STATISTICAL INFERENCE FOR ULTRAHIGH DIMENSIONAL LOCATION PARAMETER BASED ON SPATIAL MEDIAN

Guanghui Cheng, Liuhua Peng and Changliang Zou

Guangzhou University, The University of Melbourne, Nankai University

Supplementary Materials

The supplementary materials consist of preliminary lemmas, proofs of main results in the paper, and additional simulation results.

Appendix A: Preliminary lemmas

We first introduce and recall some notation. For a $d_1 \times d_2$ matrix $M = (m_{j\ell})_{d_1\times d_2}$, its matrix ϱ-norm is $\|M\|_\varrho = \sup\{\|Mx\|_\varrho : \|x\|_\varrho = 1\}$. Specifically, the 1-, 2-, and ∞-norms of M are $\|M\|_1 = \max_{1\leqslant \ell\leqslant d_2}\sum_{j=1}^{d_1}|m_{j\ell}|$, $\|M\|_2 = \{\lambda_{\max}(M^\top M)\}^{1/2}$, and $\|M\|_\infty = \max_{1\leqslant j\leqslant d_1}\sum_{\ell=1}^{d_2}|m_{j\ell}|$. The Frobenius norm of M is $\|M\|_F = \{\sum_{j=1}^{d_1}\sum_{\ell=1}^{d_2} m_{j\ell}^2\}^{1/2}$.

Define a random $p\times p$ matrix $Q = n^{-1}\sum_{i=1}^n R_i^{-1}W_iW_i^\top$ such that $\mathbb{E}(Q) = \mathbb{E}(R_i^{-1}W_iW_i^\top)$, and denote $Q_{j\ell}$ as the (j,ℓ)th element of Q. Denote $\mathbb{E}^*(\cdot)$ and $\mathrm{Var}^*(\cdot)$ be the expectation and variance conditional on $X_1,\ldots,X_n$, respectively. Recall that $W_{i,j}$ is the jth element of W_i for $i=1,\ldots,n$ and $j=1,\ldots,p$; $\omega_{j\ell}$ is the (j,ℓ)th element of Ω; and Γ_j is the jth row of Γ. Finally, we will denote various positive absolute constants by $C_1, C_2, C_3, \ldots$ without mentioning this explicitly.

In this section, we present several preliminary lemmas, whose proof can be found in Appendix C of the supplementary material.

Lemma A1. *(Concentration of norms) Suppose that Conditions C1 and C3 hold with $a_0(p)\asymp p^{1-\delta}$ for some positive constant $\delta\leqslant 1/2$. Then, for sufficient large p, there exist positive constants c_1 and c_2 such that*

$$\mathbb{P}\left\{p - \epsilon p^{(1+\delta)/2} \leqslant \|U_1\|^2 \leqslant p + \epsilon p^{(1+\delta)/2}\right\} \geqslant 1 - c_1\exp\{-c_2 p^{\delta\alpha/(4\alpha+4)}\}$$

and

$$\mathbb{P}\left\{(1-\epsilon)\mathrm{tr}(\Omega) \leqslant \|\Gamma U_1\|^2 \leqslant (1+\epsilon)\mathrm{tr}(\Omega)\right\} \geqslant 1 - c_1 \exp\{-c_2 p^{\delta\alpha/(4\alpha+4)}\}$$

for any fixed $\epsilon \in (0,1)$.

Lemma A2. *Suppose that Conditions C1, C2 and C3 hold with* $a_0(p) \asymp p^{1-\delta}$ *for some positive constant* $\delta \leqslant 1/2$. *Then, for any* $i = 1, \ldots, n$,

(i) $\mathbb{E}(\|U_i\|^4) = p\mathbb{E}(U_{i,j}^4) + p(p-1)$, $\mathbb{E}(\|U_i\|^6) = p\mathbb{E}(U_{i,j}^6) + 3p(p-1)\mathbb{E}(U_{i,j}^4) + p(p-1)(p-2)$ *and*

$$\begin{aligned}\mathbb{E}(\|U_i\|^8) \quad = \quad & p\mathbb{E}(U_{i,j}^8) + 4p(p-1)\mathbb{E}(U_{i,j_1}^6) + 3p(p-1)\{\mathbb{E}(U_{i,j_1}^4)\}^2 \\ & + 3p(p-1)\mathbb{E}(U_{i,j}^4) + p(p-1)(p-2)(p-3)\,.\end{aligned}$$

In addition, $\mathbb{E}(\|U_i\|^{2k}) = p^k + O(p^{k-1})$, $\mathbb{E}(\|U\|^k) = p^{k/2} + O(p^{k/2-1})$ *for any positive integer* k.

(ii) $\mathbb{E}(\|\Gamma U_i\|^4) = p^2 + O(p^{2-\delta})$, $\mathbb{E}(\|\Gamma U_i\|^6) = p^3 + O(p^{3-\delta})$. *In addition,* $\mathbb{E}(\|\Gamma U_i\|) = p^{1/2} + O(p^{1/2-\delta})$ *and* $\mathbb{E}(\|\Gamma U_i\|^3) = p^{3/2} + O(p^{3/2-\delta})$.

(iii) $\mathbb{E}\{\|\Gamma S(U_i)\|^2\} = 1 + O(p^{-1/2})$ *and* $\mathbb{E}\{\|\Gamma S(U_i)\|^4\} = 1 + O(p^{-1/3})$.

(iv) $\mathbb{E}(\nu_i^{-k}) \lesssim \zeta_k p^{k/2}$ *for* $k = 1, 2, 3$.

Lemma A3. *Suppose Conditions C1, C2 and C3 with* $a_0(p) \asymp p^{1-\delta}$ *for some positive constant* $\delta \leqslant 1/2$ *hold. Define a random* $p \times p$ *matrix* $Q = n^{-1}\sum_{i=1}^n R_i^{-1} W_i W_i^\top$ *and let* $Q_{j\ell}$ *be the* (j,ℓ)*th element of* Q. *Then,*

(i) $|Q_{j\ell}| \lesssim \zeta_1 p^{-1}|\omega_{j\ell}| + O_p(\zeta_1 n^{-1/2}p^{-1} + \zeta_1 p^{-7/6} + \zeta_1 p^{-1-\delta/2})$.

(ii) $Q_{j\ell} = Q_{0,j\ell} + O_p(\zeta_1 p^{-7/6} + \zeta_1 p^{-1-\delta/2})$, *where* $Q_{0,j\ell}$ *is the* (j,ℓ)*th element of*

$$Q_0 = n^{-1}p^{-1/2}\sum_{i=1}^n \nu_i^{-1}\{\Gamma S(U_i)\}\{\Gamma S(U_i)\}^\top.$$

In addition, Q_0 *satisfies* $\mathrm{tr}[\mathbb{E}(Q_0^2) - \{\mathbb{E}(Q_0)\}^2] = O(n^{-1}p^{-1})$.

Lemma A4. *Suppose Conditions C1, C2 and C3 with* $a_0(p) \asymp p^{1-\delta}$ *for some positive constant* $\delta \leqslant 1/2$ *hold. Then,*

(i) $\mathbb{E}\{(\zeta_1^{-1}W_{i,j})^4\} \lesssim \bar{M}^2$ *and* $\mathbb{E}\{(\zeta_1^{-1}W_{i,j})^2\} \gtrsim \underline{m}$ *for all* $i = 1, \ldots, n$ *and* $j = 1, \ldots, p$.

(ii) $\|\zeta_1^{-1}W_{i,j}\|_{\psi_\alpha} \lesssim \bar{B}$ *for all* $i = 1, \ldots, n$ *and* $j = 1, \ldots, p$.

(iii) $\mathbb{E}(W_{i,j}^2) = p^{-1}\omega_{jj} + O(p^{-1-\delta/2})$ *for* $j = 1, \ldots, p$ *and* $\mathbb{E}(W_{i,j}^2) = p^{-1}\omega_{j\ell} + O(p^{-1-\delta/2})$ *for* $1 \leqslant j \neq \ell \leqslant p$.

(iv) if $\log p = o(n^{1/3})$, $\left|n^{-1/2}\sum_{i=1}^{n}\zeta_1^{-1}W_i\right|_\infty = O_p\{\log^{1/2}(np)\}$ and $\left|n^{-1}\sum_{i=1}^{n}(\zeta_1^{-1}W_i)^2\right|_\infty = O_p(1)$.

Lemma A5. *Suppose the conditions of Theorem 2 hold, then*

$$n^{1/2}\tilde{\boldsymbol{\theta}}_n = n^{-1/2}\zeta_1^{-1}\sum_{i=1}^{n} Z_i W_i + \tilde{C}_n\,, \tag{S.1}$$

where $|\tilde{C}_n|_\infty = O_p\{n^{-1/4}\log^{1/2}(np) + p^{-(1/6\wedge\delta/2)}\log^{1/2}(np)\}$.

The following lemma is Nazarov's inequality, and its proof can be found in Chernozhukov et al. (2017).

Lemma A6 (Nazarov's inequality)**.** *Let* $Y_0 = (Y_{0,1}, \ldots, Y_{0,p})^\top$ *be a centered Gaussian random vector in* $\mathbb{R}^p$ *and* $\mathbb{E}(Y_{0,j}^2) \geqslant b$ *for all* $j = 1, \ldots, p$ *and some constant* $b > 0$*, then for every* $y \in \mathbb{R}^p$ *and* $a > 0$, $\mathbb{P}(Y_0 \leqslant y + a) - \mathbb{P}(Y_0 \leqslant y) \lesssim a\log^{1/2}(p)$.

Appendix B: Proof of main results

Proof of Lemma 1. As $\boldsymbol{\theta}$ is a location parameter, we assume $\boldsymbol{\theta} = 0$ without loss of generality. Then, $W_i = S(X_i) = \|X_i\|^{-1}X_i = \|\Gamma U_i\|^{-1}\Gamma U_i$ for $i = 1, \ldots, n$. The sample spatial median $\hat{\boldsymbol{\theta}}_n$ satisfies

$$\sum_{i=1}^{n} S(X_i - \hat{\boldsymbol{\theta}}_n) = \sum_{i=1}^{n}\frac{X_i - \hat{\boldsymbol{\theta}}_n}{\|X_i - \hat{\boldsymbol{\theta}}_n\|} = \sum_{i=1}^{n}\frac{W_i - R_i^{-1}\hat{\boldsymbol{\theta}}_n}{\|W_i - R_i^{-1}\hat{\boldsymbol{\theta}}_n\|} = 0\,,$$

which is is equivalent to

$$n^{-1}\sum_{i=1}^{n}(W_i - R_i^{-1}\hat{\boldsymbol{\theta}}_n)(1 - 2R_i^{-1}W_i^\top\hat{\boldsymbol{\theta}}_n + R_i^{-2}\|\hat{\boldsymbol{\theta}}_n\|^2)^{-1/2} = 0$$

as $W_i^\top W_i = 1$.

Under Condition C2, $\zeta_k = \mathbb{E}(R_i^{-k}) = O(p^{-k/2})$ for $k = 1, 2, 3, 4$. In addition, Lemma A3 indicates that $Q_{j\ell} = Q_{0,j\ell} + O_p(\zeta_1 p^{-7/6} + \zeta_1 p^{-1-\delta/2})$, where $Q_{0,j\ell}$ is the (j,ℓ)th element of $Q_0 = n^{-1}p^{-1/2}\sum_{i=1}^{n}\nu_i\{\Gamma S(U_i)\}\{\Gamma S(U_i)\}^\top$. In addition, Q_0 satisfies $\mathrm{tr}[\mathbb{E}(Q_0^2) - \{\mathbb{E}(Q_0)\}^2] = O(n^{-1}p^{-1})$. Thus, from the similar procedure as in the proof of Lemma

1.2 of Cheng et al. (2019), we can show that

$$\|\hat{\boldsymbol{\theta}}_n\| = O_p(\zeta_1^{-1} n^{-1/2}).$$

Then, for $i = 1, \ldots, n$, we have $|R_i^{-1} W_i^T \hat{\boldsymbol{\theta}}_n| \leqslant R_i^{-1} \|\hat{\boldsymbol{\theta}}_n\| = O_p(n^{-1/2})$ and $R_i^{-2} ||\hat{\boldsymbol{\theta}}_n||^2 = O_p(n^{-1})$. By the first-order Taylor expansion, the above equation can be rewritten as

$$n^{-1} \sum_{i=1}^{n} (W_i - R_i^{-1} \hat{\boldsymbol{\theta}}_n)(1 + R_i^{-1} W_i^\top \hat{\boldsymbol{\theta}}_n - 2^{-1} R_i^{-2} \|\hat{\boldsymbol{\theta}}_n\|^2 + \delta_{1i}) = 0\,, \tag{S.2}$$

where $\delta_{1i} = O_p\{(R_i^{-1} W_i^\top \hat{\boldsymbol{\theta}}_n - 2^{-1} R_i^{-2} \|\hat{\boldsymbol{\theta}}_n\|^2)^2\} = O_p(n^{-1})$. By Markov's inequality, for any $\varepsilon > 0$,

$$\begin{aligned} \mathbb{P}\left(\max_{1 \leqslant i \leqslant n} R_i^{-1} \geqslant \varepsilon \zeta_1 n^{1/4} \right) &= \mathbb{P}\left(\max_{1 \leqslant i \leqslant n} R_i^{-4} \geqslant \varepsilon^4 \zeta_1^4 n \right) \\ &\leqslant \mathbb{E}\left(\max_{1 \leqslant i \leqslant n} R_i^{-4} \right) / (\varepsilon^4 \zeta_1^4 n) \leqslant n \mathbb{E}(R_i^{-4}) / (\varepsilon^4 \zeta_1^4 n) \lesssim \varepsilon^{-4}\,, \end{aligned}$$

where the last inequality is due to Condition C2. Thus, $\max_{1 \leqslant i \leqslant n} R_i^{-2} = O_p(\zeta_1^2 n^{1/2})$, and consequently, $\max_{1 \leqslant i \leqslant n} \delta_{1i} = O_p(\|\hat{\boldsymbol{\theta}}_n\|^2 \max_{1 \leqslant i \leqslant n} R_i^{-2}) = O_p(n^{-1/2})$. Rewrite (S.2) as

$$\begin{aligned} & n^{-1} \sum_{i=1}^{n} (1 - 2^{-1} R_i^{-2} \|\hat{\boldsymbol{\theta}}_n\|^2 + \delta_{1i}) W_i + n^{-1} \sum_{i=1}^{n} R_i^{-1} (W_i^\top \hat{\boldsymbol{\theta}}_n) W_i \\ = \; & n^{-1} \sum_{i=1}^{n} R_i^{-1} (1 - 2^{-1} R_i^{-2} \|\hat{\boldsymbol{\theta}}_n\|^2 + \delta_{1i}) \hat{\boldsymbol{\theta}}_n + n^{-1} \sum_{i=1}^{n} R_i^{-2} (W_i^\top \hat{\boldsymbol{\theta}}_n) \hat{\boldsymbol{\theta}}_n\,, \end{aligned}$$

which implies

$$n^{-1} \sum_{i=1}^{n} (1 - 2^{-1} R_i^{-2} \|\hat{\boldsymbol{\theta}}_n\|^2 + \delta_{1i}) W_i + n^{-1} \sum_{i=1}^{n} R_i^{-1} (W_i^\top \hat{\boldsymbol{\theta}}_n) W_i = n^{-1} \sum_{i=1}^{n} R_i^{-1} (1 + \delta_{1i} + \delta_{2i}) \hat{\boldsymbol{\theta}}_n\,, \tag{S.3}$$

where $\delta_{2i} = R_i^{-1} W_i^\top \hat{\boldsymbol{\theta}}_n - 2^{-1} R_i^{-2} \|\hat{\boldsymbol{\theta}}_n\|^2 = O_p(\delta_{1i}^{1/2})$ satisfies $\max_{1 \leqslant i \leqslant n} \delta_{2i} = O_p(n^{-1/4})$. It is straightforward to check that $n^{-1} \sum_{i=1}^{n} R_i^{-1} (W_i^\top \hat{\boldsymbol{\theta}}_n) W_i = n^{-1} \sum_{i=1}^{n} R_i^{-1} W_i W_i^\top \hat{\boldsymbol{\theta}}_n = Q \hat{\boldsymbol{\theta}}_n$. From Lemma A3, $|Q_{j\ell}| \lesssim \zeta_1 p^{-1} |\omega_{j\ell}| + O_p(\zeta_1 n^{-1/2} p^{-1} + \zeta_1 p^{-7/6} + \zeta_1 p^{-1-\delta/2})$, and this implies that $|Q \hat{\boldsymbol{\theta}}_n|_\infty \leqslant \|Q\|_1 |\hat{\boldsymbol{\theta}}_n|_\infty \lesssim \zeta_1 p^{-1} \|\Omega\|_1 |\hat{\boldsymbol{\theta}}_n|_\infty + O_p(\zeta_1 n^{-1/2} +$

$\zeta_1 p^{-1/6} + \zeta_1 p^{-\delta/2})|\hat{\boldsymbol{\theta}}_n|_\infty$.

According to Lemma A4, we have that $|n^{-1}\sum_{i=1}^n \zeta_1^{-1}W_i|_\infty = O_p\{n^{-1/2}\log^{1/2}(np)\}$. Then, $\left|\zeta_1^{-1}n^{-1}\sum_{i=1}^n \delta_{1i}W_i\right|_\infty^2 \leqslant \left|n^{-1}\sum_{i=1}^n (\zeta_1^{-1}W_i)^2\right|_\infty \left(n^{-1}\sum_{i=1}^n \delta_{1i}^2\right) \lesssim O_p(n^{-2})$. In addition, we have that $|\zeta_1^{-1}n^{-1}\sum_{i=1}^n R_i^{-2}\|\hat{\boldsymbol{\theta}}_n\|^2 W_i|_\infty \lesssim O_p(n^{-1})$. Regarding equation (S.3) and the fact that $\zeta_1^{-1}n^{-1}\sum_{i=1}^n R_i^{-1} = 1 + O_p(n^{-1/2})$, we obtain

$$
\begin{aligned}
|\hat{\boldsymbol{\theta}}_n|_\infty &\lesssim \left|\zeta_1^{-1}n^{-1}\sum_{i=1}^n W_i\right|_\infty + \zeta_1^{-1}|Q\hat{\boldsymbol{\theta}}_n|_\infty \\
&\lesssim p^{-1}a_0(p)|\hat{\boldsymbol{\theta}}_n|_\infty + O_p(n^{-1/2} + p^{-(1/6\wedge\delta/2)})|\hat{\boldsymbol{\theta}}_n|_\infty + O_p\{n^{-1/2}\log^{1/2}(np)\}\,.
\end{aligned}
$$

Thus, we conclude that

$$
|\hat{\boldsymbol{\theta}}_n|_\infty = O_p\{n^{-1/2}\log^{1/2}(np)\}
$$

as $a_0(p) \asymp p^{1-\delta}$. In addition, we have $|\zeta_1^{-1}Q\hat{\boldsymbol{\theta}}_n|_\infty = O_p\{n^{-1/2}p^{-(1/6\wedge\delta/2)}\log^{1/2}(np) + n^{-1}\log^{1/2}(np)\}$ and $n^{-1}\sum_{i=1}^n R_i^{-1}(1 + \delta_{1i} + \delta_{2i}) = \zeta_1\{1 + O_p(n^{-1/4})\}$. Finally, we can write

$$
n^{1/2}(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}) = n^{-1/2}\zeta_1^{-1}\sum_{i=1}^n W_i + C_n\,,
$$

where C_n satisfies $|C_n|_\infty = O_p\{n^{-1/4}\log^{1/2}(np) + p^{-(1/6\wedge\delta/2)}\log^{1/2}(np)\}$. □

Proof of Theorem 1. Let $L_{n,p} = n^{-1/4}\log^{1/2}(np) + p^{-(1/6\wedge\delta/2)}\log^{1/2}(np)$. Then, for any sequence $\eta_n \to \infty$ and any $t \in \mathbb{R}^p$,

$$
\begin{aligned}
&\mathbb{P}\left\{n^{1/2}(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}) \leqslant t\right\} = \mathbb{P}\left(n^{-1/2}\zeta_1^{-1}\sum_{i=1}^n W_i + C_n \leqslant t\right) \\
\leqslant\ &\mathbb{P}\left(n^{-1/2}\zeta_1^{-1}\sum_{i=1}^n W_i \leqslant t + \eta_n L_{n,p}\right) + \mathbb{P}(|C_n|_\infty > \eta_n L_{n,p})\,.
\end{aligned}
$$

According to Lemma A4, $\mathbb{E}\{(\zeta_1^{-1}W_{i,j})^4\} \lesssim \overline{M}^2$, $\mathbb{E}\{(\zeta_1^{-1}W_{i,j})^2\} \gtrsim \underline{m}$, and $\|\zeta_1^{-1}W_{i,j}\|_{\psi_\alpha} \lesssim \bar{B}$ for all $i = 1,\ldots,n$ and $j = 1,\ldots,p$. According to the Gaussian approximation for independent partial sums in Koike (2021), let

$G \sim N(0, \zeta_1^{-2}\mathbb{B})$ with $\mathbb{B} = \mathbb{E}(W_1 W_1^\top)$, we have

$$\begin{aligned} & \mathbb{P}\left(n^{-1/2}\zeta_1^{-1}\sum_{i=1}^{n} W_i \leqslant t + \eta_n L_{n,p}\right) \\ \leqslant\ & \mathbb{P}(G \leqslant t + \eta_n L_{np}) + O\left(\left\{n^{-1}\log^5(np)\right\}^{1/6}\right) \\ \leqslant\ & \mathbb{P}(G \leqslant t) + O\{\eta_n L_{np}\log^{1/2}(p)\} + O\left(\left\{n^{-1}\log^5(np)\right\}^{1/6}\right), \end{aligned}$$

where the last inequality is from Nazarov's inequality in Lemma A6. It is also worth noting that the order $O\left(\left\{n^{-1}\log^5(np)\right\}^{1/6}\right)$ is improved to $O\left(\left\{n^{-1}\log^5(np)\right\}^{1/4}\right)$ in Chernozhukov et al. (2019). Thus,

$$\begin{aligned} \mathbb{P}\left\{n^{1/2}(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}) \leqslant t\right\} \leqslant\ & \mathbb{P}(G \leqslant t) + O\{\eta_n L_{np}\log^{1/2}(p)\} + O\left(\left\{n^{-1}\log^5(np)\right\}^{1/6}\right) \\ & + \mathbb{P}(|C_n|_\infty > \eta_n L_{n,p}). \end{aligned}$$

On the other hand, we also have

$$\begin{aligned} \mathbb{P}\left\{n^{1/2}(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}) \leqslant t\right\} \geqslant\ & \mathbb{P}(G \leqslant t) - O\{\eta_n L_{np}\log^{1/2}(p)\} - O\left(\left\{n^{-1}\log^5(np)\right\}^{1/6}\right) \\ & - \mathbb{P}(|C_n|_\infty > \eta_n L_{n,p}), \end{aligned}$$

where $\mathbb{P}(|C_n|_\infty > \eta_n L_{n,p}) \to 0$ as $n \to \infty$ according to Lemma 1.

