Supplement on "Structural Testing of High-Dimensional Correlation Matrices"

Tingting Zou, Guangren Yang, Ruitao Lin, Guo-Liang Tian and Shurong Zheng

Jilin University, Jinan University,

The University of Texas MD Anderson Cancer Center,

Southern University of Science and Technology and Northeast Normal University

This supplement consists of the detailed proofs of Lemma 1, Theorems 1, 2, 3, 5, and Corollary 3 in "Structural Testing of High-Dimensional Correlation Matrices". We also present the simulation results when the observations are generated from the Gamma population.

S1 Detailed proof of Lemma 1

This section is devoted to presenting the detailed proof of Lemma 1. Let $\mathbf{e}_\ell$ denote the ℓth column of $p \times p$ dimensional identity matrix $\mathbf{I}_p$, $\mathbf{\Gamma} = [\mathrm{diag}(\mathbf{\Sigma})]^{-1/2}\mathbf{\Sigma}^{1/2}$, and $\mathbf{C} = (c_{ij})_{i,j=1}^p$ be a $p \times p$ dimensional matrix with $c_{ij} = 2r_{ij}^3 + \beta_w r_{ij} \sum_{k=1}^p (\mathbf{e}_i^T\mathbf{\Gamma}\mathbf{e}_k)^2(\mathbf{e}_j^T\mathbf{\Gamma}\mathbf{e}_k)^2$. In order to simplify the expressions of mean and variance-covariance functions in the CLT given by Lemma 1, we define the following quantites

$$f_1(\mathbf{R}, i, j) = 2r_{ij}^2 + \beta_w \sum_{k=1}^p (\mathbf{e}_i^T\mathbf{\Gamma}\mathbf{e}_k)^2(\mathbf{e}_j^T\mathbf{\Gamma}\mathbf{e}_k)^2,$$

$$f_2(\mathbf{R}, \mathbf{D}_2) = 2\mathrm{tr}(\mathbf{R}\mathbf{D}_2)^2 + \beta_w \sum_{k=1}^p \mathbf{e}_k^T\mathbf{R}\mathbf{D}_2\mathbf{R}\mathbf{D}_2\mathbf{e}_k \sum_{\ell=1}^p (\mathbf{e}_k^T\mathbf{\Gamma}\mathbf{e}_\ell)^4,$$

$$f_3(\mathbf{R}, \mathbf{D}_2) = 2\mathrm{tr}(\mathbf{R}\mathbf{D}_2) + \beta_w \sum_{k=1}^p \mathbf{e}_k^T\mathbf{R}\mathbf{D}_2\mathbf{e}_k \sum_{\ell=1}^p (\mathbf{e}_k^T\mathbf{\Gamma}\mathbf{e}_\ell)^4,$$

$$f_4(\mathbf{R}, \mathbf{D}_1, i) = 2\mathbf{e}_i^T\mathbf{R}\mathbf{D}_1\mathbf{R}\mathbf{e}_i + \beta_w \sum_{k=1}^p (\mathbf{e}_i^T\mathbf{\Gamma}\mathbf{e}_k)^2\mathbf{e}_k^T\mathbf{\Gamma}^T\mathbf{D}_1\mathbf{\Gamma}\mathbf{e}_k,$$

$$f_5(\mathbf{R},\mathbf{D}_2) = 2\mathrm{tr}(\mathbf{R}\mathbf{D}_2)^3 + \beta_w \sum_{k=1}^{p} \mathbf{e}_k^T\mathbf{\Gamma}^T\mathbf{D}_2\mathbf{R}\mathbf{D}_2\mathbf{\Gamma}\mathbf{e}_k\mathbf{e}_k^T\mathbf{\Gamma}^T\mathbf{D}_2\mathbf{\Gamma}\mathbf{e}_k,$$

$$f_6(\mathbf{R},\mathbf{D}_2,i) = 2\mathbf{e}_i^T(\mathbf{R}\mathbf{D}_2)^2\mathbf{R}\mathbf{e}_i + \beta_w \sum_{k=1}^{p} (\mathbf{e}_i^T\mathbf{\Gamma}\mathbf{e}_k)^2\mathbf{e}_k^T\mathbf{\Gamma}^T\mathbf{D}_2\mathbf{R}\mathbf{D}_2\mathbf{\Gamma}\mathbf{e}_k,$$

$$f_7(\mathbf{R},\mathbf{D}_2,i) = 2\mathbf{e}_i^T\mathbf{R}\mathbf{D}_2\mathbf{R}\mathbf{e}_i + \beta_w \sum_{k=1}^{p} (\mathbf{e}_i^T\mathbf{\Gamma}\mathbf{e}_k)^2\mathbf{e}_k^T\mathbf{\Gamma}^T\mathbf{D}_2\mathbf{\Gamma}\mathbf{e}_k,$$

$$f_8(\mathbf{R},\mathbf{D}_1,\mathbf{D}_2) = 2\mathrm{tr}(\mathbf{R}\mathbf{D}_1(\mathbf{R}\mathbf{D}_2)^2) + \beta_w \sum_{k=1}^{p} \mathbf{e}_k^T\mathbf{\Gamma}^T\mathbf{D}_1\mathbf{\Gamma}\mathbf{e}_k\mathbf{e}_k^T\mathbf{\Gamma}^T\mathbf{D}_2\mathbf{R}\mathbf{D}_2\mathbf{\Gamma}\mathbf{e}_k,$$

$$f_9(\mathbf{R},\mathbf{D}_1,\mathbf{D}_2) = 2\mathrm{tr}(\mathbf{R}\mathbf{D}_1\mathbf{R}\mathbf{D}_2) + \beta_w \sum_{k=1}^{p} \mathbf{e}_k^T\mathbf{\Gamma}^T\mathbf{D}_1\mathbf{\Gamma}\mathbf{e}_k\mathbf{e}_k^T\mathbf{\Gamma}^T\mathbf{D}_2\mathbf{\Gamma}\mathbf{e}_k.$$

