Spatial-Sign based Maxsum Test for High Dimensional Location Parameters

Jixuan Liu, Long Feng, Ping Zhao and Zhaojun Wang

School of Statistics and Data Science, KLMDASR, LEBPS, and LPMC, Nankai University

Supplementary Material

In this supplement, we will provide the proofs of the lemmas, theorems given in paper and two real data applications. Recall that $\mathbf{D} = \text{diag}\{d_1^2, d_2^2, \cdots, d_p^2\}$. For $i = 1, 2, \cdots, n$, $\boldsymbol{U}_i = U(\mathbf{D}^{-1/2}(\boldsymbol{X}_i - \boldsymbol{\theta}))$ and $R_i = \|\mathbf{D}^{-1/2}(\boldsymbol{X}_i - \boldsymbol{\theta})\|$ as the scale-invariant spatial-sign and radius of $\boldsymbol{X}_i - \boldsymbol{\theta}$, where $U(\boldsymbol{X}) = \boldsymbol{X}/\|\boldsymbol{X}\|\mathbb{I}(\boldsymbol{X} \neq 0)$ is the multivariate sign function of $\boldsymbol{X}$, with $\mathbb{I}(\cdot)$ being the indicator function. The moments of R_i is defined as $\zeta_k = \mathbb{E}\left(R_i^{-k}\right)$ for k=1,2,3,4.

The $\mathbf{D}$-estimated version $\boldsymbol{U}_i$ and R_i is denoted as $\hat{R}_i = \|\hat{\mathbf{D}}^{-1/2}(\boldsymbol{X}_i - \boldsymbol{\theta})\|$ and $\hat{U}_i = \|\hat{\mathbf{D}}^{-1/2}(\boldsymbol{X}_i - \boldsymbol{\theta})/\|\hat{\mathbf{D}}^{-1/2}(\boldsymbol{X}_i - \boldsymbol{\theta})\|$, respectively, $i = 1, 2, \cdots, n$. We restate the Assumptions given in the main text.

Assumption 1. $W_{i,1}, \ldots, W_{i,p}$ are i.i.d. symmetric random variables with

$\mathbb{E}\left(W_{i,j}\right)=0, \mathbb{E}\left(W_{i,j}^2\right)=1$, and $\|W_{i,j}\|_{\psi_\alpha} \leqslant c_0$ with some constant $c_0>0$ and $1 \leqslant \alpha \leqslant 2$.

Assumption 2. The moments $\zeta_k=\mathbb{E}\left(R_i^{-k}\right)$ for $k=1,2,3,4$ exist for large enough p. In addition, there exist two positive constants $\underline{b}$ and $\bar{B}$ such that $\underline{b} \leqslant \limsup_p \mathbb{E}\left(R_i/\sqrt{p}\right)^{-k} \leqslant \bar{B}$ for $k=1,2,3,4$.

Assumption 3. The shape matrix $\mathbf{R}=\mathbf{D}^{-1/2}\boldsymbol{\Gamma}\boldsymbol{\Gamma}^\top\mathbf{D}^{-1/2}=\left(\sigma_{j\ell}\right)_{p\times p}$ satisfies $\operatorname{tr}(\mathbf{R})=p$ and $\max_{j=1,\cdots,p}\sum_{\ell=1}^p|\sigma_{j\ell}| \leqslant a_0(p)$. In addition, $\liminf_{p\to\infty}\min_{j=1,2,\cdots,p} d_j > \underline{d}$ for some constant $\underline{d}>0$, where $\mathbf{D}=\operatorname{diag}\{d_1^2, d_2^2, \cdots, d_p^2\}$.

Assumption 4. Let $\mathbf{R}=\left(\sigma_{ij}\right)_{1\le i,j\le p}$. For some $\varrho\in(0,1)$, assume $|\sigma_{ij}|\le\varrho$ for all $1\le i<j\le p$ and $p\ge 2$. Suppose $\{\delta_p; p\ge 1\}$ and $\{\kappa_p; p\ge 1\}$ are positive constants with $\delta_p=o(1/\log p)$ and $\kappa=\kappa_p\to 0$ as $p\to\infty$. For $1\le i\le p$, define $B_{p,i}=\{1\le j\le p; |\sigma_{ij}|\ge\delta_p\}$ and $C_p=\{1\le i\le p; |B_{p,i}|\ge p^\kappa\}$. We assume that $|C_p|/p\to 0$ as $p\to\infty$.

Assumption 5. Variables $\{\boldsymbol{X}_1,\ldots,\boldsymbol{X}_n\}$ in the n-th row are independently and identically distributed (i.i.d.) from p-variate elliptical distribution with density functions $\det(\boldsymbol{\Sigma})^{-1/2} g\left(\left\|\boldsymbol{\Sigma}^{-1/2}(\mathbf{x}-\boldsymbol{\theta})\right\|\right)$ where $\boldsymbol{\theta}$'s are the symmetry centers and $\boldsymbol{\Sigma}$'s are the positive definite symmetric $p\times p$ scatter matrices.

Assumption 6. $\operatorname{tr}\left(\mathbf{R}^4\right)=o\left\{\operatorname{tr}^2\left(\mathbf{R}^2\right)\right\}$.

Assumption 7. (i) $\operatorname{tr}(\mathbf{R}^2) - p = o(n^{-1}p^2)$, (ii) $n^{-2}p^2/\operatorname{tr}(\mathbf{R}^2) = O(1)$ and $\log p = o(n)$.

Assumption 8. There exist $C > 0$ and $\varrho \in (0,1)$ so that $\max_{1\le i<j\le p} |\sigma_{ij}| \le \varrho$ and $\max_{1\le i\le p}\sum_{j=1}^{p}\sigma_{ij}^2 \le (\log p)^C$ for all $p \ge 3$; $p^{-1/2}(\log p)^C \ll \lambda_{\min}(\mathbf{R}) \le \lambda_{\max}(\mathbf{R}) \ll \sqrt{p}(\log p)^{-1}$ and $\lambda_{\max}(\mathbf{R})/\lambda_{\min}(\mathbf{R}) = O(p^{\tau})$ for some $\tau \in (0, 1/4)$.

S1 Proof of main lemmas

Lemma 1. *Under Assumption 1, we have* $\mathbb{E}\{U(\boldsymbol{W}_i)^\top \mathbf{M} U(\boldsymbol{W}_i)\}^2 = O\{p^{-2}\mathrm{tr}(\mathbf{M}^\top\mathbf{M})\}$.

Proof. By Cauchy inequality and Assumption 1, we have

$$\mathbb{E}\{U(\boldsymbol{W}_i)_l^2 U(\boldsymbol{W}_i)_k^2\} \le \frac{1}{p^2}\mathbb{E}\{\sum_{s=1}^{p}\sum_{t=1}^{p} U(\boldsymbol{W}_i)_s^2 U(\boldsymbol{W}_i)_t^2\} = p^{-2}$$

$$\mathbb{E}\{U(\boldsymbol{W}_i)_l^4\} \le \frac{1}{p}\mathbb{E}\{\sum_{s=1}^{p} U(\boldsymbol{W}_i)_s^4\} \le \frac{1}{p}\mathbb{E}\{\sum_{s=1}^{p}\sum_{t=1}^{p} U(\boldsymbol{W}_i)_s^2 U(\boldsymbol{W}_i)_t^2\} = p^{-1},$$

and

$$\mathbb{E}\left\{U(\boldsymbol{W}_i)_l U(\boldsymbol{W}_i)_k U(\boldsymbol{W}_i)_s U(\boldsymbol{W}_i)_t\right\} \le \sqrt{E\left\{U(\boldsymbol{W}_i)_l^2 U(\boldsymbol{W}_i)_k^2\right\} E\left\{U(\boldsymbol{W}_i)_s^2 U(\boldsymbol{W}_i)_t^2\right\}}.$$

By the Cauchy inequality,

$$\sum_{i,k,s,t} a_{lk}a_{st} \le \sqrt{\sum_{l,k} a_{lk}^2 \sum_{s,t} a_{st}^2} \le \sqrt{\sum_{l,k}^{p} a_{lk}^2 \sum_{s,t}^{p} a_{st}^2} = \mathrm{tr}(\mathbf{M}^\top\mathbf{M}).$$

Thus, we get

$$
\begin{aligned}
&E\left[\left\{U(\boldsymbol{W}_i)^\top \mathbf{M}\mathbf{U}(\mathbf{W_i})\right\}^2\right] \\
=&\sum_{l\neq k=1}^{p}\sum_{s\neq t=1}^{p} a_{lk}a_{st}\mathbb{E}\left\{U(\boldsymbol{W}_i)_l U(\boldsymbol{W}_i)_k U(\boldsymbol{W}_i)_s U(\boldsymbol{W}_i)_t\right\} + \sum_{l=1}^{p}\sum_{s=1}^{p} a_{ll}a_{ss}\mathbb{E}\left\{U(\boldsymbol{W}_i)_l^2 U(\boldsymbol{W}_i)_s^2\right\} \\
\leq& p^{-2}\frac{p^4-p^2}{p^4}\mathrm{tr}(\mathbf{M}^\top\mathbf{M}) + p^{-1}\frac{p^2}{p^4}\mathrm{tr}(\mathbf{M}^\top\mathbf{M}) = O\{p^{-2}\mathrm{tr}(\mathbf{M}^\top\mathbf{M})\}.
\end{aligned}
$$

Lemma 2. *Under Assumptions 1 and 7, we have* $\max_{j=1,2,\cdots p}(\hat{d}_j - d_j) = O_p\{n^{-1/2}(\log p)^{1/2}\}$.

Proof. The proof of this lemma is along the same lines as the proof of Lemma A.2. in Feng et al. (2016), but differs in that the assumptions about the model in this paper are more general, with different constraints controlling the correlation matrix R.

Denote $\boldsymbol{\eta} = (\boldsymbol{\theta}^\top, d_1, d_2, \cdots, d_p)^\top$ and $\hat{\boldsymbol{\eta}}$ as the estimated version. By first-order Taylor expansion, we have

$$
\begin{aligned}
U(\mathbf{D}^{-1/2}(\boldsymbol{X}_i - \boldsymbol{\theta})) &= \frac{\mathbf{D}^{-1/2}\boldsymbol{\Gamma}U(\boldsymbol{W}_i)}{\{1 + U(\boldsymbol{W}_i)^\top(\mathbf{R}-\mathbf{I}_p)U(\boldsymbol{W}_i)\}^{1/2}} \\
&= \mathbf{D}^{-1/2}\boldsymbol{\Gamma}U(\boldsymbol{W}_i) + C_1 U(\boldsymbol{W}_i)^\top(\mathbf{R}-\mathbf{I}_p)U(\boldsymbol{W}_i)\mathbf{D}^{-1/2}\boldsymbol{\Gamma}U(\boldsymbol{W}_i),
\end{aligned}
\tag{S1.1}
$$

where C_1 is a bounded random variable between -0.5 and $-0.5\{1+U(\boldsymbol{W}_i)^\top(\mathbf{R}-\mathbf{I}_p)U(\boldsymbol{W}_i)\}^{-3/2}$. By Cauchy inequality and Lemma 1 and Assumption 7, we get

$$\mathbb{E}\left\{U(\mathbf{D}^{-1/2}(\boldsymbol{X}_i-\boldsymbol{\theta})\right\} \le C_2\left[\mathbb{E}\{U(\boldsymbol{W}_i)^\top(\mathbf{R}-\mathbf{I}_p)U(\boldsymbol{W}_i)\}^2\mathbb{E}\{\mathbf{D}^{-1/2}(\boldsymbol{X}_i-\boldsymbol{\theta})\}^2\right]^{1/2}$$
$$= O(p^{-1}\sqrt{\mathrm{tr}(\mathbf{R}^2)-p}) = o(n^{-1/2}).$$

Similarly, we can show that

$$\mathbb{E}\left[\mathrm{diag}\left\{U(\mathbf{D}^{-1/2}(\boldsymbol{X}_i-\boldsymbol{\theta}))U(\mathbf{D}^{-1/2}(\boldsymbol{X}_i-\boldsymbol{\theta}))^\top\right\}-\frac{1}{p}\mathbf{I}_p\right] \le O(n^{-1/2}),$$

by first-order Taylor expansion for $U(\mathbf{D}^{-1/2}(\boldsymbol{X}_i-\boldsymbol{\theta}))U(\mathbf{D}^{-1/2}(\boldsymbol{X}_i-\boldsymbol{\theta}))^\top$, Cauchy inequality and Lemma 1. The above two equations define the functional equation for each component of $\boldsymbol{\eta}$,

$$T_j(F,\eta_j) = o_p(n^{-1/2}), \tag{S1.2}$$

where F represent the distribution of $\boldsymbol{X}$, $\boldsymbol{\eta} = (\eta_1,\cdots,\eta_{2p})$. Similar with Hettmansperger and Randles (2002), the linearisation of this equation shows

$$n^{1/2}(\hat{\eta}_j-\eta_j) = -\mathbf{H}_j^{-1}n^{1/2}\left\{T_j(F_n,\eta_j)-T_j(F,\eta_j)\right\}+o_p(1),$$

where F_n represents the empirical distribution function based on $\boldsymbol{X}_1,\boldsymbol{X}_2,\cdots,\boldsymbol{X}_n$, $\mathbf{H}_j$ is the corresponding Hessian matrix of the functional defined in Equation S1.2 and

$$T(F_n,\boldsymbol{\eta}) =$$
$$\left(n^{-1}\sum_{j=1}^{n}U(\mathbf{D}^{-1/2}(\boldsymbol{X}_i-\boldsymbol{\theta}))^\top, \mathrm{vec}[\mathrm{diag}\{n^{-1}U(\mathbf{D}^{-1/2}(\boldsymbol{X}_i-\boldsymbol{\theta}))U(\mathbf{D}^{-1/2}(\boldsymbol{X}_i-\boldsymbol{\theta}))^\top-\frac{1}{p}\mathbf{I}_p\}]\right).$$

Thus, for each $\hat{d}_j$ we have

$$\sqrt{n}(\hat{d}_j - d_j) \xrightarrow{d} N(0, \sigma_{d,j}^2).$$

where $\sigma_{d,j}^2$ is the corresponding asymptotic variance. Define $\sigma_{d,max} = \max_{1\le j\le p} \sigma_{d,j}$. As $p \to \infty$,

$$\begin{aligned}
&\mathbb{P}\{\max_{j=1,2,\cdots,p}(\hat{d}_j - d_j) > \sqrt{2}\sigma_{d,max} n^{-1/2}(\log p)^{1/2}\} \\
&\le \sum_{j=1}^{p} \mathbb{P}\{\sqrt{n}(\hat{d}_j - d_j) > \sqrt{2}\sigma_{d,max}(\log p)^{1/2}\} \\
&= \sum_{j=1}^{p} [1 - \Phi\{\sqrt{2}\sigma_{d,max}\sigma_{d,j}^{-1}(\log p)^{1/2}\}] \le p[1 - \Phi\{(2\log p)^{1/2}\}] \\
&\le \frac{p}{\sqrt{2\pi \log p}} \exp(-\log p) = (4\pi)^{-1/2}(\log p)^{-1/2} \to 0,
\end{aligned}$$

which means that $\max_{j=1,2,\cdots p}(\hat{d}_j - d_j) = O_p\{n^{-1/2}(\log p)^{1/2}\}$.

Lemma 3. *Suppose the assumptions in Lemma 2 hold, then* $\hat{\zeta}_1 \xrightarrow{p} \zeta_1$.

