqrt{C\log(\sigma/\varepsilon)}. \tag{F.6}$$

In addition, we can also lower bound x_ε by observing that

$$\begin{aligned}\mathbb{E}\left[X^+\right] &= \mathbb{E}\left[X^+ \mathbb{1}_{X\le x_\varepsilon}\right] + \mathbb{E}\left[X \mathbb{1}_{X> x_\varepsilon}\right] = \mathbb{E}\left[X^+ \mathbb{1}_{X\le x_\varepsilon}\right] + \lim_{x\to x_\varepsilon^+} f(x)\\ &\le x_\varepsilon + \varepsilon \le x_\varepsilon + \frac{1}{2}c_{\mathsf{lb}}\sigma,\end{aligned}$$

which taken collectively with (F.4) yields

$$x_\varepsilon \ge \frac{1}{2}c_{\mathsf{lb}}\sigma. \tag{F.7}$$

With these calculations in place, we define $\widetilde{X}^+$ as follows:

- If $\lim_{x\searrow x_\varepsilon} f(x) = f(x_\varepsilon)$, we immediately know that $f(x_\varepsilon)=\varepsilon$. Then we can set

  $$\widetilde{X}^+ \coloneqq X^+ \mathbb{1}_{X^+<x_\varepsilon}.$$

  This construction gives $\widetilde{X}^+ \le x_\varepsilon \le \sigma\sqrt{C\log(\sigma/\varepsilon)}$ and

  $$\mathbb{E}\left[\widetilde{X}^+\right] = \mathbb{E}\left[X^+\right] - f(x_\varepsilon) = \mathbb{E}\left[X^+\right] - \varepsilon.$$

  We can also derive from (F.7) that

  $$\mathbb{P}\left(\widetilde{X}^+ \ne X^+\right) = \mathbb{P}\left(X^+ \ge x_\varepsilon\right) \le \frac{\mathbb{E}\left[X^+ \mathbb{1}_{X^+\ge x_\varepsilon}\right]}{x_\varepsilon} = \frac{f(x_\varepsilon)}{x_\varepsilon} = \frac{\varepsilon}{x_\varepsilon} \le \frac{2\varepsilon}{c_{\mathsf{lb}}\sigma}.$$

- If $\lim_{x\searrow x_\varepsilon} f(x) < f(x_\varepsilon)$, we know that

  $$\mathbb{E}\left[X^+ \mathbb{1}_{X^+=x_\varepsilon}\right] = f(x_\varepsilon) - \lim_{x\searrow x_\varepsilon} > 0. \tag{F.8}$$

  Then one can set

  $$\widetilde{X}^+ \coloneqq X^+ \mathbb{1}_{X^+<x_\varepsilon} + X \mathbb{1}_{X^+=x_\varepsilon} Q,$$

  where Q is a Bernoulli random variable (independent of X) with parameter

  $$q = \frac{f(x_\varepsilon) - \varepsilon}{f(x_\varepsilon) - \lim_{x\searrow x_\varepsilon}},$$

  i.e., $\mathbb{P}(Q=1) = 1-\mathbb{P}(Q=0) = q$. This construction gives $\widetilde{X}^+ \le x_\varepsilon \le \sigma\sqrt{C\log(\sigma/\varepsilon)}$ and

  $$\begin{aligned}\mathbb{E}\left[\widetilde{X}^+\right] &= \mathbb{E}\left[X^+ \mathbb{1}_{X^+<x_\varepsilon}\right] + q\mathbb{E}\left[X \mathbb{1}_{X^+=x_\varepsilon}\right] \quad (F.9)\\ &= \mathbb{E}\left[X^+\right] - f(x_\varepsilon) + q\left[f(x_\varepsilon) - \lim_{x\to x_\varepsilon^+} f(x)\right]\\ &= \mathbb{E}\left[X^+\right] - \varepsilon. \quad (F.10)\end{aligned}$$

  In addition, we have

  $$\mathbb{P}\left(\widetilde{X}^+ \ne X^+\right) = \mathbb{P}\left(X^+ > x_\varepsilon\right) + \mathbb{P}\left(X^+ = x_\varepsilon, Q=0\right) = \mathbb{P}\left(X^+ > x_\varepsilon\right) + (1-q)\,\mathbb{P}\left(X^+ = x_\varepsilon\right)$$