Lemma 1. *Under Assumptions A-B-C, if the spectral norms of $\mathbf{D}_1$ and $\mathbf{D}_2$ are uniformly bounded in p, then we have*

$$[\Lambda_n(\mathbf{D}_1,\mathbf{D}_2)]^{-1/2} \begin{pmatrix} \mathrm{tr}(\widehat{\mathbf{R}}_n\mathbf{D}_1) - \nu_1(\mathbf{D}_1) \\ \mathrm{tr}(\widehat{\mathbf{R}}_n\mathbf{D}_2\widehat{\mathbf{R}}_n\mathbf{D}_2) - \nu_2(\mathbf{D}_2) \end{pmatrix} \xrightarrow{d} N_2\left((0,0)^T, \mathbf{I}_2\right),$$

where

$$\begin{aligned}
\nu_1(\mathbf{D}_1) &= \mathrm{tr}(\mathbf{R}\mathbf{D}_1) - n^{-1}\Big[2\mathrm{tr}(\mathbf{R}\mathbf{D}_1) + \beta_w \sum_{k=1}^{p}\sum_{\ell=1}^{p}(\mathbf{e}_k^T\mathbf{\Gamma}\mathbf{e}_\ell)^3\mathbf{e}_k^T\mathbf{D}_1\mathbf{\Gamma}\mathbf{e}_\ell\Big] \\
&\quad + \frac{1}{4n}\mathrm{tr}(\mathbf{C}\mathbf{D}_1) + \frac{3}{4n}\Big[2\mathrm{tr}(\mathbf{R}\mathbf{D}_1) + \beta_w \sum_{k=1}^{p}\mathbf{e}_k^T\mathbf{R}\mathbf{D}_1\mathbf{e}_k\sum_{\ell=1}^{p}(\mathbf{e}_k^T\mathbf{\Gamma}\mathbf{e}_\ell)^4\Big], \\
\nu_2(\mathbf{D}_2) &= \frac{n-3}{n^2}\mathrm{tr}^2(\mathbf{R}\mathbf{D}_2) - 4n^{-1}\sum_{k=1}^{p}\mathbf{e}_k^T\mathbf{R}\mathbf{D}_2\mathbf{e}_k\mathbf{e}_k^T\mathbf{R}\mathbf{D}_2\mathbf{R}\mathbf{e}_k \\
&\quad + \frac{n-3}{n}\mathrm{tr}(\mathbf{R}\mathbf{D}_2)^2 + \beta_w n^{-1}\sum_{k=1}^{p}(\mathbf{e}_k^T\mathbf{\Gamma}^T\mathbf{D}_2\mathbf{\Gamma}\mathbf{e}_k)^2 \\
&\quad - 2\beta_w n^{-1}\mathrm{tr}(\mathbf{R}\mathbf{D}_2)n^{-1}\sum_{k=1}^{p}\sum_{\ell=1}^{p}(\mathbf{e}_k^T\mathbf{\Gamma}\mathbf{e}_\ell)^3\mathbf{e}_k^T\mathbf{D}_2\mathbf{\Gamma}\mathbf{e}_\ell \\
&\quad - 2\beta_w n^{-1}\sum_{k=1}^{p}\sum_{\ell=1}^{p}(\mathbf{e}_k^T\mathbf{\Gamma}\mathbf{e}_\ell)^3\mathbf{e}_k^T\mathbf{D}_2\mathbf{R}\mathbf{D}_2\mathbf{\Gamma}\mathbf{e}_\ell \\
&\quad - 2\beta_w n^{-1}\sum_{k=1}^{p}\sum_{\ell=1}^{p}(\mathbf{e}_k^T\mathbf{\Gamma}\mathbf{e}_\ell)^2\mathbf{e}_k^T\mathbf{D}_2\mathbf{\Gamma}\mathbf{e}_\ell\mathbf{e}_k^T\mathbf{R}\mathbf{D}_2\mathbf{\Gamma}\mathbf{e}_\ell \\
&\quad + \frac{3}{2n}f_2(\mathbf{R},\mathbf{D}_2) + \frac{3\mathrm{tr}(\mathbf{R}\mathbf{D}_2)}{2n^2}f_3(\mathbf{R},\mathbf{D}_2)
\end{aligned}$$

$$
\begin{aligned}
&+\frac{1}{2n}\sum_{i=1}^{p}\sum_{j=1}^{p}\mathbf{e}_i^T\mathbf{R}\mathbf{e}_j\mathbf{e}_i^T\mathbf{D}_2\mathbf{R}\mathbf{D}_2\mathbf{e}_j f_1(\mathbf{R},i,j)\\
&+\frac{1}{2n}\sum_{i=1}^{p}\sum_{j=1}^{p}\mathbf{e}_i^T\mathbf{D}_2\mathbf{e}_j\mathbf{e}_i^T\mathbf{R}\mathbf{D}_2\mathbf{R}\mathbf{e}_j f_1(\mathbf{R},i,j)\\
&+\frac{\operatorname{tr}(\mathbf{R}\mathbf{D}_2)}{2n^2}\sum_{i=1}^{p}\sum_{j=1}^{p}\mathbf{e}_i^T\mathbf{D}_2\mathbf{e}_j\mathbf{e}_i^T\mathbf{R}\mathbf{e}_j f_1(\mathbf{R},i,j)\\
&+\frac{1}{2n}\sum_{i=1}^{p}\sum_{j=1}^{p}\mathbf{e}_i^T\mathbf{R}\mathbf{D}_2\mathbf{e}_j\mathbf{e}_i^T\mathbf{D}_2\mathbf{R}\mathbf{e}_j f_1(\mathbf{R},i,j),
\end{aligned}
$$

and $\Lambda_n(\mathbf{D}_1,\mathbf{D}_2)=\begin{pmatrix}\sigma_{11}(\mathbf{D}_1) & \sigma_{12}(\mathbf{D}_1,\mathbf{D}_2)\\ \sigma_{21}(\mathbf{D}_1,\mathbf{D}_2) & \sigma_{22}(\mathbf{D}_2)\end{pmatrix}$ *with*