Proof. Denote $\hat{\boldsymbol{\mu}} = \hat{\boldsymbol{\theta}} - \boldsymbol{\theta}$.

$$\begin{aligned}
\|\hat{\mathbf{D}}^{-1/2}(\boldsymbol{X}_i - \hat{\boldsymbol{\theta}})\| =& \|\mathbf{D}^{-1/2}(\boldsymbol{X}_i - \boldsymbol{\theta})\| \\
&\{1 + R_i^{-2}\|(\hat{\mathbf{D}}^{-1/2} - \mathbf{D}^{-1/2})(\boldsymbol{X}_i - \boldsymbol{\theta})\|^2 \\
&+ R_i^{-2}\|\hat{\mathbf{D}}^{-1/2}\hat{\boldsymbol{\mu}}\|^2 + 2R_i^{-2}\boldsymbol{U}_i^\top(\hat{\mathbf{D}}^{-1/2} - \mathbf{D}^{-1/2})\mathbf{D}^{1/2}\boldsymbol{U}_i \\
&- 2R_i^{-1}\boldsymbol{U}_i^\top\hat{\mathbf{D}}^{-1/2}\hat{\boldsymbol{\mu}} - 2R_i^{-1}\boldsymbol{U}_i\mathbf{D}^{1/2}(\hat{\mathbf{D}}^{-1/2} - \mathbf{D}^{-1/2})\hat{\mathbf{D}}^{-1/2}\hat{\boldsymbol{\mu}}\}^{1/2}.
\end{aligned}$$

According to the proof and conclusion in Lemma 2, we can show that $R_i^{-2}\|(\hat{\mathbf{D}}^{-1/2}-\mathbf{D}^{-1/2})(\boldsymbol{X}_i-\boldsymbol{\theta})\|^2 = O_p\{n^{-1/2}(\log p)^{1/2}\} = o_p(1)$ and $R_i^{-2}\|\hat{\mathbf{D}}^{-1/2}\hat{\boldsymbol{\mu}}\|^2 =$

$O_p(n^{-1}) = o_p(1)$ and by the Cauchy inequality, the other parts are also $o_p(1)$. So,

$$n^{-1}\sum_{i=1}^{n}\left\|\hat{\mathbf{D}}^{-1/2}\left(\boldsymbol{X}_i-\hat{\boldsymbol{\theta}}\right)\right\|^{-1}=\left(n^{-1}\sum_{i=1}^{n}\left\|\mathbf{D}^{-1/2}\left(\boldsymbol{X}_i-\boldsymbol{\theta}\right)\right\|^{-1}\right)(1+o_p(1)).$$

Obviously, $\mathbb{E}\left(n^{-1}\sum_{i=1}^n R_i^{-1}\right)=\zeta_i$ and $\mathrm{Var}\left(n^{-1}\zeta_i^{-1}\sum_{i=1}^n R_i^{-1}\right)=O\left(n^{-1}\right)$. Finally, the proof is completed.

Lemma 4. *Suppose Assumptions 1-3 holds with $a_0(p)\asymp p^{1-\delta}$ for some positive constant $\delta\leq 1/2$. Define a random $p\times p$ matrix $\hat{\mathbf{Q}}=n^{-1}\sum_{i=1}^n \hat{R}_i^{-1}\hat{\boldsymbol{U}}_i\hat{\boldsymbol{U}}_i^\top$ and let $\hat{\mathbf{Q}}_{jl}$ be the (j,l)th element of $\hat{\mathbf{Q}}$. Then,*

$$\begin{aligned}\left|\hat{\mathbf{Q}}_{j\ell}\right| &\lesssim \zeta_1 p^{-1}\left|\sigma_{j\ell}\right| \\ &\quad + O_p\left(\zeta_1 n^{-1/2}p^{-1}+\zeta_1 p^{-7/6}+\zeta_1 p^{-1-\delta/2}+\zeta_1 n^{-1/2}(\log p)^{1/2}(p^{-5/2}+p^{-1-\delta/2})\right).\end{aligned}$$

Proof. Recall that $\hat{\mathbf{Q}}=\frac{1}{n}\sum_{i=1}^n \hat{R}_i^{-1}\hat{\boldsymbol{U}}_i\hat{\boldsymbol{U}}_i^\top$, then,

$$\begin{aligned}\hat{\mathbf{Q}}_{jl}=&\frac{1}{n}\sum_{i=1}^{n}\hat{R}_i^{-1}\hat{\boldsymbol{U}}_{i,j}\hat{\boldsymbol{U}}_{i,l}\\ =&\frac{1}{n}\sum_{i=1}^{n}v_i^{-1}(\hat{\mathbf{D}}^{-1/2}\boldsymbol{\Gamma}_j\boldsymbol{W}_i)(\hat{\mathbf{D}}^{-1/2}\boldsymbol{\Gamma}_l\boldsymbol{W}_i)^\top\|\hat{\mathbf{D}}^{-1/2}\boldsymbol{\Gamma}\boldsymbol{W}_i\|^{-3}\\ =&\hat{\mathbf{D}}^{-1/2}\mathbf{D}^{1/2}\left\{\frac{1}{n}\sum_{i=1}^{n}v_i^{-1}(\mathbf{D}^{-1/2}\boldsymbol{\Gamma}_j\boldsymbol{W}_i)(\mathbf{D}^{-1/2}\boldsymbol{\Gamma}_l\boldsymbol{W}_i)^\top\|\hat{\mathbf{D}}^{-1/2}\boldsymbol{\Gamma}\boldsymbol{W}_i\|^{-3}\right\}\hat{\mathbf{D}}^{-1/2}\mathbf{D}^{1/2}.\end{aligned}$$

We first consider the middle term,

$$
\begin{aligned}
&\frac{1}{n}\sum_{i=1}^{n} v_i^{-1}(\mathbf{D}^{-1/2}\boldsymbol{\Gamma}_j\boldsymbol{W}_i)(\mathbf{D}^{-1/2}\boldsymbol{\Gamma}_l\boldsymbol{W}_i)^\top\|\hat{\mathbf{D}}^{-1/2}\boldsymbol{\Gamma}\boldsymbol{W}_i\|^{-3}\\
=&\frac{1}{n}\sum_{i=1}^{n} v_i^{-1}(\mathbf{D}^{-1/2}\boldsymbol{\Gamma}_j\boldsymbol{W}_i)(\mathbf{D}^{-1/2}\boldsymbol{\Gamma}_l\boldsymbol{W}_i)^\top\left\{\|\hat{\mathbf{D}}^{-1/2}\boldsymbol{\Gamma}\boldsymbol{W}_i\|^{-3}-\|\mathbf{D}^{-1/2}\boldsymbol{\Gamma}\boldsymbol{W}_i\|^{-3}\right\}\\
&+\frac{1}{n}\sum_{i=1}^{n} v_i^{-1}(\mathbf{D}^{-1/2}\boldsymbol{\Gamma}_j\boldsymbol{W}_i)(\mathbf{D}^{-1/2}\boldsymbol{\Gamma}_l\boldsymbol{W}_i)^\top\left\{\|\mathbf{D}^{-1/2}\boldsymbol{\Gamma}\boldsymbol{W}_i\|^{-3}-p^{-3/2}\right\}\\
&+n^{-1}p^{-3/2}\sum_{i=1}^{n} v_i^{-1}(\mathbf{D}^{-1/2}\boldsymbol{\Gamma}_j\boldsymbol{W}_i)(\mathbf{D}^{-1/2}\boldsymbol{\Gamma}_l\boldsymbol{W}_i)^\top.
\end{aligned}
\tag{S1.3}
$$

The first part in Equation S1.3:

$$
\begin{aligned}
&\mathbb{E}[\frac{1}{n}\sum_{i=1}^{n} v_i^{-1}(\mathbf{D}^{-1/2}\boldsymbol{\Gamma}_j\boldsymbol{W}_i)(\mathbf{D}^{-1/2}\boldsymbol{\Gamma}_l\boldsymbol{W}_i)^\top\left\{\|\hat{\mathbf{D}}^{-1/2}\boldsymbol{\Gamma}\boldsymbol{W}_i\|^{-3}-\|\mathbf{D}^{-1/2}\boldsymbol{\Gamma}\boldsymbol{W}_i\|^{-3}\right\}]\\
=&\mathbb{E}[v_i^{-1}(\mathbf{D}^{-1/2}\boldsymbol{\Gamma}_j\boldsymbol{W}_i)(\mathbf{D}^{-1/2}\boldsymbol{\Gamma}_l\boldsymbol{W}_i)^\top\left\{\|\hat{\mathbf{D}}^{-1/2}\boldsymbol{\Gamma}\boldsymbol{W}_i\|^{-3}-\|\mathbf{D}^{-1/2}\boldsymbol{\Gamma}\boldsymbol{W}_i\|^{-3}\right\}]\\
=&\mathbb{E}[v_i^{-1}(\mathbf{D}^{-1/2}\boldsymbol{\Gamma}_j\boldsymbol{W}_i)(\mathbf{D}^{-1/2}\boldsymbol{\Gamma}_l\boldsymbol{W}_i)^\top\\
&\qquad\left\{\|\hat{\mathbf{D}}^{-1/2}\boldsymbol{\Gamma}\boldsymbol{W}_i\|^{-3}\|\mathbf{D}^{-1/2}\Gamma W_i\|^{-3}\left(\|\mathbf{D}^{-1/2}\boldsymbol{\Gamma}\boldsymbol{W}_i\|^{3}-\|\hat{\mathbf{D}}^{-1/2}\boldsymbol{\Gamma}\boldsymbol{W}_i\|^{3}\right)\right\}].
\end{aligned}
\tag{S1.4}
$$

To compute the order of Equation S1.4, we consider $\|\hat{\mathbf{D}}^{-1/2}\boldsymbol{\Gamma}\boldsymbol{W}_i\|^k - \|\mathbf{D}^{-1/2}\boldsymbol{\Gamma}\boldsymbol{W}_i\|^k$ for $k = 1, 2, \cdots$.

Firstly, for $k = 2$, By the Lemma 2 and Assumption 3, we can see that,

$$
\max_{i=1,2,\cdots,p}(\frac{d_i}{\hat{d}_i}-1)=\max_{i=1,2,\cdots,p}\frac{d_i-\hat{d}_i}{\hat{d}_i}=O_p(n^{-1/2}(\log p)^{1/2}).
$$

So, for $\|\hat{\mathbf{D}}^{-1/2}\mathbf{\Gamma}\boldsymbol{W}_i\|^2$,

$$
\begin{aligned}
&\|\hat{\mathbf{D}}^{-1/2}\mathbf{\Gamma}\boldsymbol{W}_i\|^2\\
=&\|(\hat{\mathbf{D}}^{-1/2}\mathbf{D}^{1/2}-\mathbf{I}_p)\mathbf{D}^{-1/2}\mathbf{\Gamma}\boldsymbol{W}_i+\mathbf{D}^{-1/2}\mathbf{\Gamma}\boldsymbol{W}_i\|^2\\
=&\|\mathbf{D}^{-1/2}\mathbf{\Gamma}\boldsymbol{W}_i\|^2+\|(\hat{\mathbf{D}}^{-1/2}\mathbf{D}^{1/2}-\mathbf{I}_p)\mathbf{D}^{-1/2}\mathbf{\Gamma}\boldsymbol{W}_i\|^2+\boldsymbol{W}_i\mathbf{\Gamma}^\top\mathbf{D}^{-1/2}(\hat{\mathbf{D}}^{-1/2}\mathbf{D}^{1/2}\mathbf{I}_p)\mathbf{D}^{-1/2}\mathbf{\Gamma}\boldsymbol{W}_i\\
\leq&\|\mathbf{D}^{-1/2}\mathbf{\Gamma}\boldsymbol{W}_i\|^2\left\{1+\max_{i=1,2,\cdots,p}(\frac{d_i}{\hat{d}_i}-1)^2+\max_{i=1,2,\cdots,p}(\frac{d_i}{\hat{d}_i}-1)\right\}\\
:=&\|\mathbf{D}^{-1/2}\mathbf{\Gamma}\boldsymbol{W}_i\|^2\left(1+H\right),
\end{aligned}
$$

where $H:=\max_i(\frac{d_i}{\hat{d}_i}-1)^2+\max_i(\frac{d_i}{\hat{d}_i}-1)=O_p\{n^{-1/2}(\log p)^{1/2}\}$.

For all integer k,

$$
\begin{aligned}
\|\hat{\mathbf{D}}^{-1/2}\mathbf{\Gamma}\boldsymbol{W}_i\|^k=&\|(\hat{\mathbf{D}}^{-1/2}\mathbf{D}^{1/2}-\mathbf{I}_p)\mathbf{D}^{-1/2}\mathbf{\Gamma}\boldsymbol{W}_i+\mathbf{D}^{-1/2}\mathbf{\Gamma}\boldsymbol{W}_i\|^k\\
=&\left\{\|(\hat{\mathbf{D}}^{-1/2}\mathbf{D}^{1/2}-\mathbf{I}_p)\mathbf{D}^{-1/2}\mathbf{\Gamma}\boldsymbol{W}_i+\mathbf{D}^{-1/2}\mathbf{\Gamma}\boldsymbol{W}_i\|^2\right\}^{k/2}\\
\leq&\|\mathbf{D}^{-1/2}\mathbf{\Gamma}\boldsymbol{W}_i\|^k\left(1+H\right)^{k/2}\\
:=&\|\mathbf{D}^{-1/2}\mathbf{\Gamma}\boldsymbol{W}_i\|^k\left(1+H_k\right),
\end{aligned}
\tag{S1.5}
$$

where H_k is defined as $H_k=(1+H)^{k/2}-1=O_p\{n^{-1/2}(\log p)^{1/2}\}$.

Thus, from the proof of Lemma A3. in Cheng et al. (2023), Equation

S1.4 equals

$$
\begin{aligned}
&\mathbb{E}[\frac{1}{n}\sum_{i=1}^{n} v_i^{-1}(\mathbf{D}^{-1/2}\mathbf{\Gamma}_j\boldsymbol{W}_i)(\mathbf{D}^{-1/2}\mathbf{\Gamma}_l\boldsymbol{W}_i)^\top\left\{\|\hat{\mathbf{D}}^{-1/2}\mathbf{\Gamma}\boldsymbol{W}_i\|^{-3}-\|\mathbf{D}^{-1/2}\mathbf{\Gamma}\boldsymbol{W}_i\|^{-3}\right\}]\\
=&\mathbb{E}[v_i^{-1}(\mathbf{D}^{-1/2}\mathbf{\Gamma}_j\boldsymbol{W}_i)(\mathbf{D}^{-1/2}\mathbf{\Gamma}_l\boldsymbol{W}_i)^\top\left\{\|\hat{\mathbf{D}}^{-1/2}\mathbf{\Gamma}\boldsymbol{W}_i\|^{-3}-\|\mathbf{D}^{-1/2}\mathbf{\Gamma}\boldsymbol{W}_i\|^{-3}\right\}]\\
=&\mathbb{E}[v_i^{-1}(\mathbf{D}^{-1/2}\mathbf{\Gamma}_j\boldsymbol{W}_i)(\mathbf{D}^{-1/2}\mathbf{\Gamma}_l\boldsymbol{W}_i)^\top\|\mathbf{D}^{-1/2}\mathbf{\Gamma}\boldsymbol{W}_i\|^{-3}H_3]\\
=&\mathbb{E}[v_i^{-1}(\mathbf{D}^{-1/2}\mathbf{\Gamma}_j\boldsymbol{W}_i)(\mathbf{D}^{-1/2}\mathbf{\Gamma}_l\boldsymbol{W}_i)^\top(\|\mathbf{D}^{-1/2}\mathbf{\Gamma}\boldsymbol{W}_i\|^{-3}-p^{-3/2})H_3]\\
&\quad+p^{-3/2}\mathbb{E}[v_i^{-1}(\mathbf{D}^{-1/2}\mathbf{\Gamma}_j\boldsymbol{W}_i)(\mathbf{D}^{-1/2}\mathbf{\Gamma}_l\boldsymbol{W}_i)^\top H_3]\\
\lesssim&n^{-1/2}(\log p)^{1/2}\zeta_1 p^{-5/2}(1+p^{3/2-\delta/2})\\
=&\zeta_1 n^{-1/2}(\log p)^{1/2}(p^{-5/2}+p^{-1-\delta/2}).
\end{aligned}
$$

The second and last part in Equation S1.3:

From the proof of Lemma A3. in Cheng et al. (2023), we can conclude,

$$
\frac{1}{n}\sum_{i=1}^{n} v_i^{-1}(\mathbf{D}^{-1/2}\mathbf{\Gamma}_j\boldsymbol{W}_i)(\mathbf{D}^{-1/2}\mathbf{\Gamma}_l\boldsymbol{W}_i)^\top\left\{\|\mathbf{D}^{-1/2}\mathbf{\Gamma}\boldsymbol{W}_i\|^{-3}-p^{-3/2}\right\}=O_p(\zeta_1 p^{-1-\delta/2}),
$$

and

$$
n^{-1}p^{-3/2}\sum_{i=1}^{n} v_i^{-1}(\mathbf{D}^{-1/2}\mathbf{\Gamma}_j\boldsymbol{W}_i)(\mathbf{D}^{-1/2}\mathbf{\Gamma}_l\boldsymbol{W}_i)^\top\lesssim\zeta_1 p^{-1}\left|\sigma_{j\ell}\right|+O_p\left(\zeta_1 n^{-1/2}p^{-1}+\zeta_1 p^{-7/6}\right).
$$

It follows that,

$$
\mathbf{Q}_{jl}=\left\{n^{-1}p^{-3/2}\sum_{i=1}^{n} v_i^{-1}(\mathbf{D}^{-1/2}\mathbf{\Gamma}_j\boldsymbol{W}_i)(\mathbf{D}^{-1/2}\mathbf{\Gamma}_l\boldsymbol{W}_i)^\top+O_p(A_n)\right\}(1+O_p(n^{-1}\log p)),
$$

where $A_n=\zeta_1 n^{-1/2}(\log p)^{1/2}(p^{-5/2}+p^{-1-\delta/2})+\zeta_1 p^{-1-\delta/2}$.