$$
\begin{aligned}
&\overset{\text{(i)}}{\leq} \frac{\mathbb{E}\left[X^+ \mathbb{1}_{X^+ > x_\varepsilon}\right]}{x_\varepsilon} + (1-q)\frac{\mathbb{E}\left[X^+ \mathbb{1}_{X^+ = x_\varepsilon}\right]}{x_\varepsilon} \\
&= \frac{\mathbb{E}\left[X^+\right] - \mathbb{E}\left[X^+ \mathbb{1}_{X^+ < x_\varepsilon}\right] - q\mathbb{E}\left[X^+ \mathbb{1}_{X^+ = x_\varepsilon}\right]}{x_\varepsilon} \overset{\text{(ii)}}{=} \frac{\mathbb{E}\left[X^+\right] - \mathbb{E}\left[\widetilde{X}^+\right]}{x_\varepsilon} \\
&\overset{\text{(iii)}}{=} \frac{\varepsilon}{x_\varepsilon} \overset{\text{(iv)}}{\leq} \frac{2\varepsilon}{c_{\mathsf{lb}}\sigma}.
\end{aligned}
$$

Here, (i) holds since $\mathbb{E}[X^+ \mathbb{1}_{X^+ > x_\varepsilon}] \geq x_\varepsilon \mathbb{P}(X^+ > x_\varepsilon)$ and $\mathbb{E}[X^+ \mathbb{1}_{X^+ = x_\varepsilon}] = x_\varepsilon \cdot \mathbb{P}(X^+ = x_\varepsilon)$; (ii) follows from (F.9); (iii) is a consequence of (F.10); and (iv) follows from (F.7).

We have thus constructed a random variable $\widetilde{X}^+$ that satisfies: (i) $\widetilde{X}^+$ equals either X^+ or 0; (ii) $\mathbb{P}(\widetilde{X}^+ \neq X^+) \leq 2\varepsilon/(c_{\mathsf{lb}}\sigma)$; (iii) $\mathbb{E}[\widetilde{X}^+] = \mathbb{E}[X^+] - \varepsilon$; and (iv) $0 \leq \widetilde{X}^+ \leq \sigma\sqrt{C\log(\sigma/\varepsilon)}$. Similarly, we can also construct another random variable $\widetilde{X}^-$ satisfying: (i) $\widetilde{X}^-$ equals either X^- or 0; (ii) $\mathbb{P}(\widetilde{X}^- \neq X^-) \leq 2\varepsilon/(c_{\mathsf{lb}}\sigma)$; (iii) $\mathbb{E}[\widetilde{X}^-] = \mathbb{E}[X^-] - \varepsilon$; and (iv) $0 \leq \widetilde{X}^- \leq \sigma\sqrt{C\log(\sigma/\varepsilon)}$. We shall then construct $\widetilde{X}$ as follows

$$
\widetilde{X} \coloneqq \widetilde{X}^+ - \widetilde{X}^-.
$$

Step 3: verifying the advertised properties of $\widetilde{X}$. To finish up, we can check that the following properties are satisfied:

1. $\widetilde{X}$ has mean zero, namely,
$$
\mathbb{E}\left[\widetilde{X}\right] = \mathbb{E}\left[\widetilde{X}^+\right] - \mathbb{E}\left[\widetilde{X}^-\right] = \mathbb{E}\left[X^+\right] - \varepsilon - \mathbb{E}\left[X^-\right] + \varepsilon = \mathbb{E}\left[X\right] = 0.
$$

2. $\widetilde{X}$ is identical to X with high probability, namely,
$$
\mathbb{P}\left(X \neq \widetilde{X}\right) = \mathbb{P}\left(X^+ \neq \widetilde{X}^+\right) + \mathbb{P}\left(X^- \neq \widetilde{X}^-\right) \leq \frac{4\varepsilon}{c_{\mathsf{lb}}\sigma}.
$$

3. $\widetilde{X}$ is a bounded random variable in the sense that $|\widetilde{X}| \leq \sigma\sqrt{C\log(\sigma/\varepsilon)} \lesssim \sigma\sqrt{\log(\sigma/\varepsilon)}$.

4. The variance of $\widetilde{X}$ is close to σ^2 in the sense that
$$
\mathsf{var}(\widetilde{X}) = \mathbb{E}\left[\widetilde{X}^2\right] = \mathbb{E}\left[X^2\right] - \mathbb{E}\left[X^2 \mathbb{1}_{X \neq \widetilde{X}}\right] = \left(1 + O\left(\sqrt{\varepsilon/\sigma}\right)\right)\sigma^2,
$$
where the last relation holds due to the following observation
$$
\mathbb{E}\left[X^2 \mathbb{1}_{X \neq \widetilde{X}}\right] \overset{\text{(i)}}{\leq} \left(\mathbb{E}\left[X^4\right]\right)^{\frac{1}{2}} \left(\mathbb{P}\left(X \neq \widetilde{X}\right)\right)^{\frac{1}{2}} \overset{\text{(ii)}}{\lesssim} \sigma^2\sqrt{\frac{\varepsilon}{\sigma}}.
$$
Here, (i) invokes Cauchy-Schwarz, whereas (ii) is valid due to standard properties of sub-Gaussian random variables (Vershynin, 2018, Proposition 2.5.2).