$$
\begin{aligned}
\sigma_{11}(\mathbf{D}_1) &= n^{-1}\Big[2\operatorname{tr}(\mathbf{R}\mathbf{D}_1)^2+\beta_w\sum_{k=1}^{p}(\mathbf{e}_k^T\boldsymbol{\Gamma}^T\mathbf{D}_1\boldsymbol{\Gamma}\mathbf{e}_k)^2\Big]\\
&\quad +n^{-1}\sum_{i=1}^{p}\sum_{j=1}^{p}\mathbf{e}_i^T\mathbf{R}\mathbf{D}_1\mathbf{e}_i\mathbf{e}_j^T\mathbf{R}\mathbf{D}_1\mathbf{e}_j f_1(\mathbf{R},i,j)\\
&\quad -2n^{-1}\sum_{i=1}^{p}\mathbf{e}_i^T\mathbf{R}\mathbf{D}_1\mathbf{e}_i f_4(\mathbf{R},\mathbf{D}_1,i),\\
\sigma_{22}(\mathbf{D}_2) &= 4n^{-1}\Big[2\operatorname{tr}(\mathbf{R}\mathbf{D}_2)^4+\beta_w\sum_{k=1}^{p}(\mathbf{e}_k^T\boldsymbol{\Gamma}^T\mathbf{D}_2\mathbf{R}\mathbf{D}_2\boldsymbol{\Gamma}\mathbf{e}_k)^2\Big]\\
&\quad +4[n^{-1}\operatorname{tr}(\mathbf{R}\mathbf{D}_2)]^2n^{-1}\Big[2\operatorname{tr}(\mathbf{R}\mathbf{D}_2)^2+\beta_w\sum_{k=1}^{p}(\mathbf{e}_k^T\boldsymbol{\Gamma}^T\mathbf{D}_2\boldsymbol{\Gamma}\mathbf{e}_k)^2\Big]\\
&\quad +4\big[n^{-1}\operatorname{tr}(\mathbf{R}\mathbf{D}_2)^2\big]^2+8n^{-2}\operatorname{tr}(\mathbf{R}\mathbf{D}_2)f_5(\mathbf{R},\mathbf{D}_2)\\
&\quad +4n^{-1}\sum_{i=1}^{p}\sum_{j=1}^{p}\mathbf{e}_i^T(\mathbf{R}\mathbf{D}_2)^2\mathbf{e}_i\mathbf{e}_j^T(\mathbf{R}\mathbf{D}_2)^2\mathbf{e}_j f_1(\mathbf{R},i,j)\\
&\quad +8n^{-1}\operatorname{tr}(\mathbf{R}\mathbf{D}_2)n^{-1}\sum_{i=1}^{p}\sum_{j=1}^{p}\mathbf{e}_i^T\mathbf{R}\mathbf{D}_2\mathbf{e}_i\mathbf{e}_j^T(\mathbf{R}\mathbf{D}_2)^2\mathbf{e}_j f_1(\mathbf{R},i,j)\\
&\quad +4[n^{-1}\operatorname{tr}(\mathbf{R}\mathbf{D}_2)]^2n^{-1}\sum_{i=1}^{p}\sum_{j=1}^{p}\mathbf{e}_i^T\mathbf{R}\mathbf{D}_2\mathbf{e}_i\mathbf{e}_j^T\mathbf{R}\mathbf{D}_2\mathbf{e}_j f_1(\mathbf{R},i,j)\\
&\quad -8n^{-1}\sum_{i=1}^{p}\mathbf{e}_i^T(\mathbf{R}\mathbf{D}_2)^2\mathbf{e}_i f_6(\mathbf{R},\mathbf{D}_2,i)
\end{aligned}
$$

$$
\begin{aligned}
& -8n^{-1}\mathrm{tr}(\mathbf{R}\mathbf{D}_2)n^{-1}\sum_{i=1}^{p}\mathbf{e}_i^T\mathbf{R}\mathbf{D}_2\mathbf{e}_i f_6(\mathbf{R},\mathbf{D}_2,i) \\
& -8n^{-1}\mathrm{tr}(\mathbf{R}\mathbf{D}_2)n^{-1}\sum_{i=1}^{p}\mathbf{e}_i^T(\mathbf{R}\mathbf{D}_2)^2\mathbf{e}_i f_7(\mathbf{R},\mathbf{D}_2,i) \\
& -8[n^{-1}\mathrm{tr}(\mathbf{R}\mathbf{D}_2)]^2 n^{-1}\sum_{i=1}^{p}\mathbf{e}_i^T\mathbf{R}\mathbf{D}_2\mathbf{e}_i f_7(\mathbf{R},\mathbf{D}_2,i),
\end{aligned}
$$

$$
\begin{aligned}
\sigma_{12}(\mathbf{D}_1,\mathbf{D}_2) \;=\;& 2n^{-1}f_8(\mathbf{R},\mathbf{D}_1,\mathbf{D}_2)+2n^{-2}\mathrm{tr}(\mathbf{R}\mathbf{D}_2)f_9(\mathbf{R},\mathbf{D}_1,\mathbf{D}_2) \\
& -2n^{-1}\sum_{i=1}^{p}\mathbf{e}_i^T(\mathbf{R}\mathbf{D}_2)^2\mathbf{e}_i f_4(\mathbf{R},\mathbf{D}_1,i) \\
& -2n^{-1}\mathrm{tr}(\mathbf{R}\mathbf{D}_2)n^{-1}\sum_{i=1}^{p}\mathbf{e}_i^T\mathbf{R}\mathbf{D}_2\mathbf{e}_i f_4(\mathbf{R},\mathbf{D}_1,i) \\
& -2n^{-1}\sum_{i=1}^{p}\mathbf{e}_i^T\mathbf{R}\mathbf{D}_1\mathbf{e}_i f_6(\mathbf{R},\mathbf{D}_2,i) \\
& -2n^{-1}\mathrm{tr}(\mathbf{R}\mathbf{D}_2)n^{-1}\sum_{i=1}^{p}\mathbf{e}_i^T\mathbf{R}\mathbf{D}_1\mathbf{e}_i f_7(\mathbf{R},\mathbf{D}_2,i) \\
& +2n^{-1}\sum_{i=1}^{p}\sum_{j=1}^{p}\mathbf{e}_i^T\mathbf{R}\mathbf{D}_1\mathbf{e}_i\mathbf{e}_j^T(\mathbf{R}\mathbf{D}_2)^2\mathbf{e}_j f_1(\mathbf{R},i,j) \\
& +2n^{-1}\mathrm{tr}(\mathbf{R}\mathbf{D}_2)n^{-1}\sum_{i=1}^{p}\sum_{j=1}^{p}\mathbf{e}_i^T\mathbf{R}\mathbf{D}_1\mathbf{e}_i\mathbf{e}_j^T\mathbf{R}\mathbf{D}_2\mathbf{e}_j f_1(\mathbf{R},i,j), \\
\sigma_{21}(\mathbf{D}_1,\mathbf{D}_2) \;=\;& \sigma_{12}(\mathbf{D}_1,\mathbf{D}_2).
\end{aligned}
$$

Recall that the definitions of the sample covariance and correlation matrices are

$$
\widehat{\mathbf{\Sigma}}_n = n^{-1}\sum_{k=1}^{n}(\mathbf{x}_k-\bar{\mathbf{x}})(\mathbf{x}_k-\bar{\mathbf{x}})^T, \quad \widehat{\mathbf{R}}_n = [\mathrm{diag}(\widehat{\mathbf{\Sigma}}_n)]^{-1/2}\widehat{\mathbf{\Sigma}}_n[\mathrm{diag}(\widehat{\mathbf{\Sigma}}_n)]^{-1/2}.
$$

Note that the correlation coefficient is invariant under shifting and positive scaling of random variables. Let $\mathbf{y}_k=\mathbf{\Gamma}\mathbf{w}_k$, $\bar{\mathbf{y}}=n^{-1}\sum_{k=1}^{n}\mathbf{y}_k$, and

$$
\widehat{\mathbf{\Sigma}}_n^* = n^{-1}\sum_{k=1}^{n}(\mathbf{y}_k-\bar{\mathbf{y}})(\mathbf{y}_k-\bar{\mathbf{y}})^T,
$$

then under Assumption A, that is, $\mathbf{x}_k$ has the independent component structure, we have

$$
\widehat{\mathbf{R}}_n = [\mathrm{diag}(\widehat{\mathbf{\Sigma}}_n^*)]^{-1/2}\widehat{\mathbf{\Sigma}}_n^*[\mathrm{diag}(\widehat{\mathbf{\Sigma}}_n^*)]^{-1/2}.
$$

Although a slight abuse of notions, for simplicity, in the following we still use $\widehat{\mathbf{\Sigma}}_n$ to represent the sample covariance matrix generated by $\mathbf{y}_1, \ldots, \mathbf{y}_n$, that is, we denote

$$\widehat{\mathbf{\Sigma}}_n = n^{-1} \sum_{k=1}^{n} (\mathbf{y}_k - \bar{\mathbf{y}})(\mathbf{y}_k - \bar{\mathbf{y}})^T.$$

in the following sections.