Thus,

$$|\mathbf{Q}_{j\ell}| \lesssim \zeta_1 p^{-1} |\sigma_{j\ell}| + O_p\left\{\zeta_1 n^{-1/2} p^{-1} + \zeta_1 p^{-7/6} + \zeta_1 p^{-1-\delta/2} + \zeta_1 n^{-1/2} (\log p)^{1/2} (p^{-5/2} + p^{-1-\delta/2})\right\}.$$

Lemma 5. *Suppose Assumptions 1-3 hold with* $a_0(p) \asymp p^{1-\delta}$ *for some positive constant* $\delta \leq 1/2$*. Then, if* $\log p = o(n^{1/3})$,

(i)$\|n^{-1}\sum_{i=1}^n \zeta_1^{-1}\hat{\boldsymbol{U}}_i\|_\infty = O_p\left\{n^{-1/2}\log^{1/2}(np)\right\}$,

(ii)$\|\zeta_1^{-1} n^{-1}\sum_{i=1}^n \delta_{1i}\hat{\boldsymbol{U}}_i\|_\infty = O_p(n^{-1})$.

Proof. For any $j \in \{1, 2, \cdots, p\}$,

$$\begin{aligned}\hat{\boldsymbol{U}}_{ij} - \boldsymbol{U}_{ij} &= \frac{\|\mathbf{D}^{-1/2}X_i\|}{\|\hat{\mathbf{D}}^{-1/2}X_i\|}\frac{d_j}{\hat{d}_j}\boldsymbol{U}_{ij} - \boldsymbol{U}_{ij} \\ &\leq (1+H)(1+H)\boldsymbol{U}_{ij} - \boldsymbol{U}_{ij} \\ &= H_u \boldsymbol{U}_{ij},\end{aligned} \tag{S1.6}$$

where $H_u = O_p(H^2 + 2H) = O_p\{n^{-1/2}(\log p)^{1/2}\}$. Thus, $\hat{\boldsymbol{U}}_i - \boldsymbol{U}_i = H_u \boldsymbol{U}_i$.

(i)By Equation S1.6, we have,

$$\begin{aligned}\left\|n^{-1}\sum_{i=1}^n \zeta_1^{-1}\hat{\boldsymbol{U}}_i\right\|_\infty &= \left\|n^{-1}\sum_{i=1}^n \zeta_1^{-1}(1+H_u)\boldsymbol{U}_i\right\|_\infty \\ \leq& |1+H_u|\left\|n^{-1}\sum_{i=1}^n \zeta_1^{-1}\boldsymbol{U}_i\right\|_\infty = O_p\left\{n^{-1/2}\log^{1/2}(np)\right\}.\end{aligned}$$

.

(ii)Similarly,

$$\begin{aligned}\left|\zeta_1^{-1}n^{-1}\sum_{i=1}^n \delta_{1i}\hat{\boldsymbol{U}}_i\right|_\infty &\leq |1+H_u|\left\|\zeta_1^{-1}n^{-1}\sum_{i=1}^n \delta_{1i}\boldsymbol{U}_i\right\|_\infty \\ \leq& O_p\{n^{-1}(1+n^{-1/2}\log^{1/2}p)\} = O_p(n^{-1}).\end{aligned}$$

S1.1 Proof of Lemma 1(Bahadur representation)

Proof. As $\boldsymbol{\theta}$ is a location parameter, we assume $\boldsymbol{\theta} = 0$ without loss of generality. Then $\boldsymbol{U}_i$ can be written as $\boldsymbol{U}_i = \mathbf{D}^{-1/2}\boldsymbol{X}_i/\|\mathbf{D}^{-1/2}\boldsymbol{X}_i\| = \mathbf{D}^{-1/2}\boldsymbol{\Gamma}\boldsymbol{W}_i/\|\mathbf{D}^{-1/2}\boldsymbol{\Gamma}\boldsymbol{W}_i\|$ for $i = 1, 2, \cdots, n$. The estimator $\hat{\boldsymbol{\theta}}$ satisfies $\sum_{i=1}^n U(\hat{\mathbf{D}}^{-1/2}(\boldsymbol{X}_i - \hat{\boldsymbol{\theta}})) = 0$, which is is equivalent to

$$\frac{1}{n}\sum_{i=1}^n(\hat{\boldsymbol{U}}_i - \hat{R}_i^{-1}\hat{\mathbf{D}}^{-1/2}\hat{\boldsymbol{\theta}})(1 - 2\hat{R}_i^{-1}\hat{\boldsymbol{U}}_i^\top\hat{\mathbf{D}}^{-1/2}\hat{\boldsymbol{\theta}} + \hat{R}_i^{-2}\hat{\boldsymbol{\theta}}^\top\hat{\mathbf{D}}^{-1}\hat{\boldsymbol{\theta}})^{-1/2} = 0.$$

From the proof of lemma A.3 in Feng et al. (2016), we can see that $\|\hat{\boldsymbol{\theta}}\| = O_p(\zeta_1^{-1}n^{-1/2})$. By the first-order Taylor expansion, the above equation can be rewritten as:

$$n^{-1}\sum_{i=1}^n\left(\hat{\boldsymbol{U}}_i - \hat{R}_i^{-1}\hat{\mathbf{D}}^{-1/2}\hat{\boldsymbol{\theta}}\right)\left(1 + \hat{R}_i^{-1}\hat{\boldsymbol{U}}_i^\top\hat{\mathbf{D}}^{1/2}\hat{\boldsymbol{\theta}} - 2^{-1}\hat{R}_i^{-2}\left\|\hat{\mathbf{D}}^{-1/2}\hat{\boldsymbol{\theta}}\right\|^2 + \delta_{1i}\right) = 0,$$

where $\delta_{1i} = O_p\left\{\left(\hat{R}_i^{-1}\hat{\boldsymbol{U}}_i^\top\hat{\mathbf{D}}^{1/2}\hat{\boldsymbol{\theta}} - 2^{-1}\hat{R}_i^{-2}\left\|\hat{\mathbf{D}}^{-1/2}\hat{\boldsymbol{\theta}}\right\|^2\right)^2\right\} = O_p\left(n^{-1}\right)$, which implies

$$\begin{aligned}&\frac{1}{n}\sum_{i=1}^n(1 - \frac{1}{2}\hat{R}_i^{-2}\hat{\boldsymbol{\theta}}^\top\hat{\mathbf{D}}^{-1}\hat{\boldsymbol{\theta}} + \delta_{1i})\hat{\boldsymbol{U}}_i + \frac{1}{n}\sum_{i=1}^n\hat{R}_i^{-1}(\hat{\boldsymbol{U}}_i^\top\hat{\mathbf{D}}^{-1/2}\hat{\boldsymbol{\theta}})\hat{U}_i\\ =&\frac{1}{n}\sum_{i=1}^n(1 + \delta_{1i} + \delta_{2i})\hat{R}_i^{-1}\hat{\mathbf{D}}^{-1/2}\hat{\boldsymbol{\theta}},\end{aligned} \tag{S1.7}$$

where $\delta_{2i} = O_p(\hat{R}_i^{-1}\hat{\boldsymbol{U}}_i^\top\hat{\mathbf{D}}^{1/2}\hat{\boldsymbol{\theta}} - 2^{-1}\hat{R}_i^{-2}\left\|\hat{\mathbf{D}}^{-1/2}\hat{\boldsymbol{\theta}}\right\|^2) = O_p(\delta_{1i}^{1/2})$.

Similar with Cheng et al. (2023), by Assumption 2 and Markov inequality, we have that: $\max R_i^{-2} = O_p(\zeta_1^2 n^{1/2})$, $\max\delta_{1i} = O_p(\|\hat{\mathbf{D}}^{-1/2}\boldsymbol{\theta}\|^2\max\hat{R}_i^{-2} = O_p(n^{-1/2})$ and $\max\delta_{2i} = O_p(n^{-1/4})$.

Considering the second term in Equation S1.7,

$$\frac{1}{n}\sum_{i=1}^{n}\hat{R}_i^{-1}(\hat{U}_i^\top\hat{\mathbf{D}}^{-1/2}\hat{\boldsymbol{\theta}})\hat{U}_i=\frac{1}{n}\sum_{i=1}^{n}\hat{R}_i^{-1}(\hat{U}_i\hat{U}_i^\top\hat{\mathbf{D}}^{-1/2})\hat{\boldsymbol{\theta}}=\hat{Q}\hat{\mathbf{D}}^{-1/2}\hat{\boldsymbol{\theta}},$$

where $\hat{\boldsymbol{Q}}=\frac{1}{n}\sum_{i=1}^{n}\hat{R}_i^{-1}\hat{\boldsymbol{U}}_i\hat{\boldsymbol{U}}_i^\top$. From Lemma 4 we acquire,

$$|\mathbf{Q}_{j\ell}|\lesssim\zeta_1 p^{-1}|\sigma_{j\ell}|$$
$$+O_p\left\{\zeta_1 n^{-1/2}p^{-1}+\zeta_1 p^{-7/6}+\zeta_1 p^{-1-\delta/2}+\zeta_1 n^{-1/2}(\log p)^{1/2}(p^{-5/2}+p^{-1-\delta/2})\right\},$$

and this implies that,

$$\begin{aligned}&\left\|\mathbf{Q}\hat{\mathbf{D}}^{-1/2}\hat{\boldsymbol{\theta}}\right\|_\infty\\ \leqslant&\|\mathbf{Q}\|_1\left\|\hat{\mathbf{D}}^{-1/2}\hat{\boldsymbol{\theta}}\right\|_\infty\\ \lesssim&\zeta_1 p^{-1}\|\mathbf{R}\|_1\left\|\hat{\mathbf{D}}^{-1/2}\hat{\boldsymbol{\theta}}\right\|_\infty\\ +&O_p\left\{\zeta_1 n^{-1/2}p^{-1}+\zeta_1 p^{-7/6}+\zeta_1 p^{-1-\delta/2}+\zeta_1 n^{-1/2}(\log p)^{1/2}(p^{-5/2}+p^{-1-\delta/2})\right\}\left\|\hat{\mathbf{D}}^{-1/2}\hat{\boldsymbol{\theta}}\right\|_\infty\\ &O_p\left\{\zeta_1 n^{-1/2}p^{-1}+\zeta_1 p^{-7/6}+\zeta_1 p^{-1-\delta/2}+\zeta_1 n^{-1/2}(\log p)^{1/2}p^{-5/2}\right\}\left\|\hat{\mathbf{D}}^{-1/2}\hat{\boldsymbol{\theta}}\right\|_\infty.\end{aligned}\tag{S1.8}$$

According to Lemma 5, $\left\|n^{-1}\sum_{i=1}^{n}\zeta_1^{-1}\hat{\boldsymbol{U}}_i\right\|_\infty=O_p\left\{n^{-1/2}\log^{1/2}(np)\right\}$ and $\left\|\zeta_1^{-1}n^{-1}\sum_{i=1}^{n}\delta_{1i}\hat{\boldsymbol{U}}_i\right\|_\infty=O_p(n^{-1})$. In addition, we obtain,

$$\left\|\zeta_1^{-1}n^{-1}\sum_{i=1}^{n}\hat{R}_i^{-2}\|\hat{\mathbf{D}}^{-1/2}\hat{\boldsymbol{\theta}}\|^2\hat{\boldsymbol{U}}_i\right\|_\infty\leq|1+H_u|\left\|\zeta_1^{-1}n^{-1}\sum_{i=1}^{n}\hat{R}_i^{-2}\|\hat{\mathbf{D}}^{-1/2}\hat{\boldsymbol{\theta}}\|^2\boldsymbol{U}_i\right\|_\infty$$
$$\lesssim O_p(n^{-1})[1+O_p\{n^{-3/2}(\log p)^{1/2}\}]=O_p(n^{-1}).$$

Considering the third term :

Using the fact that $\zeta_1^{-1}n^{-1}\sum_{i=1}^{n}R_i^{-1}=1+O_p(n^{-1/2})$ and Equation

S1.5, we have

$$\begin{aligned}
&\frac{1}{n}\zeta_1^{-1}\sum_{i=1}^{n}\hat{R}_i^{-1}\\
=&\frac{1}{n}\zeta_1^{-1}\sum_{i=1}^{n}R_i^{-1}[1+O_p\{n^{-1/2}(\log p)^{1/2}\}]\\
=&\left\{1+O_p(n^{-1/2})\right\}[1+O_p\{n^{-1/2}(\log p)^{1/2}\}]\\
=&1+O_p\{n^{-1/2}(\log p)^{1/2}\}
\end{aligned}.$$

We final obtain:

$$\begin{aligned}
\hat{\mathbf{D}}^{-1/2}\hat{\boldsymbol{\theta}}\Big|_\infty \lesssim& \left\|\zeta_1^{-1}n^{-1}\sum_{i=1}^{n}\hat{\boldsymbol{U}}_i\right\|_\infty + \zeta_1^{-1}\left\|\mathbf{Q}\hat{\mathbf{D}}^{-1/2}\hat{\boldsymbol{\theta}}\right\|_\infty\\
\lesssim& p^{-1}a_0(p)\left\|\hat{\mathbf{D}}^{-1/2}\hat{\boldsymbol{\theta}}\right\|_\infty + O_p\left\{n^{-1/2}\log^{1/2}(np)\right\}\\
&+O_p\left\{n^{-1/2}+p^{-(1/6\wedge\delta/2)}+n^{-1/2}(\log p)^{1/2}p^{-3/2}\right\}\left\|\hat{\mathbf{D}}^{-1/2}\hat{\boldsymbol{\theta}}\right\|_\infty.
\end{aligned}$$

Thus we conclude that:

$$\left\|\hat{\mathbf{D}}^{-1/2}\hat{\boldsymbol{\theta}}\right\|_\infty = O_p\{n^{-1/2}\log^{1/2}(np)\},$$

as $a_0(p) \asymp p^{1-\delta}$.

In addition, we have

$$\begin{aligned}
&\left\|\zeta_1^{-1}\mathbf{Q}\hat{\mathbf{D}}^{-1/2}\hat{\boldsymbol{\theta}}\right\|_\infty\\
=&O_p\left\{n^{-1}\log^{1/2}(np)+n^{-1/2}p^{-(1/6\wedge\delta/2)}\log^{1/2}(np)+n^{-1}(\log p)^{1/2}p^{-3/2}\log^{1/2}(np)\right\},
\end{aligned}$$

and

$$
\begin{aligned}
& n^{-1}\sum_{i=1}^{n}\hat{R}_i^{-1}\left(1+\delta_{1i}+\delta_{2i}\right) \\
=&\zeta_1\left\{1+O_p\left(n^{-1/4}\right)\right\}\left[1+O_p\{n^{-1/2}(\log p)^{1/2}\}\right] \\
=&\zeta_1\left[1+O_p\{n^{-1/4}+n^{-1/2}(\log p)^{1/2}\}\right].
\end{aligned}
$$

Finally, we can write

$$
n^{1/2}\hat{\mathbf{D}}^{-1/2}(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta}) = n^{-1/2}\zeta_1^{-1}\sum_{i=1}^{n}\boldsymbol{U}_i + C_n, \tag{S1.9}
$$

where

$$
\begin{aligned}
\|C_n\|_\infty =&O_p\{n^{-1/2}\log^{1/2}(np)+p^{-(1/6\wedge\delta/2)}\log^{1/2}(np)+n^{-1/2}(\log p)^{1/2}p^{-3/2}\log^{1/2}(np)\} \\
&+O_p\{n^{-1/4}\log^{1/2}(np)+n^{-1/2}(\log p)^{1/2}\log^{1/2}(np)\} \\
=&O_p\{n^{-1/4}\log^{1/2}(np)+p^{-(1/6\wedge\delta/2)}\log^{1/2}(np)+n^{-1/2}(\log p)^{1/2}\log^{1/2}(np)\}.
\end{aligned}
$$

S1.2 Proof of Lemma 2(Gaussian approximation)

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

$$
\begin{aligned}
\mathbb{P}(n^{1/2}\hat{\mathbf{D}}^{-1/2}(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta})\le t) &= \mathbb{P}(n^{-1/2}\zeta_1^{-1}\sum_{i=1}^{n}\boldsymbol{U}_i + C_n \le t) \\
&\le \mathbb{P}(n^{-1/2}\zeta_1^{-1}\sum_{i=1}^{n}\boldsymbol{U}_i \le t+\eta_n L_{n,p}) + \mathbb{P}(\|C_n\|_\infty > \eta_n L_{n,p}).
\end{aligned}
$$

According to Lemma 7, $\mathbb{E}\{(\zeta_1^{-1}\boldsymbol{U}_{i,j})^4\}\lesssim \bar{M}^2$ and $\mathbb{E}\{(\zeta_1^{-1}\boldsymbol{U}_{i,j})^2\}\gtrsim \underline{m}$ for all $i=1,2,\cdots,n$, $j=1,2,\cdots,p$, and the Gaussian approximation for independent partial sums in Koike (2021), let $\boldsymbol{G}\sim N\left(0,\zeta_1^{-2}\boldsymbol{\Sigma}_u\right)$ with $\boldsymbol{\Sigma}_u =$

$\mathbb{E}\left(\boldsymbol{U}_1\boldsymbol{U}_1^\top\right)$, we have

$$\mathbb{P}(n^{1/2}\zeta_1^{-1}\sum_{i=1}^{n}\boldsymbol{U}_i \le t+\eta_n L_{n,p}) \le \mathbb{P}(\boldsymbol{G}\le t+\eta_n L_{n,p}) + O[\{n^{-1}\log^5(np)\}^{1/6}]$$

$$\le \mathbb{P}(\boldsymbol{G}\le t) + O\{\eta_n L_{n,p}\log^{1/2}(p)\} + O[\{n^{-1}\log^5(np)\}^{1/6}],$$

where the second inequality holds from Nazarov's inequality in Lemma 8.