Then, if $\log p = o(n^{1/5})$ and $\log n = o(p^{1/3 \wedge \delta})$, with sufficiently slow $\eta_n \to \infty$, we have

$$\sup_{t \in \mathbb{R}^p} \left|\mathbb{P}\{n^{1/2}(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}) \leqslant t\} - \mathbb{P}(G \leqslant t)\right| \to 0.$$

We obtain immediately from Corollary 5.1 in Chernozhukov et al. (2017) that

$$\rho_n(\mathcal{A}^{\mathrm{re}}) = \sup_{A \in \mathcal{A}^{\mathrm{re}}} \left|\mathbb{P}\{n^{1/2}(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}) \in A\} - \mathbb{P}\left(G \in A\right)\right| \to 0,$$

which leads to the conclusion of this theorem. □

Proof of Theorem 2. Let $\tilde{X}_i = X_i - \hat{\boldsymbol{\theta}}_n$ and $\tilde{R}_i = \|\tilde{X}_i\|$ for $i = 1, \ldots, n$. According to Lemma A5,

$$n^{1/2}\tilde{\boldsymbol{\theta}}_n = n^{-1/2}\zeta_1^{-1}\sum_{i=1}^{n} Z_i W_i + \tilde{C}_n\,,$$

where $\tilde{C}_n$ satisfies $|\tilde{C}_n|_\infty = O_p\{n^{-1/4}\log^{1/2}(np) + p^{-(1/6\wedge\delta/2)}\log^{1/2}(np)\}$.

Denote $\overline{W}_n = n^{-1}\sum_{i=1}^n W_i$ and rewrite

$$n^{1/2}\tilde{\boldsymbol{\theta}}_n = n^{-1/2}\zeta_1^{-1}\sum_{i=1}^{n} Z_i(W_i - \overline{W}_n) + \left(n^{-1/2}\zeta_1^{-1}\sum_{i=1}^{n} Z_i\right)\overline{W}_n + \tilde{C}_n,$$

where

$$\left|\left(n^{-1/2}\zeta_1^{-1}\sum_{i=1}^{n} Z_i\right)\overline{W}_n\right|_\infty \leqslant \zeta_1^{-1}\left|n^{-1/2}\sum_{i=1}^{n} Z_i\right|\left|\overline{W}_n\right|_\infty \lesssim n^{-1/2}\log^{1/2}(np)$$

according to Lemma A4 (iii).

It is clear that $\mathbb{E}^*\left\{n^{-1/2}\zeta_1^{-1}\sum_{i=1}^n Z_i(W_i - \overline{W}_n)\right\} = 0$. Let $\hat{\mathbb{B}} = n^{-1}\sum_{i=1}^n W_i W_i^\top$, then

$$\mathrm{Var}^*\left\{n^{-1/2}\zeta_1^{-1}\sum_{i=1}^{n} Z_i(W_i - \overline{W}_n)\right\} = \zeta_1^{-1}\hat{\mathbb{B}} - \zeta_1^{-2}\overline{W}_n\overline{W}_n^\top.$$

Denote $\mathbb{B}_{j\ell}$ and $\hat{\mathbb{B}}_{j\ell}$ be the (j,ℓ)th element of $\mathbb{B}$ and $\hat{\mathbb{B}}$, respectively. In addition, denote $\overline{W}_{n,j}$ as the jth element of $\overline{W}_n$. Define

$$\Delta_n = \max_{1\leqslant j,\ell\leqslant p}\left|\zeta_1^{-2}\hat{\mathbb{B}}_{j\ell} - \zeta_1^{-2}\overline{W}_{n,j}\overline{W}_{n,\ell} - \zeta_1^{-2}\mathbb{B}_{j\ell}\right|,$$

then

$$\Delta_n \quad\leqslant\quad \Delta_{n1} + \max_{1\leqslant j,\ell\leqslant p}\left|\zeta_1^{-2}\overline{W}_{n,j}\overline{W}_{n,\ell}\right| \lesssim \Delta_{n1} + n^{-1}\log(np),$$

where

$$\Delta_{n1} = \max_{1\leqslant j,\ell\leqslant p}|\zeta_1^{-2}\hat{\mathbb{B}}_{j\ell} - \zeta_1^{-2}\mathbb{B}_{j\ell}| = \max_{1\leqslant j,\ell\leqslant p}\left|n^{-1}\zeta_1^{-2}\sum_{i=1}^{n}\{W_{i,j}W_{i,\ell} - \mathbb{E}(W_{i,j}W_{i,\ell})\}\right|.$$

From the properties of the ψ_α norm, it holds that

$$\begin{aligned}
\left\|\max_{1\leqslant i\leqslant n;1\leqslant j,\ell\leqslant p}|\zeta_1^{-2}W_{i,j}W_{i,\ell}|\right\|_{\psi_{\alpha/2}} &\lesssim \left\|\max_{1\leqslant i\leqslant n,1\leqslant j\leqslant p}|\zeta_1^{-2}W_{i,j}|^2\right\|_{\psi_{\alpha/2}} \\
&= \zeta_1^{-2}\left\|\max_{1\leqslant i\leqslant n,1\leqslant j\leqslant p}|W_{i,j}|\right\|_{\psi_\alpha}^2 \\
&\lesssim \log^2(np)\,.
\end{aligned}$$

Let $J_n = \max_{1\leqslant i\leqslant n;1\leqslant j,\ell\leqslant p}\zeta_1^{-2}|W_{i,j}W_{i,\ell} - \mathbb{E}(W_{i,j}W_{i,\ell})|$, and

$$\begin{aligned}
\sigma_n^2 &= \max_{1\leqslant j,\ell\leqslant p}\zeta_1^{-2}\sum_{i=1}^{n}\mathbb{E}\{W_{i,j}W_{i,\ell} - \mathbb{E}(W_{i,j}W_{i,\ell})\}^2 \\
&\lesssim \max_{1\leqslant j,\ell\leqslant p}\zeta_1^{-2}\sum_{i=1}^{n}\mathbb{E}\{|W_{i,j}W_{i,\ell}|^2\} \lesssim n\,.
\end{aligned}$$

It also follows that

$$\|J_n\|_{\psi_{\alpha/2}} \lesssim \zeta_1^{-2}\left\|\max_{1\leqslant i\leqslant n;1\leqslant j,\ell\leqslant p}|W_{ij},W_{i,\ell}|\right\|_{\psi_{\alpha/2}} + \max_{1\leqslant i\leqslant n;1\leqslant j,\ell\leqslant p}\zeta_1^{-2}\mathbb{E}(|W_{i,j}W_{i,\ell}|) \lesssim \log^2(np)\,.$$

By Lemma E.1 in Chernozhukov et al. (2017), it holds that

$$\begin{aligned}
\mathbb{E}(\Delta_{n1}) &\lesssim n^{-1}\left[\sigma_n\log^{1/2}(p) + \{\mathbb{E}(J_n^2)\}^{1/2}\log p\right] \\
&\lesssim n^{-1}\left\{n^{1/2}\log^{1/2}(p) + \log^{1/\alpha+1}(np)\right\} \\
&\lesssim n^{-1/2}\log^{1/2}(np)\,.
\end{aligned}$$

Then applying Lemma E.2 in Chernozhukov et al. (2017) with $\eta = 1$ and $\beta = \alpha/2$, we obtain that

$$\mathbb{P}(\Delta_{n1} \geqslant 2\mathbb{E}(\Delta_n) + t) \lesssim \exp\left(-C_1 n t^2\right) + 3\exp\left\{-C_2\{tn\log^{-2/\alpha}(np)\}^{\alpha/2}\right\}.$$

Thus, there exist a constant C_1 depends on δ such that

$$\mathbb{P}\left\{\Delta_{n1} > C_1 n^{-1/2}\log^{1/2}(np)\right\} \lesssim p^{-\delta} \to 0\,.$$

From the multiplier bootstrap theorem and Gaussian comparison in Chernozhukov et al. (2017) and Koike (2021),

$$\begin{aligned}&\sup_{t\in\mathbb{R}^p}\left|\mathbb{P}^*\left\{n^{-1/2}\zeta_1^{-1}\sum_{i=1}^n Z_i(W_i - \overline{W}_n) \leqslant t\right\} - \mathbb{P}(G \leqslant t)\right| \\ \lesssim\ & \Delta_n^{1/2}\log(p) + \{n^{-1}\log^5(np)\}^{1/4}\,,\end{aligned}$$

on $\{\Delta_n \lesssim n^{-1/2}\log^{1/2}(np)\}$, which occurs with probability $1 - p^{-\delta}$.

Finally, similar to the proof of Theorem 1, we can show that under Conditions C2 and C3 with $a_0(p) \asymp p^{1-\delta}$, if $\log p = o(n^{1/5})$ and $\log n = o(p^{1/3\wedge\delta})$, we have

$$\sup_{A\in\mathcal{A}^{\mathrm{re}}}\left|\mathbb{P}\{n^{1/2}(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}) \in A\} - \mathbb{P}^*(n^{1/2}\tilde{\boldsymbol{\theta}}_n \in A)\right| \to 0$$

in probability, which completes the proof of this theorem. □

Proof of Theorem 3. Theorems 1 and 2 indicates that there exists a positive sequence $\beta_{n,p} \to 0$ as $n, p \to \infty$ such that

$$\sup_{t\in\mathbb{R}}\left|\mathbb{P}(n^{1/2}|\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}|_\infty \leqslant t) - \mathbb{P}(|G|_\infty \leqslant t)\right| \leqslant \beta_{n,p}/2$$

and

$$\sup_{t\in\mathbb{R}}\left|\mathbb{P}(n^{1/2}|\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}|_\infty \leqslant t) - \mathbb{P}^*(n^{1/2}|\tilde{\boldsymbol{\theta}}_n|_\infty \leqslant t)\right| \leqslant \beta_{n,p}$$

with probability approaching one when $n \to \infty$. Letting $q_{1-\tau}$ be the $(1-\tau)$th quantile of $n^{1/2}|\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}|_\infty$, that is, $q_{1-\tau} = \inf\{u \in \mathbb{R} : \mathbb{P}(n^{1/2}|\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}|_\infty > u) \leqslant \tau\}$. Then,

$$\mathbb{P}^*(n^{1/2}|\tilde{\boldsymbol{\theta}}_n|_\infty \leqslant q_{1-\tau+\beta_{n,p}}) \geqslant \mathbb{P}(n^{1/2}|\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}|_\infty \leqslant q_{1-\tau+\beta_{n,p}}) - \beta_{n,p} \geqslant 1-\tau\,,$$

with probability approaching one as $n \to \infty$. On the other hand, it holds with the same probability that

$$\begin{aligned}
& \mathbb{P}^*(n^{1/2}|\tilde{\boldsymbol{\theta}}_n|_\infty \leqslant q_{1-\tau-3\beta_{n,p}}) \\
\leqslant\ & \mathbb{P}(n^{1/2}|\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}|_\infty \leqslant q_{1-\tau-3\beta_{n,p}}) + \beta_{n,p} \\
=\ & \mathbb{P}(n^{1/2}|\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}|_\infty \leqslant q_{1-\tau-3\beta_{n,p}} - n^{-1/6}) + \beta_{n,p} \\
& + \mathbb{P}(n^{1/2}|\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}|_\infty \leqslant q_{1-\tau-3\beta_{n,p}}) \\
& - \mathbb{P}(n^{1/2}|\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}|_\infty \leqslant q_{1-\tau-3\beta_{n,p}} - n^{-1/6}) \\
<\ & 1-\tau-2\beta_{n,p} + \mathbb{P}(n^{1/2}|\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}|_\infty \leqslant q_{1-\tau-3\beta_{n,p}}) \\
& - \mathbb{P}(n^{1/2}|\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}|_\infty \leqslant q_{1-\tau-3\beta_{n,p}} - n^{-1/6})\,,
\end{aligned}$$

where $\mathbb{P}(n^{1/2}|\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}|_\infty \leqslant q_{1-\tau-3\beta_{n,p}}) - \mathbb{P}(n^{1/2}|\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}|_\infty \leqslant q_{1-\tau-3\beta_{n,p}} - n^{-1/6})$ can be bounded by

$$\begin{aligned}
& \Big|\mathbb{P}(n^{1/2}|\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}|_\infty \leqslant q_{1-\tau-3\beta_{n,p}}) \\
& - \mathbb{P}(n^{1/2}|\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}|_\infty \leqslant q_{1-\tau-3\beta_{n,p}} - n^{-1/6})\Big| \\
\leqslant\ & \Big|\mathbb{P}(|G|_\infty \leqslant q_{1-\tau-3\beta_{n,p}}) - \mathbb{P}(|G|_\infty \leqslant q_{1-\tau-3\beta_{n,p}} - n^{-1/6})\Big| \\
& + \Big|\mathbb{P}(n^{1/2}|\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}|_\infty \leqslant q_{1-\tau-3\beta_{n,p}}) - \mathbb{P}(|G|_\infty \leqslant q_{1-\tau-3\beta_{n,p}})\Big| \\
& + \Big|\mathbb{P}(n^{1/2}|\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}|_\infty \leqslant q_{1-\tau-3\beta_{n,p}} - n^{-1/6}) \\
& \qquad - \mathbb{P}(|G|_\infty \leqslant q_{1-\tau-3\beta_{n,p}} - n^{-1/6})\Big| \\
\leqslant\ & \Big|\mathbb{P}(|G|_\infty \leqslant q_{1-\tau-3\beta_{n,p}}) - \mathbb{P}(|G|_\infty \leqslant q_{1-\tau-3\beta_{n,p}} - n^{-1/6})\Big| + \beta_{n,p} \\
\leqslant\ & C_1 \left\{n^{-1} \log^5(np)\right\}^{1/6} + \beta_{n,p},
\end{aligned}$$

for some positive constant C_1, where the last inequality follows from the Nazarov's inequality. Choosing $C_1\left\{n^{-1}\log^5(np)\right\}^{1/6}\leqslant\beta_{n,p}$, we obtain

$$\mathbb{P}^*(n^{1/2}|\tilde{\boldsymbol{\theta}}_n|_\infty\leqslant q_{1-\tau-3\beta_{n,p}})<1-\tau$$

with probability approaching one. It follows that

$$\mathbb{P}(q_{1-\tau-3\beta_{n,p}}<q^B_{1-\tau}\leqslant q_{1-\tau+\beta_{n,p}})\to 1,\quad\text{as } n,p\to\infty\,.$$

Therefore,

$$\begin{aligned}\mathbb{P}(n^{1/2}|\hat{\boldsymbol{\theta}}_n-\boldsymbol{\theta}|_\infty>q^B_{1-\tau}) &\leqslant \mathbb{P}(n^{1/2}|\hat{\boldsymbol{\theta}}_n-\boldsymbol{\theta}|_\infty>q_{1-\tau-3\beta_{n,p}})+\mathbb{P}(q^B_{1-\tau}\leqslant q_{1-\tau-3\beta_{n,p}})\\ &\leqslant \tau+3\beta_{n,p}+o(1) \qquad\qquad \text{(S.4)}\end{aligned}$$

and

$$\begin{aligned}&\mathbb{P}(n^{1/2}|\hat{\boldsymbol{\theta}}_n-\boldsymbol{\theta}|_\infty>q^B_{1-\tau})\\ \geqslant\ &\mathbb{P}(n^{1/2}|\hat{\boldsymbol{\theta}}_n-\boldsymbol{\theta}|_\infty>q_{1-\tau+\beta_{n,p}})-\mathbb{P}(q^B_{1-\tau}>q_{1-\tau+\beta_{n,p}})\\ \geqslant\ &\mathbb{P}(n^{1/2}|\hat{\boldsymbol{\theta}}_n-\boldsymbol{\theta}|_\infty>q_{1-\tau+\beta_{n,p}}-n^{-1/6})-o(1)\\ &+\mathbb{P}(n^{1/2}|\hat{\boldsymbol{\theta}}_n-\boldsymbol{\theta}|_\infty>q_{1-\tau+\beta_{n,p}})\\ &-\mathbb{P}(n^{1/2}|\hat{\boldsymbol{\theta}}_n-\boldsymbol{\theta}|_\infty>q_{1-\tau+\beta_{n,p}}-n^{-1/6})\\ \geqslant\ &\tau-2\beta_{n,p}-C_2\left\{n^{-1}\log^5(np)\right\}^{1/6}\geqslant\tau-3\beta_{n,p}.\end{aligned}$$

for some positive constant C_2, where the second last inequality follows from the Nazarov's inequality and the last inequality is from choosing $\beta_{n,p}\geqslant C_2\left\{n^{-1}\log^5(np)\right\}^{1/6}$. Finally, as $\beta_{n,p}\to 0$,

$$\mathbb{P}(n^{1/2}|\hat{\boldsymbol{\theta}}_n-\boldsymbol{\theta}|_\infty\geqslant q^B_{1-\tau})-\tau\to 0,$$

which completes the proof of this theorem. □

Proof of Theorem 4. Without loss of generality, we assume $\boldsymbol{\theta}_0 = 0$. Rewrite the test statistic as $T_n = n^{1/2}|\hat{\boldsymbol{\theta}}_n|_\infty$, and let $T_n^c = n^{1/2}|\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}|_\infty$, which has the same distribution of T_n under H_0. Then, it holds that

$$T_n \geqslant n^{1/2}|\boldsymbol{\theta}|_\infty - T_n^c\,.$$

Therefore, the power of the test based on T_n satisfies

$$\begin{aligned}\mathbb{P}(T_n > q_{1-\tau}^B \mid H_1) &\geqslant \mathbb{P}(n^{1/2}|\boldsymbol{\theta}|_\infty - T_n^c \geqslant q_{1-\tau}^B \mid H_1)\\ &= \mathbb{P}(T_n^c \leqslant n^{1/2}|\boldsymbol{\theta}|_\infty - q_{1-\tau}^B \mid H_1)\end{aligned}$$

Under the conditions of Theorem 2, there exists a positive sequence $\beta_{n,p} \to 0$ as $n, p \to \infty$, satisfies

$$\sup_{t\in\mathbb{R}} |\mathbb{P}(T_n^c > t \mid H_1) - \mathbb{P}(|G|_\infty > t \mid H_1)| \leqslant \beta_{n,p}, \tag{S.5}$$

where $G \sim N(0, \zeta_1^{-2}\mathbb{B})$. Letting $q_{1-\tau}$ be the $(1-\tau)$th quantile of T_n^C and $q_{1-\tau}^G$ be the $(1-\tau)$th quantile of $|G|_\infty$. Choosing $t = q_{1-\tau+2\beta_{n,p}}^G$ in equation (S.5), we obtain that $|\mathbb{P}(T_n^c > q_{1-\tau+2\beta_{n,p}}^G \mid H_1) - \tau + 2\beta_{n,p}| \leqslant \beta_{n,p}$ and $\mathbb{P}(T_n^c > q_{1-\tau+2\beta_{n,p}}^G \mid H_1) \leqslant \tau - \beta_{n,p}$, which implies that $q_{1-\tau+\beta_{n,p}} \leqslant q_{1-\tau+2\beta_{n,p}}^G$.

Note that $q_{1-\tau}^B$ is the $(1-\tau)$th quantile of $n^{1/2}|\tilde{\boldsymbol{\theta}}_n|_\infty$ conditional on $X_1, \ldots, X_n$. By carrying out similar procedure as in the proof of equation (S.4), we can show that

$$\mathbb{P}(T_n^c > n^{1/2}|\boldsymbol{\theta}|_\infty - q_{1-\tau}^B \mid H_1) \leqslant \mathbb{P}(T_n^c > n^{1/2}|\boldsymbol{\theta}|_\infty - q_{1-\tau+\beta_{n,p}} \mid H_1) + o(1)\,. \tag{S.6}$$

It follows that

$$\mathbb{P}(T_n^c > n^{1/2}|\boldsymbol{\theta}|_\infty - q_{1-\tau}^B \mid H_1) \leqslant \mathbb{P}(T_n^c > n^{1/2}|\boldsymbol{\theta}|_\infty - q_{1-\tau+2\beta_{n,p}}^G \mid H_1) + o(1)\,.$$

For $|G|_\infty$, for any $t>0$, applying the tail bound of normal random variable, it holds that

$$\mathbb{P}(|G|_\infty > t) \leqslant 2p\exp\{-2t^2/\{\max_{j\leqslant p}\mathrm{var}\{G_j\}\}\} \leqslant 2p\exp\{-C_2t^2\}\,,$$

where $C_2 = 2\bar{B}^{-1}\bar{M}^{-1}$. Choosing $t = C_2^{-1/2}\log^{1/2}(2p/(\tau-2\beta_{n,p}))$, we arrive at

$$\mathbb{P}(|G|_\infty > C_2^{-1/2}\log^{1/2}(2p/(\tau-2\beta_{n,p}))) \leqslant \tau - 2\beta_{n,p}\,,$$

which leads to

$$q^G_{1-\tau+2\beta_{n,p}} \leqslant C_2^{-1/2}\log^{1/2}\{2p/(\tau-2\beta_{n,p})\}$$

for sufficiently large n and fix level τ. It follows that $q^G_{1-\tau+2\beta_{n,p}} \leqslant 2C_2^{-1/2}\log^{1/2}\{np\}$. Then, if $|\boldsymbol{\theta}|_\infty \geqslant Cn^{-1/2}\log^{1/2}(np)$ for a large enough constant $C = 2C_2^{-1}C_3$, which depends on constants $\bar{B}$ and $\bar{M}$, it holds with constant $C_3 \geqslant 1$ that

$$\begin{aligned}
& \mathbb{P}(T_n > q^B_{1-\tau} \mid H_1) \\
\geqslant\ & \mathbb{P}(T_n^c \leqslant n^{1/2}|\boldsymbol{\theta}|_\infty - q^G_{1-\tau+2\beta_{n,p}} \mid H_1) + o(1) \\
\geqslant\ & \mathbb{P}\{T_n^c \leqslant 2C_2^{-1/2}C_3\log^{1/2}(np) \mid H_1\} + o(1) \\
\geqslant\ & \mathbb{P}\{|G|_\infty \leqslant 2C_2^{-1/2}C_3\log^{1/2}(np) \mid H_1\} - \beta_{n,p} + o(1) \\
\geqslant\ & 1 - n^{-2C_3} - \beta_{n,p} + o(1)\,.
\end{aligned}$$