S1.1 Truncation, centralization and rescaling

For $k = 1, \ldots, n$, denote $\breve{\mathbf{w}}_k = (\breve{w}_{1k}, \ldots, \breve{w}_{pk})^T$ with $\breve{w}_{\ell k} = w_{\ell k} I\{|w_{\ell k}| < \eta_n \sqrt{n}\}$, where $\eta_n = (\log n)^{-(1+\epsilon)/2}$, $\breve{\mathbf{y}}_k = \mathbf{\Gamma} \breve{\mathbf{w}}_k$, $\bar{\breve{\mathbf{y}}} = n^{-1} \sum_{k=1}^{n} \breve{\mathbf{y}}_k$,

$$\breve{\mathbf{\Sigma}}_n = n^{-1} \sum_{k=1}^{n} (\breve{\mathbf{y}}_k - \bar{\breve{\mathbf{y}}})(\breve{\mathbf{y}}_k - \bar{\breve{\mathbf{y}}})^T \quad \text{and} \quad \breve{\mathbf{R}}_n = [\mathrm{diag}(\breve{\mathbf{\Sigma}}_n)]^{-1/2} \breve{\mathbf{\Sigma}}_n [\mathrm{diag}(\breve{\mathbf{\Sigma}}_n)]^{-1/2}.$$

According to the truncation lemma (Lemma 2) in Noureddine (2009), we have

$$\mathrm{P}(\widehat{\mathbf{R}}_n \neq \breve{\mathbf{R}}_n, i.o.) = 0.$$

Define $\tilde{\mathbf{w}}_k = (\tilde{w}_{1k}, \ldots, \tilde{w}_{pk})^T$ with $\tilde{w}_{\ell k} = (\breve{w}_{\ell k} - \mathrm{E}\breve{w}_{\ell k}) / \sqrt{\mathrm{E}(\breve{w}_{\ell k} - \mathrm{E}\breve{w}_{\ell k})^2}$. Also define $\tilde{\mathbf{y}}_k$, $\bar{\tilde{\mathbf{y}}}$, $\tilde{\mathbf{\Sigma}}_n$, and $\tilde{\mathbf{R}}_n$ as the analogues of $\mathbf{y}_k$, $\bar{\mathbf{y}}$, $\widehat{\mathbf{\Sigma}}_n$, and $\widehat{\mathbf{R}}_n$ with $\mathbf{w}_k$ replaced by $\tilde{\mathbf{w}}_k$. Based on (1.6) in Yin et al. (2023), for any $1 \leq \ell \leq p, 1 \leq k \leq n$, we have

$$|\mathrm{E}\breve{w}_{\ell k}| = o(n^{-3/2}), \quad \mathrm{E}(\breve{w}_{\ell k} - \mathrm{E}\breve{w}_{\ell k})^2 - 1 = o(n^{-1}), \tag{S1.1}$$

$$\left(\sqrt{\mathrm{E}(\breve{w}_{\ell k} - \mathrm{E}\breve{w}_{\ell k})^2} - 1\right) / \sqrt{\mathrm{E}(\breve{w}_{\ell k} - \mathrm{E}\breve{w}_{\ell k})^2} = o(n^{-1}). \tag{S1.2}$$

Since $\mathbf{R} = \mathbf{\Gamma}\mathbf{\Gamma}^T$, we have

$$\begin{aligned} \|\breve{\mathbf{\Sigma}}_n - \tilde{\mathbf{\Sigma}}_n\| &= \|\mathbf{\Gamma}(n^{-1} \sum_{k=1}^{n} \breve{\mathbf{w}}_k \breve{\mathbf{w}}_k^T - n^{-1} \sum_{k=1}^{n} \tilde{\mathbf{w}}_k \tilde{\mathbf{w}}_k^T) \mathbf{\Gamma}^T - \mathbf{\Gamma}(\bar{\breve{\mathbf{w}}} \bar{\breve{\mathbf{w}}}^T - \bar{\tilde{\mathbf{w}}} \bar{\tilde{\mathbf{w}}}^T) \mathbf{\Gamma}^T\| \\ &\leq \|\mathbf{R}\| \Big(\|n^{-1} \sum_{k=1}^{n} \breve{\mathbf{w}}_k \breve{\mathbf{w}}_k^T - n^{-1} \sum_{k=1}^{n} \tilde{\mathbf{w}}_k \tilde{\mathbf{w}}_k^T\| + \|\bar{\breve{\mathbf{w}}} \bar{\breve{\mathbf{w}}}^T - \bar{\tilde{\mathbf{w}}} \bar{\tilde{\mathbf{w}}}^T\| \Big), \end{aligned}$$

where $\|\cdot\|$ denotes the spectral norm of a matrix. According to (1.7) in Yin et al. (2023), $\|n^{-1}\sum_{k=1}^{n}\check{\mathbf{w}}_k\check{\mathbf{w}}_k^T - n^{-1}\sum_{k=1}^{n}\tilde{\mathbf{w}}_k\tilde{\mathbf{w}}_k^T\| = o_{a.s.}(n^{-1})$. Note that

$$\bar{\tilde{\mathbf{w}}} = \frac{1}{\sqrt{\mathrm{E}(\breve{w}_{\ell k} - \mathrm{E}\breve{w}_{\ell k})^2}}\bar{\check{\mathbf{w}}} - \frac{|\mathrm{E}\breve{w}_{\ell k}|}{\sqrt{\mathrm{E}(\breve{w}_{\ell k} - \mathrm{E}\breve{w}_{\ell k})^2}}\mathbf{1}_p,$$
$$\bar{\check{\mathbf{w}}} - \bar{\tilde{\mathbf{w}}} = \frac{\sqrt{\mathrm{E}(\breve{w}_{\ell k} - \mathrm{E}\breve{w}_{\ell k})^2} - 1}{\sqrt{\mathrm{E}(\breve{w}_{\ell k} - \mathrm{E}\breve{w}_{\ell k})^2}}\bar{\check{\mathbf{w}}} + \frac{|\mathrm{E}\breve{w}_{\ell k}|}{\sqrt{\mathrm{E}(\breve{w}_{\ell k} - \mathrm{E}\breve{w}_{\ell k})^2}}\mathbf{1}_p,$$