Thus,

$$\mathbb{P}(n^{1/2}\hat{\mathbf{D}}^{-1/2}(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta})\le t) \le \mathbb{P}(\boldsymbol{G}\le t) + O\{\eta_n L_{n,p}\log^{1/2}(p)\} + O[\{n^{-1}\log^5(np)\}^{1/6}]$$

$$+ \mathbb{P}(|C_n|_\infty > \eta_n l_{n,p}).$$

On the other hand, we have

$$\mathbb{P}(n^{1/2}\hat{\mathbf{D}}^{-1/2}(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta})\le t) \ge \mathbb{P}(\boldsymbol{G}\le t) - O\{\eta_n L_{n,p}\log^{1/2}(p)\} - O[\{n^{-1}\log^5(np)\}^{1/6}]$$

$$- \mathbb{P}(\|C_n\|_\infty > \eta_n l_{n,p}).$$

where $\mathbb{P}(\|C_n\|_\infty > \eta_n l_{n,p}) \to 0$ as $n\to\infty$ by Lemma S1.1.

Then we have that, if $\log p = o(n^{1/5})$ and $\log n = o(p^{1/3\wedge\delta})$,

$$\sup_{t\in\mathbb{R}^p}|\mathbb{P}(n^{1/2}\hat{\mathbf{D}}^{-1/2}(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta})\le t) - \mathbb{P}(\boldsymbol{G}\le t)| \to 0.$$

Further,

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

by the Corollary 5.1 in Chernozhukov et al. (2017).

S1.3 Proof of Lemma 3(Variance approximation)

Proof. $\mathbb{E}Z_j^2 = \zeta_1^{-2}E(R_i^2)^{-1} \le \bar{B}$ by Assumption 2 and $\mathbb{E}(\max_{1\le j\le p} Z_j) \asymp$

$\sqrt{\log p + \log\log p}$ by Theorem 2 in Feng et al. (2022). Let $\Delta_0 = \max_{1\le j,k\le p} |p(\mathbb{E}U_1U_1^\top)_{j,k} - R_{j,k}|$, by Lemma 7,

$$\Delta_0 = \max_{1\le j,k\le p} |p(\mathbb{E}\boldsymbol{U}_1\boldsymbol{U}_1^\top)_{j,k} - \mathbf{R}_{j,k}| = O(p^{-\delta/2}).$$

According to Lemma 9 ,we have

$$\sup_{t\in\mathbb{R}} |\mathbb{P}(\|\boldsymbol{Z}\|_\infty \leqslant t) - \mathbb{P}(\|\boldsymbol{G}\|_\infty \leqslant t)| \leqslant C' n^{-1/3}\{1\vee \log(np)\}^{2/3} \to 0.$$

S2 Proof of main results

S2.1 Proof of Theorem 1(Limit distribution of maxima)

Proof.

$$\begin{aligned}
\widetilde{\rho}_n &= \sup_{t\in\mathbb{R}} \left|\mathbb{P}\left(n^{1/2}\left\|\hat{\mathbf{D}}^{-1/2}(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta})\right\|_\infty \leqslant t\right) - \mathbb{P}(\|\boldsymbol{Z}\|_\infty \leqslant t)\right| \\
&\le \sup_{t\in\mathbb{R}} \left|\mathbb{P}\left(n^{1/2}\left\|\hat{\mathbf{D}}^{-1/2}(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta})\right\|_\infty \leqslant t\right) - \mathbb{P}(\|\boldsymbol{G}\|_\infty \leqslant t)\right| \\
&\quad + \sup_{t\in\mathbb{R}} |\mathbb{P}(\|\boldsymbol{G}\|_\infty \leqslant t) - \mathbb{P}(\|\boldsymbol{Z}\|_\infty \leqslant t)| \\
&\to 0.
\end{aligned}$$

The last step holds from Lemma 2 and 3 in subsection S1.2 and S1.3.

S2.2 Proof of Theorem 2(Exact limit distribution of maxima)

Proof. According to the Theorem 2 in Feng et al. (2022), we have

$$\mathbb{P}(p\zeta_1^2 \max_{1\le i\le p} Z_i^2 - 2\log p + \log\log p \le x) \to F(x) = \exp\left\{-\frac{1}{\sqrt{\pi}}e^{-x/2}\right\},$$

a cdf of the Gumbel distribution, as $p \to \infty$. Thus, according to Lemma 3 and Theorem 1 in subsection S2.1.

$$
\begin{aligned}
&|\mathbb{P}(T_{MAX} - 2\log p + \log\log p \le x) - F(x)| \\
\le & \left|\mathbb{P}(\zeta_1^2 \hat{\zeta}_1^{-2} T_{MAX} - 2\log p + \log\log p \le x) - F(x)\right| + o(1) \\
\le & \left|\mathbb{P}(\zeta_1^2 \hat{\zeta}_1^{-2} T_{MAX} - 2\log p + \log\log p \le x) - \mathbb{P}(p\zeta_1^2 \max_{1\le i\le p} Z_i^2 - 2\log p + \log\log p \le x)\right| \\
& + \left|\mathbb{P}(p\zeta_1^2 \max_{1\le i\le p} Z_i^2 - 2\log p + \log\log p \le x) - F(x)\right| + o(1) \to 0,
\end{aligned}
$$

for any $x \in \mathbb{R}$.

S2.3 Proof of Theorem 3(Consistency for max-type test)

Proof. Recall that $u_p(y) = y + 2\log p - \log\log p$, $T_{MAX} = n\|\hat{\mathbf{D}}^{-1/2}\hat{\boldsymbol{\theta}}\|_\infty^2 \hat{\zeta}_1^2 p\left(1 - n^{-1/2}\right)$. In order to make the following proof process briefly, we abbreviate $u_p(0)$ to u_p, define $\widetilde{q}_{1-\alpha} = (\max\{q_{1-\alpha} + u_p, 0\})^{1/2} = O_p[\{\log p - 2\log\log(1-\alpha)^{-1}\}^{1/2}]$, $T = T_{MAX}^{1/2} = n^{1/2}\|\hat{\mathbf{D}}^{-1/2}\hat{\boldsymbol{\theta}}\|_\infty \hat{\zeta}_1 p^{1/2}(1-n^{-1/2})^{1/2}$ and $T^c = n^{1/2}\|\hat{\mathbf{D}}^{-1/2}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})\|_\infty \hat{\zeta}_1 p^{1/2}(1-n^{-1/2})^{1/2}$, which has the same distribution of T under H_0.

It is clear that, $T \ge n^{1/2}\|\hat{\mathbf{D}}^{-1/2}\boldsymbol{\theta}\|_\infty \hat{\zeta}_1 p^{1/2}(1-n^{-1/2})^{1/2} - T^c$. Combined with Assumption 2 and Lemma 2, we get

$$
\begin{aligned}
&\mathbb{P}(T_{MAX} - u_p \geq q_{1-\alpha} \mid H_1) \\
\geq &\mathbb{P}(n^{1/2} \|\hat{\mathbf{D}}^{-1/2}\boldsymbol{\theta}\|_\infty \hat{\zeta}_1 p^{1/2}(1-n^{-1/2})^{1/2} - T^c \geq \widetilde{q}_{1-\alpha} \mid H_1) \\
= &\mathbb{P}(T^c \leq n^{1/2} \|\hat{\mathbf{D}}^{-1/2}\boldsymbol{\theta}\|_\infty \hat{\zeta}_1 p^{1/2}(1-n^{-1/2})^{1/2} - \widetilde{q}_{1-\alpha} \mid H_1) \\
\geq &\mathbb{P}(T^c \leq n^{1/2} \left(\|\mathbf{D}^{-1/2}\boldsymbol{\theta}\|_\infty - \|(\hat{\mathbf{D}}^{-1/2} - \mathbf{D}^{-1/2})\boldsymbol{\theta}\|_\infty\right) \hat{\zeta}_1 p^{1/2}(1-n^{-1/2})^{1/2} - \widetilde{q}_{1-\alpha} \mid H_1) \\
\geq &\mathbb{P}(T^c \leq n^{1/2} \|\mathbf{D}^{-1/2}\boldsymbol{\theta}\|_\infty [1 + O_p\{n^{1/2}\log^{1/2}(np)\}] \hat{\zeta}_1 p^{1/2}(1-n^{-1/2})^{1/2} - \widetilde{q}_{1-\alpha} \mid H_1) \to 1,
\end{aligned}
$$

if $\|\boldsymbol{\theta}\|_\infty \geq \widetilde{C} n^{-1/2}\{\log p - 2\log\log(1-\alpha)^{-1}\}^{1/2}$ for some large enough constant $\widetilde{C}$.

The last inequality holds since

$$
\begin{aligned}
\|(\hat{\mathbf{D}}^{-1/2} - \mathbf{D}^{-1/2})\boldsymbol{\theta}\|_\infty = \max_{i=1,2,\cdots,p} \frac{\hat{d}_i - d_i}{\hat{d}_i d_i}\theta_i &\leq \max_{i=1,2,\cdots,p} |1 - \frac{d_i}{\hat{d}_i}| \|\mathbf{D}^{-1/2}\boldsymbol{\theta}\|_\infty \\
&\leq O_p\{n^{-1/2}\log^{1/2}(np)\} \|\mathbf{D}^{-1/2}\boldsymbol{\theta}\|_\infty.
\end{aligned}
$$

S2.4 Proof of Theorem 5(Without elliptical distribution)

Theorem 4 is the special case of Theorem 5 with $\boldsymbol{\theta} = 0$, so we only need to show that Theorem 5 holds.

Proof. The following proof is based on the idea of the proof in article Feng and Sun (2016), with modifications in some equations. We restate the equations in Feng and Sun (2016) on $U(\hat{\mathbf{D}}_{ij}^{-1/2}\boldsymbol{X}_i)^\top U(\hat{\mathbf{D}}_{ij}^{-1/2}\boldsymbol{X}_i)$. By the definition, we have

$$
\begin{aligned}
&\frac{2}{n(n-1)} \sum_{1 \leq i<j \leq n} U\left(\hat{\mathbf{D}}_{ij}^{-1/2} \boldsymbol{X}_i\right)^{\top} U\left(\hat{\mathbf{D}}_{ij}^{-1/2} \boldsymbol{X}_j\right) \\
=&\frac{2}{n(n-1)} \sum_{1 \leq i<j \leq n}\left\{\boldsymbol{U}_i+R_i^{-1} \mathbf{D}^{-1/2} \boldsymbol{\theta}+\left(\hat{\mathbf{D}}_{ij}^{-1/2} \mathbf{D}^{1/2}-\mathbf{I}_p\right) \boldsymbol{U}_i\right\}^{\top} \\
&\times\left\{\boldsymbol{U}_j+R_j^{-1} \mathbf{D}^{-1/2} \boldsymbol{\theta}+\left(\hat{\mathbf{D}}_{ij}^{-1/2} \mathbf{D}^{1/2}-\mathbf{I}_p\right) \boldsymbol{U}_j\right\}\left(1+\alpha_{ij}\right)^{-1/2}\left(1+\alpha_{ji}\right)^{-1/2} \\
=&\frac{2}{n(n-1)} \sum_{1 \leq i<j \leq n} \boldsymbol{U}_i^{\top} \boldsymbol{U}_j+\frac{2}{n(n-1)} \sum_{1 \leq i<j \leq n} R_i^{-1} R_j^{-1} \boldsymbol{\theta}^{\top} \mathbf{D}^{-1} \boldsymbol{\theta} \\
&+\frac{2}{n(n-1)} \sum_{1 \leq i<j \leq n} \boldsymbol{U}_i^{\top} \boldsymbol{U}_j\left\{\left(1+\alpha_{ij}\right)^{-1/2}\left(1+\alpha_{ji}\right)^{-1/2}-1\right\} \\
&+\frac{4}{n(n-1)} \sum_{1 \leq i<j \leq n} \boldsymbol{U}_i^{\top}\left(\hat{\mathbf{D}}_{ij}^{-1/2} \mathbf{D}^{1/2}-\mathbf{I}_p\right) \boldsymbol{U}_j\left(1+\alpha_{ij}\right)^{-1/2}\left(1+\alpha_{ji}\right)^{-1/2} \\
&+\frac{2}{n(n-1)} \sum_{1 \leq i<j \leq n} \boldsymbol{U}_i^{\top}\left(\hat{\mathbf{D}}_{ij}^{-1/2} \mathbf{D}^{1/2}-\mathbf{I}_p\right)^2 \boldsymbol{U}_j\left(1+\alpha_{ij}\right)^{-1/2}\left(1+\alpha_{ji}\right)^{-1/2} \\
&+\frac{2}{n(n-1)} \sum_{1 \leq i<j \leq n} R_j^{-1} \boldsymbol{\theta}^{\top} \mathbf{D}^{-1} \boldsymbol{\theta}\left\{\left(1+\alpha_{ij}\right)^{-1/2}\left(1+\alpha_{ji}\right)^{-1/2}-1\right\} \\
&+\frac{4}{n(n-1)} \sum_{1 \leq i<j \leq n} \boldsymbol{U}_i^{\top} \mathbf{D}^{-1/2} \boldsymbol{\theta}\left(1+\alpha_{ij}\right)^{-1/2}\left(1+\alpha_{ji}\right)^{-1/2} \\
&+\frac{2}{n(n-1)} \sum_{1 \leq i<j \leq n} \boldsymbol{U}_i^{\top}\left(\hat{\mathbf{D}}_{ij}^{-1/2} \mathbf{D}^{1/2}-\mathbf{I}_p\right) \mathbf{D}^{-1/2} \boldsymbol{\theta}\left(1+\alpha_{ij}\right)^{-1/2}\left(1+\alpha_{ji}\right)^{-1/2} \\
:=&\frac{2}{n(n-1)} \sum_{1 \leq i<j \leq n} \boldsymbol{U}_i^{\top} \boldsymbol{U}_j+\frac{2}{n(n-1)} \sum_{1 \leq i<j \leq n} R_i^{-1} R_j^{-1} \boldsymbol{\theta}^{\top} \mathbf{D}^{-1} \boldsymbol{\theta} \\
&+A_{n1}+A_{n2}+A_{n3}+A_{n4}+A_{n5}+A_{n6}.
\end{aligned}
$$

where

$$
\begin{aligned}
\alpha_{ij}=&2 R_i^{-1} \boldsymbol{U}_i^{\top}\left(\hat{\mathbf{D}}_{ij}^{-1/2}-\mathbf{D}^{-1/2}\right) \boldsymbol{X}_i+R_i^{-2}\left\|\left(\hat{\mathbf{D}}_{ij}^{-1/2}-\mathbf{D}^{-1/2}\right) \boldsymbol{X}_i\right\|^2 \\
&+2 R_i^{-1} \mathbf{D}^{-1/2} \boldsymbol{\theta}+R_i^{-2} \boldsymbol{\theta}^{\top} \mathbf{D}^{-1} \boldsymbol{\theta}.
\end{aligned}
$$

Note that $R_i^{-1}\boldsymbol{U}_i^\top(\hat{\mathbf{D}}_{ij}^{-1/2}-\mathbf{D}^{-1/2})\boldsymbol{X}_i=\boldsymbol{U}_i^\top(\hat{\mathbf{D}}_{ij}^{-1/2}\mathbf{D}^{1/2}-\mathbf{I}_p)\boldsymbol{U}_i+R_i^{-1}\boldsymbol{U}_i^\top(\hat{\mathbf{D}}_{ij}^{-1/2}-\mathbf{D}^{-1/2})\boldsymbol{\theta}=O_p\{n^{-1/2}(\log p)^{1/2}\}$ and $R_i^{-2}\|(\hat{\mathbf{D}}_{ij}^{-1/2}-\mathbf{D}^{-1/2})\boldsymbol{X}_i\|^2=O_p\left(n^{-1}\log p\right)$ by Lemma 2. By Assumption 7 and H_1, $R_i^{-1}\mathbf{D}^{-1/2}\boldsymbol{\theta}=O_p(\sigma_n^{1/2})=O_p\left(n^{-1}\right)$ and $R_i^{-2}\boldsymbol{\theta}^\top\mathbf{D}^{-1}\boldsymbol{\theta}=O_p\left(\sigma_n\right)=O_p\left(n^{-2}\right)$ where $\sigma_n^2=2\operatorname{tr}\left(\mathbf{R}^2\right)/n(n-1)p^2$. So $\alpha_{ij}=O_p\{n^{-1/2}(\log p)^{1/2}\}$.