We complete the proof of this theorem. □

Proof of Theorem 5. Recall that $\hat{\zeta}_1 = n^{-1}\sum_{i=1}^n \|X_i - \hat{\boldsymbol{\theta}}_n\|^{-1}$. It has been shown in the proof of Lemma A5 that

$$\|X_i - \hat{\boldsymbol{\theta}}_n\|^{-1} = R_i^{-1}\left(1 + R_i^{-1}W_i^\top\hat{\boldsymbol{\theta}}_n - 2^{-1}R_i^{-2}\|\hat{\boldsymbol{\theta}}_n\|^2 + \tilde{\delta}_{1i}\right),$$

where $\tilde{\delta}_{1i}$ satisfies $\tilde{\delta}_{1i} = O_p(n^{-1})$ and $\max_{1\leqslant i\leqslant n} \tilde{\delta}_{1i} = O_p(n^{-1/2})$. Thus,

$$\begin{aligned}\hat{\zeta}_1 &= n^{-1}\sum_{i=1}^{n} R_i^{-1}\left(1 + R_i^{-1}W_i^{\top}\hat{\boldsymbol{\theta}}_n - 2^{-1}R_i^{-2}\|\hat{\boldsymbol{\theta}}_n\|^2 + \tilde{\delta}_{1i}\right)\\ &= n^{-1}\sum_{i=1}^{n} R_i^{-1}(1+\tilde{\delta}_{3i}),\end{aligned}$$

where $\tilde{\delta}_{3i}$ satisfies $\tilde{\delta}_{3i} = O_p(n^{-1/2})$ and $\max_{1\leqslant i\leqslant n} \tilde{\delta}_{3i} = O_p(n^{-1/4})$. By the fact that $n^{-1}\sum_{i=1}^{n} R_i^{-1} = \zeta_1 + O_p(\zeta_1 n^{-1/2})$, we conclude that

$$\hat{\zeta}_1/\zeta_1 - 1 = O_p(n^{-1/2}). \tag{S.7}$$

Let $\tilde{W}_i = (X_i - \hat{\boldsymbol{\theta}}_n)/\|X_i - \hat{\boldsymbol{\theta}}_n\|$ for $i = 1,\ldots,n$. From the proof of Lemma A5,

$$\tilde{W}_i = (W_i - R_i^{-1}\hat{\boldsymbol{\theta}}_n)(1+\tilde{\delta}_{2i}) = W_i + W_i\tilde{\delta}_{2i} - R_i^{-1}\hat{\boldsymbol{\theta}}_n(1+\tilde{\delta}_{2i}),$$

where $\tilde{\delta}_{2i}$ satisfies $\tilde{\delta}_{2i} = O_p(n^{-1/2})$ and $\max_{1\leqslant i\leqslant n} \tilde{\delta}_{2i} = O_p(n^{-1/4})$. Let $\tilde{W}_{i,j}$ be the jth component of $\tilde{W}_i$, then

$$\begin{aligned}\hat{\mathbb{B}}_{jj} &= n^{-1}\sum_{i=1}^{n}\tilde{W}_{i,j}^2\\ &= n^{-1}\sum_{i=1}^{n} W_{i,j}^2\{1+O_p(\tilde{\delta}_{2i})\} + n^{-1}\sum_{i=1}^{n} R_i^{-1}W_{i,j}\hat{\theta}_{n,j}\{1+O_p(\tilde{\delta}_{2i})\}\\ &\quad + n^{-1}\sum_{i=1}^{n} R_i^{-2}\hat{\theta}_{n,j}^2\{1+O_p(\tilde{\delta}_{2i})\},\end{aligned}$$

where $\max_{1\leqslant j\leqslant p}|n^{-1}\sum_{i=1}^{n} W_{i,j}^2\{1+O_p(\tilde{\delta}_{2i})\}/\mathbb{B}_{jj} - 1| = O_p\{n^{-1/4}\log^{1/2}(np)\}$,

$$\begin{aligned}&\max_{1\leqslant j\leqslant p}\left|n^{-1}\sum_{i=1}^{n} R_i^{-2}\hat{\theta}_{n,j}^2\{1+O_p(\tilde{\delta}_{2i})\}\right|\\ \lesssim\ &\left|n^{-1}\sum_{i=1}^{n} R_i^{-2}\right|\max_{1\leqslant j\leqslant p}|\hat{\theta}_{n,j}^2|\\ =\ &O_p\{\zeta_1^2 n^{-1}\log^{1/2}(np)\}\end{aligned}$$

and

$$\begin{aligned}
&\max_{1\leqslant j\leqslant p}\left|n^{-1}\sum_{i=1}^{n}R_i^{-1}W_{i,j}\hat{\theta}_{n,j}\{1+O_p(\tilde{\delta}_{2i})\}\right| \\
\lesssim\ & \left(n^{-1}\sum_{i=1}^{n}R_i^{-2}\right)^{1/2}\max_{1\leqslant j\leqslant p}\left\{\left(n^{-1}\sum_{i=1}^{n}W_{i,j}^2\right)^{1/2}\right\}\max_{1\leqslant j\leqslant p}|\hat{\theta}_{n,j}| \\
=\ & O_p\{\zeta_1 p^{-1/2}n^{-1/2}\log^{1/2}(np)\}.
\end{aligned}$$

It follows that

$$\max_{1\leqslant j\leqslant p}\hat{\mathbb{B}}_{jj}/\mathbb{B}_{jj}=1+O_p\{n^{-1/4}\log^{1/2}(np)\}. \tag{S.8}$$

Recall $L_{n,p}=n^{-1/4}\log^{1/2}(np)+p^{-(1/6\wedge\delta/2)}\log^{1/2}(np)$ and $\mathcal{A}^{\mathrm{re}}=\{\prod_{j=1}^{p}[a_j,b_j]:-\infty\leqslant a_j\leqslant b_j\leqslant\infty, j=1,\ldots,p\}$ be the class of rectangles in $\mathbb{R}^p$. Under Conditions C1–C3 with $a_0(p)\asymp p^{1-\delta}$ for some positive constant $\delta\leqslant 1/2$, if $\log p=o(n^{1/5})$ and $\log n=o(p^{1/3\wedge\varepsilon})$, we have

$$\sup_{A\in\mathcal{A}^{\mathrm{re}}}\left|\mathbb{P}\left\{\sqrt{n}(\hat{\boldsymbol{\theta}}_n-\boldsymbol{\theta})\in A\right\}-\mathbb{P}(G\in A)\right|\leqslant O\{\eta_n L_{np}\log^{1/2}(p)\}+O\left(\left\{n^{-1}\log^5(np)\right\}^{1/6}\right)+\mathbb{P}(|C_n|_\infty>\eta_n L_{n,p})\to 0\,,$$

where $G\sim N(0,\zeta_1^{-2}\mathbb{B})$ with $\mathbb{B}=\mathbb{E}\{W_1W_1^\top\}$. Moreover, for any $\hat{\theta}_{nj}-\theta_{0,j}$, with choosing $a_i=-\infty$ and $b_i=\infty$ for all $1\leqslant i\leqslant p$ and $i\neq j$, we obtain

$$\begin{aligned}
&\max_{1\leqslant j\leqslant p}\sup_{t\in\mathbb{R}}\left|\mathbb{P}\left(\sqrt{n}(\hat{\theta}_{nj}-\theta_{0,j})\leqslant t\right)-\mathbb{P}\{N(0,\zeta_1^{-2}\mathbb{B}_{jj})\leqslant t\}\right| \\
&\quad\leqslant O\{\eta_n L_{np}\log^{1/2}(p)\}+O\left(\left\{n^{-1}\log^5(np)\right\}^{1/6}\right)+\mathbb{P}(|C_n|_\infty>\eta_n L_{n,p})\to 0\,.
\end{aligned}$$

From (S.7) and (S.8), we have $\max_{1\leqslant j\leqslant p}\left|s_{n,j}^2/\zeta_1^{-2}\mathbb{B}_{jj}-1\right|=O_p(n^{-1/4}\log^{1/2}np)$. Since

$$\left|\max_{1\leqslant j\leqslant p}\left|\frac{\sqrt{n}(\hat{\theta}_{nj}-\theta_{0,j})}{s_{n,j}}\right|-\max_{1\leqslant j\leqslant p}\left|\frac{\sqrt{n}(\hat{\theta}_{nj}-\theta_{0,j})}{\sqrt{\zeta_1^{-2}\mathbb{B}_{jj}}}\right|\right|\leqslant\max_{1\leqslant j\leqslant p}\left|\frac{\sqrt{n}(\hat{\theta}_{nj}-\theta_{0,j})}{\sqrt{\zeta_1^{-2}\mathbb{B}_{jj}}}\right|\max_{1\leqslant j\leqslant p}\left|\frac{\sqrt{\zeta_1^{-2}\mathbb{B}_{jj}}}{s_{n,j}}-1\right|=O_p(n^{-1/4}\log(np))\,,$$

then, for all $1 \leqslant j \leqslant p$,

$$\mathbb{P}\biggl(\frac{\sqrt{n}(\hat{\theta}_{nj}-\theta_{0,j})}{s_{n,j}}>t\biggr)-\mathbb{P}\biggl(\frac{\sqrt{n}(\hat{\theta}_{nj}-\theta_{0,j})}{\sqrt{\zeta_1^{-2}\mathbb{B}_{jj}}}>t-\epsilon\biggr)\leqslant\mathbb{P}\biggl\{\max_{1\leqslant j\leqslant p}\biggl|\frac{\sqrt{n}(\hat{\theta}_{nj}-\theta_{0,j})}{\sqrt{\zeta_1^{-2}\mathbb{B}_{jj}}}\biggr|\max_{1\leqslant j\leqslant p}\biggl|\frac{\sqrt{\zeta_1^{-2}\mathbb{B}_{jj}}}{s_{n,j}}-1\biggr|\geqslant\epsilon\biggr\}.$$

As a result, by choosing $\epsilon = Cn^{-1/4}\log(np) \to 0$ with sufficient slow $C \to \infty$, the reverse inequality can be shown using a similar argument, then, $\max_{1\leqslant j\leqslant p}\sup_{t\in\mathbb{R}}\left|\mathbb{P}\left(\sqrt{n}s_{n,j}^{-1}(\hat{\theta}_{nj}-\theta_{0,j})\leqslant t\right)-\Phi(t)\right| \to 0$. Hence, if $\log p = o(n^{1/5})$ and $\log n = o(p^{1/3\wedge\delta})$, we have

$$\sup_{0\leqslant x\leqslant 2^{1/2}\log^{1/2}(np)}\left|\mathbb{P}\left\{n^{1/2}(\hat{\theta}_{n,j}-\theta_j)/s_{n,j}>x\right\}-\{1-\Phi(x)\}\right|\to 0\,. \tag{S.9}$$

Let $\bar{T}_{n,j} = n^{-1/2}\sum_{i=1}^n W_{i,j}/\{n^{-1}\sum_{i=1}^n W_{i,j}^2 - (n^{-1}\sum_{i=1}^n W_{i,j})^2\}^{1/2}$, and by the Lemma 6.1 in Liu and Shao (2014), it holds that for any $\epsilon > 0$ and some $\vartheta > 0$,

$$\mathbb{E}\biggl(\sum_{j\in\mathcal{H}_0}\{\mathbb{I}\{|\bar{T}_{n,j}|\geqslant t\}-\mathbb{P}(|\bar{T}_{n,j}|)\geqslant t\}\biggr)^2\lesssim p_0^2\mathcal{G}_\kappa^2(t)\biggl(\frac{1}{p_0\mathcal{G}_\kappa(t)}+\frac{\exp(r+\epsilon)t^2/(1+r)}{p^{1-\eta}}+\log p^{-1-\vartheta}\biggr)\,, \tag{S.10}$$

where $\mathcal{G}_\kappa(t)$ is some function depends on $\kappa_j = \mathbb{E}|W_{i,j}|^3$ such that $\mathcal{G}_\kappa(t) \geqslant \mathcal{G}(t) = 2\{1-\Phi(t)\}$ for all $t \in \mathbb{R}$, and $\mathcal{G}_\kappa(t) = \mathcal{G}(t)\{1+o(1)\}$ uniformly over $0 \leqslant t \leqslant 2^{1/2}\log^{1/2}(p)$ with $\log p = o(n^{1/5})$. Define the sequence $z_0 < z_1 \ldots, < z_{d_p} = 1$, define $z_0 = b_p/p, \ldots, z_j = b_p/p + b_p^{2/3}\exp^{j^u}/p$, for $0 < u < 1$, and $d_p = \left(\log\frac{p-b_p}{b_p^{2/3}}\right)^{1/u}$ with any $b_p = o(p) \to \infty$, then it holds that with S.9 and S.10 for any $\epsilon > 0$, and talking $\frac{1}{1+\vartheta} < u < 1$,

$$\sum_{k=1}^{d_p}\mathbb{P}\left(\left|\frac{\sum_{j\in H_0}\mathbb{I}\{|\bar{T}_{n,j}|\geqslant t_k\}}{p_0\mathcal{G}_\kappa(t_k)}-1\right|>C\right)\leqslant C^{-2}\biggl(\frac{1}{(1-p^{\varpi-1})b_p+}+\frac{1}{b_p^{2/3}}\sum_{k=1}^{d_p}\exp^{-k^u}+d_pp^{-1+\eta+\frac{2(r+\epsilon)}{1+r}}+d_p\log p^{-1-\vartheta}\biggr), \tag{S.11}$$

then we have by the Markov inequality

$$\max_{0\leqslant k\leqslant d_p}\left|\frac{\sum_{j\in H_0}\mathbb{I}\{|\bar{T}_{n,j}|\geqslant t_k\}}{p_0\mathcal{G}_\kappa(t_k)}-1\right|=o_p(1)\,,$$

Then based on Equation (13) of Liu and Shao (2014), as $n \to \infty$, with Condition C4, and $p_0 \geqslant p(1-p^{\varpi-1}) \to \infty$ for some $0 < \varpi < 1$, and some $0 < \eta < (1-r)/(1+r)$, it holds that

$$\sup_{0 \leqslant t \leqslant \mathcal{G}_\kappa^{-1}(b_p/p)} \left| \frac{\sum_{j \in \mathcal{H}_0} \mathbb{I}\{|\bar{T}_{n,j}| \geqslant t\}}{p_0 \mathcal{G}_\kappa(t)} - 1 \right| = o_p(1)\,,$$

Then, with enough large n, p, as long as $|\mathcal{H}| \to \infty$ and $|\mathcal{H}| > 2/\tau$, we have

$$\tau|\mathcal{H}|/p \geqslant 2/p = 2\exp\{-(2^{1/2}\log^{1/2} p)^2/2\} \geqslant 2\{1 - \Phi(2^{1/2}\log^{1/2} p)\} = \mathcal{G}(2^{1/2}\log^{1/2} p)\,.$$

It follows that $\mathcal{G}^{-1}(\tau|\mathcal{H}|/p) \leqslant 2^{1/2}\log^{1/2} p$, consequently,

$$\sup_{0 \leqslant t \leqslant \mathcal{G}^{-1}(\tau|\mathcal{H}|/p)} \left| \frac{\sum_{j \in \mathcal{H}_0} \mathbb{I}\{|\bar{T}_{n,j}| \geqslant t\}}{p_0 \mathcal{G}(t)} - 1 \right| = o_p(1)\,.$$

Let $T'_{n,j} = n^{1/2}(\hat{\theta}_{n,j} - \theta_j)/s_{n,j}$, we obtain $\max_{1 \leqslant j \leqslant p} |T'_{n,j} - \bar{T}_{n,j}| = o_p\{\log^{-1/2}(p)\}$ with some careful calculations. With similar procedure to Page 84 of Belloni et al. (2018), it holds that

$$\sup_{0 \leqslant t \leqslant \mathcal{G}^{-1}(\tau|\mathcal{H}|/p)} \left| \frac{\sum_{j \in \mathcal{H}_0} \mathbb{I}\{|T'_{n,j}| \geqslant t\}}{p_0 \mathcal{G}(t)} - 1 \right| = o_p(1)\,. \tag{S.12}$$

The B-H method with $P_1, \ldots, P_p$ is equivalent to the following procedure: reject H_{0j}, if only if $P_j \leqslant \hat{t}_0$, where

$$\hat{t}_0 = \sup\left\{0 \leqslant t \leqslant 1 : t \leqslant \frac{\tau \max\{\sum_{j=1}^p \mathbb{I}(P_j \leqslant t)\}}{p}\right\}.$$

Then we have $\hat{t}_0 = \frac{\tau \max\{\sum_{j=1}^p \mathbb{I}(P_j \leqslant \hat{t}_0), 1\}}{p}$, and $\tau|\mathcal{H}|/p \geqslant \mathcal{G}(2^{1/2}\log^{1/2} p)$. Set $t = \mathcal{G}^{-1}(\tau|\mathcal{H}|/p)$, then $t \leqslant 2^{1/2}\log^{1/2} p$ with probability tends to 1. Under H_{0j}, it holds that $T_{n,j} = T'_{n,j} = n^{1/2}(\hat{\theta}_{n,j} - \theta_{0,j})/s_{n,j}$, Thus, we have

$$\begin{aligned}
\mathcal{G}(t) = \frac{\tau|\mathcal{H}|}{p} &\leqslant \frac{\tau \max\{\sum_{j=1}^p \mathbb{I}(|T_{n,j}| \geqslant 2^{1/2}\log^{1/2} p), 1\}}{p} \\
&\leqslant \frac{\tau \max\{\sum_{j=1}^p \mathbb{I}(|T_{n,j}| \geqslant t), 1\}}{p}\,,
\end{aligned}$$

where the second inequality implied by (B.29) of Belloni et al. (2018), which is denoted as

$$\mathbb{P}\Big(\sum_{j=1}^{p}|T_{n,j}| \geqslant \sqrt{2\log p} \geqslant \mathcal{H}\Big) \to 1\,.$$

It implies that $\mathbb{P}(\hat{t}_0 \geqslant \tau|\mathcal{H}|/p) \to 1$ with $\hat{t}_0 = \mathcal{G}(\hat{t})$, and together with (S.12), we have

$$\frac{\sum_{j\in\mathcal{H}_0}\mathbb{I}\{|T'_{n,j}| \geqslant \hat{t}\}}{p_0\mathcal{G}(\hat{t})} = \frac{\sum_{j\in\mathcal{H}_0}\mathbb{I}(|T_{n,j}| \geqslant \hat{t})}{p_0\mathcal{G}(\hat{t})} \xrightarrow{P} 1\,,$$

which is equivalent to

$$\frac{\sum_{j\in\mathcal{H}_0}\mathbb{I}(P_j \leqslant \hat{t}_0)}{p_0\hat{t}_0} \xrightarrow{P} 1\,.$$

Finally,

$$\mathrm{FDP}_M = \frac{\sum_{j\in\mathcal{H}_0}\mathbb{I}(P_j \leqslant \hat{t}_0)}{\max\{\sum_{j=1}^{p}\mathbb{I}(P_j \leqslant \hat{t}_0), 1\}} = \frac{\sum_{j\in\mathcal{H}_0}\mathbb{I}(P_j \leqslant \hat{t}_0)}{p\hat{t}_0/\tau} \xrightarrow{P} \frac{\tau p_0}{p}$$

as $n \to \infty$, since FDP_M is bounded between 0 and 1, it follows that $\mathrm{FDP}_M \to \frac{\tau p_0}{p}$, which completes the proof of this theorem.

□

Appendix C: Proof of preliminary lemmas

In this section, we present proofs of preliminary lemmas in Appendix A.

Proof of Lemma A1. As the components of U_1 are independent and standardized, simple calculations yield $\mathbb{E}(\|U_1\|^2) = p$ and

$$\mathbb{E}(\|\Gamma U_1\|^2) = \mathbb{E}(U_1^\top\Gamma^\top\Gamma U_1) = \mathrm{tr}\{\Gamma^\top\Gamma\mathbb{E}(U_1U_1^\top)\} = \mathrm{tr}(\Omega)\,.$$

Under Condition C1, the components of $U_1 = (U_{1,1}, \ldots, U_{1,p})^\top$ are independent sub-exponential random variables such that $\max_{1\leqslant j\leqslant p}\|U_{1,j}\|_{\psi_\alpha} \leqslant c_0$. Applying the concentration inequality in the proof of Lemma S2.1 in (Wang et al., 2015), for every $t \geqslant 0$,

$$\mathbb{P}\left(\left|\|U_1\|^2 - p\right| \geqslant t\right) \leqslant C_1 \exp\left\{-C_2\left(p^{-1}t^2\right)^{\alpha/(4\alpha+4)}\right\}. \tag{S.13}$$

and

$$\mathbb{P}\left\{\left|\|\Gamma U_1\|^2 - \mathrm{tr}(\Omega)\right| \geqslant t\right\} \leqslant C_1 \exp\left[-C_2\left\{\frac{t^2}{\mathrm{tr}(\Omega^2)}\right\}^{\alpha/(4\alpha+4)}\right]. \tag{S.14}$$

For any fixed $0 < \epsilon < 1$, let

$$\mathcal{A}_1 = \{p - \epsilon p^{(1+\delta)/2} \leqslant \|U_1\|^2 \leqslant p + \epsilon p^{(1+\delta)/2}\}$$

and

$$\mathcal{A}_2 = \{(1-\epsilon)\mathrm{tr}(\Omega) \leqslant \|\Gamma U_1\|^2 \leqslant (1+\epsilon)\mathrm{tr}(\Omega)\}.$$

Taking $t = \epsilon p^{(1+\delta)/2}$ in (S.13) and $t = \epsilon \mathrm{tr}(\Omega)$ in (S.14), we have

$$\mathbb{P}(\mathcal{A}_1) \geqslant 1 - C_1 \exp\left\{-C_2(\epsilon^2 p^\delta)^{\alpha/(4\alpha+4)}\right\}$$

and

$$\mathbb{P}(\mathcal{A}_2) \geqslant 1 - C_1 \exp\left[-C_2\left\{\frac{\epsilon^2 \mathrm{tr}^2(\Omega)}{\mathrm{tr}(\Omega^2)}\right\}^{\alpha/(4\alpha+4)}\right].$$

Under Condition C3,

$$\mathrm{tr}(\Omega^2) = \sum_{j=1}^p \sum_{\ell=1}^p \omega_{j\ell}^2 \leqslant \overline{M} p \max_{1 \leqslant \ell \leqslant p} \sum_{j=1}^p |\omega_{j\ell}| \leqslant \overline{M} p a_0(p).$$

Since $\mathrm{tr}(\Omega) = p$ and $a_0(p) \asymp p^{1-\delta}$, we conclude that

$$\frac{\mathrm{tr}^2(\Omega)}{\mathrm{tr}(\Omega^2)} \geqslant \frac{p^2}{\overline{M} p a_0(p)} \asymp p^{\delta}\,.$$

Consequently, for some positive constants c_1 and c_2, we get that

$$\mathbb{P}(\mathcal{A}_1) \geqslant 1 - c_1 \exp\{-c_2 p^{\delta\alpha/(4\alpha+4)}\}\,.$$

and

$$\mathbb{P}(\mathcal{A}_2) \geqslant 1 - c_1 \exp\{-c_2 p^{\delta\alpha/(4\alpha+4)}\}$$

for sufficient large p. Thus, we finish the proof of this lemma. □

Proof of Lemma A2. (i) As the components of $U_i = (U_{i,1}, \ldots, U_{i,p})^\top$ are i.i.d. standardized sub-exponential random variables, simple algebra yields

$$\begin{aligned}
\mathbb{E}(\|U_i\|^4) &= \mathbb{E}\left\{\left(\sum_{j=1}^{p} U_{i,j}^2\right)^2\right\} \\
&= \sum_{j=1}^{p} \mathbb{E}(U_{i,j}^4) + \sum_{1\leqslant j_1\neq j_2\leqslant p} \mathbb{E}(U_{i,j_1}^2)\mathbb{E}(U_{i,j_2}^2) \\
&= p\mathbb{E}(U_{i,j}^4) + p(p-1)
\end{aligned}$$

and

$$\begin{aligned}
\mathbb{E}(\|U_i\|^6) &= \sum_{j=1}^{p} \mathbb{E}(U_{i,j}^6) + 3\sum_{1\leqslant j_1\neq j_2\leqslant p} \mathbb{E}(U_{i,j_1}^4)\mathbb{E}(U_{i,j_2}^2) \\
&\quad + \sum_{1\leqslant j_1\neq j_2\neq j_3\leqslant p} \mathbb{E}(U_{i,j_1}^2)\mathbb{E}(U_{i,j_2}^2)\mathbb{E}(U_{i,j_3}^2) \\
&= p\mathbb{E}(U_{i,j}^6) + 3p(p-1)\mathbb{E}(U_{i,j}^4) + p(p-1)(p-2)\,.
\end{aligned}$$

In addition,

$$\begin{aligned}\mathbb{E}(\|U_i\|^8) &= \sum_{j=1}^{p}\mathbb{E}(U_{i,j}^8) + 4\sum_{1\leqslant j_1\neq j_2\leqslant p}\mathbb{E}(U_{i,j_1}^6)\mathbb{E}(U_{i,j_2}^2) \\ &\quad +3\sum_{1\leqslant j_1\neq j_2\leqslant p}\mathbb{E}(U_{i,j_1}^4)\mathbb{E}(U_{i,j_2}^4) \\ &\quad +6\sum_{1\leqslant j_1\neq j_2\neq j_3\leqslant p}\mathbb{E}(U_{i,j_1}^4)\mathbb{E}(U_{i,j_2}^2)\mathbb{E}(U_{i,j_3}^2) \\ &\quad +\sum_{1\leqslant j_1\neq j_2\neq j_3\neq j_4\leqslant p}\mathbb{E}(U_{i,j_1}^2)\mathbb{E}(U_{i,j_2}^2)\mathbb{E}(U_{i,j_3}^2)\mathbb{E}(U_{i,j_4}^2) \\ &= p\mathbb{E}(U_{i,j}^8) + 4p(p-1)\mathbb{E}(U_{i,j_1}^6) + 3p(p-1)\{\mathbb{E}(U_{i,j_1}^4)\}^2 \\ &\quad +3p(p-1)\mathbb{E}(U_{i,j}^4) + p(p-1)(p-2)(p-3)\,.\end{aligned}$$

The result of $\mathbb{E}(\|U_i\|^{2k}) = p^k + O(p^{k-1})$ for any positive integer k can be checked by

$$\begin{aligned}\mathbb{E}(\|U_i\|^{2k}) &= \sum_{1\leqslant j_1\neq\cdots\neq j_k\leqslant p}\mathbb{E}(U_{i,j_1}^2)\times\cdots\times\mathbb{E}(U_{i,j_k}^2)\{1+O(p^{-1})\} \\ &= p^k + O(p^{k-1})\,.\end{aligned}$$

Moreover, by the fact that $\{1+u-(u-1)^2\}/2 \leqslant u^{1/2} \leqslant (1+u)/2$ for all $u \geqslant 0$, we can get that $\mathbb{E}(\|U\|^k) = p^{k/2} + O(p^{k/2-1})$ for all positive integer k.