combining the above equations with (S1.1) and (S1.2), we obtain

$$\|\bar{\check{\mathbf{w}}}\bar{\check{\mathbf{w}}}^T - \bar{\tilde{\mathbf{w}}}\bar{\tilde{\mathbf{w}}}^T\| \leq \|\bar{\check{\mathbf{w}}}(\bar{\check{\mathbf{w}}} - \bar{\tilde{\mathbf{w}}})^T\| + \|(\bar{\check{\mathbf{w}}} - \bar{\tilde{\mathbf{w}}})\bar{\tilde{\mathbf{w}}}^T\| = o_p(n^{-1}),$$

the last equality follows from the fact that $\bar{\check{\mathbf{w}}}^T\bar{\check{\mathbf{w}}} = O_p(1)$. Therefore, we have

$$\|\check{\mathbf{\Sigma}}_n - \tilde{\mathbf{\Sigma}}_n\| = o_p(n^{-1}). \tag{S1.3}$$

Let $\boldsymbol{\gamma}_\ell$ be the transpose of the ℓth row of the matrix $\mathbf{\Gamma}$, it can be verified that

$$\begin{aligned}\max_{\ell=1,\dots,p}\|\mathbf{e}_\ell^T\check{\mathbf{\Sigma}}_n\mathbf{e}_\ell - \mathbf{e}_\ell^T\tilde{\mathbf{\Sigma}}_n\mathbf{e}_\ell\| \leq \max_{\ell=1,\dots,p}\Big(&\|\boldsymbol{\gamma}_\ell^T(n^{-1}\sum_{k=1}^{n}\check{\mathbf{w}}_k\check{\mathbf{w}}_k^T - n^{-1}\sum_{k=1}^{n}\tilde{\mathbf{w}}_k\tilde{\mathbf{w}}_k^T)\boldsymbol{\gamma}_\ell\| \\ &+ \|\boldsymbol{\gamma}_\ell^T(\bar{\check{\mathbf{w}}}\bar{\check{\mathbf{w}}}^T - \bar{\tilde{\mathbf{w}}}\bar{\tilde{\mathbf{w}}}^T)\boldsymbol{\gamma}_\ell\|\Big) = o_p(n^{-1}).\end{aligned} \tag{S1.4}$$

Combining (S1.3) and (S1.4), we get

$$\begin{aligned}\|\check{\mathbf{R}}_n - \tilde{\mathbf{R}}_n\| &= \|\mathrm{diag}^{-1/2}(\check{\mathbf{\Sigma}}_n)\check{\mathbf{\Sigma}}_n\mathrm{diag}^{-1/2}(\check{\mathbf{\Sigma}}_n) - \mathrm{diag}^{-1/2}(\tilde{\mathbf{\Sigma}}_n)\tilde{\mathbf{\Sigma}}_n\mathrm{diag}^{-1/2}(\tilde{\mathbf{\Sigma}}_n)\| \\ &\leq \|\big(\mathrm{diag}^{-1/2}(\check{\mathbf{\Sigma}}_n) - \mathrm{diag}^{-1/2}(\tilde{\mathbf{\Sigma}}_n)\big)\check{\mathbf{\Sigma}}_n\mathrm{diag}^{-1/2}(\check{\mathbf{\Sigma}}_n)\| \\ &\quad + \|\mathrm{diag}^{-1/2}(\tilde{\mathbf{\Sigma}}_n)\big(\check{\mathbf{\Sigma}}_n - \tilde{\mathbf{\Sigma}}_n\big)\mathrm{diag}^{-1/2}(\check{\mathbf{\Sigma}}_n)\| \\ &\quad + \|\mathrm{diag}^{-1/2}(\tilde{\mathbf{\Sigma}}_n)\tilde{\mathbf{\Sigma}}_n\big(\mathrm{diag}^{-1/2}(\check{\mathbf{\Sigma}}_n) - \mathrm{diag}^{-1/2}(\tilde{\mathbf{\Sigma}}_n)\big)\| = o_p(n^{-1}).\end{aligned}$$

It follows that

$$|\mathrm{tr}(\check{\mathbf{R}}_n\mathbf{D}_1) - \mathrm{tr}(\tilde{\mathbf{R}}_n\mathbf{D}_1)| = \sum_{\ell=1}^{p}|\lambda_\ell\big((\check{\mathbf{R}}_n - \tilde{\mathbf{R}}_n)\mathbf{D}_1\big)| \leq p\|\check{\mathbf{R}}_n - \tilde{\mathbf{R}}_n\|\|\mathbf{D}_1\| = o_p(1).$$

Similarly, we have $|\mathrm{tr}(\check{\mathbf{R}}_n\mathbf{D}_2\check{\mathbf{R}}_n\mathbf{D}_2) - \mathrm{tr}(\tilde{\mathbf{R}}_n\mathbf{D}_2\tilde{\mathbf{R}}_n\mathbf{D}_2)| = o_p(1)$. Therefore, without loss of generality, in the following proof, we assume that $w_{\ell k}$ satisfies that $|w_{\ell k}| \leq \eta_n\sqrt{n}$, $\mathrm{E}w_{\ell k} = 0$ and $\mathrm{E}w_{\ell k}^2 = 1$.

S1.2 Some primary steps

We decompose $\mathrm{tr}\big(\widehat{\mathbf{R}}_n\mathbf{D}_1\big)$ and $\mathrm{tr}\big(\widehat{\mathbf{R}}_n\mathbf{D}_2\widehat{\mathbf{R}}_n\mathbf{D}_2\big)$ into the sum of several terms in this section, and then analyze the terms in the following sections. Since for $\ell = 1, \dots, p$,

$$\hat{\sigma}_{\ell\ell} = n^{-1}\sum_{k=1}^{n}(\mathbf{e}_\ell^T\boldsymbol{\Gamma}\mathbf{w}_k)^2 - (n^{-1}\sum_{k=1}^{n}\mathbf{e}_\ell^T\boldsymbol{\Gamma}\mathbf{w}_k)^2,$$

then we have

$$\max_{1\leq\ell\leq p}|\hat{\sigma}_{\ell\ell} - 1| \leq \max_{1\leq\ell\leq p}\left|n^{-1}\sum_{k=1}^{n}(\mathbf{e}_\ell^T\boldsymbol{\Gamma}\mathbf{w}_k)^2 - 1\right| + \max_{1\leq\ell\leq p}\left|n^{-1}\sum_{k=1}^{n}\mathbf{e}_\ell^T\boldsymbol{\Gamma}\mathbf{w}_k\right|^2.$$