Similarly, we will show that $A_{n1}=o_p\left(\sigma_n\right)$. Under some calculations, we get $\mathbb{E}\{(\boldsymbol{U}_i^\top\boldsymbol{U}_j)^2\}=\operatorname{tr}(\boldsymbol{\Sigma}_u^2)$. By Lemma 7, we find that $\boldsymbol{\Sigma}_{u,i,j}=p^{-1}\sigma_{i,j}+O(p^{-1-\delta/2})$. Thus we have,

$$
\begin{aligned}
\operatorname{tr}(\boldsymbol{\Sigma}_u^2)&=\sum_{i=1}^p\sum_{j=1}^p\boldsymbol{\Sigma}_{u,i,j}^2=\sum_{i=1}^p\sum_{j=1}^p\left(p^{-2}\sigma_{i,j}^2+\sigma_{i,j}O(p^{-2-\delta/2})\right)\\
&=\sum_{i=1}^p\sum_{j=1}^p p^{-2}\sigma_{i,j}^2+\sum_{p^{-\delta/2}=O(\sigma_{ij})}\sigma_{i,j}O(p^{-2-\delta/2})\\
&+\sum_{\sigma_{ij}\in[C_1\frac{O\{\operatorname{tr}(\mathbf{R}^2)\}}{p^{2-\delta/2}},C_2p^{-\delta/2}]}\sigma_{i,j}O(p^{-2-\delta/2})+\sum_{\sigma_{ij}=O(\frac{O\{\operatorname{tr}(\mathbf{R}^2)\}}{p^{2-\delta/2}})}\sigma_{i,j}O(p^{-2-\delta/2})\\
&=p^{-2}\operatorname{tr}(\mathbf{R}^2)\{1+O(1)\}+O(p^{-2-\delta/2})\frac{p^{2-\delta/2}}{O\{\operatorname{tr}(\mathbf{R}^2)\}}o(\frac{p^2}{n})+p^{-2}O\{\operatorname{tr}(\mathbf{R}^2)\}\\
&=O\{p^{-2}\operatorname{tr}(\mathbf{R}^2)\}.
\end{aligned}
\tag{S2.1}
$$

By the Cauchy inequality,

$$
\begin{aligned}
\mathbb{E}\left(A_{n1}^2\right)&=O\left(n^{-4}\right)\sum_{i<j}\mathbb{E}\left[\boldsymbol{U}_i^T\boldsymbol{U}_j\left\{\left(1+\alpha_{ij}\right)^{-1/2}\left(1+\alpha_{ji}\right)^{-1/2}-1\right\}\right]^2\\
&\le O\left(n^{-2}\right)\mathbb{E}\left(\boldsymbol{U}_i^T\boldsymbol{U}_j\right)^2\mathbb{E}\left\{\left(1+\alpha_{ij}\right)^{-1/2}\left(1+\alpha_{ji}\right)^{-1/2}-1\right\}^2\\
&=O\left(n^{-3}\log p\left[p^{-2}\operatorname{tr}(\mathbf{R}^2)+O(p^{-2-\delta/2})\{p+po(\frac{p}{n^{1/2}})\}\right]\right)=o\left(\sigma_n^2\right).
\end{aligned}
$$

$$A_{n2} = \frac{4}{n(n-1)} \sum_{i<j} \boldsymbol{U}_i^\top \left(\hat{\mathbf{D}}_{ij}^{-1/2}\mathbf{D}^{1/2} - \mathbf{I}_p\right) \boldsymbol{U}_j$$
$$+ \frac{4}{n(n-1)} \sum_{i<j} \boldsymbol{U}_i^\top \left(\hat{\mathbf{D}}_{ij}^{-1/2}\mathbf{D}^{1/2} - \mathbf{I}_p\right) \boldsymbol{U}_j \left\{(1+\alpha_{ij})^{-1/2}(1+\alpha_{ji})^{-1/2} - 1\right\}$$
$$:= G_{n1} + G_{n2}.$$

By Lemma 2 and Equation S2.1,

$$\mathbb{E}[\{\boldsymbol{U}_i^\top \left(\hat{\mathbf{D}}_{ij}^{-1/2}\mathbf{D}^{1/2} - \mathbf{I}_p\right) \boldsymbol{U}_j\}^2] \le O\{n^{-1}\log p \ \mathrm{tr}(\boldsymbol{\Sigma}_u^2)\} = o\{p^{-2}\mathrm{tr}(\mathbf{R}^2)\}.$$

Then we obtain $G_{n1} = o_p(\sigma_n)$. Similar to A_{n1}, we can show $G_{n2} = o_p(\sigma_n)$. Taking the same procedure as A_{n2}, we can obtain $A_{n3} = o_p(\sigma_n)$. Similarly to the processing of Equation S1.1, we get

$$\frac{2}{n(n-1)} \sum_{1\le i<j\le n} \boldsymbol{U}_i^\top \boldsymbol{U}_j = \frac{2}{n(n-1)} \sum_{1\le i<j\le n} \{\mathbf{D}^{-1/2}\boldsymbol{\Gamma} U(\boldsymbol{W}_i)\}^\top \mathbf{D}^{-1/2}\boldsymbol{\Gamma} U(\boldsymbol{W}_j) + o_p(\sigma_n).$$

We replace the Lemma 1 in Feng and Sun (2016) by Lemma 7, and final acquire

$$\sqrt{\frac{n(n-1)p^2}{2\,\mathrm{tr}\,(\mathbf{R}^2)}} \frac{2}{n(n-1)} \sum_{1\le i<j\le n} \{\mathbf{D}^{-1/2}\boldsymbol{\Gamma} U(\boldsymbol{W}_i)\}^\top \mathbf{D}^{-1/2}\boldsymbol{\Gamma} U(\boldsymbol{W}_j) \xrightarrow{d} N(0,1).$$

S2.5 Proof of Theorem 6(Asymptotically independent under H_0)

Proof. To prove T_{SUM} and T_{MAX} are asymptotically independent, it suffices to show that: Under H_0,

$$\mathbb{P}\left(\frac{T_{SUM}}{\sigma_n} \le x, T_{MAX} - 2\log p + \log\log p \le y\right) \to \Phi(x)\exp\left\{-\frac{1}{\sqrt{\pi}}e^{-y/2}\right\}. \tag{S2.2}$$

Let $u_p(y) = y + 2\log p - \log\log p$, and we rewrite Equation S2.2 as

$$\mathbb{P}(\frac{T_{SUM}}{\sigma_n} \le x, T_{MAX} \le u_p(y)) \to \Phi(x)\exp\left\{-\frac{1}{\sqrt{\pi}}e^{-y/2}\right\}. \tag{S2.3}$$

From the proof of Theorem 2 in Feng and Sun (2016), we acquire

$$T_{SUM} = \frac{2}{n(n-1)}\sum\sum_{i<j}\boldsymbol{U}_i^\top\boldsymbol{U}_j + o_p(\sigma_n), \tag{S2.4}$$

and it's easy to find that $\sigma_n^2 = \frac{2}{n(n-1)p} + o(\frac{1}{n^3})$ according to Assumption 7.

Combined with Lemma 1 in subsection S1.1, it suffice to show,

$$\begin{aligned} &\mathbb{P}(\frac{\frac{2}{n(n-1)}\sum\sum_{i<j}\boldsymbol{U}_i^\top\boldsymbol{U}_j}{\sigma_n} + o_p(1) \le x \\ &\quad , p\left\|n^{-1/2}\sum_{i=1}^n\boldsymbol{U}_i\right\|_\infty^2 + O_p(L_{n,p}) \le u_p(y)) \\ &\to \Phi(x)\exp\left\{-\frac{1}{\sqrt{\pi}}e^{-y/2}\right\}. \end{aligned} \tag{S2.5}$$

We next prove that,

$$\begin{aligned} &\mathbb{P}(\sqrt{\frac{n}{n-1}}\left(\frac{\|\sqrt{\frac{p}{n}}\sum_{i=1}^n\boldsymbol{U}_i\|_2^2 - p}{\sqrt{2\mathrm{tr}(\mathbf{R}^2)}}\right) \le x, \left\|\sqrt{\frac{p}{n}}\sum_{i=1}^n\boldsymbol{U}_i\right\|_\infty^2 \le u_p(y)) \\ &\to \Phi(x)\exp\left\{-\frac{1}{\sqrt{\pi}}e^{-y/2}\right\}. \end{aligned} \tag{S2.6}$$

When Equation S2.6 holds, combined with $O_p(L_{n,p}) = o_p(1)$, Equation S2.5 holds obviously, which means that the independence of T_{SUM} and T_{MAX} follows.

Proof of Equation S2.6: From the Theorem 2 in Feng et al. (2022), the Equation S2.6 holds if $\boldsymbol{U}_i$ follows the normal distribution. We then investigate the non-normal case. Let $\boldsymbol{\xi}_i = \boldsymbol{U}_i \in R^p, i = 1, 2, \cdots, n$. For $\boldsymbol{z} =$

$(z_1, \cdots, z_q)^\top \in R^q$, we consider a smooth approximation of the maximum function, namely,

$$F_\beta(\boldsymbol{z}) := \beta^{-1} \log\{\sum_{j=1}^{q} \exp(\beta z_j)\},$$

where $\beta > 0$ is the smoothing parameter that controls the level of approximation. An elementary calculation shows that for all $z \in R^q$,

$$0 \le F_\beta(\boldsymbol{z}) - \max_{1\le j\le q} z_j \le \beta^{-1} \log q.$$

Define $\sigma_S^2 = 2n^2 \text{ tr } (\mathbf{R}^2)$,

$$\begin{aligned} W(\boldsymbol{x}_1, \cdots, \boldsymbol{x}_n) &= \frac{\|\sqrt{\frac{p}{n}} \sum_{i=1}^{n} \boldsymbol{x}_i\|_2^2 - p}{\sqrt{2\text{tr}(\mathbf{R}^2)}} \\ &= \frac{p\sum_{i\neq j} \boldsymbol{x}_i^\top \boldsymbol{x}_j}{\sqrt{2n^2 \text{ tr } (\mathbf{R}^2)}} := \frac{p\sum_{i\neq j} \boldsymbol{x}_i^\top \boldsymbol{x}_j}{\sigma_S}, \\ V(\boldsymbol{x}_1, \cdots, \boldsymbol{x}_n) &= \beta^{-1} \log\{\sum_{j=1}^{p} \exp(\beta\sqrt{\frac{p}{n}} \sum_{i=1}^{n} \boldsymbol{x}_{i,j})\}. \end{aligned}$$

By setting $\beta = n^{1/8} \log n$, Equation S2.6 is equivalent to

$$P(W(\boldsymbol{\xi}_1, \cdots, \boldsymbol{\xi}_p) \le x, V(\boldsymbol{\xi}_1, \cdots, \boldsymbol{\xi}_p) \le u_p(y)) \to \Phi(x) \exp\{-\exp(y)\}. \tag{S2.7}$$

Suppose $\{\boldsymbol{Y}_1, \boldsymbol{Y}_2, \cdots, \boldsymbol{Y}_n\}$ are sample from $N(0, \mathbb{E}(\boldsymbol{U}_1^\top \boldsymbol{U}_1))$, and independent with $\boldsymbol{U}_1, \cdots, \boldsymbol{U}_n$ (or write as $\boldsymbol{\xi}_1, \cdots, \boldsymbol{\xi}_n$). The key idea is to show that: $(W(\boldsymbol{\xi}_1, \cdots, \boldsymbol{\xi}_n), V(\boldsymbol{\xi}_1, \cdots, \boldsymbol{\xi}_n))$ has the same limiting distribution as $(W(\boldsymbol{Y}_1, \cdots, \boldsymbol{Y}_n), V(\boldsymbol{Y}_1, \cdots, \boldsymbol{Y}_n))$.

Let $l_b^2(\mathbb{R})$ denote the class of bounded functions with bounded and continuous derivatives up to order 3. It is known that a sequence of randon variables $\{Z_n\}_{n=1}^{\infty}$ converges weakly to a random variable Z if and only if for every $f \in l_b^3(\mathbb{R})$, $\mathbb{E}\{f(Z_n)\} \to \mathbb{E}\{f(Z)\}$.

It suffices to show that:

$$\mathbb{E}\{f(W(\boldsymbol{\xi}_1,\cdots,\boldsymbol{\xi}_n),V(\boldsymbol{\xi}_1,\cdots,\boldsymbol{\xi}_n))\}-\mathbb{E}\{f(W(\boldsymbol{Y}_1,\cdots,\boldsymbol{Y}_n),V(\boldsymbol{Y}_1,\cdots,\boldsymbol{Y}_n))\}\to 0,$$

for every $f \in l_b^3(\mathbb{R}^2)$ as $(n,p)\to\infty$.