5. By construction, we can see that for any $t \geq 0$,
$$
\mathbb{P}\left(\left|\widetilde{X}\right| \geq t\right) \leq \mathbb{P}\left(|X| \geq t\right) \leq 2\exp\left(-\frac{t^2}{\widetilde{c}\sigma^2}\right)
$$
holds for some absolute constant $\widetilde{c} > 0$, where the last relation follows from Vershynin (2018, Proposition 2.5.2). By invoking the definition of sub-Gaussian random variables (Vershynin, 2018, Definition 2.5.6) as well as standard properties of sub-Gaussian random variables (Vershynin, 2018, Proposition 2.5.2), we can conclude that $\widetilde{X}$ is sub-Gaussian obeying $\|\widetilde{X}\|_{\psi_2} \lesssim \sigma$.

By taking $\varepsilon = \delta c_{\mathsf{lb}}\sigma/4$ for any $\delta \in (0, 2)$, we establish the desired result.

G Extensions and additional discussions

In this section, we briefly discuss a couple of potential extensions of our algorithms and theory, and provide a few technical remarks about our settings and assumptions.

G.1 Relaxing the Gaussian design

In fact, all results in this paper can be generalized to the following setting:

- the samples are generated such that
$$\boldsymbol{x}_j = \boldsymbol{U}^\star (\boldsymbol{\Lambda}^\star)^{1/2} \boldsymbol{f}_j, \qquad 1 \le j \le n,$$
where $\boldsymbol{f}_1, \dots, \boldsymbol{f}_n$ are independent sub-Gaussian random vectors in $\mathbb{R}^r$ with independent entries satisfying
$$\mathbb{E}\left[f_{i,j}\right] = 0, \qquad \mathbb{E}\left[f_{i,j}^2\right] = 1, \qquad \mathbb{E}\left[f_{i,j}^4\right] = M_4, \qquad \text{and} \qquad \|f_{i,j}\|_{\psi_2} = O(1)$$
for any $1 \le i \le n$ and $1 \le j \le r$.

Our distributional theory for the principal subspace $\boldsymbol{U}^\star$ (cf. Theorem 1) continues to hold with

$$\begin{aligned}\boldsymbol{\Sigma}_{U,l}^\star &\coloneqq \left(\frac{1-p}{np}\left\|\boldsymbol{U}_{l,\cdot}^\star \boldsymbol{\Sigma}^\star\right\|_2^2 + \frac{\omega_l^{\star 2}}{np}\right)(\boldsymbol{\Sigma}^\star)^{-2} + (M_4 - 3)\frac{1-p}{np}\mathsf{diag}\left\{\boldsymbol{U}_{l,\cdot}^{\star\top}\boldsymbol{U}_{l,\cdot}^\star\right\} \\ &\quad + \frac{2(1-p)}{np}\boldsymbol{U}_{l,\cdot}^{\star\top}\boldsymbol{U}_{l,\cdot}^\star + (\boldsymbol{\Sigma}^\star)^{-2}\boldsymbol{U}^{\star\top}\mathsf{diag}\left\{\left[d_{l,i}^\star\right]_{1\le i\le d}\right\}\boldsymbol{U}^\star(\boldsymbol{\Sigma}^\star)^{-2} \end{aligned} \tag{G.1}$$

where

$$\begin{aligned} d_{l,i}^\star &\coloneqq \frac{1}{np^2}\left[\omega_l^{\star 2} + (1-p)\left\|\boldsymbol{U}_{l,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2^2\right]\left[\omega_i^{\star 2} + (1-p)\left\|\boldsymbol{U}_{i,\cdot}^\star\boldsymbol{\Sigma}^\star\right\|_2^2\right] \\ &\quad + \frac{2(1-p)^2}{np^2}S_{l,i}^{\star 2} + (M_4-3)\sum_{s=1}^r \sigma_s^{\star 4}U_{l,s}^{\star 2}U_{i,s}^{\star 2}. \end{aligned}$$