(ii) Write $\Lambda_{j\ell} = \sum_{j_1=1}^{p}\Gamma_{j_1 j}\Gamma_{j_1\ell}$ as the (j,ℓ)th element of $\Gamma^\top\Gamma$, then

$$\begin{aligned}\mathbb{E}(\|\Gamma U_i\|^4) &= \mathbb{E}\left\{\left(\sum_{j=1}^{p}\sum_{\ell=1}^{p}\Lambda_{j\ell}U_{i,j}U_{i,\ell}\right)^2\right\} \\ &= \sum_{j=1}^{p}\Lambda_{jj}^2\mathbb{E}(U_{i,j}^4) + 2\sum_{1\leqslant j_1\neq j_2\leqslant p}\Lambda_{j_1j_2}^2\mathbb{E}(U_{i,j_1}^2)\mathbb{E}(U_{i,j_2}^2) \\ &\quad + \sum_{1\leqslant j_1\neq j_2\leqslant p}\Lambda_{j_1j_1}\Lambda_{j_2j_2}\mathbb{E}(U_{i,j_1}^2)\mathbb{E}(U_{i,j_2}^2) \\ &= \mathbb{E}(U_{i,j}^4)\sum_{j=1}^{p}\Lambda_{jj}^2 + 2\sum_{1\leqslant j_1\neq j_2\leqslant p}\Lambda_{j_1j_2}^2 + \sum_{1\leqslant j_1\neq j_2\leqslant p}\Lambda_{j_1j_1}\Lambda_{j_2j_2} \\ &= \left(\sum_{j=1}^{p}\Lambda_{jj}\right)^2 + \{\mathbb{E}(U_{i,j}^4)-1\}\sum_{j=1}^{p}\Lambda_{jj}^2 + 2\sum_{1\leqslant j_1\neq j_2\leqslant p}\Lambda_{j_1j_2}^2 \\ &= \{\mathrm{tr}(\Omega)\}^2 + O\{\mathrm{tr}(\Omega^2)\}\end{aligned}$$

as $\sum_{j=1}^{p}\Lambda_{jj}^2 + \sum_{1\leqslant j_1\neq j_2\leqslant p}\Lambda_{j_1j_2}^2 = \sum_{j=1}^{p}\sum_{\ell=1}^{p}\Lambda_{j\ell}^2 = \mathrm{tr}(\Omega^2)$ and $\mathrm{tr}(\Omega^2) \lesssim p^{2-\delta}$ based on Condition C3. Similarly, we can show that

$$\begin{aligned}\mathbb{E}(\|\Gamma U_i\|^6) &= \sum_{1\leqslant j_1\neq j_2\neq j_3\leqslant p}(\Lambda_{j_1j_1}\Lambda_{j_2j_2}\Lambda_{j_3j_3}+\Lambda_{j_1j_2}^2\Lambda_{j_3j_3}\\&\qquad +\Lambda_{j_1j_2}\Lambda_{j_1j_3}\Lambda_{j_2j_3})\mathbb{E}(U_{i,j_1}^2)\mathbb{E}(U_{i,j_2}^2)\mathbb{E}(U_{i,j_3}^2)\{1+O(p^{-1})\}\\&= p^3+O(p^{3-\delta})\end{aligned}$$

and $\mathbb{E}(\|\Gamma U_i\|^{12})=p^6+O(p^{6-\delta})$.

Similar to the proof of part (i), the result $\mathbb{E}(\|\Gamma U_i\|)=p^{1/2}+O(p^{1/2-\delta})$ and $\mathbb{E}(\|\Gamma U_i\|^3)=p^{3/2}+O(p^{3/2-\delta})$ are directly consequences of $\mathbb{E}(\|\Gamma U_i\|^2)=p$, $\mathbb{E}(\|\Gamma U_i\|^4)=p^2+O(p^{2-\delta})$, $\mathbb{E}(\|\Gamma U_i\|^6)=p^3+O(p^{3-\delta})$, $\mathbb{E}(\|\Gamma U_i\|^{12})=p^6+O(p^{6-\delta})$ and $\{1+u-(u-1)^2\}/2\leqslant u^{1/2}\leqslant(1+u)/2$ for all $u\geqslant 0$.

(iii) Now we consider $\mathbb{E}\{\|\Gamma S(U_i)\|^2\}$. For $i=1,\dots,n$, let

$$\mathcal{A}_{1i}=\{p-\epsilon p^{(1+\delta)/2}\leqslant\|U_i\|^2\leqslant p+\epsilon p^{(1+\delta)/2}\}$$

for a fixed $0<\epsilon<1$. According to Lemma A1 and the fact that $\|\Gamma U_i\|^2\leqslant\mathrm{tr}(\Omega)\|U_i\|^2$,

$$\begin{aligned}\mathbb{E}\{\|\Gamma S(U_i)\|^2\} &= \mathbb{E}(\|\Gamma U_i\|^2\|U_i\|^{-2})\\&= p^{-1}\mathbb{E}\{\|\Gamma U_i\|^2\}+\mathbb{E}\left\{\|\Gamma U_i\|^2\left(\|U_i\|^{-2}-p^{-1}\right)\right\}\\&= 1+\mathbb{E}\left\{\|\Gamma U_i\|^2\left(\|U_i\|^{-2}-p^{-1}\right)\right\},\end{aligned}$$

where

$$
\begin{aligned}
& \mathbb{E}\left\{\|\Gamma U_i\|^2\left(\|U_i\|^{-2}-p^{-1}\right)\right\} \\
\leqslant\ & p^{-1}\mathbb{E}\left(\|\Gamma U_i\|^2\|U_i\|^{-2}\left|\|U_i\|^2-p\right|\right) \\
=\ & p^{-1}\mathbb{E}\left\{\|\Gamma U_i\|^2\|U_i\|^{-2}\left|\|U_i\|^2-p\right|\mathbb{I}(\mathcal{A}_{1i})\right\}+p^{-1}\mathbb{E}\left\{\|\Gamma U_i\|^2\|U_i\|^{-2}\left|\|U_i\|^2-p\right|\mathbb{I}(\mathcal{A}_{1i}^c)\right\} \\
\leqslant\ & p^{-1}\{p-\epsilon p^{(1+\delta)/2}\}^{-1}\mathbb{E}\left(\|\Gamma U_i\|^2\left|\|U_i\|^2-p\right|\right)+p^{-1}\mathrm{tr}(\Omega)\mathbb{E}\left\{\left|\|U_i\|^2-p\right|\mathbb{I}(\mathcal{A}_{1i}^c)\right\} \\
\leqslant\ & p^{-1}\{p-\epsilon p^{(1+\delta)/2}\}^{-1}\left\{\mathbb{E}(\|\Gamma U_i\|^4)\right\}^{1/2}\left\{\mathbb{E}(\left|\|U_i\|^2-p\right|^2)\right\}^{1/2}+\left\{\mathbb{E}(\left|\|U_i\|^2-p\right|^2)\right\}^{1/2}\{\mathbb{P}(\mathcal{A}_{1i}^c)\}^{1/2} \\
\leqslant\ & p^{-1}(p-\epsilon p^{1-\delta})^{-1}\{p^2+O(p^{2-\delta})\}^{1/2}\times O(p^{1/2})+O(p^{1/2})\times c_1^{1/2}\exp\{-c_2p^{\delta\alpha/(4\alpha+4)}/2\} \\
=\ & O(p^{-1/2})\,.
\end{aligned}
$$

It follows that $\mathbb{E}\{\|\Gamma S(U_i)\|^2\}=1+O(p^{-1/2})$.

Similarly, the last result follows from

$$
\begin{aligned}
\mathbb{E}\{\|\Gamma S(U_i)\|^4\} &= p^{-2}\mathbb{E}\{\|\Gamma U_i\|^4\}+\mathbb{E}\left\{\|\Gamma U_i\|^4\left(\|U_i\|^{-4}-p^{-2}\right)\right\} \\
&= 1+O(p^{-\delta})+\mathbb{E}\left\{\|\Gamma U_i\|^4\left(\|U_i\|^{-4}-p^{-2}\right)\right\}\,,
\end{aligned}
$$

where

$$
\begin{aligned}
& \mathbb{E}\left\{\|\Gamma U_i\|^4\left(\|U_i\|^{-4}-p^{-2}\right)\right\} \\
\leqslant\ & p^{-2}\mathbb{E}\left(\|\Gamma U_i\|^4\|U_i\|^{-4}\left|\|U_i\|^4-p^2\right|\right) \\
=\ & p^{-2}\mathbb{E}\left\{\|\Gamma U_i\|^4\|U_i\|^{-4}\left|\|U_i\|^4-p^2\right|\mathbb{I}(\mathcal{A}_{1i})\right\}+p^{-2}\mathbb{E}\left\{\|\Gamma U_i\|^4\|U_i\|^{-4}\left|\|U_i\|^4-p^2\right|\mathbb{I}(\mathcal{A}_{1i}^c)\right\} \\
\leqslant\ & p^{-2}(p-\epsilon p^{1-\delta})^{-2}\mathbb{E}\left(\|\Gamma U_i\|^4\left|\|U_i\|^4-p^2\right|\right)+p^{-2}\{\mathrm{tr}(\Omega)\}^2\mathbb{E}\left\{\left|\|U_i\|^4-p^2\right|\mathbb{I}(\mathcal{A}_{1i}^c)\right\} \\
\leqslant\ & p^{-2}(p-\epsilon p^{1-\delta})^{-2}\left\{\mathbb{E}(\|\Gamma U_i\|^6)\right\}^{2/3}\left\{\mathbb{E}(\left|\|U_i\|^4-p^2\right|^3)\right\}^{1/3}+\left\{\mathbb{E}(\left|\|U_i\|^4-p^2\right|^2)\right\}^{1/2}\{\mathbb{P}(\mathcal{A}_{1i}^c)\}^{1/2} \\
\leqslant\ & p^{-2}(p-\epsilon p^{1-\delta})^{-2}\times O(p^2)\times O(p^{3/2})+O(p^{3/2})\times c_1^{1/2}\exp\{-c_2p^{\delta\alpha/(4\alpha+4)}/2\} \\
=\ & O(p^{-1/3})\,.
\end{aligned}
$$

(iv) as ν_i and $S(U_i)$ are independent,

$$\begin{aligned}
&\mathbb{E}(\nu_i^{-1})\mathbb{E}\{\|\Gamma S(U_i)\|^{-1}\} = \mathbb{E}(\nu_i^{-1}\|\Gamma U_i\|^{-1}\|U_i\|) = \mathbb{E}(R_i^{-1}\|U_i\|)\\
= \;& \mathbb{E}\{R_i^{-1}\|U_i\|\mathbb{I}(\mathcal{A}_{1i})\} + \mathbb{E}\{R_i^{-1}\|U_i\|\mathbb{I}(\mathcal{A}_{1i}^c)\}\\
\leqslant \;& \{p+\epsilon p^{(1+\delta)/2}\}^{1/2}\mathbb{E}\{R_i^{-1}\mathbb{I}(\mathcal{A}_{1i})\} + \{\mathbb{E}(R_i^{-4})\}^{1/4}\{\mathbb{E}\|U_i\|^4\}^{1/4}\{\mathbb{P}(\mathcal{A}_{1i}^c)\}^{1/2}\\
\lesssim \;& \{p+\epsilon p^{(1+\delta)/2}\}^{1/2}\mathbb{E}(R_i^{-1}) + \zeta_4^{1/4}\times p^{1/2}\times c_1^{1/2}\exp\{-c_2 p^{\delta\alpha/(4\alpha+4)}/2\}\\
\lesssim \;& \zeta_1 p^{1/2}\,,
\end{aligned}$$

and

$$\begin{aligned}
\mathbb{E}(\nu_i^{-2})\mathbb{E}\{\|\Gamma S(U_i)\|^{-2}\} &= \mathbb{E}\{R_i^{-2}\|U_i\|^2\mathbb{I}(\mathcal{A}_{1i})\} + \mathbb{E}\{R_i^{-2}\|U_i\|^2\mathbb{I}(\mathcal{A}_{1i}^c)\}\\
&\leqslant \{p+\epsilon p^{(1+\delta)/2}\}\mathbb{E}\{R_i^{-2}\mathbb{I}(\mathcal{A}_{1i})\} + \{\mathbb{E}(R_i^{-4})\}^{1/2}\{\mathbb{E}\|U_i\|^6\}^{1/3}\{\mathbb{P}(\mathcal{A}_{1i}^c)\}^{1/6}\\
&\lesssim \{p+\epsilon p^{(1+\delta)/2}\}\mathbb{E}(R_i^{-2}) + \zeta_4^{1/2}\times p\times c_1^{1/6}\exp\{-c_2 p^{\delta\alpha/(4\alpha+4)}/6\}\\
&\lesssim \zeta_2 p\,,
\end{aligned}$$

In addition, we also have

$$\begin{aligned}
\mathbb{E}(\nu_i^{-3})\mathbb{E}\{\|\Gamma S(U_i)\|^{-3}\} &= \mathbb{E}\{R_i^{-3}\|U_i\|^3\mathbb{I}(\mathcal{A}_{1i})\} + \mathbb{E}\{R_i^{-3}\|U_i\|^3\mathbb{I}(\mathcal{A}_{1i}^c)\}\\
&\leqslant \{p+\epsilon p^{(1+\delta)/2}\}^{3/2}\mathbb{E}\{R_i^{-3}\mathbb{I}(\mathcal{A}_{1i})\} + \{\mathbb{E}(R_i^{-4})\}^{3/4}\{\mathbb{E}\|U_i\|^{18}\}^{1/6}\{\mathbb{P}(\mathcal{A}_{1i}^c)\}^{1/12}\\
&\lesssim \{p+\epsilon p^{(1+\delta)/2}\}^{3/2}\mathbb{E}(R_i^{-3}) + \zeta_4^{3/4}\times p^{3/2}\times c_1^{1/12}\exp\{-c_2 p^{\delta\alpha/(4\alpha+4)}/12\}\\
&\lesssim \zeta_3 p^{3/2}\,.
\end{aligned}$$

By Cauchy-Schwarz inequality and Jensen's inequality, we can show that

$$[\mathbb{E}\{\|\Gamma S(U_i)\|^{-1}\}]^{-1} \leqslant \mathbb{E}\{\|\Gamma S(U_i)\|\} \leqslant [\mathbb{E}\{\|\Gamma S(U_i)\|^2\}]^{1/2} = 1 + O(p^{-1/2})\,,$$

$$[\mathbb{E}\{\|\Gamma S(U_i)\|^{-2}\}]^{-1} \leqslant \mathbb{E}\{\|\Gamma S(U_i)\|^2\} = 1 + O(p^{-1/2})\,,$$

and

$$[\mathbb{E}\{\|\Gamma S(U_i)\|^{-3}\}]^{-1} \leqslant \mathbb{E}\{\|\Gamma S(U_i)\|^3\} \leqslant [\mathbb{E}\{\|\Gamma S(U_i)\|^4\}]^{3/4} = 1 + O(p^{-1/3}).$$

Then, the results of this part follows immediately. We finish the proof of this lemma. □

Proof of Lemma A3. (i) For $i = 1, \ldots, n$, let $\mathcal{A}_{2i} = \{(1-\epsilon)\mathrm{tr}(\Omega) \leqslant \|\Gamma U_i\|^2 \leqslant (1+\epsilon)\mathrm{tr}(\Omega)\}$ for a fixed $0 < \epsilon < 1$. Recall that Γ_j is the jth row of Γ and $W_{i,j} = \Gamma_j U_i / \|\Gamma U_i\|$, then

$$\begin{aligned} Q_{j\ell} &= n^{-1}\sum_{i=1}^n R_i^{-1} W_{i,j} W_{i,\ell} = n^{-1}\sum_{i=1}^n \nu_i^{-1}(\Gamma_j U_i)(\Gamma_\ell U_i)\|\Gamma U_i\|^{-3} \\ &= n^{-1}p^{-3/2}\sum_{i=1}^n \nu_i^{-1}(\Gamma_j U_i)(\Gamma_\ell U_i) \\ &\quad + n^{-1}\sum_{i=1}^n \nu_i^{-1}(\Gamma_j U_i)(\Gamma_\ell U_i)\left(\|\Gamma U_i\|^{-3} - p^{-2/3}\right), \end{aligned}$$

where the last term satisfies

$$\begin{aligned} &\left|\mathbb{E}\left\{n^{-1}\sum_{i=1}^n \nu_i^{-1}(\Gamma_j U_i)(\Gamma_\ell U_i)\left(\|\Gamma U_i\|^{-3} - p^{-2/3}\right)\right\}\right| \\ \leqslant\ & p^{-3/2}\mathbb{E}\left\{\nu_i^{-1}\left|(\Gamma_j U_i)(\Gamma_\ell U_i)\right|\|\Gamma U_i\|^{-3}\left|\|\Gamma U_i\|^3 - p^{3/2}\right|\right\} \\ =\ & p^{-3/2}\mathbb{E}\left\{R_i^{-1}\left|(\Gamma_j U_i)(\Gamma_\ell U_i)\right|\|\Gamma U_i\|^{-2}\left|\|\Gamma U_i\|^3 - p^{3/2}\right|\right\} \\ =\ & p^{-3/2}\mathbb{E}\left\{R_i^{-1}\left|(\Gamma_j U_i)(\Gamma_\ell U_i)\right|\|\Gamma U_i\|^{-2}\left|\|\Gamma U_i\|^3 - p^{3/2}\right|\mathbb{I}(\mathcal{A}_{2i})\right\} \\ & + p^{-3/2}\mathbb{E}\left\{R_i^{-1}\left|(\Gamma_j U_i)(\Gamma_\ell U_i)\right|\|\Gamma U_i\|^{-2}\left|\|\Gamma U_i\|^3 - p^{3/2}\right|\mathbb{I}(\mathcal{A}_{2i}^c)\right\} \\ \leqslant\ & (1-\epsilon)^{-1}p^{-5/2}\mathbb{E}\left\{R_i^{-1}\left|(\Gamma_j U_i)(\Gamma_\ell U_i)\right|\left|\|\Gamma U_i\|^3 - p^{3/2}\right|\mathbb{I}(\mathcal{A}_{2i})\right\} \\ & + p^{-3/2}\mathbb{E}\left\{R_i^{-1}\left|\|\Gamma U_i\|^3 - p^{3/2}\right|\mathbb{I}(\mathcal{A}_{2i}^c)\right\} \\ \lesssim\ & p^{-5/2}\{\mathbb{E}(R_i^{-4})\}^{1/4}\left[\mathbb{E}\left\{\left|(\Gamma_j U_i)(\Gamma_\ell U_i)\right|^4\right\}\right]^{1/4}\left\{\mathbb{E}\left(\left|\|\Gamma U_i\|^3 - p^{3/2}\right|^2\right)\right\}^{1/2} \\ & + p^{-3/2}\{\mathbb{E}(R_i^{-4})\}^{1/4}\left\{\mathbb{E}\left(\left|\|\Gamma U_i\|^3 - p^{3/2}\right|^2\right)\right\}^{1/2}\{\mathbb{P}(\mathcal{A}_{2i}^c)\}^{1/4} \\ \lesssim\ & \zeta_1 p^{-1-\delta/2}. \end{aligned}$$

It follows that

$$Q_{j\ell} = n^{-1}p^{-3/2}\sum_{i=1}^{n}\nu_i^{-1}(\Gamma_j U_i)(\Gamma_\ell U_i) + O_p(\zeta_1 p^{-1-\delta/2})\,.$$

For $i = 1, \ldots, n$, let $\mathcal{A}_{1i} = \{p - \epsilon p^{(1+\delta)/2} \leqslant \|U_i\|^2 \leqslant p + \epsilon p^{(1+\delta)/2}\}$ for a fixed $0 < \epsilon < 1$. According to Lemma A1,

$$\begin{aligned}
& \mathbb{E}\left[\left\{\Gamma_j S(U_i)S(U_i)^\top \Gamma_\ell^\top\right\}^2\right] = \mathbb{E}\left\{\|U_i\|^{-4}(\Gamma_j U_i U_i^\top \Gamma_\ell^\top)^2\right\} \\
= \;& \mathbb{E}\left\{\|U_i\|^{-4}(\Gamma_j U_i U_i^\top \Gamma_\ell^\top)^2 \mathbb{I}(\mathcal{A}_{1i})\right\} + \mathbb{E}\left\{\|U_i\|^{-4}(\Gamma_j U_i U_i^\top \Gamma_\ell^\top)^2 \mathbb{I}(\mathcal{A}_{1i}^c)\right\} \\
\lesssim \;& \{p - \epsilon p^{(1+\delta)/2}\}^{-2}\mathbb{E}\left\{(\Gamma_j U_i U_i^\top \Gamma_\ell^\top)^2\right\} + p^2\mathbb{P}(\mathcal{A}_{1i}^c) \\
\lesssim \;& \{p - \epsilon p^{(1+\delta)/2}\}^{-2} + p^2 \times c_1 \exp\{-c_2 p^{\delta\alpha/(4\alpha+4)}\} \\
\lesssim \;& p^{-2}\,.
\end{aligned}$$

Then, we can show that

$$\begin{aligned}
& n^{-1}p^{-3/2}\sum_{i=1}^{n}\nu_i^{-1}(\Gamma_j U_i)(\Gamma_\ell U_i) \\
= \;& n^{-1}p^{-1/2}\sum_{i=1}^{n}\nu_i^{-1}\Gamma_j S(U_i)S(U_i)^\top \Gamma_\ell^\top + O_p(\zeta_1 p^{-7/6})\,,
\end{aligned}$$

where the last equality is indicated by

$$\begin{aligned}
& \mathbb{E}|p^{-3/2}\nu_i^{-1}\Gamma_j S(U_i)S(U_i)^\top \Gamma_\ell^\top(\|U_i\|^2 - p)| \\
\lesssim \;& p^{-3/2}\{\mathbb{E}(\nu_i^{-3})\}^{1/3}\left(\mathbb{E}\left[\left\{\Gamma_j S(U_i)S(U_i)^\top \Gamma_\ell^\top\right\}^2\right]\right)^{1/2}\left[\mathbb{E}\left\{(\|U_i\|^2 - p)^6\right\}\right]^{1/6} \\
\lesssim \;& \zeta_1 p^{-7/6}\,.
\end{aligned}$$

Thus, we obtain that

$$Q_{j\ell} = n^{-1}p^{-1/2}\sum_{i=1}^{n}\nu_i^{-1}\Gamma_j S(U_i)S(U_i)^\top \Gamma_\ell^\top + O_p(\zeta_1 p^{-7/6} + \zeta_1 p^{-1-\delta/2})\,.$$

As ν_i and $S(U_i)$ are independent with each other, we have

$$\mathbb{E}\left\{n^{-1}p^{-1/2}\sum_{i=1}^{n}\nu_i^{-1}\Gamma_j S(U_i)S(U_i)^\top\Gamma_\ell^\top\right\} = p^{-1/2}\mathbb{E}(\nu_i^{-1})\mathbb{E}\left\{\Gamma_j S(U_i)S(U_i)^\top\Gamma_\ell^\top\right\},$$

where $\mathbb{E}(\nu_i^{-1}) \lesssim p^{1/2}\zeta_1$ from Lemma A2.