We introduce $\widetilde{W}_d = W(\boldsymbol{\xi}_1,\cdots,\boldsymbol{\xi}_{d-1},\boldsymbol{Y}_d,\cdots,\boldsymbol{Y}_n)$ and $\widetilde{V}_d = V(\boldsymbol{\xi}_1,\cdots,\boldsymbol{\xi}_{d-1},$ $\boldsymbol{Y}_d,\cdots,\boldsymbol{Y}_n)$ for $d=1,\cdots,n+1$, $\mathcal{F}_d = \sigma\{\boldsymbol{\xi}_1,\cdots,\boldsymbol{\xi}_{d-1},\boldsymbol{Y}_{d+1},\cdots,\boldsymbol{Y}_n\}$ for $d=1,\cdots,n$. If there is no danger of confusion, we simply write $\widetilde{W}_d$ and $\widetilde{V}_d$ as W_d and V_d respectively (only for this part). Then,

$$\begin{aligned}&\mid \mathbb{E}\{f(W(\boldsymbol{\xi}_1,\cdots,\boldsymbol{\xi}_n),V(\boldsymbol{\xi}_1,\cdots,\boldsymbol{\xi}_n))\}-\mathbb{E}\{f(W(\boldsymbol{Y}_1,\cdots,\boldsymbol{Y}_n),V(\boldsymbol{Y}_1,\cdots,\boldsymbol{Y}_n))\}\mid\\ \le&\sum_{d=1}^{n}\mid \mathbb{E}\{f(W_d,V_d)-\mathbb{E}\{f(W_{d+1},V_{d+1})\}\mid.\end{aligned}$$

Let

$$W_{d,0}=\frac{2p\sum_{i<j}^{d-1}\boldsymbol{\xi}_i^{\top}\boldsymbol{\xi}_j+2p\sum_{d+1\le i<j\le n}\boldsymbol{Y}_i^{\top}\boldsymbol{Y}_j+2p\sum_{i=1}^{d-1}\sum_{j=d+1}^{n}\boldsymbol{\xi}_i^{\top}\boldsymbol{Y}_j}{\sigma_S}\in\mathcal{F}_d,$$

$$V_{d,0}=\beta^{-1}\log\{\sum_{j=1}^{p}\exp(\beta\sqrt{\frac{p}{n}}\sum_{i=1}^{d-1}\xi_{i,j}+\beta\sqrt{\frac{p}{n}}\sum_{i=d+1}^{n}Y_{i,j})\}\in\mathcal{F}_d.$$

By Taylor expansion, we have,

$$\begin{aligned}
& f\left(W_d, V_d\right)-f\left(W_{d,0}, V_{d,0}\right) \\
= & f_1\left(W_{d,0}, V_{d,0}\right)\left(W_d-W_{d,0}\right)+f_2\left(W_{d,0}, V_{d,0}\right)\left(V_d-V_{d,0}\right) \\
& +\frac{1}{2} f_{11}\left(W_{d,0}, V_{d,0}\right)\left(W_d-W_{d,0}\right)^2+\frac{1}{2} f_{22}\left(W_{d,0}, V_{d,0}\right)\left(V_d-V_{d,0}\right)^2 \\
& +\frac{1}{2} f_{12}\left(W_{d,0}, V_{d,0}\right)\left(W_d-W_{d,0}\right)\left(V_d-V_{d,0}\right) \\
& +O\left(\left|\left(V_d-V_{d,0}\right)\right|^3\right)+O\left(\left|\left(W_d-W_{d,0}\right)\right|^3\right),
\end{aligned}$$

and

$$\begin{aligned}
& f\left(W_{d+1}, V_{d+1}\right)-f\left(W_{d,0}, V_{d,0}\right) \\
= & f_1\left(W_{d,0}, V_{d,0}\right)\left(W_{d+1}-W_{d,0}\right)+f_2\left(W_{d,0}, V_{d,0}\right)\left(V_{d+1}-V_{d,0}\right) \\
& +\frac{1}{2} f_{11}\left(W_{d,0}, V_{d,0}\right)\left(W_{d+1}-W_{d,0}\right)^2+\frac{1}{2} f_{22}\left(W_{d,0}, V_{d,0}\right)\left(V_{d+1}-V_{d,0}\right)^2 \\
& +\frac{1}{2} f_{12}\left(W_{d,0}, V_{d,0}\right)\left(W_{d+1}-W_{d,0}\right)\left(V_{d+1}-V_{d,0}\right) \\
& +O\left(\left|\left(V_{d+1}-V_{d,0}\right)\right|^3\right)+O\left(\left|\left(W_{d+1}-W_{d,0}\right)\right|^3\right),
\end{aligned}$$

where for $f := f(x,y)$, $f_1(x,y)=\frac{\partial f}{\partial x}$, $f_2(x,y)=\frac{\partial f}{\partial y}$, $f_{11}(x,y)=\frac{\partial f^2}{\partial^2 x}$, $f_{22}(x,y)=\frac{\partial f^2}{\partial^2 y}$ and $f_{12}(x,y)=\frac{\partial f^2}{\partial x \partial y}$.

We first consider $W_d, W_{d+1}, W_{d,0}$ and notice that,

$$\begin{aligned}
W_d-W_{d,0} & =\frac{p \sum_{i=1}^{d-1} \boldsymbol{\xi}_i^{\top} \boldsymbol{Y}_d+p \sum_{i=d+1}^n \boldsymbol{Y}_i^{\top} \boldsymbol{Y}_d}{\sigma_S}, \\
W_{d+1}-W_{d,0} & =\frac{p \sum_{i=1}^{d-1} \boldsymbol{\xi}_i^{\top} \boldsymbol{\xi}_d+p \sum_{i=d+1}^n \boldsymbol{Y}_i^{\top} \boldsymbol{\xi}_d}{\sigma_S}.
\end{aligned}$$

Due to $\mathbb{E}\left(\boldsymbol{\xi}_t\right)=\mathbb{E}\left(\boldsymbol{Y}_t\right)=0$ and $\mathbb{E}\left(\boldsymbol{\xi}_t \boldsymbol{\xi}_t^{\top}\right)=\mathbb{E}\left(\boldsymbol{Y}_t \boldsymbol{Y}_t^{\top}\right)$, it can be verified that,

$$\mathbb{E}\left(W_d - W_{d,0} \mid \mathcal{F}_d\right) = \mathbb{E}\left(W_{d+1} - W_{d,0} \mid \mathcal{F}_d\right) \text{ and}$$

$$\mathbb{E}\left((W_d - W_{d,0})^2 \mid \mathcal{F}_d\right) = \mathbb{E}\left((W_{d+1} - W_{d,0})^2 \mid \mathcal{F}_d\right).$$

Hence,

$$\mathbb{E}\left\{f_1\left(W_{d,0}, V_{d,0}\right)\left(W_d - W_{d,0}\right)\right\} = \mathbb{E}\left\{f_1\left(W_{d,0}, V_{d,0}\right)\left(W_{d+1} - W_{d,0}\right)\right\} \text{ and}$$

$$\mathbb{E}\left\{f_{11}\left(W_{d,0}, V_{d,0}\right)\left(W_d - W_{d,0}\right)^2\right\} = \mathbb{E}\left\{f_{11}\left(W_{d,0}, V_{d,0}\right)\left(W_{d+1} - W_{d,0}\right)^2\right\}.$$

Next we consider $V_d - V_{d,0}$. Let $z_{d,0,j} = \sqrt{p/n}\sum_{i=1}^{d-1}\xi_{i,j} + \sqrt{p/n}\sum_{i=d+1}^{n} Y_{i,j}$, $z_{d,j} = z_{d,0,j} + n^{-1/2}\sqrt{p}Y_{d,j}$, $z_{d+1,j} = z_{d,0,j} + n^{-1/2}\sqrt{p}\xi_{d,j}$. By Taylor expansion, we have that:

$$\begin{aligned}
&V_d - V_{d,0} \\
=&\sum_{l=1}^{n}\partial_l F_\beta\left(\boldsymbol{z}_{d,0}\right)\left(z_{d,l} - z_{d,0,l}\right) + \frac{1}{2}\sum_{l=1}^{n}\sum_{k=1}^{n}\partial_k\partial_l F_\beta\left(\boldsymbol{z}_{d,0}\right)\left(z_{d,l} - z_{d,0,l}\right)\left(z_{d,k} - z_{d,0,k}\right) \\
&+ \frac{1}{6}\sum_{l=1}^{n}\sum_{k=1}^{n}\sum_{v=1}^{n}\partial_v\partial_k\partial_l F_\beta\left(\boldsymbol{z}_{d,0} + \delta\left(\boldsymbol{z}_d - \boldsymbol{z}_{d,0}\right)\right)\left(z_{d,l} - z_{d,0,l}\right)\left(z_{d,k} - \boldsymbol{z}_{d,0,k}\right)\left(\boldsymbol{z}_{d,v} - \boldsymbol{z}_{d,0,v}\right),
\end{aligned} \tag{S2.8}$$

for some $\delta \in (0,1)$. Again, due to $\mathbb{E}\left(\boldsymbol{\xi}_t\right) = \mathbb{E}\left(\boldsymbol{Y}_t\right) = 0$ and $\mathbb{E}\left(\boldsymbol{\xi}_t\boldsymbol{\xi}_t^\top\right) = \mathbb{E}\left(\boldsymbol{Y}_t\boldsymbol{Y}_t^\top\right)$, we can verify that

$$\mathbb{E}\left\{z_{d,l} - z_{d,0,l} \mid \mathcal{F}_d\right\} = \mathbb{E}\left\{z_{d+1,l} - z_{d,0,l} \mid \mathcal{F}_d\right\} \text{ and}$$

$$\mathbb{E}\left\{\left(z_{d,l} - z_{d,0,l}\right)^2 \mid \mathcal{F}_d\right\} = \mathbb{E}\left\{\left(z_{d+1,l} - z_{d,0,l}\right)^2 \mid \mathcal{F}_d\right\}.$$

By Lemma A.2 in Chernozhukov et al. (2013), we have,

$$\left|\sum_{l=1}^{n}\sum_{k=1}^{n}\sum_{v=1}^{n}\partial_v\partial_k\partial_l F_\beta\left(\boldsymbol{z}_{d,0} + \delta\left(\boldsymbol{z}_d - \boldsymbol{z}_{d,0}\right)\right)\right| \leq C\beta^2,$$

for some positive constant C.

By Lemma 7, we have that: $\left\|\zeta_1^{-1} U_{i,j}\right\|_{\psi_\alpha} \lesssim \bar{B}$, for all $i = 1, \ldots, n$ and $j = 1, \ldots, p$, which means $\mathbb{P}(| \sqrt{p}\xi_{i,j} | \geq t) \leq 2\exp(-(ct\sqrt{p}/\zeta_1)^\alpha) \lesssim 2\exp(-(ct)^\alpha)$, $\mathrm{P}\left(\max_{1\leq i\leq n}\left|\sqrt{p}\xi_{ij}\right| > C\log(n)\right) \to 0$ and since $\sqrt{p}Y_{tj} \sim N(0,1)$, $\mathrm{P}\left(\max_{1\leq i\leq n}\left|\sqrt{p}Y_{ij}\right| > C\log(n)\right) \to 0$. Hence,

$$\left|\frac{1}{6}\sum_{l=1}^{n}\sum_{k=1}^{n}\sum_{v=1}^{n}\partial_v\partial_k\partial_l F_\beta\left(\boldsymbol{z}_{d,0}+\delta\left(\boldsymbol{z}_d-\boldsymbol{z}_{d,0}\right)\right)\left(\boldsymbol{z}_{d,l}-\boldsymbol{z}_{d,0,l}\right)\left(\boldsymbol{z}_{d,k}-\boldsymbol{z}_{d,0,k}\right)\left(\boldsymbol{z}_{d,v}-\boldsymbol{z}_{d,0,v}\right)\right|$$
$$\leq C\beta^2 n^{-3/2}\log^3 n,$$
$$\left|\frac{1}{6}\sum_{l=1}^{n}\sum_{k=1}^{n}\sum_{v=1}^{n}\partial_v\partial_k\partial_l F_\beta\left(\boldsymbol{z}_{d+1,0}+\delta\left(\boldsymbol{z}_{d+1}-\boldsymbol{z}_{d,0}\right)\right)\left(\boldsymbol{z}_{d+1,l}-\boldsymbol{z}_{d,0,l}\right)\left(\boldsymbol{z}_{d+1,k}-\boldsymbol{z}_{d,0,k}\right)\left(\boldsymbol{z}_{d+1,v}-\boldsymbol{z}_{d,0,v}\right)\right|$$
$$\leq C\beta^2 n^{-3/2}\log^3 n,$$

holds with probability approaching one. Consequently, we have that: with probability one,

$$\left|\mathrm{E}\left\{f_2\left(W_{d,0},V_{d,0}\right)\left(V_d-V_{d,0}\right)\right\}-\mathrm{E}\left\{f_2\left(W_{d,0},V_{d,0}\right)\left(V_{d+1}-V_{d,0}\right)\right\}\right| \leq C\beta^2 n^{-3/2}\log^3 n.$$

Similarly, it can be verified that,

$$\left|\mathrm{E}\left\{f_{22}\left(W_{d,0},V_{d,0}\right)\left(V_d-V_{d,0}\right)^2\right\}-\mathrm{E}\left\{f_{22}\left(W_{d,0},V_{d,0}\right)\left(V_{d+1}-V_{d,0}\right)^2\right\}\right| \leq C\beta^2 n^{-3/2}\log^3 n,$$

and

$$\left|\mathrm{E}\left\{f_{12}\left(W_{d,0},V_{d,0}\right)\left(W_d-W_{d,0}\right)\left(V_d-V_{d,0}\right)\right\}-\mathrm{E}\left\{f_{12}\left(W_{d,0},V_{d,0}\right)\left(W_{d+1}-W_{d,0}\right)\left(V_{d+1}-V_{d,0}\right)\right\}\right|$$
$$\leq C\beta^2 n^{-3/2}\log^3 n.$$

By Equation S2.8, $\mathrm{E}\left(|V_d - V_{d,0}|^3\right) = O\left(n^{-3/2}\log^3 n\right)$. For $\mathbb{E}\left(|W_d - W_{d,0}|^3\right)$, we first calculate $\mathbb{E}\left\{(W_d - W_{d,0})^4\right\}$, then it's easy to get the order for 3-order term.

$$\begin{aligned}\mathbb{E}\left\{(W_d - W_{d,0})^4\right\} &= \mathbb{E}\left(\frac{p\sum_{i=1}^{d-1}\boldsymbol{\xi}_i^\top \boldsymbol{Y}_d + p\sum_{i=d+1}^{n}\boldsymbol{Y}_i^\top \boldsymbol{Y}_d}{\sigma_S}\right)^4 \\ &= \frac{p^4}{2n^4\{\mathrm{tr}\ (\mathbf{R}^2)\}^2}\mathbb{E}\left(\sum_{i=1}^{d-1}\boldsymbol{\xi}_i^\top \boldsymbol{Y}_d + \sum_{i=d+1}^{n}\boldsymbol{Y}_i^\top \boldsymbol{Y}_d\right)^4.\end{aligned} \tag{S2.9}$$

We consider the binomial expansion term and calculate them separately in Equation S2.9:

$$\begin{aligned}(i) =&\mathbb{E}(\sum_{i=d+1}^{n}\boldsymbol{Y}_i^\top \boldsymbol{Y}_d)^4,\ (ii) = \mathbb{E}\{(\sum_{i=1}^{d-1}\boldsymbol{\xi}_i^\top \boldsymbol{Y}_d)(\sum_{i=d+1}^{n}\boldsymbol{Y}_i^\top \boldsymbol{Y}_d)^3\},\\ (iii) =&\mathbb{E}\{(\sum_{i=1}^{d-1}\boldsymbol{\xi}_i^\top \boldsymbol{Y}_d)^2(\sum_{i=d+1}^{n}\boldsymbol{Y}_i^\top \boldsymbol{Y}_d)^2\},\\ (iv) =&\mathbb{E}\{(\sum_{i=1}^{d-1}\boldsymbol{\xi}_i^\top \boldsymbol{Y}_d)^3(\sum_{i=d+1}^{n}\boldsymbol{Y}_i^\top \boldsymbol{Y}_d)\},\ (v) = \mathbb{E}\{(\sum_{i=1}^{d-1}\boldsymbol{\xi}_i^\top \boldsymbol{Y}_d)^4\}.\end{aligned} \tag{S2.10}$$

Since $\mathbb{E}\boldsymbol{Y}_i = \mathbb{E}\boldsymbol{\xi}_i = 0$, we easily find that Equation S2.10 -(ii)(iv) equal to 0. Next we can get the following equations for Equation S2.10-(iii) after some straightforward calculations.

$$
\begin{aligned}
\mathbb{E}\{(\sum_{i=1}^{d-1} \boldsymbol{\xi}_i^\top \boldsymbol{Y}_d)^2 (\sum_{i=d+1}^{n} \boldsymbol{Y}_i^\top \boldsymbol{Y}_d)^2\} =& \mathbb{E}[\mathbb{E}[(\sum_{i=1}^{d-1} \boldsymbol{\xi}_i^\top \boldsymbol{Y}_d)^2 (\sum_{i=d+1}^{n} \boldsymbol{Y}_i^\top \boldsymbol{Y}_d)^2 \mid \boldsymbol{Y}_d]] \\
=& \mathbb{E}[(d-1)(n-d)(\boldsymbol{Y}_d^\top \boldsymbol{\Sigma}_u \boldsymbol{Y}_d)^2] \\
=& \mathbb{E}[(d-1)(n-d)((\boldsymbol{\Sigma}_u^{-1/2}\boldsymbol{Y}_d)^\top \boldsymbol{\Sigma}_u^2 (\boldsymbol{\Sigma}_u^{-1/2}\boldsymbol{Y}_d))^2] \\
=& (d-1)(n-d)2\text{tr }(\boldsymbol{\Sigma}_u^4) \\
\leq& (d-1)(n-d)O\{\text{tr}(\boldsymbol{\Sigma}_u^2)^2\}.
\end{aligned} \tag{S2.11}
$$

By some properties for standard normal random variable, the last inequality holds with some simple calculations shown below.