In addition, the distributional theory for the spiked covariance matrix $\boldsymbol{S}^\star$ (cf. Theorem 3) also holds with

$$\begin{aligned} v_{i,j}^\star &\coloneqq \frac{2-p}{np}S_{i,i}^\star S_{j,j}^\star + \frac{4-3p}{np}S_{i,j}^{\star 2} + \frac{1}{np}\left(\omega_i^{\star 2}S_{j,j}^\star + \omega_j^{\star 2}S_{i,i}^\star\right) + (M_4-3)\frac{2(1-p)}{np}\sum_{s=1}^r \sigma_s^{\star 4}U_{i,s}^{\star 2}U_{j,s}^{\star 2} \\ &\quad + \frac{1}{np^2}\sum_{k=1}^d\left\{\left[\omega_i^{\star 2} + (1-p)S_{i,i}^\star\right]\left[\omega_k^{\star 2} + (1-p)S_{k,k}^\star\right] + 2(1-p)^2 S_{i,k}^{\star 2}\right\}\left(\boldsymbol{U}_{k,\cdot}^\star\boldsymbol{U}_{j,\cdot}^{\star\top}\right)^2 \\ &\quad + \frac{1}{np^2}\sum_{k=1}^d\left\{\left[\omega_j^{\star 2} + (1-p)S_{j,j}^\star\right]\left[\omega_k^{\star 2} + (1-p)S_{k,k}^\star\right] + 2(1-p)^2 S_{j,k}^{\star 2}\right\}\left(\boldsymbol{U}_{k,\cdot}^\star\boldsymbol{U}_{i,\cdot}^{\star\top}\right)^2 \\ &\quad + (M_4-3)\frac{1}{np^2}\sum_{k=1}^d\sum_{s=1}^r \sigma_s^{\star 4}U_{k,s}^{\star 2}\left[U_{i,s}^{\star 2}\left(\boldsymbol{U}_{k,\cdot}^\star\boldsymbol{U}_{j,\cdot}^{\star\top}\right)^2 + U_{j,s}^{\star 2}\left(\boldsymbol{U}_{k,\cdot}^\star\boldsymbol{U}_{i,\cdot}^{\star\top}\right)^2\right] \end{aligned} \tag{G.2}$$

for $i \neq j$, and

$$\begin{aligned} v_{i,i}^\star &\coloneqq \frac{12-9p}{np}S_{i,i}^{\star 2} + \frac{4}{np}\omega_i^{\star 2}S_{i,i}^\star + (M_4-3)\frac{4(1-p)}{np}\sum_{s=1}^r \sigma_s^{\star 4}U_{i,s}^{\star 4} \\ &\quad + \frac{4}{np^2}\sum_{k=1}^d\left\{\left[\omega_i^{\star 2} + (1-p)S_{i,i}^\star\right]\left[\omega_k^{\star 2} + (1-p)S_{k,k}^\star\right] + 2(1-p)^2 S_{i,k}^{\star 2}\right\}\left(\boldsymbol{U}_{k,\cdot}^\star\boldsymbol{U}_{i,\cdot}^{\star\top}\right)^2 \\ &\quad + (M_4-3)\frac{4}{np^2}\sum_{k=1}^d\sum_{s=1}^r \sigma_s^{\star 4}U_{k,s}^{\star 2}U_{i,s}^{\star 2}\left(\boldsymbol{U}_{k,\cdot}^\star\boldsymbol{U}_{i,\cdot}^{\star\top}\right)^2. \end{aligned} \tag{G.3}$$

It is noteworthy that the fourth-moment M_4 plays an important role in the above variance calculation. One can then estimate $\boldsymbol{\Sigma}_{U,l}^{\star}$ (resp. $v_{i,j}^{\star}$) using the plug-in method similar to Algorithm 3 (resp. Algorithm 4), which naturally leads to fine-grained confidence regions for $\boldsymbol{U}^{\star}$ and entrywise confidence intervals for $\boldsymbol{S}^{\star}$. Note that the theoretical guarantees in this paper are proved without exploiting the Gaussian design (we only require basic sub-Gaussian properties and the information about the fourth moment), and hence the distributional theory and the validity of confidence regions/intervals under sub-Gaussian design can be established in an almost identical manner; for this reason, we omit the details for the sake of brevity.

G.2 The necessity of incoherence condition

Careful readers might wonder whether the incoherence condition (cf. Assumption 1) is necessary for our algorithm designs and theoretrial guarantees. As has been pointed out in the low-rank matrix estimation literature, the incoherence condition plays a crucial role in enabling reliable estimation in the presence of missing data (Candès and Tao, 2010; Chi et al., 2019; Keshavan et al., 2010). To make it more precise, we state below a theorem that captures the fundamental relation between the incoherence parameter μ and the information-theoretic sampling limit.