By Lemma 5.7 in Yin and Ma (2022), we have

$$\max_{1\leq\ell\leq p}\left|n^{-1}\sum_{k=1}^{n}(\mathbf{e}_\ell^T\boldsymbol{\Gamma}\mathbf{w}_k)^2 - 1\right| = O_p(n^{-1/2}).$$

For any $\epsilon > 0$, by Markov's inequality, we get

$$P\left(\max_{1\leq\ell\leq p}n^{1/3}\left|n^{-1}\sum_{k=1}^{n}\mathbf{e}_\ell^T\boldsymbol{\Gamma}\mathbf{w}_k\right|^2 > \epsilon\right) \leq \epsilon^{-2}\sum_{\ell=1}^{p}n^{2/3}\mathrm{E}\left|n^{-1}\sum_{k=1}^{n}\mathbf{e}_\ell^T\boldsymbol{\Gamma}\mathbf{w}_k\right|^4 = o(1).$$

It follows that

$$\max_{1\leq\ell\leq p}|\hat{\sigma}_{\ell\ell} - 1| = o_p(n^{-1/3}). \tag{S1.5}$$

Moreover, by the Taylor expansion, we have

$$(\hat{\sigma}_{\ell\ell})^{-1/2} = 1 - \frac{1}{2}(\hat{\sigma}_{\ell\ell} - 1) + \frac{3}{8}(\hat{\sigma}_{\ell\ell} - 1)^2 - \frac{5}{16}(\sigma_{\ell\ell}^*)^{-7/2}(\hat{\sigma}_{\ell\ell} - 1)^3, \tag{S1.6}$$

where $\sigma_{\ell\ell}^*$ lies between 1 and $\hat{\sigma}_{\ell\ell}$. Since $\sigma_{\ell\ell}^* \geq \min(1, \hat{\sigma}_{\ell\ell})$ for all ℓ, based on (S1.5), we have

$$\max_{1\leq\ell\leq p}|(\sigma_{\ell\ell}^*)^{-7/2}(\hat{\sigma}_{\ell\ell} - 1)^3| = o_p(n^{-1}). \tag{S1.7}$$

Combining (S1.6) and (S1.7), we have

$$(\hat{\sigma}_{\ell\ell})^{-1/2} = 1 - \frac{1}{2}(\hat{\sigma}_{\ell\ell} - 1) + \frac{3}{8}(\hat{\sigma}_{\ell\ell} - 1)^2 + o_p(n^{-1}), \tag{S1.8}$$

where $o_p(n^{-1})$ is uniform for all $\ell = 1, ..., p$. Since

$$\begin{aligned}
\widehat{\mathbf{R}}_n &= \big[\mathrm{diag}(\widehat{\mathbf{\Sigma}}_n)\big]^{-1/2}\widehat{\mathbf{\Sigma}}_n\big[\mathrm{diag}(\widehat{\mathbf{\Sigma}}_n)\big]^{-1/2} \\
&= \widehat{\mathbf{\Sigma}}_n + \big[\mathrm{diag}^{-1/2}(\widehat{\mathbf{\Sigma}}_n) - \mathbf{I}_p\big]\widehat{\mathbf{\Sigma}}_n\mathrm{diag}^{-1/2}(\widehat{\mathbf{\Sigma}}_n) + \widehat{\mathbf{\Sigma}}_n\big[\mathrm{diag}^{-1/2}(\widehat{\mathbf{\Sigma}}_n) - \mathbf{I}_p\big],
\end{aligned}$$

where $\mathrm{diag}(\widehat{\mathbf{\Sigma}}_n) = \mathrm{diag}(\hat{\sigma}_{11}, ..., \hat{\sigma}_{pp})$. Then by (S1.5) and (S1.8), we have

$$\begin{aligned}
\mathrm{tr}\big(\widehat{\mathbf{R}}_n\mathbf{D}_1\big) &= \mathrm{tr}\big(\widehat{\mathbf{\Sigma}}_n\mathbf{D}_1\big) + \mathrm{tr}\big[\big(\mathrm{diag}^{-1/2}(\widehat{\mathbf{\Sigma}}_n) - \mathbf{I}_p\big)\widehat{\mathbf{\Sigma}}_n\mathrm{diag}^{-1/2}(\widehat{\mathbf{\Sigma}}_n)\mathbf{D}_1\big] \\
&\quad + \mathrm{tr}\big[\widehat{\mathbf{\Sigma}}_n\big(\mathrm{diag}^{-1/2}(\widehat{\mathbf{\Sigma}}_n) - \mathbf{I}_p\big)\mathbf{D}_1\big] \\
&= \mathrm{tr}\big(\widehat{\mathbf{\Sigma}}_n\mathbf{D}_1\big) - \mathrm{tr}\big[\big(\mathrm{diag}(\widehat{\mathbf{\Sigma}}_n) - \mathbf{I}_p\big)\widehat{\mathbf{\Sigma}}_n\mathbf{D}_1\big] + \frac{3}{4}\mathrm{tr}\big[\big(\mathrm{diag}(\widehat{\mathbf{\Sigma}}_n) - \mathbf{I}_p\big)^2\widehat{\mathbf{\Sigma}}_n\mathbf{D}_1\big] \\
&\quad + \frac{1}{4}\mathrm{tr}\big[\big(\mathrm{diag}(\widehat{\mathbf{\Sigma}}_n) - \mathbf{I}_p\big)\widehat{\mathbf{\Sigma}}_n\big(\mathrm{diag}(\widehat{\mathbf{\Sigma}}_n) - \mathbf{I}_p\big)\mathbf{D}_1\big] + o_p(1)
\end{aligned} \tag{S1.9}$$