According to Lemma A1 and regarding that $\Gamma_j\Gamma_\ell^\top = \omega_{j\ell}$,

$$\begin{aligned}
& \mathbb{E}\left\{\Gamma_j S(U_i)S(U_i)^\top\Gamma_\ell^\top\right\} \\
= \; & \mathbb{E}\left(\Gamma_j U_i U_i^\top\Gamma_\ell^\top\|U_i\|^{-2}\right) = p^{-1}\mathbb{E}\left(\Gamma_j U_i U_i^\top\Gamma_\ell^\top\right) + \mathbb{E}\left\{\Gamma_j U_i U_i^\top\Gamma_\ell^\top\left(\|U_i\|^{-2}-p^{-1}\right)\right\} \\
= \; & p^{-1}\omega_{j\ell} + \mathbb{E}\left\{\Gamma_j U_i U_i^\top\Gamma_\ell^\top\left(\|U_i\|^{-2}-p^{-1}\right)\right\} \\
\leqslant \; & p^{-1}|\omega_{j\ell}| + \mathbb{E}\left(\left|\Gamma_j U_i U_i^\top\Gamma_\ell^\top\right|\left|\|U_i\|^{-2}-p^{-1}\right|\right) \\
= \; & p^{-1}|\omega_{j\ell}| + p^{-1}\mathbb{E}\left(\left|\Gamma_j U_i U_i^\top\Gamma_\ell^\top\right|\|U_i\|^{-2}\left|\|U_i\|^2-p\right|\right) \\
= \; & p^{-1}|\omega_{j\ell}| + p^{-1}\mathbb{E}\left\{\left|\Gamma_j U_i U_i^\top\Gamma_\ell^\top\right|\|U_i\|^{-2}\left|\|U_i\|^2-p\right|\mathbb{I}(\mathcal{A}_{1i})\right\} \\
& + p^{-1}\mathbb{E}\left\{\left|\Gamma_j U_i U_i^\top\Gamma_\ell^\top\right|\|U_i\|^{-2}\left|\|U_i\|^2-p\right|\mathbb{I}(\mathcal{A}_{1i}^c)\right\} \\
\lesssim \; & p^{-1}|\omega_{j\ell}| + \{p^2-\epsilon p^{(3+\delta)/2}\}^{-1}\mathbb{E}\left(\left|\Gamma_j U_i U_i^\top\Gamma_\ell^\top\right|\left|\|U_i\|^2-p\right|\right) + \mathbb{E}\left\{\left|\|U_i\|^2-p\right|\mathbb{I}(\mathcal{A}_{1i}^c)\right\} \\
\leqslant \; & p^{-1}|\omega_{j\ell}| + \{p^2-\epsilon p^{(3+\delta)/2}\}^{-1}\left[\mathbb{E}\left\{\left(\Gamma_j U_i U_i^\top\Gamma_\ell^\top\right)^2\right\}\right]^{1/2}\left[\mathbb{E}\left\{\left(\|U_i\|^2-p\right)^2\right\}\right]^{1/2} \\
& + \left[\mathbb{E}\left\{\left(\|U_i\|^2-p\right)^2\right\}\right]^{1/2}\{\mathbb{P}(\mathcal{A}_{1i}^c)\}^{1/2} \\
\leqslant \; & p^{-1}|\omega_{j\ell}| + O(p^{-3/2}) + O(p^{1/2})\times c_1^{1/2}\exp\{-c_2 p^{\delta\alpha/(4\alpha+4)}/2\} \\
\lesssim \; & p^{-1}|\omega_{j\ell}| + O(p^{-3/2}),
\end{aligned}$$

where the second last inequality is due to

$$\begin{aligned}
\mathbb{E}\left\{\left(\|U_i\|^2-p\right)^2\right\} &= \mathbb{E}(\|U_i\|^4 - 2p\|U_i\|^2 + p^2) \\
&= p\mathbb{E}(U_{i,j}^4) + p(p-1) - 2p^2 + p^2 \\
&= O(p).
\end{aligned}$$

Thus, it follows that

$$\mathbb{E}\left\{n^{-1}p^{-1/2}\sum_{i=1}^{n}\nu_i^{-1}\Gamma_j S(U_i)S(U_i)^\top\Gamma_\ell^\top\right\}\lesssim\zeta_1 p^{-1}|\omega_{jj}|+O(\zeta_1 p^{-3/2})\,.$$

Furthermore, as $\mathbb{E}(\nu_i^{-2})\lesssim p\zeta_2$, we can conclude that

$$\begin{aligned}
&\mathrm{Var}\left\{n^{-1}p^{-1/2}\sum_{i=1}^{n}\nu_i^{-1}\Gamma_j S(U_i)S(U_i)^\top\Gamma_\ell^\top\right\}\\
=\;& n^{-1}p^{-1}\mathbb{E}(\nu_i^{-2})\mathbb{E}\left[\left\{\Gamma_j S(U_i)S(U_i)^\top\Gamma_\ell^\top\right\}^2\right]\\
&-n^{-1}p^{-1}\{\mathbb{E}(\nu_i^{-1})\}^2\left[\mathbb{E}\left\{\Gamma_j S(U_i)S(U_i)^\top\Gamma_\ell^\top\right\}\right]^2\\
\lesssim\;& \zeta_1^2 n^{-1}p^{-2}.
\end{aligned}$$

It follows from the Chebychev's inequality that

$$\left|n^{-1}p^{-1/2}\sum_{i=1}^{n}\nu_i^{-1}\Gamma_j S(U_i)S(U_i)^\top\Gamma_\ell^\top\right|\lesssim\zeta_1 p^{-1}|\omega_{j\ell}|+O_p(\zeta_1 n^{-1/2}p^{-1}+\zeta_1 p^{-3/2})\,.$$

Finally, we arrive at $|Q_{j,\ell}|\lesssim\zeta_1 p^{-1}|\omega_{j\ell}|+O_p(\zeta_1 n^{-1/2}p^{-1}+\zeta_1 p^{-7/6}+\zeta_1 p^{-1-\delta/2})$.

(ii) From the proof of part (i), we know that $Q_{j\ell}=Q_{0,j\ell}+O_p(\zeta_1 p^{-7/6}+\zeta_1 p^{-1-\delta/2})$, where $Q_{0,j\ell}$ is the (j,ℓ)th component of the random matrix $Q_0=n^{-1}p^{-1/2}\sum_{i=1}^{n}\nu_i^{-1}\{\Gamma S(U_i)\}\{\Gamma S(U_i)\}^\top$. In addition, $\mathbb{E}\left\{\Gamma_j S(U_i)S(U_i)^\top\Gamma_\ell^\top\right\}\lesssim p^{-1}|\omega_{j\ell}|+O(p^{-3/2})$. It follows that

$$\begin{aligned}
\mathrm{tr}\left\{\left(\mathbb{E}\left[\{\Gamma S(U_i)\}\{\Gamma S(U_i)\}^\top\right]\right)^2\right\} &= \sum_{j=1}^{p}\sum_{\ell=1}^{p}\left[\mathbb{E}\left\{\Gamma_j S(U_i)S(U_i)^\top\Gamma_\ell^\top\right\}\right]^2\\
&\lesssim p^{-2}\sum_{j=1}^{p}\sum_{\ell=1}^{p}|\omega_{j\ell}|^2+p^{-5/2}\sum_{j=1}^{p}\sum_{\ell=1}^{p}|\omega_{j\ell}|+p^{-1}\\
&\lesssim p^{-1}a_0(p)+p^{-3/2}a_0(p)+p^{-1}\\
&\lesssim p^{-\delta}\,.
\end{aligned}$$

This implies that

$$\mathrm{tr}[\{\mathbb{E}(Q_0)\}^2] = p^{-1}\{\mathbb{E}(\nu_i^{-1})\}^2\mathrm{tr}\left\{\left(\mathbb{E}\left[\{\Gamma S(U_i)\}\{\Gamma S(U_i)\}^\top\right]\right)^2\right\} \quad \lesssim \quad p^{-1-\delta}$$

and

$$\begin{aligned}
\mathbb{E}\{\mathrm{tr}(Q_0^2)\} &= n^{-1}p^{-1}\mathrm{tr}\left(\mathbb{E}\left[\nu_i^{-2}\{\Gamma S(U_i)\}\{\Gamma S(U_i)\}^\top\{\Gamma S(U_i)\}\{\Gamma S(U_i)\}^\top\right]\right) \\
&\quad +(1-n^{-1})p^{-1}\mathrm{tr}\left\{\left(\mathbb{E}\left[\nu_i^{-1}\{\Gamma S(U_i)\}\{\Gamma S(U_i)\}^\top\right]\right)^2\right\} \\
&= n^{-1}p^{-1}\mathbb{E}(\nu_i^{-2})\mathbb{E}\left\{\|\Gamma S(U_i)\|^4\right\} \\
&\quad +(1-n^{-1})p^{-1}\{\mathbb{E}(\nu_i^{-1})\}^2\mathrm{tr}\left\{\left(\mathbb{E}\left[\{\Gamma S(U_i)\}\{\Gamma S(U_i)\}^\top\right]\right)^2\right\} \\
&= O(n^{-1}p^{-1}) + \mathrm{tr}[\{\mathbb{E}(Q_0)\}^2](1-n^{-1})\,.
\end{aligned}$$

Thus, we have

$$\mathrm{tr}[\mathbb{E}(Q_0^2) - \{\mathbb{E}(Q_0)\}^2] = O(n^{-1}p^{-1})\,.$$

We complete the proof of this lemma. □

Proof of Lemma A4. Recall that Γ_j is the jth row of Γ, and denote $\Gamma_{j\ell}$ to be the (j,ℓ)th element of Γ, then $\Gamma_j U_i = \sum_{\ell=1}^p \Gamma_{j\ell}U_{i,\ell}$. It is noticed that $\omega_{j\ell} = \sum_{j_1=1}^p \Gamma_{jj_1}\Gamma_{\ell j_1}$, then

$$\mathrm{Var}(\Gamma_j U_i) = \sum_{\ell=1}^p \Gamma_{j\ell}^2 = \omega_{jj}$$

and

$$\begin{aligned}
\mathbb{E}\{(\Gamma_j U_i)^4\} &= \mathbb{E}\left\{\left(\sum_{\ell=1}^p \Gamma_{j\ell}U_{i,\ell}\right)^4\right\} = \sum_{\ell=1}^p \Gamma_{j\ell}^4\mathbb{E}(U_{i,\ell}^4) + 6\sum_{1\leqslant \ell_1\neq \ell_2\leqslant p}\Gamma_{j\ell_1}^2\Gamma_{j\ell_2}^2\mathbb{E}(U_{i,\ell_1}^2)\mathbb{E}(U_{i,\ell_2}^2) \\
&\lesssim \omega_{jj}^2\,.
\end{aligned}$$

(i) For $i = 1, \ldots, n$, let $\mathcal{A}_{2i} = \{(1-\epsilon)\mathrm{tr}(\Omega) \leqslant \|\Gamma U_i\|^2 \leqslant (1+\epsilon)\mathrm{tr}(\Omega)\}$ for a fixed $0 < \epsilon < 1$, then

$$\mathbb{P}(\mathcal{A}_{2i}) \geqslant 1 - c_1 \exp\{-c_2 p^{\delta\alpha/(4\alpha+4)}\}$$

according to the proof of Lemma A1. It follows that

$$\begin{aligned}
\mathbb{E}(W_{i,j}^4) &= \mathbb{E}\{\|\Gamma U_i\|^{-4}(\Gamma_j U_i)^4\} \\
&= \mathbb{E}\{\|\Gamma U_i\|^{-4}(\Gamma_j U_i)^4 \mathbb{I}(\mathcal{A}_{2i})\} + \mathbb{E}\{\|\Gamma U_i\|^{-4}(\Gamma_j U_i)^4 \mathbb{I}(\mathcal{A}_{2i}^c)\} \\
&\leqslant \{(1-\epsilon)\mathrm{tr}(\Omega)\}^{-2}\mathbb{E}\{(\Gamma_j U_i)^4\} + \mathbb{P}(\mathcal{A}_{2i}^c) \\
&\lesssim \omega_{jj}^2\{(1-\epsilon)\mathrm{tr}(\Omega)\}^{-2} + c_1 \exp\{-c_2 p^{\delta\alpha/(4\alpha+4)}\} \\
&\lesssim \omega_{jj}^2\{\mathrm{tr}(\Omega)\}^{-2}
\end{aligned}$$

and

$$\begin{aligned}
\mathbb{E}(W_{i,j}^2) &\geqslant \mathbb{E}\{\|\Gamma U_i\|^{-2}(\Gamma_j U_i)^2 \mathbb{I}(\mathcal{A}_{2i})\} \\
&\geqslant \{(1+\epsilon)\mathrm{tr}(\Omega)\}^{-1}\mathbb{E}\{(\Gamma_j U_i)^2 \mathbb{I}(\mathcal{A}_{2i})\} \\
&= \{(1+\epsilon)\mathrm{tr}(\Omega)\}^{-1}\mathbb{E}\{(\Gamma_j U_i)^2\} - \{(1+\epsilon)\mathrm{tr}(\Omega)\}^{-1}\mathbb{E}\{(\Gamma_j U_i)^2 \mathbb{I}(\mathcal{A}_{2i}^c)\} \\
&\geqslant \{(1+\epsilon)\mathrm{tr}(\Omega)\}^{-1}\mathbb{E}\{(\Gamma_j U_i)^2\} - \{(1+\epsilon)\mathrm{tr}(\Omega)\}^{-1}[\mathbb{E}\{(\Gamma_j U_i)^4\}]^{1/2}\{\mathbb{P}(\mathcal{A}_{2i}^c)\}^{1/2} \\
&\gtrsim \omega_{jj}\{(1+\epsilon)\mathrm{tr}(\Omega)\}^{-1} - \{(1+\epsilon)\mathrm{tr}(\Omega)\}^{-1} \times \omega_{jj} \times c_1^{1/2} \exp\{-c_2 p^{\delta\alpha/(4\alpha+4)}/2\} \\
&\gtrsim \omega_{jj}\{\mathrm{tr}(\Omega)\}^{-1},
\end{aligned}$$

from which we conclude that

$$\mathbb{E}\{(\zeta_1^{-1}W_{i,j})^4\} \lesssim \zeta_1^{-4}p^{-2}\omega_{jj}^2 \lesssim \overline{M}^2$$

and

$$\mathbb{E}\{(\zeta_1^{-1}W_{i,j})^2\} \gtrsim \zeta_1^{-2}p^{-1}\omega_{jj} \gtrsim \underline{m}.$$

(ii) Similar to the proof of part (i), for any $\varrho \geqslant 1$,

$$\begin{aligned}
\mathbb{E}\left\{|\zeta_1^{-1}W_{i,j}|^{\varrho}\right\} &= \mathbb{E}\left\{|\zeta_1^{-1}W_{i,j}|^{\varrho}\mathbb{I}(\mathcal{A}_{1i})\right\} + \mathbb{E}\left\{|\zeta_1^{-1}W_{i,j}|^{\varrho}\mathbb{I}(\mathcal{A}_{1i}^c)\right\} \\
&\lesssim \zeta_1^{-\varrho}\{\mathrm{tr}(\Omega)\}^{-\varrho/2}\mathbb{E}\{|\Gamma_j U_i|^{\varrho}\} + \zeta_1^{-\varrho}\mathbb{P}(\mathcal{A}_{1i}^c) \\
&\lesssim \mathbb{E}\{|\Gamma_j U_i|^{\varrho}\} + p^{\varrho/2}\exp\{-c_2 p^{\delta\alpha/(4\alpha+4)}\}.
\end{aligned}$$

Since $\max_{1\leqslant j\leqslant p}\|U_{i,j}\|_{\psi_\alpha} \leqslant c_0$ for some constant c_0, we have $\|\Gamma_j U_i\|_{\psi_\alpha} \lesssim c_0$ according to Lemma B.4 in Koike (2021). Then, we known that $\mathbb{E}\{|\Gamma_j U_i|^{\varrho}\} \lesssim \varrho^{\varrho/\alpha}$ for any $\varrho \geqslant 1$ by the equivalent sub-exponential properties (Koike, 2021). Therefore,

$$\mathbb{E}\left\{|\zeta_1^{-1}W_{i,j}|^{\varrho}\right\} \lesssim \varrho^{\varrho/\alpha}$$

for any $\varrho \geqslant 1$ for sufficient large p, which indicates that $\zeta_1^{-1}W_{i,j}$ is sub-exponential, and thus $\|\zeta_1^{-1}W_{i,j}\|_{\psi_\alpha} \lesssim \bar{B}$.

(iii) By simple algebra,

$$\begin{aligned}
\mathbb{E}(W_{i,j}^2) &= p^{-1}\mathbb{E}\{(\Gamma_j U_i)^2\} + \mathbb{E}\{(\Gamma_j U_i)^2(\|\Gamma U_i\|^{-2} - p^{-1})\} \\
&= p^{-1}\omega_{jj} + \mathbb{E}\{(\Gamma_j U_i)^2(\|\Gamma U_i\|^{-2} - p^{-1})\},
\end{aligned}$$

where $\mathbb{E}\{(\Gamma_j U_i)^2(\|\Gamma U_i\|^{-2}-p^{-1})\}$ satisfies

$$\begin{aligned}
&\left|\mathbb{E}\{(\Gamma_j U_i)^2(\|\Gamma U_i\|^{-2}-p^{-1})\}\right| \\
\leqslant\ & p^{-1}\mathbb{E}\{(\Gamma_j U_i)^2\|\Gamma U_i\|^{-2}|\|\Gamma_j U_i\|^2-p|\} \\
=\ & p^{-1}\mathbb{E}\{(\Gamma_j U_i)^2\|\Gamma U_i\|^{-2}|\|\Gamma_j U_i\|^2-p|\mathbb{I}(\mathcal{A}_{2i})\}+p^{-1}\mathbb{E}\{(\Gamma_j U_i)^2\|\Gamma U_i\|^{-2}|\|\Gamma_j U_i\|^2-p|\mathbb{I}(\mathcal{A}_{2i}^c)\} \\
\leqslant\ & p^{-1}\{(1-\epsilon)\mathrm{tr}(\Omega)\}^{-1}\mathbb{E}\{(\Gamma_j U_i)^2|\|\Gamma_j U_i\|^2-p|\}+p^{-1}\mathbb{E}\{\|\Gamma_j U_i\|^2-p|\mathbb{I}(\mathcal{A}_{2i}^c)\} \\
\leqslant\ & p^{-2}(1-\epsilon)^{-1}[\mathbb{E}\{(\Gamma_j U_i)^4\}]^{1/2}\{\mathbb{E}(|\|\Gamma_j U_i\|^2-p|^2)\}^{1/2} \\
& +p^{-1}\{\mathbb{E}(|\|\Gamma_j U_i\|^2-p|^2)\}^{1/2}\{\mathbb{P}(\mathcal{A}_{2i}^c)\}^{1/2} \\
\lesssim\ & p^{-2}\times p^{1-\delta/2}+p^{-1}\times p^{1-\delta/2}\times c_1^{1/2}\exp\{-c_2 p^{\delta\alpha/(4\alpha+4)}/2\} \\
\lesssim\ & p^{-1-\delta/2}.
\end{aligned}$$

In addition, for $1\leqslant j\neq \ell\leqslant p$, we have

$$\begin{aligned}
\mathbb{E}(W_{i,j}W_{i,\ell}) &= p^{-1}\mathbb{E}\{(\Gamma_j U_i)(\Gamma_\ell U_i)\}+\mathbb{E}\{(\Gamma_j U_i)(\Gamma_\ell U_i)(\|\Gamma U_i\|^{-2}-p^{-1})\} \\
&= p^{-1}\omega_{j\ell}+\mathbb{E}\{(\Gamma_j U_i)(\Gamma_\ell U_i)(\|\Gamma U_i\|^{-2}-p^{-1})\},
\end{aligned}$$

where $\mathbb{E}\{(\Gamma_j U_i)(\Gamma_\ell U_i)(\|\Gamma U_i\|^{-2} - p^{-1})\}$ satisfies

$$\begin{aligned}
&\big|\mathbb{E}\{(\Gamma_j U_i)(\Gamma_\ell U_i)(\|\Gamma U_i\|^{-2} - p^{-1})\}\big| \\
\leqslant\ & p^{-1}\mathbb{E}\{|(\Gamma_j U_i)(\Gamma_\ell U_i)|\|\Gamma U_i\|^{-2}|\|\Gamma_j U_i\|^2 - p|\} \\
=\ & p^{-1}\mathbb{E}\{|(\Gamma_j U_i)(\Gamma_\ell U_i)|\|\Gamma U_i\|^{-2}|\|\Gamma_j U_i\|^2 - p|\mathbb{I}(\mathcal{A}_{2i})\} \\
& + p^{-1}\mathbb{E}\{|(\Gamma_j U_i)(\Gamma_\ell U_i)|\|\Gamma U_i\|^{-2}|\|\Gamma_j U_i\|^2 - p|\mathbb{I}(\mathcal{A}_{2i}^c)\} \\
\leqslant\ & p^{-1}\{(1-\epsilon)\mathrm{tr}(\Omega)\}^{-1}\mathbb{E}\{|(\Gamma_j U_i)(\Gamma_\ell U_i)||\|\Gamma_j U_i\|^2 - p|\} + p^{-1}\mathbb{E}\{\|\Gamma_j U_i\|^2 - p|\mathbb{I}(\mathcal{A}_{2i}^c)\} \\
\leqslant\ & p^{-2}(1-\epsilon)^{-1}[\mathbb{E}\{|(\Gamma_j U_i)(\Gamma_\ell U_i)|^2\}]^{1/2}\{\mathbb{E}(|\|\Gamma_j U_i\|^2 - p|^2)\}^{1/2} \\
& + p^{-1}\{\mathbb{E}(|\|\Gamma_j U_i\|^2 - p|^2)\}^{1/2}\{\mathbb{P}(\mathcal{A}_{2i}^c)\}^{1/2} \\
\lesssim\ & p^{-2} \times p^{1-\delta/2} + p^{-1} \times p^{1-\delta/2} \times c_1^{1/2}\exp\{-c_2 p^{\delta\alpha/(4\alpha+4)}/2\} \\
\lesssim\ & p^{-1-\delta/2}.
\end{aligned}$$