Theorem 19. *Consider $r = 1$ and any incoherence parameter $\mu \ll d$. Suppose that $n \geq d$, and $p < (1-\varepsilon)\sqrt{\frac{\mu}{nd}}$ for an arbitrarily small constant $0 < \varepsilon < 1/4$. Generate Ω according to the random sampling model. Then with probability at least 0.9, there exist $(\boldsymbol{u}_1^{\star}, \boldsymbol{f}_1)$ and $(\boldsymbol{u}_2^{\star}, \boldsymbol{f}_2)$ that satisfy, for both $i = 1, 2$,*

1. *$\boldsymbol{u}_i^{\star} \in \mathbb{R}^d$ is a fixed unit vector and is μ-incoherent: $\|\boldsymbol{u}_i^{\star}\|_{\infty} \leq \sqrt{\mu/d}$,*
2. *$\boldsymbol{f}_i \in \mathcal{N}(\boldsymbol{0}, \boldsymbol{I}_n)$ is an isotropic Gaussian random vector,*

such that $\min \|\boldsymbol{u}_1^{\star} \pm \boldsymbol{u}_2^{\star}\|_2 \asymp 1$ but $\mathcal{P}_{\Omega}(\boldsymbol{u}_1^{\star}\boldsymbol{f}_1^{\top}) = \mathcal{P}_{\Omega}(\boldsymbol{u}_2^{\star}\boldsymbol{f}_2^{\top})$.

This theorem shows that even in the simplest noiseless setting with $r = 1$: if $p < (1-\varepsilon)\sqrt{\frac{\mu}{nd}}$, then with high probability one cannot distinguish $\boldsymbol{u}_1^{\star}$ and $\boldsymbol{u}_2^{\star}$ given only observations $\mathcal{P}_{\Omega}(\boldsymbol{u}_1^{\star}\boldsymbol{f}_1^{\top})$. Consequently, in order to ensure identifiability, one needs to require the total sample size ndp^2 to exceed the order of μ. In other words, this theorem demonstrates how the incoherence parameter μ dictates the difficulty of the problem: when the eigenvectors are spiky, we require more samples to reliably estimate the principal subspace. Additionally, it is worth pointing out that the dependency w.r.t. μ in our theory is very likely to be suboptimal (e.g., Theorem 11 requires the sample size to exceed the order of μ^4). This might be improvable via more refined analysis, and we leave it to future investigation.

G.2.1 Proof of Theorem 19

Without loss of generality, assume that d/μ is an integer. Let

$$(\boldsymbol{u}_1^{\star})_i = \begin{cases} \sqrt{\frac{\mu}{d}}, & \text{if } 1 \leq i \leq \frac{d}{\mu}, \\ 0, & \text{otherwise} \end{cases} \qquad \text{and} \qquad \boldsymbol{f}_1 \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{I}_n).$$

Consider a random bipartite graph $\mathcal{G}$ generated by taking two disjoint vertex sets $\mathcal{U}$ and $\mathcal{V}$ with $|\mathcal{U}| = d/\mu$ and $|\mathcal{V}| = n$, and connecting each $u \in \mathcal{U}$ and $v \in \mathcal{V}$ independently with probability p. One can check that there is an equivalence between the edge set of $\mathcal{G}$ and a subset Ω' of the subsampled index set $\Omega \subseteq [d] \times [n]$ defined as

$$\Omega' \coloneqq \left\{(i,j) : 1 \leq i \leq \frac{d}{\mu}, (i,j) \in \Omega\right\}.$$

More precisely, if one connects the i-th vertex of $\mathcal{U}$ and the j-th vertex of $\mathcal{V}$ if and only if $(i,j) \in \Omega'$, then the resulting random graph has the same distribution as $\mathcal{G}$. In view of Johansson (2012, Theorem 6), when $p < (1-\varepsilon)\sqrt{\frac{\mu}{nd}}$, for any $k > 0$ the probability that there is no connected component in $\mathcal{G}$ with at least k vertices in $\mathcal{U}$ is at least

$$1 - \frac{d}{\mu(1-\varepsilon)} \exp\left(-\frac{1}{6}k\varepsilon^3\right).$$