(i)

$$
\begin{aligned}
\text{tr }(\boldsymbol{\Sigma}_u^4) = \|\boldsymbol{\Sigma}_u^2\|_F^2 \leq (\|\boldsymbol{\Sigma}_u\|_F \|\boldsymbol{\Sigma}_u\|_F)^2 \\
= \|\boldsymbol{\Sigma}_u\|_F^4 = \text{tr }(\boldsymbol{\Sigma}_u^2)^2.
\end{aligned} \tag{S2.12}
$$

(ii)If $\boldsymbol{X}, \boldsymbol{Y} \overset{i.i.d.}{\sim} N(0, \mathbf{I}_p)$, then

$$
\begin{aligned}
&\mathbb{E}(\boldsymbol{X}^\top \mathbf{A} \boldsymbol{X})^2 = 2\text{tr}(\mathbf{A}^2) - \text{tr}^2(\mathbf{A}) \leq \text{tr}(\mathbf{A}^2), \\
&\mathbb{E}(\boldsymbol{X}^\top \mathbf{A} \boldsymbol{Y})^4 = \mathbb{E}[\mathbb{E}[(\boldsymbol{Y}^\top \mathbf{A} \boldsymbol{X} \boldsymbol{X}^\top \mathbf{A} \boldsymbol{Y})^2 \mid \boldsymbol{X}]] \leq 2\mathbb{E}[\text{tr}(\mathbf{A} \boldsymbol{X} \boldsymbol{X}^\top \mathbf{A})^2] \\
&\qquad = 2\mathbb{E}[(\boldsymbol{X}^\top \mathbf{A}^2 \boldsymbol{X})^2] \leq 4\text{tr}(\mathbf{A}^4).
\end{aligned} \tag{S2.13}
$$

For Equation S2.10-(1), according to $\sum_{i=d+1}^n \boldsymbol{Y}_i \sim N(0, (n-d)\boldsymbol{\Sigma}_u)$ and

Equation S2.12-S2.13, we have,

$$\begin{aligned}\mathbb{E}(\sum_{i=d+1}^{n}\boldsymbol{Y}_i^{\top}\boldsymbol{Y}_d)^4 =&\mathbb{E}\{(\frac{1}{\sqrt{(n-d)}}\boldsymbol{\Sigma}_u^{-1/2}\sum_{i=d+1}^{n}\boldsymbol{Y}_i)^{\top}(\sqrt{n-d}\boldsymbol{\Sigma}_u)(\boldsymbol{\Sigma}_u^{-1/2}\boldsymbol{Y}_d)\}^4\\ \leq&\text{tr }((n-d)^2\boldsymbol{\Sigma}_u^4) = (n-d)^2 O\{\text{tr}(\boldsymbol{\Sigma}_u^2)^2\}.\end{aligned} \tag{S2.14}$$

Similar with Equation S2.14, for Equation S2.10-(v),

$$\mathbb{E}\{(\sum_{i=1}^{d-1}\boldsymbol{\xi}_i^{\top}\boldsymbol{Y}_d)^4\} \leq (d-1)^2\text{tr }(\boldsymbol{\Sigma}_u^4) \leq (d-1)^2 O\{\text{tr}(\boldsymbol{\Sigma}_u^2)^2\}. \tag{S2.15}$$

Thus, in combining with the Equation S2.1,

$$\begin{aligned}&\mathbb{E}\left\{(W_d - W_{d,0})^4\right\}\\ =&\frac{p^4}{2n^4\{\text{tr }(\mathbf{R}^2)\}^2}\mathbb{E}\left(\sum_{i=1}^{d-1}\boldsymbol{\xi}_i^{\top}\boldsymbol{Y}_d + \sum_{i=d+1}^{n}\boldsymbol{Y}_i^{\top}\boldsymbol{Y}_d\right)^4\\ \leq&\frac{p^4}{2n^4\{\text{tr }(\mathbf{R}^2)\}^2}\{(d-1)(n-d)+(n-d)^2+(d-1)^2\}O\{\text{tr}(\boldsymbol{\Sigma}_u^2)^2\}\\ \leq&\frac{p^4}{2n^4(\text{tr }(\mathbf{R}^2))^2}n^2 O\{\text{tr}(\boldsymbol{\Sigma}_u^2)^2\} = O(\frac{1}{n^2}).\end{aligned}$$

By Jensen's inequality , we get

$$\sum_{d=1}^{n}\mathbb{E}\left|W_d - W_{d,0}\right|^3 \leq \sum_{d=1}^{n}\left\{\mathbb{E}\left(W_d - W_{d,0}\right)^4\right\}^{3/4} \leq C'n^{-1/2},$$

for some positive constant C', Combining all facts together, we conclude that

$$\sum_{d=1}^{n}\left|\mathbb{E}\left\{f\left(W_d, V_d\right)\right\} - \mathbb{E}\left\{f\left(W_{d+1}, V_{d+1}\right)\right\}\right| \leq C\beta^2 n^{-1/2}\log^3 n + C'n^{-1/2} \to 0,$$

as $(n,p)\to\infty$. The conclusion follows.

S2.6 Proof of Theorem 7(asymptotically independent under H_1)

Proof. From the proof of Theorem 2 in Feng and Sun (2016), we can find that

$$T_{SUM} = \frac{2}{n(n-1)} \sum \sum_{i<j} \boldsymbol{U}_i^\top \boldsymbol{U}_j + \zeta_1^2 \boldsymbol{\theta}^\top \mathbf{D}^{-1} \boldsymbol{\theta} + o_p(\sigma_n),$$

and according to Lemma 1 with minor modifications, we get the Bahadur representation in L^∞ norm,

$$n^{1/2}\mathbf{D}^{-1/2}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}) = n^{-1/2}\zeta_1^{-1} \sum_{i=1}^{n} (\boldsymbol{U}_i + \zeta_1 \mathbf{D}^{-1/2}\boldsymbol{\theta}) + C_n.$$

Similar with the proof in Theorem 6, it's suffice to show the result holds for normal version, i.e. it suffice to show that:

$$\|\sqrt{\frac{p}{n}} \sum_{i=1}^{n} \boldsymbol{Y}_i\|^2 \text{ and } \|\sqrt{\frac{p}{n}} \sum_{i=1}^{n} (\boldsymbol{Y}_i + \zeta_1 \mathbf{D}^{-1/2}\boldsymbol{\theta})\|_\infty^2,$$

are asymptotic independent, where $\{\boldsymbol{Y}_1, \boldsymbol{Y}_2, \cdots, \boldsymbol{Y}_n\}$ are sample from $N(0, \mathbb{E}\boldsymbol{U}_1^\top \boldsymbol{U}_1)$.

Denote $\sqrt{p/n}\sum_{i=1}^n \boldsymbol{Y}_i := \boldsymbol{\varphi} = (\varphi_1, \cdots, \varphi_p)^\top, \boldsymbol{\varphi}_{\mathcal{A}} = (\varphi_{j_1}, \cdots, \varphi_{j_d})^\top$, and $\boldsymbol{\varphi}_{\mathcal{A}^c} = (\varphi_{j_{d+1}}, \cdots, \varphi_{j_p})^\top$, where $\mathcal{A} = \{j_1, j_2, \cdots, j_d\}$. Then, $S = \|\boldsymbol{\varphi}\|^2 = \|\boldsymbol{\varphi}_{\mathcal{A}}\|^2 + \|\boldsymbol{\varphi}_{\mathcal{A}^c}\|^2$, $M = \|\boldsymbol{\varphi} + \sqrt{np}\zeta_1 \mathbf{D}^{-1/2}\boldsymbol{\theta}\|_\infty = \max_{i\in\mathcal{A}}(\varphi_i + \sqrt{np}\zeta_1 \mathbf{D}^{-1/2}\boldsymbol{\theta}) + \max_{i\in\mathcal{A}^c} \varphi_i$. From the proof of Theorem 6 in subsection S2.5, we know that $\|\boldsymbol{\varphi}_{\mathcal{A}^c}\|^2$ and $\max_{i\in\mathcal{A}^c} \varphi_i$ are asymptotically independent. Hence, it suffice to show that $\|\boldsymbol{\varphi}_{\mathcal{A}^c}\|^2$ is asymptotically independent with $\boldsymbol{\varphi}_{\mathcal{A}}$.

By Lemma 10, $\boldsymbol{\varphi}_{\mathcal{A}^c}$ can be decomposed as $\boldsymbol{\varphi}_{\mathcal{A}^c} = \boldsymbol{E} + \boldsymbol{F}$, where $\boldsymbol{E} = \boldsymbol{\varphi}_{\mathcal{A}^c} - \boldsymbol{\Sigma}_{U,\mathcal{A}^c,\mathcal{A}}\boldsymbol{\Sigma}_{U,\mathcal{A},\mathcal{A}}^{-1}\boldsymbol{\varphi}_{\mathcal{A}}$, $\boldsymbol{F} = \boldsymbol{\Sigma}_{U,\mathcal{A}^c,\mathcal{A}}\boldsymbol{\Sigma}_{U,\mathcal{A},\mathcal{A}}^{-1}\boldsymbol{\varphi}_{\mathcal{A}}$, $\boldsymbol{\Sigma}_U = p\mathbb{E}\boldsymbol{U}_1\boldsymbol{U}_1^\top = p\boldsymbol{\Sigma}_u$, which fulfill the properties $\boldsymbol{E} \sim N(0, \boldsymbol{\Sigma}_{U,\mathcal{A}^c,\mathcal{A}^c} - \boldsymbol{\Sigma}_{U,\mathcal{A}^c,\mathcal{A}}\boldsymbol{\Sigma}_{U,\mathcal{A},\mathcal{A}}^{-1}\boldsymbol{\Sigma}_{U,\mathcal{A},\mathcal{A}^c})$, $\boldsymbol{F} \sim N(0, \boldsymbol{\Sigma}_{U,\mathcal{A}^c,\mathcal{A}}\boldsymbol{\Sigma}_{U,\mathcal{A},\mathcal{A}}^{-1}\boldsymbol{\Sigma}_{U,\mathcal{A},\mathcal{A}^c})$ and $\boldsymbol{E}$ and $\boldsymbol{\varphi}_{\mathcal{A}}$ are independent.

Then, we rewrite

$$\|\boldsymbol{\varphi}_{\mathcal{A}^c}\|^2 = \boldsymbol{E}^\top\boldsymbol{E} + \boldsymbol{F}^\top\boldsymbol{F} + 2\boldsymbol{E}^\top\boldsymbol{F}.$$

According the proof of lemma S.7 in Feng et al. (2022), we have that:

$$\mathbb{P}(|\boldsymbol{F}^\top\boldsymbol{F} + 2\boldsymbol{E}^\top\boldsymbol{F}| \geq \epsilon\nu_p) \leq \frac{3}{p^t} \to 0,$$

by $d = o\{\lambda_{\min}(\mathbf{R})\mathrm{tr}(\mathbf{R}^2)^{1/2}/(\log p)^C\}$, where $\nu_p = \{2\mathrm{tr}(\mathbf{R}^2)\}^{1/2}$, $t = t_p := C\epsilon/8v_p/\{\lambda_{\max}(\mathbf{R})\log p\} \to \infty$,$\epsilon_p := (\log p)^C/\{v_p\lambda_{\min}(\mathbf{R})\} \to 0$.

S3 Some useful lemmas

Lemma 6. *(Lemma A3. in Cheng et al. (2023)) Suppose Assumptions 1-3 holds with $a_0(p) \asymp p^{1-\delta}$ for some positive constant $\delta \leq 1/2$. Define a random $p \times p$ matrix $\mathbf{Q} = n^{-1}\sum_{i=1}^n R_i^{-1}\boldsymbol{U}_i\boldsymbol{U}_i^\top$ and let $\mathbf{Q}_{jl}$ be the (j,l)th element of $\mathbf{Q}$. Then,*

(i)$|\mathbf{Q}_{jl}| \lesssim \zeta_1 p^{-1}|\sigma_{jl}| + O_p(\zeta_1 n^{-1/2}p^{-1} + \zeta_1 p^{-7/6} + \zeta_1 p^{-1-\delta/2})$.

(ii)$\mathbf{Q}_{jl} = \mathbf{Q}_{0,jl} + O_p(\zeta_1 p^{-7/6} + \zeta_1 p^{-1-\delta/2})$, *where* $\mathbf{Q}_{0,jl}$ *is the* (j,l)*th*

element of

$$\mathbf{Q}_0 = n^{-1}p^{-1/2}\sum_{i=1}^{n}\nu_i^{-1}\{\mathbf{D}^{-1/2}\boldsymbol{\Gamma}U(\boldsymbol{W}_i)\}\{\mathbf{D}^{-1/2}\boldsymbol{\Gamma}U(\boldsymbol{W}_i)\}^\top.$$

In addition, $\mathbf{Q}_0$ *satisfies*

$$tr[\mathbb{E}(\mathbf{Q}_0^2) - \{\mathbb{E}(\mathbf{Q}_0)\}^2] = O(n^{-1}p^{-1}).$$

Lemma 7. *(Lemma A4. in Cheng et al. (2023))Suppose Assumptions 1-3 holds with* $a_0(p) \asymp p^{1-\delta}$ *for some positive constant* $\delta \le 1/2$*. Then,*

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

(ii) $\|\zeta_1^{-1}U_{i,j}\|_{\psi_\alpha} \lesssim \bar{B}$ *for all* $i = 1, 2, \cdots, n$ *and* $j = 1, 2, \cdots, p$*.*

(iii)$\mathbb{E}(U_{i,j}^2) = p^{-1} + O(p^{-1-\delta/2})$ *for* $j = 1, 2, \cdots, p$ *and* $\mathbb{E}(U_{i,j}U_{i,l}) = p^{-1}\sigma_{j,l} + O(p^{-1-\delta/2})$ *for* $1 \le j \ne l \le p$*.*

(iv) if $\log p = o(n^{1/3})$*,*

$$\left|n^{-1/2}\sum_{i=1}^{n}\zeta_1^{-1}\boldsymbol{U}_i\right|_\infty = O_p\{\log^{1/2}(np)\} \text{ and } \left|n^{-1}\sum_{i=1}^{n}(\zeta_1^{-1}\boldsymbol{U}_i)^2\right|_\infty = O_p(1).$$

Lemma 8. *(Nazarov's inequality) Let* $\boldsymbol{Y}_0 = (Y_{0,1}, Y_{0,2}, \cdots, Y_{0,p})^\top$ *be a centered Gaussian random vector in* $\mathbb{R}^p$ *and* $\mathbb{E}(Y_{0,j}^2) \ge b$ *for all* $j = 1, 2, \cdots, p$ *and some constant* $b > 0$*, then for every* $y \in \mathbb{R}^p$ *and* $a > 0$*,*

$$\mathbb{P}(\boldsymbol{Y}_0 \le y + a) - \mathbb{P}(\boldsymbol{Y}_0 \le y) \lesssim a\log^{1/2}(p).$$

Lemma 9. *(Theorem 2 in Chernozhukov et al. (2015)) Let* $\boldsymbol{X} = (X_1, \ldots, X_p)^\top$

and $\boldsymbol{Y} = (Y_1, \ldots, Y_p)^\top$ *be centered Gaussian random vectors in* $\mathbb{R}^p$ *with covariance matrices* $\boldsymbol{\Sigma}^X = \left(\sigma_{jk}^X\right)_{1\le j,k\le p}$ *and* $\boldsymbol{\Sigma}^Y = \left(\sigma_{jk}^Y\right)_{1\le j,k\le p}$, *respectively. In terms of p,*

$$\Delta := \max_{1\le j,k\le p} \left|\sigma_{jk}^X - \sigma_{jk}^Y\right|, \text{ and } a_p := \mathrm{E}\left\{\max_{1\le j\le p} \left(Y_j/\sigma_{jj}^Y\right)\right\}.$$

Suppose that $p \ge 2$ *and* $\sigma_{jj}^Y > 0$ *for all* $1 \le j \le p$. *Then*

$$\sup_{x\in\mathbb{R}} \left|\mathbb{P}\left(\max_{1\le j\le p} X_j \le x\right) - \mathbb{P}\left(\max_{1\le j\le p} Y_j \le x\right)\right| \\ \le C\Delta^{1/3}\left\{\left(1 \vee a_p^2 \vee \log(1/\Delta)\right\}^{1/3} \log^{1/3} p,$$

where $C > 0$ *depends only on* $\min_{1\le j\le p} \sigma_{jj}^Y$ *and* $\max_{1\le j\le p} \sigma_{jj}^Y$ *(the right side is understood to be 0 when* $\Delta = 0$ *). Moreover, in the worst case,* $a_p \le \sqrt{2\log p}$, *so that*

$$\sup_{x\in\mathbb{R}} \left|\mathrm{P}\left(\max_{1\le j\le p} X_j \le x\right) - \mathrm{P}\left(\max_{1\le j\le p} Y_j \le x\right)\right| \le C'\Delta^{1/3}\{1 \vee \log(p/\Delta)\}^{2/3},$$

where as before $C' > 0$ *depends only on* $\min_{1\le j\le p} \sigma_{jj}^Y$ *and* $\max_{1\le j\le p} \sigma_{jj}^Y$.