(iv) According to part (ii), $\zeta_1^{-1}W_1, \ldots, \zeta_1^{-1}W_n$ are i.i.d. p-dimensional random vectors satisfies $\|\zeta_1^{-1}W_{i,j}\|_{\psi_\alpha} \lesssim \bar{B}$ for all $i = 1, \ldots, n$ and $j = 1, \ldots, p$. By Lemma 2.2.2 of van der Vaart and Wellner (1996),

$$\Big\|\max_{1\leqslant i\leqslant n}\max_{1\leqslant j\leqslant p}|\zeta_1^{-1}W_{i,j}|\Big\|_{\psi_\alpha} \lesssim \log^{1/\alpha}(np)\,.$$

Similar to the proof of part (i), we can show that

$$\begin{aligned}
\mathbb{E}\{(\zeta_1^{-1}W_{i,j})^2\} &= \zeta_1^{-2}\mathbb{E}\{\|\Gamma U_i\|^{-2}(\Gamma_j U_i)^2\mathbb{I}(\mathcal{A}_{1i})\} \\
&\quad + \zeta_1^{-2}\mathbb{E}\{\|\Gamma U_i\|^{-4}(\Gamma_j U_i)^4\mathbb{I}(\mathcal{A}_{1i}^c)\} \\
&\leqslant \zeta_1^{-2}\{(1+\epsilon)\mathrm{tr}(\Omega)\}^{-1}\mathbb{E}\{(\Gamma_j U_i)^2\} + \zeta_1^{-2}\mathbb{E}\{\mathbb{I}(\mathcal{A}_{1i}^c)\} \\
&\leqslant \zeta_1^{-2}\omega_{jj}\{(1+\epsilon)\mathrm{tr}(\Omega)\}^{-1} + \zeta_1^{-2}c_1\exp\{-c_2 p^{\delta/(4+4\alpha)}\} \\
&= \zeta_1^{-2}\omega_{jj}\{(1+\epsilon)\mathrm{tr}(\Omega)\}^{-1}\{1+o(1)\}\,.
\end{aligned}$$

It follows that

$$\max_{1\leqslant j\leqslant p}\sum_{i=1}^{n}\mathbb{E}\{(\zeta_1^{-1}W_{i,j})^2\} \quad\lesssim\quad \max_{1\leqslant j\leqslant p}\sum_{i=1}^{n}\zeta_1^{-2}\omega_{jj}\{(1+\epsilon)\mathrm{tr}(\Omega)\}^{-1}\lesssim n\max_{1\leqslant j\leqslant p}\omega_{jj}\leqslant \overline{M}n\,,$$

Applying Lemma E.1 of Chernozhukov et al. (2017), it holds that with $\alpha\geqslant 1$ and $n^{-1/2}\log^{3/2}(np)\lesssim 1$,

$$\begin{aligned}\mathbb{E}\left(\left|n^{-1/2}\sum_{i=1}^{n}\zeta_1^{-1}W_i\right|_\infty\right) &\lesssim n^{-1/2}\{n^{1/2}\log^{1/2}(p)+\log^{1/\alpha}(np)\log(p)\}\\ &\lesssim \log^{1/2}(np)\,.\end{aligned}$$

From the properties of the ψ_α norm, it holds that

$$\left\|\max_{1\leqslant i\leqslant n,1\leqslant j\leqslant p}|\zeta_1^{-1}W_{i,j}|^2\right\|_{\psi_{\alpha/2}}\lesssim \log^2(np).$$

According to Lemma E.3 of Chernozhukov et al. (2017), we have that

$$\mathbb{E}\left(\left|n^{-1}\sum_{i=1}^{n}(\zeta_1^{-1}W_i)^2\right|_\infty\right) \quad\lesssim\quad n^{-1}\{\overline{M}n+\log^2(np)\log(p)\}\lesssim \overline{M}\,.$$

We finish the proof of this lemma. □

Proof of Lemma A5. Let $\tilde{X}_i = X_i - \hat{\boldsymbol{\theta}}_n$ and $\tilde{R}_i = \|\tilde{X}_i\|$ for $i=1,\ldots,n$. According to the proof of Lemma 1, $\|\hat{\boldsymbol{\theta}}_n\| = O_p(\zeta_1^{-1}n^{-1/2})$ and $\max_{1\leqslant i\leqslant n}R_i^{-1}=O_p(\zeta_1 n^{1/4})$. Then $R_i^{-1}\|\hat{\boldsymbol{\theta}}_n\|$ satisfies

$$R_i^{-1}\|\hat{\boldsymbol{\theta}}_n\| = O_p(n^{-1/2}) \quad\text{and}\quad \max_{1\leqslant i\leqslant n}R_i^{-1}\|\hat{\boldsymbol{\theta}}_n\| = O_p(n^{-1/4})\,.$$

As $\tilde{R}_i^{-1} = R_i^{-1}\|W_i - R_i^{-1}\hat{\boldsymbol{\theta}}_n\|^{-1} = R_i^{-1}\left(1-2R_i^{-1}W_i^\top\hat{\boldsymbol{\theta}}_n + R_i^{-2}\|\hat{\boldsymbol{\theta}}_n\|^2\right)^{-1/2}$, by Taylor expansion,

$$\tilde{R}_i^{-1} = R_i^{-1}\left(1+R_i^{-1}W_i^\top\hat{\boldsymbol{\theta}}_n - 2^{-1}R_i^{-2}\|\hat{\boldsymbol{\theta}}_n\|^2+\tilde{\delta}_{1i}\right),$$

where $\tilde{\delta}_{1i}$ satisfies $\tilde{\delta}_{1i} = O_p(n^{-1})$ and $\max_{1\leqslant i\leqslant n} \tilde{\delta}_{1i} = O_p(n^{-1/2})$. It follows that

$$\tilde{R}_i^{-1} = R_i^{-1}(1 + \tilde{\delta}_{2i}) \, ,$$

where $\tilde{\delta}_{2i} = R_i^{-1} W_i^\top \hat{\boldsymbol{\theta}}_n - 2^{-1} R_i^{-2} \|\hat{\boldsymbol{\theta}}_n\|^2 + \tilde{\delta}_{1i}$ satisfies $\tilde{\delta}_{2i} = O_p(n^{-1/2})$ and $\max_{1\leqslant i\leqslant n} \tilde{\delta}_{2i} = O_p(n^{-1/4})$. Thus,

$$\tilde{R}_i^{-1} = O_p(\zeta_1) \quad \text{and} \quad \max_{1\leqslant i\leqslant n} \tilde{R}_i^{-1} = O_p(\zeta_1 n^{1/4}) \, .$$

Denote $\tilde{W}_i = \tilde{X}_i/\|\tilde{X}_i\|$ for $i = 1, \ldots, n$. Then,

$$\begin{aligned} \tilde{W}_i &= \tilde{R}_i^{-1}(X_i - \hat{\boldsymbol{\theta}}_n) \\ &= R_i^{-1}(X_i - \hat{\boldsymbol{\theta}}_n)(1 + \tilde{\delta}_{2i}) \\ &= (W_i - R_i^{-1}\hat{\boldsymbol{\theta}}_n)(1 + \tilde{\delta}_{2i}) \, . \end{aligned}$$

We first show that $\|\tilde{\boldsymbol{\theta}}_n\| = O_p(\zeta_1^{-1} n^{-1/2})$. It is noticed that $\tilde{\boldsymbol{\theta}}_n$ minimizes

$$L_n^*(\boldsymbol{\beta}) = \sum_{i=1}^n \|Z_i \tilde{X}_i - \boldsymbol{\beta}\| \, ,$$

which is a strictly convex function of $\boldsymbol{\beta}$. Thus, if we can show that $L_n^*(\boldsymbol{\beta})$ has a $\zeta_1 n^{1/2}$-consistent local minimizer, then this local minimizer must be a $\zeta_1 n^{1/2}$-consistent global minimizer of $L_n^*(\boldsymbol{\beta})$. The existence of a $\zeta_1 n^{1/2}$-consistent local minimizer is implied by the fact that for an arbitrarily small $\varepsilon > 0$, there exists a constant C_0, which does not depend on n and p, such that

$$\liminf_n \mathbb{P}\left\{ \inf_{q\in\mathbb{R}^p, \ \|q\|=C_0} \ L_n^*(\zeta_1^{-1} n^{-1/2} q) > L_n^*(0) \right\} > 1 - \varepsilon, \tag{S.15}$$

Since $|Z_i| = 1$, we rewrite $\|Z_i\tilde{X}_i - \zeta_1^{-1}n^{-1/2}q\|$ as

$$\begin{aligned} & \|Z_i\tilde{X}_i - \zeta_1^{-1}n^{-1/2}q\| \\ = \quad & \tilde{R}_i\left(1 - 2\zeta_1^{-1}n^{-1/2}\tilde{R}_i^{-1}Z_iq^\top\tilde{W}_i + \zeta_1^{-2}n^{-1}\tilde{R}_i^{-2}\|q\|^2\right)^{1/2}. \end{aligned}$$

As $|\zeta_1^{-1}n^{-1/2}\tilde{R}_i^{-1}Z_iq^T\tilde{W}_i| = O_p(n^{-1/2})$ and $\zeta_1^{-2}n^{-1}\tilde{R}_{1i}^{-2}\|q\|^2 = O_p(n^{-1})$, by Taylor expansion, we obtain that

$$\begin{aligned} & \|Z_i\tilde{X}_i - \zeta_1^{-1}n^{-1/2}q\| \\ = \quad & \tilde{R}_i - \zeta_1^{-1}n^{-1/2}Z_iq^\top\tilde{W}_i + 2^{-1}\zeta_1^{-2}n^{-1}\tilde{R}_i^{-1}\|q\|^2 \\ & -2^{-1}\zeta_1^{-2}n^{-1}\tilde{R}_i^{-1}q^\top\tilde{W}_i\tilde{W}_i^\top q + O_p(\zeta_1^{-1}n^{-3/2}). \end{aligned}$$

Then,

$$\begin{aligned} & \zeta_1\left\{L_n^*(\zeta_1^{-1}n^{-1/2}q) - L_n^*(0)\right\} \\ = \quad & \zeta_1\sum_{i=1}^n\left(\|Z_i\tilde{X}_i - \zeta_1^{-1}n^{-1/2}q\| - \|\tilde{X}_i\|\right) \\ = \quad & -n^{-1/2}q^\top\left(\sum_{i=1}^n Z_i\tilde{W}_i\right) + 2^{-1}\zeta_1^{-1}n^{-1}\|q\|^2\sum_{i=1}^n\tilde{R}_i^{-1} \\ & -2^{-1}\zeta_1^{-1}n^{-1}q^\top\left(\sum_{i=1}^n\tilde{R}_i\tilde{W}_i\tilde{W}_i^\top\right)q + O_p(n^{-1/2}). \end{aligned} \tag{S.16}$$

As $\mathbb{E}^*\left(n^{-1/2}\sum_{i=1}^n Z_i\tilde{W}_i\right) = 0$ and

$$\mathbb{E}^*\left(\left\|n^{-1/2}\sum_{i=1}^n Z_i\tilde{W}_i\right\|^2\right) = n^{-1}\sum_{i=1}^n\tilde{W}_i^\top\tilde{W}_i = 1,$$

we obtain that

$$\left|n^{-1/2}q^\top\sum_{i=1}^n Z_i\tilde{W}_i\right| \leqslant \|q\|\left\|n^{-1/2}\sum_{i=1}^n Z_i\tilde{W}_i\right\| = O_p(\|q\|).$$

In the meanwhile, as $\zeta_1^{-1}n^{-1}\sum_{i=1}^n R_i^{-1}=1+O_p(n^{-1/2})$, we have

$$\begin{aligned}\zeta_1^{-1}n^{-1}\|q\|^2\sum_{i=1}^n \tilde{R}_i^{-1} &= \zeta_1^{-1}n^{-1}\|q\|^2\sum_{i=1}^n R_i^{-1}(1+\tilde{\delta}_{2i})\\ &= \|q\|^2\{1+O_p(n^{-1/4})\}\,.\end{aligned}$$

Simple algebra yields

$$\begin{aligned}&n^{-1}\sum_{i=1}^n \tilde{R}_i^{-1}\tilde{W}_i\tilde{W}_i^\top\\ =\ &n^{-1}\sum_{i=1}^n R_i^{-1}(W_i-R_i^{-1}\hat{\boldsymbol{\theta}}_n)(W_i-R_i^{-1}\hat{\boldsymbol{\theta}}_n)^\top(1+\tilde{\delta}_{2i})\\ =\ &n^{-1}\sum_{i=1}^n R_iW_iW_i^\top(1+\tilde{\delta}_{2i})-2n^{-1}\sum_{i=1}^n R_i^{-2}W_i\hat{\boldsymbol{\theta}}_n^\top(1+\tilde{\delta}_{2i})\\ &+n^{-1}\sum_{i=1} R_i^{-3}\hat{\boldsymbol{\theta}}_n\hat{\boldsymbol{\theta}}_n^\top(1+\tilde{\delta}_{2i})\,.\end{aligned}$$

Similar to the proof of Lemma 1.2 in Cheng et al. (2019) and utilizing the results on $Q=n^{-1}\sum_{i=1}^n R_i^{-1}W_iW_i^{-1}$ in Lemma A3, we can show that $n^{-1}q^\top\sum_{i=1}^n R_iW_iW_i^\top q(1+\tilde{\delta}_{2i})=O_p(\zeta_1n^{-1/2}+\zeta_1p^{-(1/6\wedge\delta/2)})$. In addition, as

$$n^{-1}\sum_{i=1}^n R_i^{-2}q^\top W_i\leqslant n^{-1}\sum_{i=1}^n R_i^{-2}\|q\|\|W_i\|=\|q\|n^{-1}\sum_{i=1}^n R_i^{-2}=O_p(\zeta_1^2)$$

and $n^{-1}\sum_{i=1}^n R_i^{-3}=\zeta_3\{1+o_p(1)\}$, we have

$$\begin{aligned}&n^{-1}q^\top\sum_{i=1}^n R_i^{-2}W_i\hat{\boldsymbol{\theta}}_n^\top q(1+\tilde{\delta}_{2i})\\ =\ &n^{-1}\sum_{i=1}^n R_i^{-2}q^\top W_i(1+\tilde{\delta}_{2i})(\hat{\boldsymbol{\theta}}_n^\top q)=O_p(\zeta_1n^{-1/2})\,.\end{aligned}$$

and

$$n^{-1}q^\top\sum_{i=1} R_i^{-3}\hat{\boldsymbol{\theta}}_n\hat{\boldsymbol{\theta}}_n^\top q(1+\tilde{\delta}_{2i})=n^{-1}\sum_{i=1} R_i^{-3}(1+\tilde{\delta}_{2i})\|q^\top\hat{\boldsymbol{\theta}}_n\|^2=O_p(\zeta_1n^{-1})\,.$$

Thus, we obtain

$$
\begin{aligned}
& 2^{-1}\zeta_1^{-1}n^{-1}\|q\|^2\sum_{i=1}^n \tilde{R}_i^{-1} + 2^{-1}\zeta_1^{-1}n^{-1}q^\top\left(\sum_{i=1}^n \tilde{R}_i\tilde{W}_i\tilde{W}_i^\top\right)q \\
= \quad & 2^{-1}\|q\|^2 + O_p(n^{-1/4}+p^{-\delta})\,.
\end{aligned}
$$

Choosing a sufficient large constant C_0, the second term dominates the first term in (S.16) and thus $\zeta_1\left\{L_n^*(\zeta_1^{-1}n^{-1/2}q) - L_n^*(0)\right\} > 0$. Hence, we have $\|\tilde{\boldsymbol{\theta}}_n\| = O_p(\zeta_1^{-1}n^{-1/2})$.

Denote $\Theta_i = Z_i\hat{\boldsymbol{\theta}}_n + \tilde{\boldsymbol{\theta}}_n$ for $i = 1,\ldots,n$. Then

$$
\max_{1\leqslant i\leqslant n}\|\Theta_i\| \leqslant \|\hat{\boldsymbol{\theta}}_n\| + \|\tilde{\boldsymbol{\theta}}_n\| = O_p(\zeta_1^{-1}n^{-1/2}).
$$

Recall that $\tilde{\boldsymbol{\theta}}_n$ satisfies

$$
\sum_{i=1}^n \frac{Z_i\tilde{X}_i - \tilde{\boldsymbol{\theta}}_n}{\|Z_i\tilde{X}_i - \tilde{\boldsymbol{\theta}}_n\|} = \sum_{i=1}^n \frac{Z_iW_i - R_i^{-1}\Theta_i}{\|Z_iW_i - R_i^{-1}\Theta_i\|} = 0\,,
$$

which is equivalently to

$$
n^{-1}\sum_{i=1}^n (Z_iW_i - R_i^{-1}\Theta_i)\left(1 - 2Z_iR_i^{-1}W_i^\top\Theta_i + R_i^{-2}\|\Theta_i\|^2\right)^{-1/2} = 0\,,
$$

where $|R_i^{-1}W_i^\top\Theta_i| = O_p(n^{-1/2})$, $R_i^{-2}\|\Theta_i\|^2 = O_p(n^{-1})$,

$$
\max_{1\leqslant i\leqslant n}|R_i^{-1}W_i^\top\Theta_i| = O_p(n^{-1/4}) \quad \text{and} \quad \max_{1\leqslant i\leqslant n}R_i^{-2}\|\Theta_i\|^2 = O_p(n^{-1/2})\,.
$$

Taylor expansion leads to

$$
n^{-1}\sum_{i=1}^n (Z_iW_i - R_i^{-1}\Theta_i)(1 + Z_iR_i^{-1}W_i^\top\Theta_i - 2R_i^{-2}\|\Theta_i\|^2 + \tilde{\delta}_{3i}) = 0
$$

where $\delta_{3i} = O_p\{(Z_i R_i^{-1} W_i^\top \Theta_i - R_i^{-2}\|\Theta_i\|^2)^2\} = O_p(n^{-1})$, and $\max_{1\leqslant i\leqslant n} \delta_{3i} = O_p(n^{-1/2})$. Then,

$$
\begin{aligned}
& n^{-1}\sum_{i=1}^n Z_i W_i(1-2R_i^{-2}\|\Theta_i\|^2+\tilde{\delta}_{3i}) + n^{-1}\sum_{i=1}^n R_i^{-1}(W_i^\top\Theta_i)W_i \\
= \;& n^{-1}\sum_{i=1}^n Z_i W_i(1-2R_i^{-2}\|\Theta_i\|^2+\tilde{\delta}_{3i}) + n^{-1}\sum_{i=1}^n Z_i R_i^{-1} W_i W_i^\top \hat{\boldsymbol{\theta}}_n \\
& + n^{-1}\sum_{i=1}^n R_i^{-1} W_i W_i^\top \tilde{\boldsymbol{\theta}}_n \\
= \;& n^{-1}\sum_{i=1}^n R_i^{-1}\Theta_i(1+\tilde{\delta}_{3i}+\tilde{\delta}_{4i}) \\
= \;& n^{-1}\sum_{i=1}^n R_i^{-1}\tilde{\boldsymbol{\theta}}_n(1+\tilde{\delta}_{3i}+\tilde{\delta}_{4i}) + n^{-1}\sum_{i=1}^n Z_i R_i^{-1}\hat{\boldsymbol{\theta}}_n(1+\tilde{\delta}_{3i}+\tilde{\delta}_{4i}),
\end{aligned}
$$

where $\tilde{\delta}_{4i} = Z_i R_i^{-1} W_i^\top \Theta_i - 2R_i^{-2}\|\Theta_i\|^2 = O_p(\tilde{\delta}_{3i}^{1/2})$ satisfies $\max_{1\leqslant i\leqslant n} \tilde{\delta}_{4i} = O_p(n^{-1/4})$.

From the proof of Lemma 1, we know that $|\hat{\boldsymbol{\theta}}|_\infty = O_p\{n^{-1/2}\log^{1/2}(np)\}$. In addition, as $\mathbb{E}^*\left(n^{-1}\sum_{i=1}^n Z_i R_i^{-1}\right) = 0$ and $\mathbb{E}^*\left\{\left(n^{-1}\sum_{i=1}^n Z_i R_i^{-1}\right)^2\right\} = n^{-2}\sum_{i=1}^n R_i^{-2} = O_p(n^{-1}\zeta_2)$, we have $n^{-1}\sum_{i=1}^n Z_i R_i^{-1} = O_p(\zeta_1 n^{-1/2})$.

As Z_i is bounded, it is straightforward to show that $|n^{-1/2}\sum_{i=1}^n Z_i W_i|_\infty = O_p\{p^{-1/2}\log^{1/2}(np)\}$ similar as in the proof of Lemma A4 (iii). Thus, similar to the proof of Lemma 1, we obtain that

$$
|\tilde{\boldsymbol{\theta}}|_\infty = O_p\{n^{-1/2}\log^{1/2}(np)\}
$$

and

$$
\begin{aligned}
& \left| n^{-1}\sum_{i=1}^n R_i^{-1} W_i W_i^\top \tilde{\boldsymbol{\theta}}_n \right|_\infty \\
= \;& O_p\{\zeta_1 n^{-1/2} p^{-(1/6\wedge\delta/2)}\log^{1/2}(np) + \zeta_1 n^{-1}\log^{1/2}(np)\}.
\end{aligned}
$$

In the meanwhile, it holds that $|n^{-1}\sum_{i=1}^n R_i^{-1}| = \zeta_1 + O_p(\zeta_1 n^{-1/2})$. Finally,

$$
n^{1/2}\tilde{\boldsymbol{\theta}}_n = n^{-1/2}\zeta_1^{-1}\sum_{i=1}^n Z_i W_i + \tilde{C}_n, \tag{S.17}
$$

and $\tilde{C}_n$ is the remainder term satisfies

$$|\tilde{C}_n|_\infty = O_p\{n^{-1/4}\log^{1/2}(np) + p^{-\delta-(1/6\wedge\delta/2)}\log^{1/2}(np)\}\,.$$

We finish the proof of this lemma. □

Appendix D: Additional simulation results

In this section, we report additional simulation results. Section D.1 shows computation efficiency of inference procedures based on spatial median in high dimensions. Section D.2 presents simulation results on SCIs for $\rho = 0.2$ and 0.5. Section D.3 reports simulations on global tests for high-dimensional location parameters.