By taking $k = 12\varepsilon^{-3}\log(d/\mu)$ and recall that $d \gg \mu$ and $\varepsilon < 1/4$, we can show that that with probability exceeding 0.9, there exists no connected component in $\mathcal{G}$ with at least $12\varepsilon^{-3}\log(d/\mu)$ vertices in $\mathcal{U}$. Denote by $\mathcal{C}_1, \dots, \mathcal{C}_K$ the collection of connected components in $\mathcal{G}$, and let $\mathcal{U}_i$ (resp. $\mathcal{V}_i$) be the set of vertices in $\mathcal{U}$ (resp. $\mathcal{V}$) that reside in $\mathcal{C}_i$. When $d \gg \mu$, we can always find a subset of $\mathcal{I} \subset [K]$ such that

$$\frac{1}{2}\left(\frac{d}{\mu} - 12\varepsilon^{-3}\log\frac{d}{\mu}\right) \le \sum_{i\in\mathcal{I}} |\mathcal{U}_j| \le \frac{1}{2}\left(\frac{d}{\mu} + 12\varepsilon^{-3}\log\frac{d}{\mu}\right).$$

Then we can proceed to construct $\boldsymbol{u}_2^\star$ and $\boldsymbol{f}_2$ in the following manner:

$$(\boldsymbol{u}_2^\star)_i = \begin{cases} (\boldsymbol{u}_1^\star)_i\,, & \text{if } 1 \le i \le \frac{d}{\mu} \text{ and } i \in \mathcal{U}_k \text{ for some } k \in \mathcal{I} \\ -(\boldsymbol{u}_2^\star)_i\,, & \text{if } 1 \le i \le \frac{d}{\mu} \text{ and } i \in \mathcal{U}_k \text{ for some } k \notin \mathcal{I} \\ 0 & \text{otherwise} \end{cases}$$

for any $1 \le i \le d$, and

$$(\boldsymbol{f}_2)_j = \begin{cases} (\boldsymbol{f}_1)_j\,, & \text{if } j \in \mathcal{V}_k \text{ for some } k \in \mathcal{I} \\ -(\boldsymbol{f}_1)_j\,, & \text{if } j \in \mathcal{V}_k \text{ for some } k \notin \mathcal{I} \end{cases}$$

for any $1 \le j \le n$. It is straightforward to check that when $d \gg \mu$, both $\boldsymbol{u}_1^\star$ and $\boldsymbol{u}_2^\star$ are μ-incoherent obeying $\min \|\boldsymbol{u}_1^\star \pm \boldsymbol{u}_2^\star\|_2 \asymp 1$, and $\boldsymbol{f}_2 \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{I}_n)$. This completes the proof.

G.3 Other observational models

This subsection discusss the possibility of extending the algorithms and theory in the current paper to accommodate another model stated below, which finds applications in medical research and wireless communication. Consider the setting where, instead of missing all entries outside the sampling set Ω, we observe pure noise without knowing Ω in advance; that is, we observe

$$y_{l,j} = \begin{cases} x_{l,j} + \eta_{l,j}, & \text{for all } (l,j) \in \Omega, \\ \eta_{l,j}, & \text{otherwise,} \end{cases}$$

or equivalently, $\boldsymbol{Y} = \mathcal{P}_\Omega(\boldsymbol{X}) + \boldsymbol{N}$ in the matrix form. Similar to (2.2) we can compute

$$\frac{1}{p^2}\mathbb{E}\left[\boldsymbol{Y}\boldsymbol{Y}^\top \mid \boldsymbol{X}\right] = \boldsymbol{X}\boldsymbol{X}^\top + \left(\frac{1}{p} - 1\right)\mathcal{P}_{\mathsf{diag}}\left(\boldsymbol{X}\boldsymbol{X}^\top\right) + \frac{n}{p^2}\mathsf{diag}\left\{\left[\omega_l^{\star 2}\right]_{1\le l\le d}\right\}.$$

Therefore, it is reasonable to employ $\mathsf{HeteroPCA}$ to attempt estimation of the principal subspace $\boldsymbol{U}^\star$ and the spiked covariance matrix $\boldsymbol{S}^\star$ under this setting, although we might need to impose stronger noise conditions due to the presence of more noise.

Note, however, that it would be difficult to perform valid statistical inference (as we did in Algorithm 3 and 4) for this setting; the reason is that it becomes fairly difficult to estimate the sampling rate p, a parameter that is required for computing an estimate for the covariance matrix $\boldsymbol{\Sigma}_{U,l}^\star$ and the variance $v_{i,j}^\star$. To further elucidate this point, consider the setting where the noise components are i.i.d. $\mathcal{N}(0,\sigma^2)$ with known $\sigma > 0$. Then this observation model can be described as follows: for each (i,j),

$$y_{l,j} \sim \begin{cases} \mathcal{N}\left(x_{l,j}, \sigma^2\right), & \text{with probability } p, & \text{(Case 1)} \\ \mathcal{N}\left(0, \sigma^2\right), & \text{with probability } 1-p. & \text{(Case 2)} \end{cases}$$

The task of estimating p boils down to estimating the portion of the entries corresponding to Case 1, which, however, is impossible in general. For instance, suppose that half of the rows of $\boldsymbol{U}^\star$ are zero. When $\boldsymbol{U}_{l,\cdot}^\star = \boldsymbol{0}$, we know that $x_{l,j} = 0$ for all $j \in [n]$. Therefore for roughly half of the entries, one cannot distinguish whether they belong to Case 1 or Case 2. This means that it is in general impossible to estimate p in an accurate manner.