Lemma 10. *(Theorem 1.2.11 in Muirhead (2009))Let* $\boldsymbol{X} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ *with invertible* $\boldsymbol{\Sigma}$, *and partition* $\boldsymbol{X}, \boldsymbol{\mu}$ *and* $\boldsymbol{\Sigma}$ *as*

$$\boldsymbol{X} = \begin{pmatrix} \boldsymbol{X}_1 \\ \boldsymbol{X}_2 \end{pmatrix}, \quad \boldsymbol{\mu} = \begin{pmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{pmatrix}, \quad \boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{pmatrix}.$$

Then $\boldsymbol{X}_2 - \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{X}_1 \sim N\left(\boldsymbol{\mu}_2 - \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_{22\cdot 1}\right)$ *and is independent of* $\boldsymbol{X}_1$, *where* $\boldsymbol{\Sigma}_{22\cdot 1} = \boldsymbol{\Sigma}_{22} - \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{12}$.

S4 Real data Application

S4.1 US stocks data

In this section, we utilize our methods to tackle a financial pricing problem. Our goal is to test whether the expected returns of all assets are equivalent to their respective risk-free returns. Let $X_{ij} \equiv R_{ij} - \mathrm{rf}_i$ denote the excess return of the jth asset at time i for $i = 1, \cdots, n$ and $j = 1, \cdots, p$, where R_{ij} is the return on asset j during period i and rf_i is the risk-free return rate of all asset during period i. We study the following pricing model

$$X_{ij} = \mu_j + \xi_{ij}, \tag{S4.1}$$

for $i = 1, \cdots, n$ and $j = 1, \cdots, p$, or, in vector form, $\mathbf{X}_i = \boldsymbol{\mu} + \boldsymbol{\xi}_i$, where $\mathbf{X}_i = (X_{i1}, \ldots, X_{ip})^\top$, $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_p)^\top$, and $\boldsymbol{\xi}_i = (\xi_{i1}, \ldots, \xi_{ip})^\top$ is the zero-mean error vector. We consider the following hypothesis

$$H_0 : \boldsymbol{\mu} = \mathbf{0} \quad \text{versus} \quad H_1 : \boldsymbol{\mu} \neq \mathbf{0}.$$

We examined the weekly return rates of stocks that are part of the S&P 500 index from January 2005 to November 2018. The weekly data were derived from the stock prices every Friday. Over time, the composition of the index changed and some stocks were introduced during this period. Therefore, we only considered a total of 424 stocks that were consistently included in the S&P 500 index throughout this period. We compiled a total of 716 weekly return rates for each stock during this period, excluding Fridays that were holidays. Given the possibility of autocorrelation in the weekly stock returns, we applied the Ljung-Box test Ljung and Box (1978) at a 0.05 level for zero autocorrelations to each stock. We retained 280 stocks for which the Ljung-Box test at a 0.05 level was not rejected. It's important to note that if we had used all 424 stocks, there might be autocorrelation between observations, which would violate our assumption and necessitate further studies.

Figure 1 show the histogram of standard deviation of those 280 securities. We found that the variances of those assets are obviously not equal. So the scalar-invariant test procedure is preferred. Thus, We apply the above six test procedures–SS-SUM,SS-MAX,SS-CC,MAX,SUM,COM on the total samples. All the tests reject the null hypothesis significantly. To evaluate the performance of our proposed tests and other competing tests for both small and large sample sizes, we randomly sampled $n = 52K$ observations from the 716 weekly returns, where K ranges from 3 to 8. This experiment was repeated 1000 times for each n value.

Table 1 presents the rejection rates of six tests. We discovered that spatial-sign based test procedures outperform mean-based test procedures. This is primarily due to the heavy-tailed nature of asset returns. Figure 2 displays Q-Q plots of the weekly return rates of some stocks in the S&P 500 index. We observed that all data deviate from a normal distribution and exhibit heavy tails. Additionally, sum-type tests perform better than max-type test procedures, mainly because the alternative is dense. Figure 3 illustrates the t-test statistic for each stock. We noticed that many t-test statistics are larger than 2, and most of them are positive. Among these tests, the SS-CC test performs the best. Although the SS-SUM outperforms the SS-MAX, the SS-MAX still retains some power in all cases. As indicated in the theoretical results in Subsection 3.2, our proposed Cauchy Combination would be more powerful than both max-type and sum-type tests in this scenario. Therefore, the application of real data also demonstrates the superiority of our proposed maxsum test procedure.

It's worth noting that the rejection of the null hypothesis, which suggests that return rates are not solely composed of risk-free rates on average, aligns with the perspectives of numerous economists. Indeed, the consideration of a non-zero excess return rate and the attempt to model it has spurred a vast amount of research on factor pricing models in finance (Sharpe, 1964; Fama and French, 1993, 2015). These models, which have many practical applications, operate under the Arbitrage Pricing Theory (Ross, 1976). Recently, numerous studies have focused on the high-dimensional alpha test under the linear factor pricing model, including works by Fan et al. (2015); Pesaran and Yamagata (2017); Feng et al. (2022); Liu et al. (2023). Notably, Liu et al. (2023) proposed a spatial-sign based sum-type test procedure for testing alpha for heavy-tailed distributions. It would be intriguing to extend the methods presented in this paper to propose a spatial sign based max-type and maxsum-type test procedures for testing alpha. This is an area that warrants further exploration.

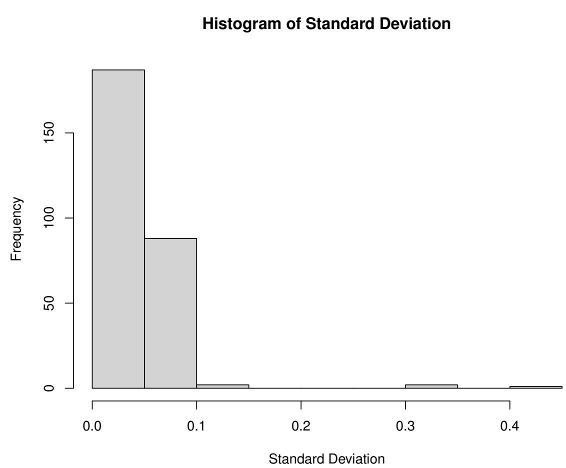

Figure 1: Histogram of standard deviation of US securities.

S4.2 Paired colon dataset

Another important application of the one-sample test discussed in this paper is assessing the mean difference between two paired samples. In this

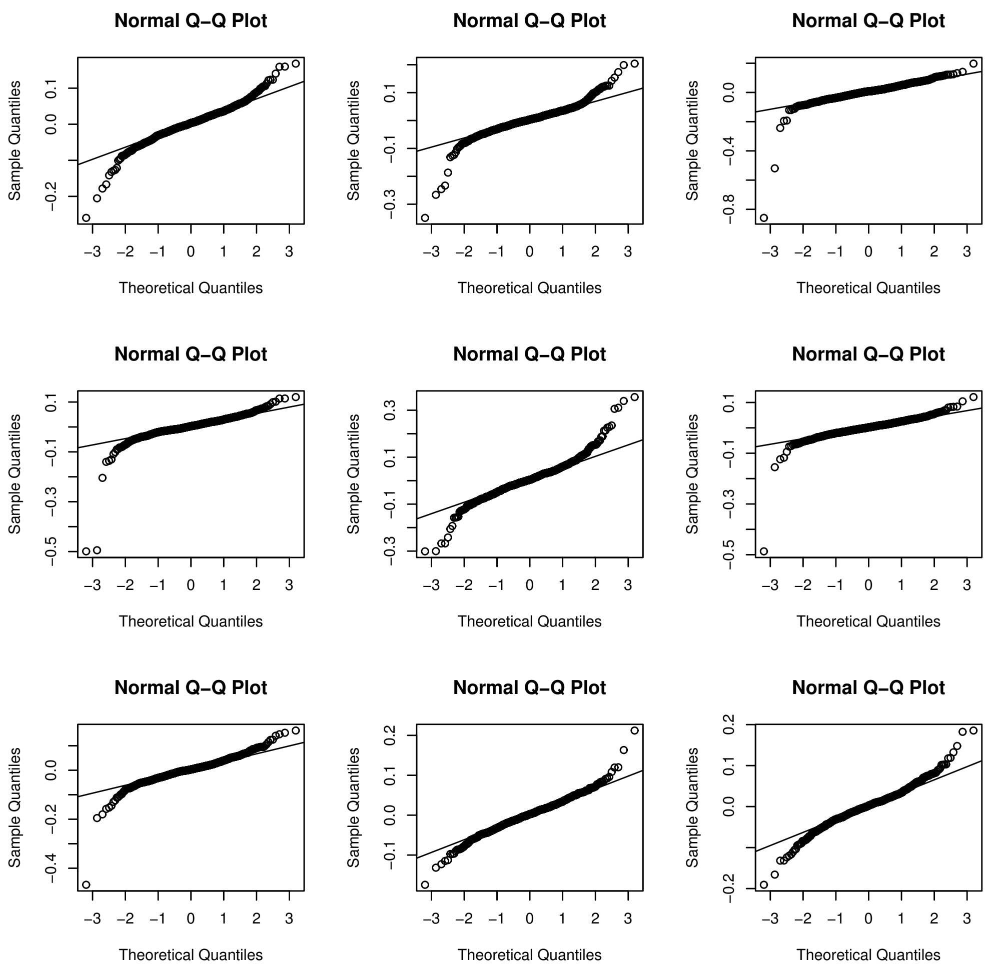

Figure 2: Q-Q plots of the weekly return rates of some stocks with heavy-tailed distributions in the S&P500 index.

section, we utilize our methods to test the mean difference for paired samples and consider the colon dataset provided by Alon et al. (1999). The colon dataset includes gene expression data from 40 colon cancer patients,

Table 1: The rejection rates of testing excess returns of the S&P stocks for $p = 280$ and $n = 52K$ with $K = 3, \cdots, 8$. For each n, we sampled 1000 data sets.

	SS-MAX	SS-SUM	SS-CC	MAX	SUM	COM
$n = 156$	0.295	0.361	0.380	0.124	0.219	0.204
$n = 208$	0.364	0.448	0.458	0.128	0.217	0.206
$n = 260$	0.424	0.542	0.556	0.140	0.276	0.246
$n = 312$	0.506	0.633	0.645	0.137	0.272	0.236
$n = 364$	0.652	0.738	0.758	0.143	0.317	0.282
$n = 416$	0.753	0.821	0.843	0.163	0.339	0.303

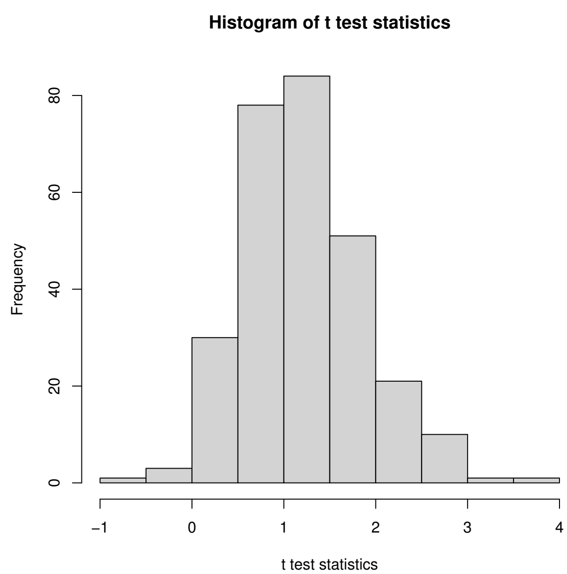

Figure 3: t test statistics of the weekly excess return rates of each stock.

comprising 22 paired samples from normal and tumor colon tissue and additional 18 samples from tumor tissue, with each sample containing 2,000 gene expressions. Our objective is to assess whether the mean gene expression levels differ between normal and tumor tissues. To streamline the analysis, we exclude unpaired samples and retain only the $n = 22$ paired normal and tumor tissue samples. To compare the methods, MAX, SUM and COM methods are also displayed. We observed that SUM test fails to reject the null hypothesis while others strongly reject if the significant level is set to $\alpha = 0.05$, see Figure 2. It is aligned with the simulation results of non-normal cases, suggesting a significant difference in gene expression between normal and tumor tissues, warranting further investigation. In addition, we find that the COM test successfully rejects the null hypothesis, whereas the SUM test does not. This indicates that methods based on the theorem of the independence between the test statistic can enhance test power while ensuring that Type I error remains controlled, particularly when data sparsity is uncertain.

Table 2: The p-values of testing the difference of gene expression levels of the normal and tumor colon tissues.

SS-MAX	SS-SUM	SS-CC	MAX	SUM	COM
6.88×10^{-5}	6.71×10^{-7}	1.33×10^{-6}	1.68×10^{-3}	9.62×10^{-2}	3.30×10^{-3}

References

Alon, U., N. Barkai, D. A. Notterman, K. Gish, S. Ybarra, D. Mack, and A. J. Levine (1999). Broad patterns of gene expression revealed by clustering analysis of tumor and normal colon tissues probed by oligonucleotide arrays. *Proceedings of the National Academy of Sciences 96*(12), 6745–6750.

Cheng, G., L. Peng, and C. Zou (2023). Statistical inference for ultrahigh dimensional location parameter based on spatial median. *arXiv preprint arXiv:2301.03126*.

Chernozhukov, V., D. Chetverikov, and K. Kato (2013). Gaussian approximations and multiplier bootstrap for maxima of sums of high-dimensional random vectors. *The Annals of Statistics 41*(6), 2786–2819.

Chernozhukov, V., D. Chetverikov, and K. Kato (2015). Comparison and anti-concentration bounds for maxima of gaussian random vectors. *Probability Theory and Related Fields 162*, 47–70.

Chernozhukov, V., D. Chetverikov, and K. Kato (2017). Central limit theorems and bootstrap in high dimensions. *Annals of probability: An official journal of the Institute of Mathematical Statistics 45*(4), 2309–2352.

Fama, E. F. and K. R. French (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics 33*(1), 3–56.

Fama, E. F. and K. R. French (2015). A five-factor asset pricing model. *Journal of Financial*

Economics 116(1), 1–22.

Fan, J., Y. Liao, and J. Yao (2015). Power enhancement in high dimensional cross-sectional tests. *Econometrica 83*(4), 1497–1541.

Feng, L., T. Jiang, X. Li, and B. Liu (2022). Asymptotic independence of the sum and maximum of dependent random variables with applications to high-dimensional tests. *arXiv preprint arXiv:2205.01638*.

Feng, L., W. Lan, B. Liu, and Y. Ma (2022). High-dimensional test for alpha in linear factor pricing models with sparse alternatives. *Journal of Econometrics 229*(1), 152–175.

Feng, L. and F. Sun (2016). Spatial-sign based high-dimensional location test. *Electronic Journal of Statistics 10*, 2420–2434.

Feng, L., C. Zou, and Z. Wang (2016). Multivariate-sign-based high-dimensional tests for the two-sample location problem. *Journal of the American Statistical Association 111*(514), 721–735.

Hettmansperger, T. P. and R. H. Randles (2002). A practical affine equivariant multivariate median. *Biometrika 89*(4), 851–860.

Koike, Y. (2021). Notes on the dimension dependence in high-dimensional central limit theorems for hyperrectangles. *Japanese Journal of Statistics and Data Science 4*, 257–297.

Liu, B., L. Feng, and Y. Ma (2023). High-dimensional alpha test of linear factor pricing models with heavy-tailed distributions. *Statistica Sinica 33*, 1389–1410.

Ljung, G. M. and G. E. Box (1978). On a measure of lack of fit in time series models. *Biometrika 65*(2), 297–303.

Muirhead, R. J. (2009). *Aspects of multivariate statistical theory.* John Wiley & Sons.

Pesaran, M. H. and T. Yamagata (2017). Testing for alpha in linear factor pricing models with a large number of securities. *Social Science Electronic Publishing*.

Ross, S. A. (1976). The arbitrage theory of capital asset pricing. *Journal of Economic Theory 13*(3), 341–360.

Sharpe, W. F. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. *The Journal of Finance 19*(3), 425–442.