D.1 Computation time

Calculating the sample spatial median is known to require solving the equation $\sum_{i=1}^n S(X_i - \hat{\boldsymbol{\theta}}_n) = 0$, which is significantly more complex than calculating the closed-form sample mean. To evaluate the computational efficiency of the algorithm proposed by Vardi and Zhang (2000), we have included a comparison of computation times between the sample spatial median and the sample mean, as well as the implementation times for both the spatial median-based and sample mean-based multiplier bootstrap methods. These comparisons are presented in Tables A1 and A2. The results indicate that for $n = 200$ and $p = 1000$, the spatial median-based multiplier bootstrap implementation takes approximately 2.4 seconds, demonstrating its computational efficiency.

Table A1: Computing time (in **milliseconds**, averaged over 1000 times) of sample spatial median and sample mean with $X \sim N(0, I_p)$ in `R` software using a AMD EPYC 7763 2.44 GHz Processor.

	$p = 100$		$p = 1000$	
	Spatial median	Mean	Spatial median	Mean
$n = 100$	0.149	0.022	0.894	0.138
$n = 200$	0.265	0.046	2.342	0.353

Table A2: Computing time (in **seconds**, averaged over 1000 times) of implementations of the spatial median-based multiplier bootstrap and the sample mean-based multiplier bootstrap in constructing simultaneous confidence intervals with $X \sim N(0, I_p)$ and the number of bootstrap iterations $B = 400$ in R software using a AMD EPYC 7763 2.44 GHz Processor.

	$p = 100$		$p = 1000$	
	Spatial median	Mean	Spatial median	Mean
$n = 100$	0.135	0.088	0.872	0.507
$n = 200$	0.208	0.113	2.402	1.591

D.2 Addition simulation results on simultaneous confidence intervals

Tables A3 reports the coverage probability and median length of the SCIs based on $\hat{\boldsymbol{\theta}}_n$ for $\rho = 0.2$ and 0.5, the results of the SCIs based on the sample mean $\bar{X}_n$ are presented in parentheses. We observe that the performance of the SCIs based on $\hat{\boldsymbol{\theta}}_n$ with $\rho = 0.2$ and 0.5 is similar to that of $\rho = 0.0$ and 0.8 in the main paper. The SCIs achieve satisfactory coverage probability, and it is much shorter than those based on $\bar{X}_n$ under the multivariate t-distribution, which is heavy-tailed.

D.3 Simulations on global tests for high-dimensional location parameters

In this section, we report the performance of the test based on T_n (Median test) for one-sample high-dimensional location parameters, and compare it with three alternative approaches: the test of Chen and Qin (2010, CQ test); the test based on T_{Mean} (Mean test) and bootstrap approximation for $\bar{X}_n$; the test of Wang et al. (2015, WPL test) based on T_{WPL}. We consider the same data generation models (I, II and III) as in Section 5.1. For $\boldsymbol{\theta}$, we set its first $\lfloor c_0 \log p \rfloor$ components as non-zero, while the other elements are all zero. c_0 is chosen from 0.5 and 1. The magnitude of non-zero entries in $\boldsymbol{\theta}$ is $\kappa(\log p/n)^{1/2}$, where κ is chosen from 0 to 5. Note that $\kappa = 0$ refers to the null hypothesis. We consider $n = 50$ or 100, and $p = 100$ and 1000 for each sample size.

Figures A1–A8 plot the empirical size ($\kappa = 0$) and power ($\kappa \neq 0$) of four (CQ, Mean, Median, and WPL) tests at the 5% significance level for Models I and II. The results of $\kappa = 0$ indicates that the empirical sizes of all these four tests are close to the nominal significance level under different case scenarios. When $\kappa \neq 0$, the power of these tests

Table A3: Coverage probability (in %) and median length of the SCIs based on $\hat{\boldsymbol{\theta}}_n$, the results of the SCIs based on $\bar{X}_n$ are in parentheses.

				$\boldsymbol{\theta} = \boldsymbol{\theta}_1$				$\boldsymbol{\theta} = \boldsymbol{\theta}_2$			
				Coverage probability		Median length		Coverage probability		Median length	
Model	ρ	n	p	90%	95%	90%	95%	90%	95%	90%	95%
I	0.2	100	100	89.8 (89.9)	94.5 (94.5)	0.65 (0.65)	0.69 (0.69)	88.8 (88.7)	94.4 (94.4)	0.65 (0.65)	0.69 (0.69)
			1000	88.7 (88.7)	94.5 (94.3)	0.77 (0.77)	0.80 (0.80)	90.0 (89.6)	94.7 (94.8)	0.77 (0.77)	0.80 (0.80)
		200	100	89.0 (88.9)	94.3 (94.1)	0.46 (0.46)	0.49 (0.49)	88.8 (88.8)	94.0 (94.2)	0.46 (0.46)	0.49 (0.49)
			1000	89.8 (89.8)	94.4 (94.4)	0.55 (0.55)	0.57 (0.57)	88.7 (89.2)	94.6 (94.3)	0.55 (0.55)	0.57 (0.57)
	0.5	100	100	89.6 (89.8)	94.5 (94.4)	0.65 (0.65)	0.69 (0.69)	88.4 (88.8)	94.0 (94.1)	0.65 (0.65)	0.69 (0.69)
			1000	88.4 (88.4)	94.3 (94.3)	0.77 (0.77)	0.80 (0.80)	87.4 (87.4)	94.1 (94.2)	0.77 (0.77)	0.80 (0.80)
		200	100	90.9 (90.9)	95.1 (95.2)	0.46 (0.46)	0.49 (0.49)	89.7 (90.0)	95.3 (95.0)	0.46 (0.46)	0.49 (0.49)
			1000	89.0 (89.0)	94.2 (94.3)	0.55 (0.55)	0.57 (0.57)	88.8 (88.6)	94.3 (94.0)	0.55 (0.55)	0.57 (0.57)
II	0.2	100	100	89.2 (88.8)	94.8 (94.2)	0.71 (1.05)	0.75 (1.12)	88.4 (88.8)	93.7 (94.3)	0.71 (1.05)	0.75 (1.11)
			1000	89.0 (89.4)	94.1 (94.8)	0.84 (1.24)	0.88 (1.30)	89.0 (88.9)	94.4 (94.6)	0.84 (1.24)	0.88 (1.30)
		200	100	90.7 (89.8)	95.3 (94.7)	0.50 (0.76)	0.53 (0.80)	89.2 (89.7)	94.0 (94.4)	0.50 (0.76)	0.53 (0.80)
			1000	88.6 (89.5)	94.2 (94.6)	0.59 (0.90)	0.62 (0.93)	89.0 (90.6)	95.0 (95.1)	0.59 (0.90)	0.62 (0.94)
	0.5	100	100	89.2 (87.9)	93.6 (93.9)	0.71 (1.05)	0.75 (1.12)	89.4 (88.6)	94.6 (94.1)	0.71 (1.05)	0.75 (1.11)
			1000	89.2 (88.9)	94.4 (94.2)	0.84 (1.24)	0.88 (1.30)	90.0 (89.4)	94.7 (94.6)	0.84 (1.25)	0.88 (1.30)
		200	100	89.4 (90.0)	94.1 (94.6)	0.50 (0.76)	0.53 (0.80)	89.7 (88.6)	95.0 (93.6)	0.50 (0.76)	0.53 (0.80)
			1000	90.0 (89.9)	95.6 (94.8)	0.59 (0.90)	0.62 (0.94)	88.8 (89.5)	93.8 (94.4)	0.59 (0.89)	0.62 (0.93)
III	0.2	100	100	89.6 (89.5)	95.0 (95.1)	0.65 (0.66)	0.69 (0.70)	89.4 (89.4)	94.6 (94.6)	0.65 (0.66)	0.69 (0.70)
			1000	89.3 (88.8)	94.5 (94.5)	0.78 (0.78)	0.82 (0.82)	90.3 (90.7)	95.0 (94.9)	0.78 (0.78)	0.82 (0.82)
		200	100	89.2 (89.0)	94.4 (94.4)	0.46 (0.46)	0.49 (0.49)	90.0 (89.6)	95.1 (95.2)	0.46 (0.46)	0.49 (0.49)
			1000	89.7 (89.7)	94.6 (94.8)	0.55 (0.55)	0.57 (0.58)	90.4 (90.6)	95.0 (95.0)	0.55 (0.55)	0.57 (0.57)
	0.5	100	100	88.9 (89.3)	94.0 (94.6)	0.65 (0.65)	0.69 (0.69)	88.0 (88.5)	94.2 (94.0)	0.65 (0.65)	0.69 (0.69)
			1000	89.1 (89.2)	94.3 (94.2)	0.78 (0.78)	0.81 (0.81)	89.2 (88.9)	94.1 (94.0)	0.78 (0.78)	0.81 (0.81)
		200	100	89.6 (89.7)	95.0 (94.4)	0.46 (0.46)	0.49 (0.49)	89.6 (89.7)	94.9 (94.4)	0.46 (0.46)	0.49 (0.49)
			1000	89.0 (89.1)	94.3 (94.4)	0.55 (0.55)	0.57 (0.57)	89.3 (89.6)	95.4 (95.0)	0.55 (0.55)	0.57 (0.57)

increases as κ increases, that is, as the signal getting stronger. For Gaussian data, the Mean test based on T_{Mean} and the Median test based on T_n have similar power performances, and they advance both the CQ test and the WPL test, which are L_2-norm type tests. In addition, when the data are from multivariate t-distribution, the Median test outperforms the Mean test, which shows the superiority of the procedure based on the sample spatial median over that based on the sample mean under heavy-tailedness. In summary, the Median test based on T_n is preferred among the four tests when the alternative is sparse and the underlying distribution is heavy-tailed.

Second, Figure A9 depicts empirical size and power of the four tests (CQ, Mean, Median, WPL) for Model III with $\rho = 0$. It can be seen that, even Model III is not a member of the elliptical distribution family, the size of the Median test can still control the size at the nominal level $\alpha = 0.05$, and this is also the case for the WPL test. We can also see that the Median test and the Mean test have better power performance than the CQ test and the WPL test, especially for $c_0 = 0.5$ when the number of non-zero element in $\boldsymbol{\theta}$ is relatively small.

Bibliography

Belloni, A., V. Chernozhukov, D. Chetverikov, C. Hansen, and K. Kato (2018). High-dimensional econometrics and generalized gmm. *arXiv*, 1806.01888.

Chen, S. X. and Y. Qin (2010). A two-sample test for high-dimensional data with applications to gene-set testing. *Ann. Statist. 38*, 808–835.

Cheng, G., B. Liu, L. Peng, B. Zhang, and S. Zheng (2019). Testing the equality of two high-dimensional spatial sign covariance matrices. *Scand. J. Statist. 46*, 257–271.

Chernozhukov, V., D. Chetverikov, and K. Kato (2017). Central limit theorems and bootstrap in high dimensions. *The Annals of Probability 45*, 2309–2352.

Chernozhukov, V., D. Chetverikov, and K. Kato (2019). Improved central limit theorem and bootstrap approximation in high dimensions. *arXiv*, 1912.10529.

Koike, Y. (2021). Notes on the dimension dependence in high-dimensional central limit theorems for hyperrectangles. *Japanese Journal of Statistics and Data Science 1*, 257–297.

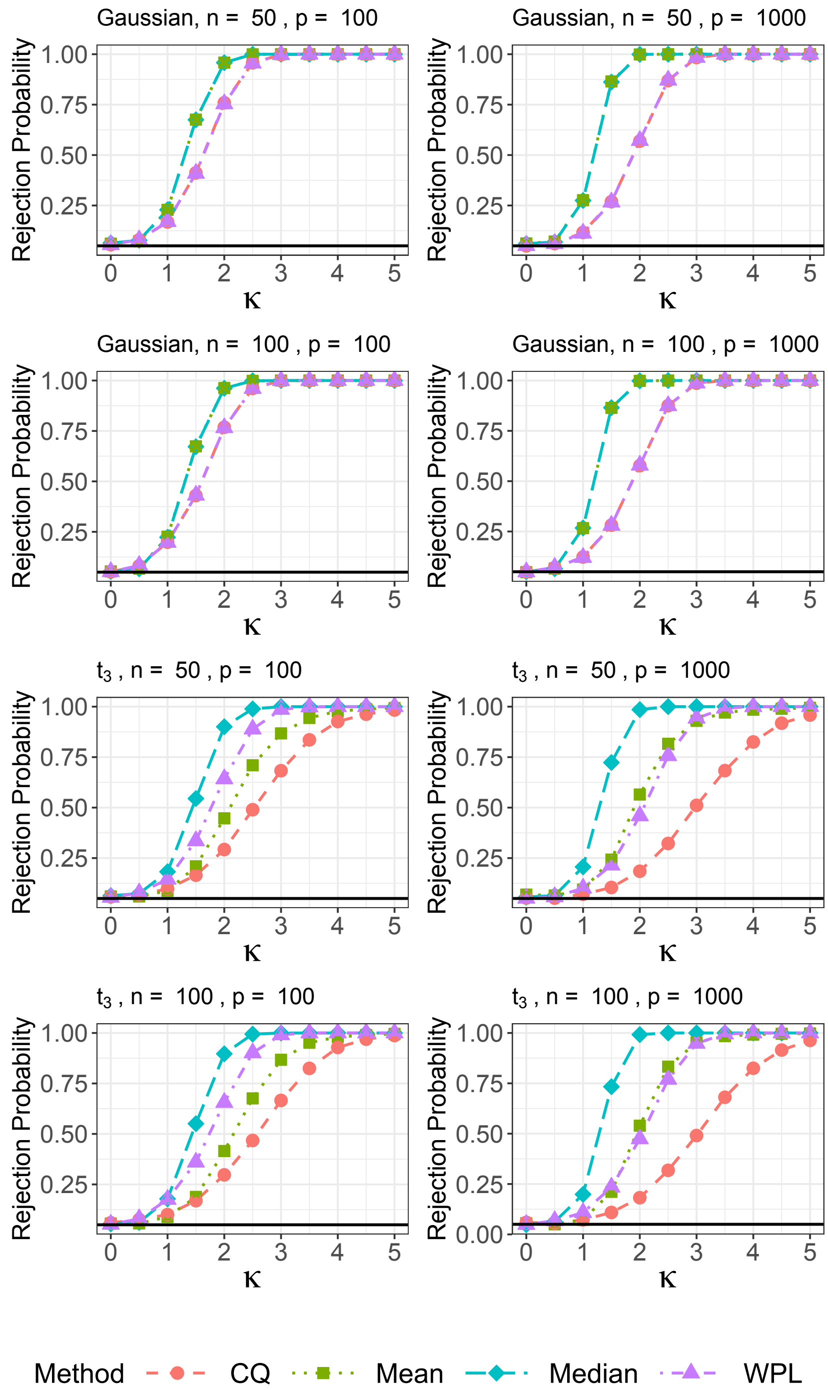

Figure A1: Empirical size and power of the four tests (CQ, Mean, Median, WPL) for Models I and II with $c_0 = 0.5$ and $\rho = 0$. The horizontal black solid line refers to the nominal 5% significance level. "Gaussian" denotes the multivariate normal distribution, and t_3 denotes the multivariate t-distribution with 3 degrees of freedom.

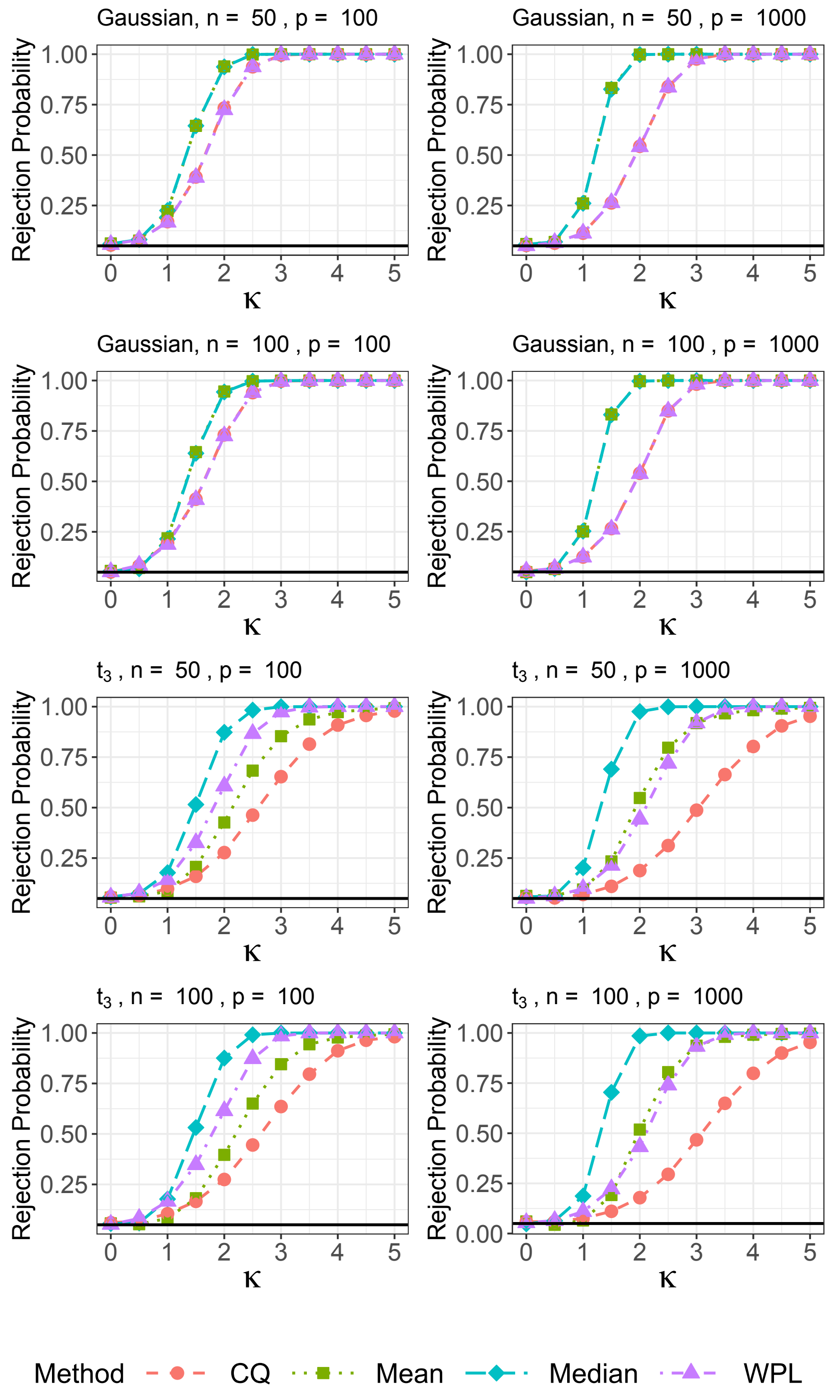

Figure A2: Empirical size and power of the four tests (CQ, Mean, Median, WPL) for Models I and II with $c_0 = 0.5$ and $\rho = 0.2$. The horizontal black solid line refers to the nominal 5% significance level. "Gaussian" denotes the multivariate normal distribution, and t_3 denotes the multivariate t-distribution with 3 degrees of freedom.

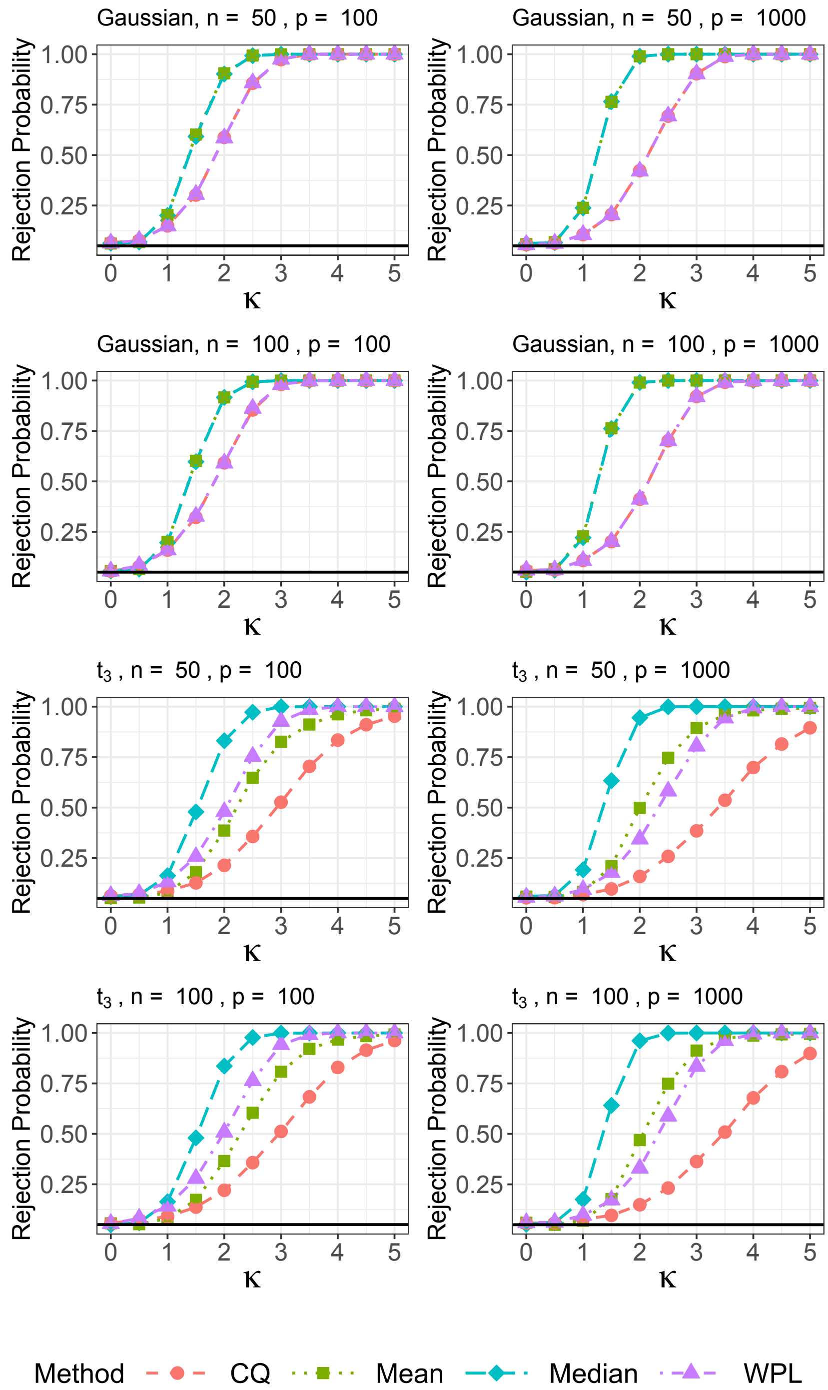

Figure A3: Empirical size and power of the four tests (CQ, Mean, Median, WPL) for Models I and II with $c_0 = 0.5$ and $\rho = 0.5$. The horizontal black solid line refers to the nominal 5% significance level. "Gaussian" denotes the multivariate normal distribution, and t_3 denotes the multivariate t-distribution with 3 degrees of freedom.