Nevertheless, when stronger noise conditions are met, we shall be able to estimate the sampling rate p reliably using the following scheme. Employing similar analysis as in Appendix E.3.2, we can show that the sum of squares of all observations $y_{l,j}$ concentrates around

$$\sum_{l=1}^{d}\sum_{j=1}^{n} y_{l,j}^2 \approx np\left\|\boldsymbol{U}^\star\boldsymbol{\Sigma}^\star\right\|_{\mathrm{F}}^2 + n\sum_{l=1}^{d}\omega_l^{\star 2}$$

with high probability. When the noise level $\omega_{\max}$ is sufficiently small, namely $\omega_{\max} \ll \sqrt{p/d}\,\sigma_r^\star$, the second term on the right hand side of the above equation becomes negligible, and hence we can faithfully estimate p by means of the following data-driven plug-in estimate:

$$\widehat{p} = \frac{1}{n\|\boldsymbol{U}\boldsymbol{\Sigma}\|_{\mathrm{F}}^2}\sum_{l=1}^{d}\sum_{j=1}^{n} y_{l,j}^2.$$

This in turn allows us to perform valid statistical inference for this setting, through the same analysis framework introduced in this paper. For example, we can still apply Theorem 5 to establish the first- and second-order approximations (3.13) and (3.14) with a different effective noise matrix $\boldsymbol{E} = p^{-1}\mathcal{P}_\Omega(\boldsymbol{X}) - \boldsymbol{X} + p^{-1}\boldsymbol{N}$, then the distributional characterization and data-driven construction of confidence regions/intervals follow naturally from the same analysis. Given that this model is not the focus of the current paper, we omit the details for the sake of brevity.

G.4 Heteroskedastic noise across rows

In this paper, we allow the noise levels to vary across different rows, but require the noise variance to be identical within each row. Naturally, one might ask whether it is feasible to further relax such assumptions and accommodate fully heterogeneous noise across all entries. Unfortunately, while it is plausible to establish similar estimation guarantees and distributional theory (characterizing the distribution of estimation errors via model parameters) for this extension, the design of data-driven inferential procedures — such as constructing confidence regions for $\boldsymbol{U}^\star$ or entrywise confidence intervals for $\boldsymbol{S}^\star$ — would be substantially more challenging. The rationale is as follows.

- We first notice that, the setting of Theorem 5 (i.e., our general second-order subspace perturbation theory for HeteroPCA) allows heteroskedastic random noise across entries (see Assumption 3). This suggests that in the PCA context, even if we assume heteroskedastic noise across entries, we can still invoke Theorem 5 to obtain our key error decomposition (3.8), which in turn allows us to establish estimation guarantees (cf. Lemmas 18 and 19) and distributional theory (cf. Theorems 1 and 3).
- When it comes to statistical inference, however, we need to reliably estimate the covariance matrix $\boldsymbol{\Sigma}_{U,l}^\star$ and $v_{i,j}^\star$ in a data-driven manner. The key difficulty is that, under our assumptions, we cannot hope to provide an entrywise consistent estimate for the ground truth data matrix $\boldsymbol{X} \in \mathbb{R}^{d\times n}$. We can only estimate the underlying spiked covariance matrix $\boldsymbol{S}^\star \in \mathbb{R}^{d\times d}$ and its associated eigenspace $\boldsymbol{U}^\star$ and eigenvalues $\boldsymbol{\Lambda}^\star$ reliably. If the noise is heteroskedastic across entries, say, $\mathsf{var}(\eta_{l,j}) = \omega_{l,j}^{\star 2}$, then it is in general impossible to faithfully estimate each noise variance $\omega_{l,j}^{\star 2}$ or their (nonlinear) functions. This makes it hard to construct confidence regions for $\boldsymbol{U}^\star$ or entrywise confidence intervals for $\boldsymbol{S}^\star$ in a data-driven manner (as done in Algorithms 3 and 4).