and

$$\begin{aligned}
&\mathrm{tr}\big(\widehat{\mathbf{R}}_n\mathbf{D}_2\widehat{\mathbf{R}}_n\mathbf{D}_2\big) \\
= \;& \mathrm{tr}\Big\{\big[\widehat{\mathbf{\Sigma}}_n\mathbf{D}_2 + \big(\mathrm{diag}^{-1/2}(\widehat{\mathbf{\Sigma}}_n) - \mathbf{I}_p\big)\widehat{\mathbf{\Sigma}}_n\mathrm{diag}^{-1/2}(\widehat{\mathbf{\Sigma}}_n)\mathbf{D}_2 + \widehat{\mathbf{\Sigma}}_n\big(\mathrm{diag}^{-1/2}(\widehat{\mathbf{\Sigma}}_n) - \mathbf{I}_p\big)\mathbf{D}_2\big]^2\Big\} \\
= \;& \mathrm{tr}\big(\widehat{\mathbf{\Sigma}}_n\mathbf{D}_2\widehat{\mathbf{\Sigma}}_n\mathbf{D}_2\big) + 2\mathrm{tr}\big[\big(\mathrm{diag}^{-1/2}(\widehat{\mathbf{\Sigma}}_n) - \mathbf{I}_p\big)\widehat{\mathbf{\Sigma}}_n\mathrm{diag}^{-1/2}(\widehat{\mathbf{\Sigma}}_n)\mathbf{D}_2\widehat{\mathbf{\Sigma}}_n\mathbf{D}_2\big] \\
& + 2\mathrm{tr}\big[\big(\mathrm{diag}^{-1/2}(\widehat{\mathbf{\Sigma}}_n) - \mathbf{I}_p\big)\mathbf{D}_2\widehat{\mathbf{\Sigma}}_n\mathbf{D}_2\widehat{\mathbf{\Sigma}}_n\big] \\
& + \mathrm{tr}\big[\big(\mathrm{diag}^{-1/2}(\widehat{\mathbf{\Sigma}}_n) - \mathbf{I}_p\big)\widehat{\mathbf{\Sigma}}_n\mathrm{diag}^{-1/2}(\widehat{\mathbf{\Sigma}}_n)\mathbf{D}_2\big(\mathrm{diag}^{-1/2}(\widehat{\mathbf{\Sigma}}_n) - \mathbf{I}_p\big)\widehat{\mathbf{\Sigma}}_n\mathrm{diag}^{-1/2}(\widehat{\mathbf{\Sigma}}_n)\mathbf{D}_2\big] \\
& + 2\mathrm{tr}\big[\big(\mathrm{diag}^{-1/2}(\widehat{\mathbf{\Sigma}}_n) - \mathbf{I}_p\big)\mathbf{D}_2\big(\mathrm{diag}^{-1/2}(\widehat{\mathbf{\Sigma}}_n) - \mathbf{I}_p\big)\widehat{\mathbf{\Sigma}}_n\mathrm{diag}^{-1/2}(\widehat{\mathbf{\Sigma}}_n)\mathbf{D}_2\widehat{\mathbf{\Sigma}}_n\big] \\
& + \mathrm{tr}\big[\big(\mathrm{diag}^{-1/2}(\widehat{\mathbf{\Sigma}}_n) - \mathbf{I}_p\big)\mathbf{D}_2\widehat{\mathbf{\Sigma}}_n\big(\mathrm{diag}^{-1/2}(\widehat{\mathbf{\Sigma}}_n) - \mathbf{I}_p\big)\mathbf{D}_2\widehat{\mathbf{\Sigma}}_n\big] \\
= \;& \mathrm{tr}\big(\widehat{\mathbf{\Sigma}}_n\mathbf{D}_2\widehat{\mathbf{\Sigma}}_n\mathbf{D}_2\big) - 2\mathrm{tr}\big[\big(\mathrm{diag}(\widehat{\mathbf{\Sigma}}_n) - \mathbf{I}_p\big)\widehat{\mathbf{\Sigma}}_n\mathbf{D}_2\widehat{\mathbf{\Sigma}}_n\mathbf{D}_2\big] \\
& + \frac{3}{2}\mathrm{tr}\big[\big(\mathrm{diag}(\widehat{\mathbf{\Sigma}}_n) - \mathbf{I}_p\big)^2\widehat{\mathbf{\Sigma}}_n\mathbf{D}_2\widehat{\mathbf{\Sigma}}_n\mathbf{D}_2\big]
\end{aligned} \tag{S1.10}$$

$$+\frac{1}{2}\mathrm{tr}\big[\big(\mathrm{diag}(\widehat{\mathbf{\Sigma}}_n)-\mathbf{I}_p\big)\widehat{\mathbf{\Sigma}}_n\big(\mathrm{diag}(\widehat{\mathbf{\Sigma}}_n)-\mathbf{I}_p\big)\mathbf{D}_2\widehat{\mathbf{\Sigma}}_n\mathbf{D}_2\big]$$
$$+\frac{1}{2}\mathrm{tr}\big[\big(\mathrm{diag}(\widehat{\mathbf{\Sigma}}_n)-\mathbf{I}_p\big)\mathbf{D}_2\big(\mathrm{diag}(\widehat{\mathbf{\Sigma}}_n)-\mathbf{I}_p\big)\widehat{\mathbf{\Sigma}}_n\mathbf{D}_2\widehat{\mathbf{\Sigma}}_n\big]$$
$$+\frac{1}{2}\mathrm{tr}\big[\big(\mathrm{diag}(\widehat{\mathbf{\Sigma}}_n)-\mathbf{I}_p\big)\widehat{\mathbf{\Sigma}}_n\mathbf{D}_2\big(\mathrm{diag}(\widehat{\mathbf{\Sigma}}_n)-\mathbf{I}_p\big)\widehat{\mathbf{\Sigma}}_n\mathbf{D}_2\big]+o_p(1).$$

S1.3 Analysis of the constant order terms

In this section, we prove that the last two terms in (S1.9) and the last four terms in (S1.10) converge in probability to their expectations, which are of the constant order. Let $\mathbf{r}_i = n^{-1/2}\mathbf{w}_i$ for $i=1,...,n$, then we have

$$\begin{aligned}
&\mathrm{E}\mathrm{tr}\big[\big(\mathrm{diag}(\widehat{\mathbf{\Sigma}}_n)-\mathbf{I}_p\big)^2\widehat{\mathbf{\Sigma}}_n\mathbf{D}_1\big]\\
=\ &\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{h=1}^{n}\sum_{k=1}^{p}\mathrm{E}[\mathbf{e}_k^T\mathbf{\Gamma}\mathbf{r}_i\mathbf{r}_i^T\mathbf{\Gamma}^T\mathbf{D}_1\mathbf{e}_k(\mathbf{e}_k^T\mathbf{\Gamma}\mathbf{r}_j\mathbf{r}_j^T\mathbf{\Gamma}^T\mathbf{e}_k-n^{-1})\\
&\quad\times(\mathbf{e}_k^T\mathbf{\Gamma}\mathbf{r}_h\mathbf{r}_h^T\mathbf{\Gamma}^T\mathbf{e}_k-n^{-1})]+o(1)\\
=\ &\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{h=1}^{n}\sum_{k=1}^{p}\mathrm{E}[\mathbf{r}_i^T\mathbf{\Gamma}^T\mathbf{D}_1\mathbf{e}_k\mathbf{e}_k^T\mathbf{\Gamma}\mathbf{r}_i(\mathbf{r}_j^T\mathbf{\Gamma}^T\mathbf{e}_k\mathbf{e}_k^T\mathbf{\Gamma}\mathbf{r}_j-n^{-1})\\
&\quad\times(\mathbf{r}_h^T\mathbf{\Gamma}^T\mathbf{e}_k\mathbf{e}_k^T\mathbf{\Gamma}\mathbf{r}_h-n^{-1})]+o(1)\\
=\ &\frac{(n-1)}{n^2}\Big[2\mathrm{tr}(\mathbf{R}\mathbf{D}_1)+\beta_w\sum_{k=1}^{p}\mathbf{e}_k^T\mathbf{R}\mathbf{D}_1\mathbf{e}_k\sum_{\ell=1}^{p}(\mathbf{e}_\ell^T\mathbf{\Gamma}^T\mathbf{e}_k\mathbf{e}_k^T\mathbf{\Gamma}\mathbf{e}_\ell)^2\Big]\\
&+\sum_{i=1}^{n}\sum_{k=1}^{p}\mathrm{E}[\mathbf{r}_i^T\mathbf{\Gamma}^T\mathbf{D}_1\mathbf{e}_k\mathbf{e}_k^T\mathbf{\Gamma}\mathbf{r}_i(\mathbf{r}_i^T\mathbf{\Gamma}^T\mathbf{e}_k\mathbf{e}_k^T\mathbf{\Gamma}\mathbf{r}_i-n^{-1})^2]+o(1)\\
=\ &\frac{(n-1)}{n^2}\Big[2\mathrm{tr}(\mathbf{R}\mathbf{D}_1)+\beta_w\sum_{k=1}^{p}\mathbf{e}_k^T\mathbf{R}\mathbf{D}_1\mathbf{e}_k\sum_{\ell=1}^{p}(\mathbf{e}_k^T\mathbf{\Gamma}\mathbf{e}_\ell)^4\Big]+o(1),
\end{aligned}$$