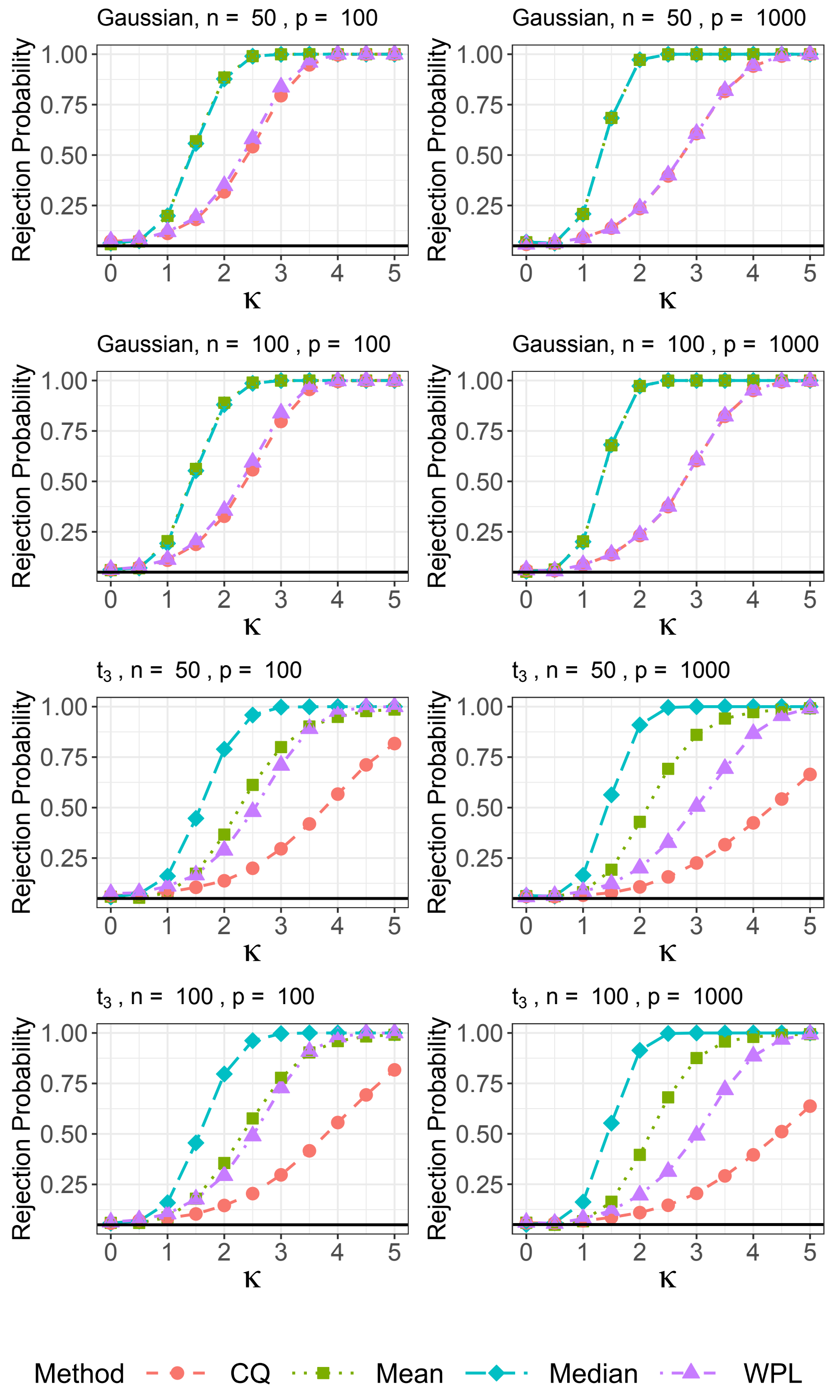

Figure A4: Empirical size and power of the four tests (CQ, Mean, Median, WPL) for Models I and II with $c_0 = 0.5$ and $\rho = 0.8$. The horizontal black solid line refers to the nominal 5% significance level. "Gaussian" denotes the multivariate normal distribution, and t_3 denotes the multivariate t-distribution with 3 degrees of freedom.

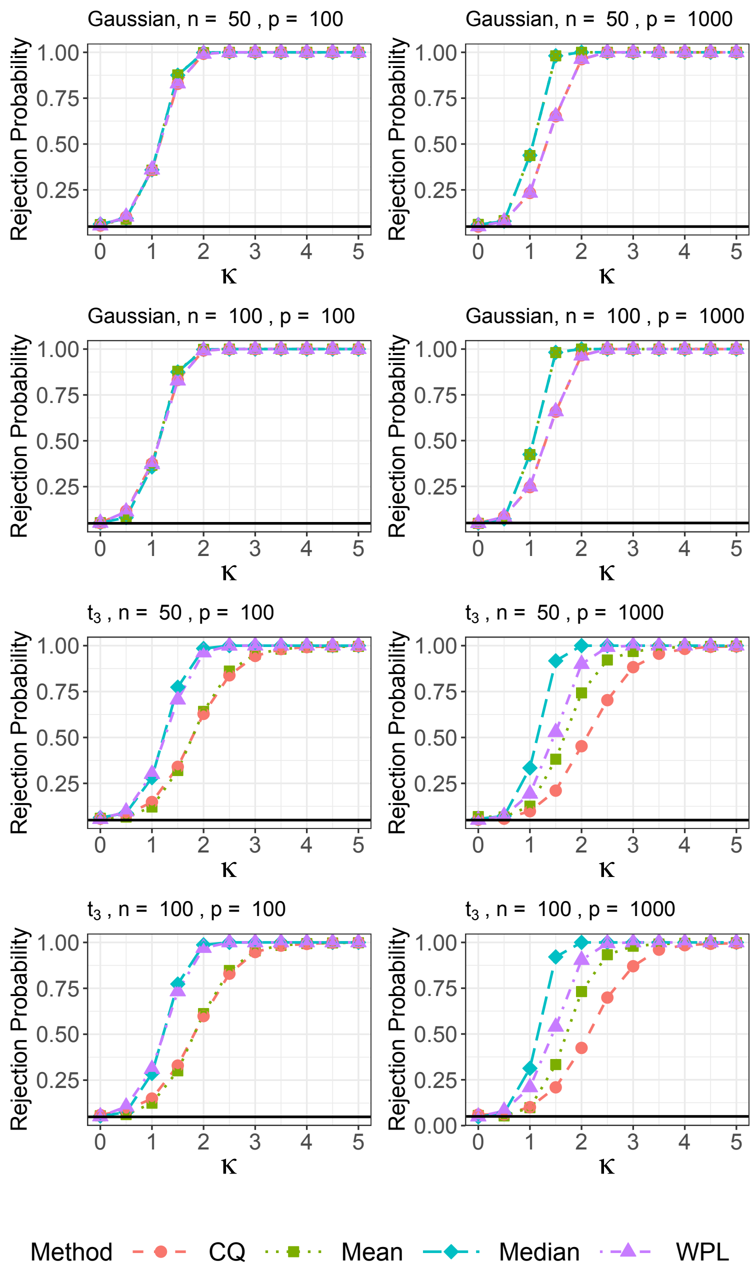

Figure A5: Empirical size and power of the four tests (CQ, Mean, Median, WPL) for Models I and II with $c_0 = 1$ and $\rho = 0$. The horizontal black line refers to the nominal 5% significance level. "Gaussian" denotes the multivariate normal distribution, and t_3 denotes the multivariate t-distribution with 3 degrees of freedom.

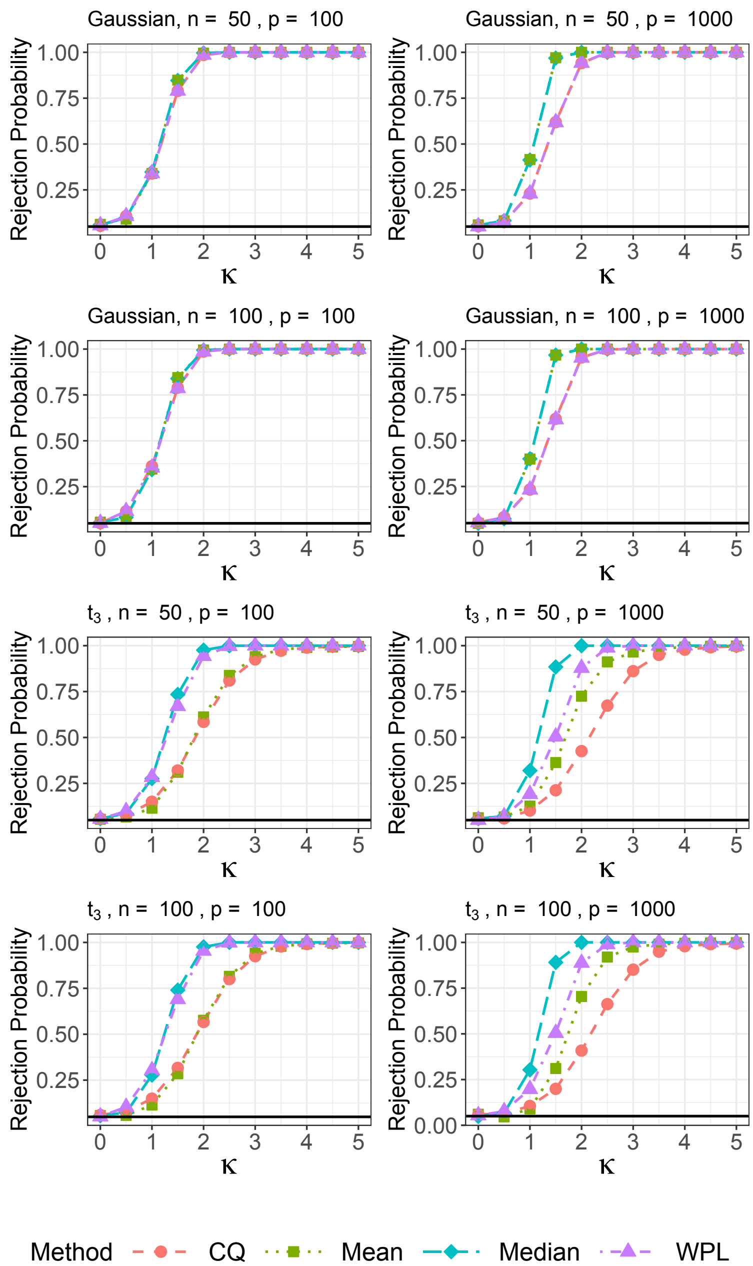

Figure A6: Empirical size and power of the four tests (CQ, Mean, Median, WPL) for Models I and II with $c_0 = 1$ and $\rho = 0.2$. The horizontal black solid line refers to the nominal 5% significance level. "Gaussian" denotes the multivariate normal distribution, and t_3 denotes the multivariate t-distribution with 3 degrees of freedom.

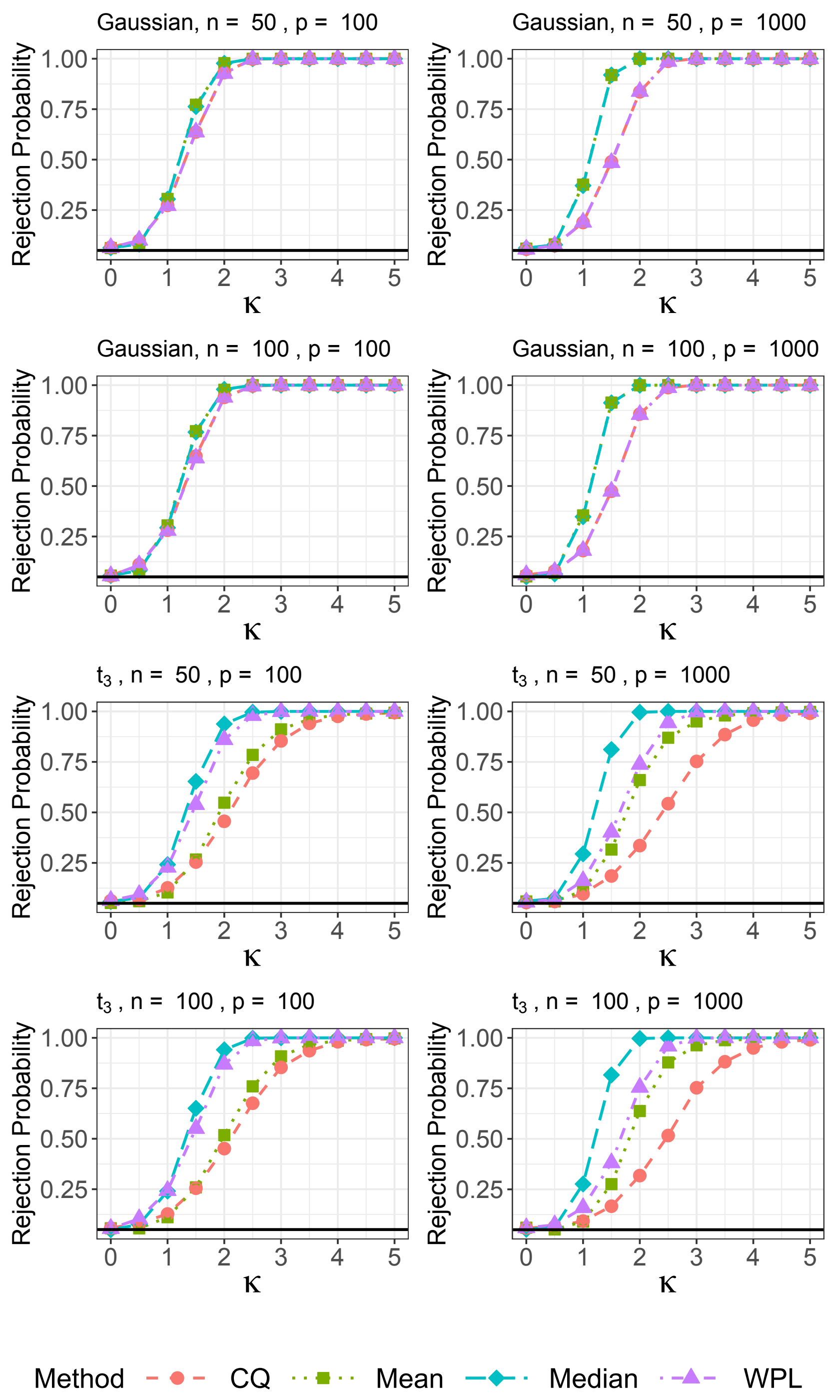

Figure A7: Empirical size and power of the four tests (CQ, Mean, Median, WPL) for Models I and II with $c_0 = 1$ and $\rho = 0.5$. The horizontal black solid line refers to the nominal 5% significance level. "Gaussian" denotes the multivariate normal distribution, and t_3 denotes the multivariate t-distribution with 3 degrees of freedom.

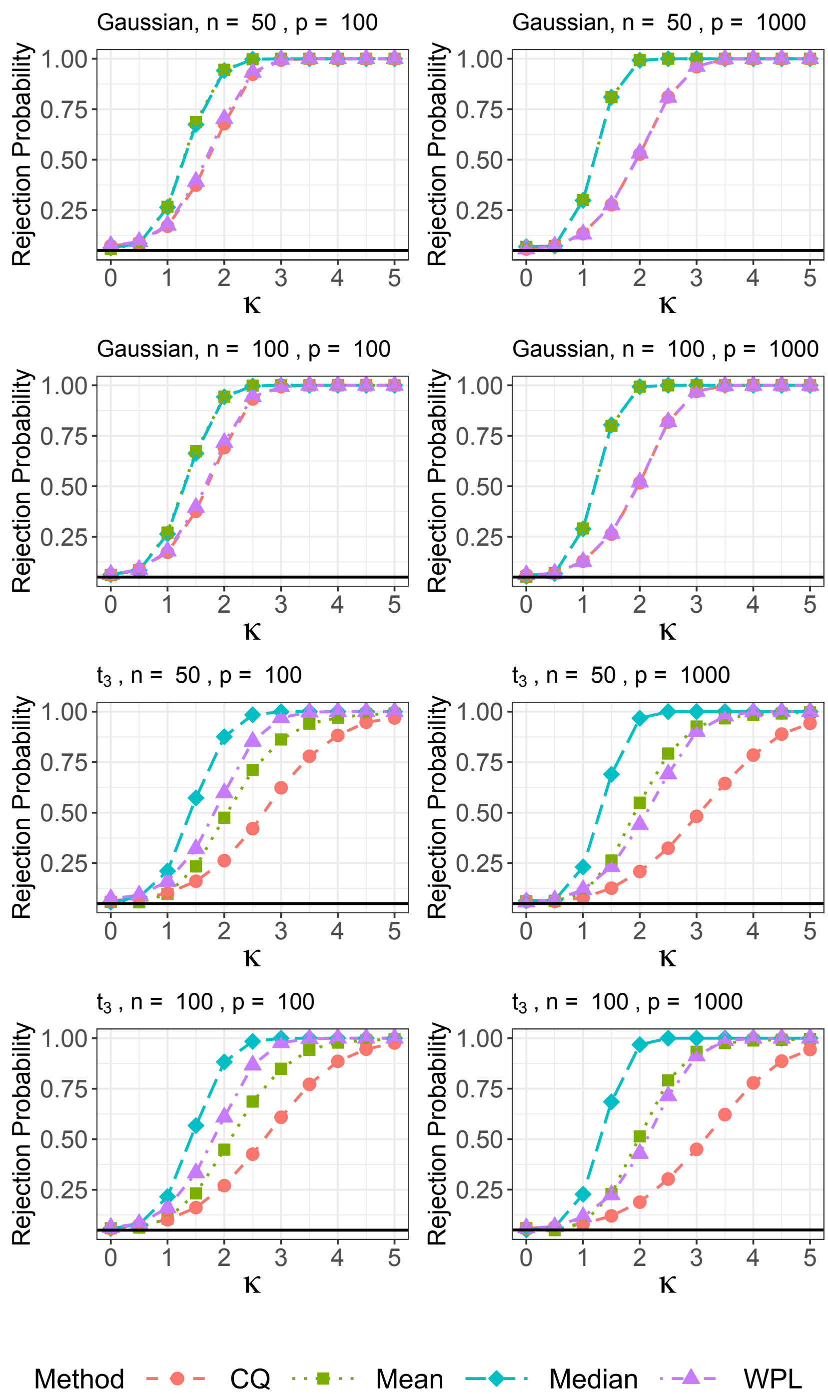

Figure A8: Empirical size and power of the four tests (CQ, Mean, Median, WPL) for Models I and II with $c_0 = 1$ and $\rho = 0.8$. The horizontal black line refers to the nominal 5% significance level. "Gaussian" denotes the multivariate normal distribution, and t_3 denotes the multivariate t-distribution with 3 degrees of freedom.

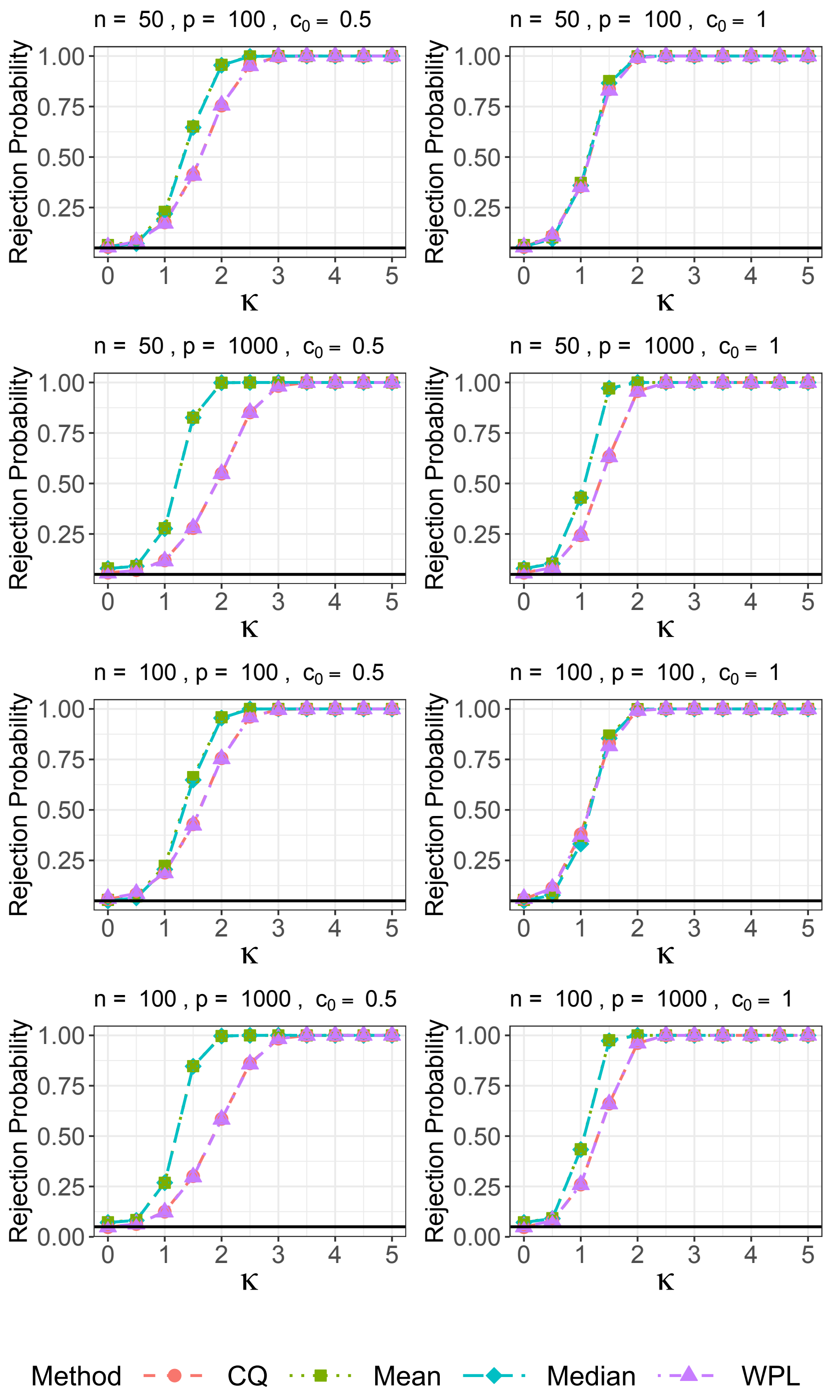

Figure A9: Empirical size and power of the four tests (CQ, Mean, Median, WPL) for Model III with $\rho = 0$. The horizontal black solid line refers to the nominal 5% significance level.

Liu, W. and Q.-M. Shao (2014). Phase transition and regularized bootstrap in large scale t-tests with false discovery rate control. *Ann. Statist. 42*, 2003–2025.

van der Vaart, A. W. and J. A. Wellner (1996). *Weak Convergence and Empirical Processes: With Applications to Statistics*. Springer.

Wang, L., B. Peng, and R. Li (2015). A high-dimensional nonparametric multivariate test for mean vector. *J. Am. Statist. Assoc. 110*, 1658–1669.