It is worth noting that several previous studies on PCA (Zhu et al., 2022) and random matrix theory (Alt et al., 2021; Latała et al., 2018) permit heterogeneous noise across entries. In the realm of statistical inference, however, the studies of heteroskedastic noise remained highly limited. For example, Bao et al. (2021); Koltchinskii et al. (2020); Xia (2021) requires the noise components to be either i.i.d. Gaussian or have identical first fourth moments. To the best of our knowledge, tolerating noise that is heterogeneous across rows, as we have in this current study, represents the mildest assumption within the framework of statistical inference for PCA.

References

Abbe, E., Fan, J., Wang, K., Zhong, Y., et al. (2020). Entrywise eigenvector analysis of random matrices with low expected rank. *Annals of Statistics*, 48(3):1452–1474.

Alt, J., Erdős, L., and Krüger, T. (2021). Spectral radius of random matrices with independent entries. *Probability and Mathematical Physics*, 2(2):221–280.

Bao, Z., Ding, X., and Wang, K. (2021). Singular vector and singular subspace distribution for the matrix denoising model. *The Annals of Statistics*, 49(1):370–392.

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. *The Annals of Statistics*, 49(2):944–967.

Candès, E. and Tao, T. (2010). The power of convex relaxation: Near-optimal matrix completion. *IEEE Transactions on Information Theory*, 56(5):2053 –2080.

Chen, J., Liu, D., and Li, X. (2020). Nonconvex rectangular matrix completion via gradient descent without $\ell_{2,\infty}$ regularization. *IEEE Transactions on Information Theory*.

Chen, L. H., Goldstein, L., and Shao, Q.-M. (2010). *Normal approximation by Stein's method*. Springer Science & Business Media.

Chen, Y., Chi, Y., Fan, J., and Ma, C. (2019a). Gradient descent with random initialization: Fast global convergence for nonconvex phase retrieval. *Mathematical Programming*, 176(1-2):5–37.

Chen, Y., Chi, Y., Fan, J., Ma, C., et al. (2021). Spectral methods for data science: A statistical perspective. *Foundations and Trends® in Machine Learning*, 14(5):566–806.

Chen, Y., Fan, J., Ma, C., and Wang, K. (2019b). Spectral method and regularized MLE are both optimal for top-K ranking. *Annals of Statistics*, 47(4):2204–2235.

Chi, Y., Lu, Y. M., and Chen, Y. (2019). Nonconvex optimization meets low-rank matrix factorization: An overview. *IEEE Transactions on Signal Processing*, 67(20):5239 – 5269.

Devroye, L., Mehrabian, A., and Reddad, T. (2018). The total variation distance between high-dimensional gaussians. *arXiv preprint arXiv:1810.08693*.

Johansson, T. (2012). The giant component of the random bipartite graph.

Keshavan, R. H., Montanari, A., and Oh, S. (2010). Matrix completion from noisy entries. *J. Mach. Learn. Res.*, 11:2057–2078.

Koltchinskii, V., Löffler, M., and Nickl, R. (2020). Efficient estimation of linear functionals of principal components. *The Annals of Statistics*, 48(1):464–490.

Latała, R., van Handel, R., and Youssef, P. (2018). The dimension-free structure of nonhomogeneous random matrices. *Inventiones mathematicae*, 214:1031–1080.

Raič, M. (2019). A multivariate Berry-Esseen theorem with explicit constants. *Bernoulli*, 25(4A):2824–2853.

Schmitt, B. A. (1992). Perturbation bounds for matrix square roots and Pythagorean sums. *Linear Algebra Appl.*, 174:215–227.

Tropp, J. A. (2015). An introduction to matrix concentration inequalities. *Found. Trends Mach. Learn.*, 8(1-2):1–230.

Tsybakov, A. B. and Zaiats, V. (2009). *Introduction to nonparametric estimation*, volume 11. Springer.

Vershynin, R. (2012). Introduction to the non-asymptotic analysis of random matrices. *Compressed Sensing, Theory and Applications*, pages 210 – 268.

Vershynin, R. (2018). *High-dimensional probability: An introduction with applications in data science*, volume 47. Cambridge university press.

Wainwright, M. (2019). *High-Dimensional Statistics: A Non-Asymptotic Viewpoint*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press.

Xia, D. (2021). Normal approximation and confidence region of singular subspaces. *Electronic Journal of Statistics*, 15(2):3798–3851.

Zhang, A. R., Cai, T. T., and Wu, Y. (2022). Heteroskedastic pca: Algorithm, optimality, and applications. *The Annals of Statistics*, 50(1):53–80.

Zhu, Z., Wang, T., and Samworth, R. J. (2022). High-dimensional principal component analysis with heterogeneous missingness. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 84(5):2000–2031.