where the last equality is from (9.9.6) of Bai and Silverstein (2010). We can also prove that

$$\mathrm{Var}\{\mathrm{tr}[(\mathrm{diag}(\widehat{\mathbf{\Sigma}}_n)-\mathbf{I}_p)^2\widehat{\mathbf{\Sigma}}_n\mathbf{D}_1]\}=o(1).$$

From Chebyshev's inequality,

$$\mathrm{tr}\big[\big(\mathrm{diag}(\widehat{\mathbf{\Sigma}}_n)-\mathbf{I}_p\big)^2\widehat{\mathbf{\Sigma}}_n\mathbf{D}_1\big] \tag{S1.11}$$

$$= \frac{(n-1)}{n^2}\Big[2\mathrm{tr}(\mathbf{R}\mathbf{D}_1) + \beta_w \sum_{\ell=1}^{p} \mathbf{e}_\ell^T \mathbf{R}\mathbf{D}_1 \mathbf{e}_\ell \sum_{k=1}^{p} (\mathbf{e}_\ell^T \boldsymbol{\Gamma} \mathbf{e}_k)^4\Big] + o_p(1).$$

Similarly,

$$\begin{aligned}
& \mathrm{tr}\big[\big(\mathrm{diag}(\widehat{\boldsymbol{\Sigma}}_n) - \mathbf{I}_p\big)\widehat{\boldsymbol{\Sigma}}_n\big(\mathrm{diag}(\widehat{\boldsymbol{\Sigma}}_n) - \mathbf{I}_p\big)\mathbf{D}_1\big] && \text{(S1.12)}\\
= \; & \frac{(n-1)}{n^2}\mathrm{tr}(\mathbf{C}\mathbf{D}_1) + o_p(1), \\
& \mathrm{tr}\big[\big(\mathrm{diag}(\widehat{\boldsymbol{\Sigma}}_n) - \mathbf{I}_p\big)^2\widehat{\boldsymbol{\Sigma}}_n\mathbf{D}_2\widehat{\boldsymbol{\Sigma}}_n\mathbf{D}_2\big] && \text{(S1.13)}\\
= \; & \frac{(n-1)(n-2)}{n^3}\Big[2\mathrm{tr}(\mathbf{R}\mathbf{D}_2\mathbf{R}\mathbf{D}_2) + \beta_w \sum_{k=1}^{p} \mathbf{e}_k^T \mathbf{R}\mathbf{D}_2\mathbf{R}\mathbf{D}_2 \mathbf{e}_k \sum_{\ell=1}^{p} (\mathbf{e}_k^T \boldsymbol{\Gamma} \mathbf{e}_\ell)^4\Big] \\
& + \frac{(n-1)\mathrm{tr}(\mathbf{R}\mathbf{D}_2)}{n^3}\Big[2\mathrm{tr}(\mathbf{R}\mathbf{D}_2) + \beta_w \sum_{k=1}^{p} \mathbf{e}_k^T \mathbf{R}\mathbf{D}_2 \mathbf{e}_k \sum_{\ell=1}^{p} (\mathbf{e}_k^T \boldsymbol{\Gamma} \mathbf{e}_\ell)^4\Big] + o_p(1), \\
& \mathrm{tr}\big[\big(\mathrm{diag}(\widehat{\boldsymbol{\Sigma}}_n) - \mathbf{I}_p\big)\widehat{\boldsymbol{\Sigma}}_n\big(\mathrm{diag}(\widehat{\boldsymbol{\Sigma}}_n) - \mathbf{I}_p\big)\mathbf{D}_2\widehat{\boldsymbol{\Sigma}}_n\mathbf{D}_2\big] && \text{(S1.14)}\\
= \; & \frac{(n-1)(n-2)}{n^3}\sum_{i=1}^{p}\sum_{j=1}^{p} \mathbf{e}_i^T\mathbf{R}\mathbf{e}_j\mathbf{e}_i^T\mathbf{D}_2\mathbf{R}\mathbf{D}_2\mathbf{e}_j\Big[2(\mathbf{e}_i^T\mathbf{R}\mathbf{e}_j)^2 + \beta_w\sum_{k=1}^{p}(\mathbf{e}_i^T\boldsymbol{\Gamma}\mathbf{e}_k)^2(\mathbf{e}_j^T\boldsymbol{\Gamma}\mathbf{e}_k)^2\Big] + o_p(1), \\
& \mathrm{tr}\big[\big(\mathrm{diag}(\widehat{\boldsymbol{\Sigma}}_n) - \mathbf{I}_p\big)\mathbf{D}_2\big(\mathrm{diag}(\widehat{\boldsymbol{\Sigma}}_n) - \mathbf{I}_p\big)\widehat{\boldsymbol{\Sigma}}_n\mathbf{D}_2\widehat{\boldsymbol{\Sigma}}_n\big] && \text{(S1.15)}\\
= \; & \frac{(n-1)(n-2)}{n^3}\sum_{i=1}^{p}\sum_{j=1}^{p} \mathbf{e}_i^T\mathbf{D}_2\mathbf{e}_j\mathbf{e}_i^T\mathbf{R}\mathbf{D}_2\mathbf{R}\mathbf{e}_j\Big[2(\mathbf{e}_i^T\mathbf{R}\mathbf{e}_j)^2 + \beta_w\sum_{k=1}^{p}(\mathbf{e}_i^T\boldsymbol{\Gamma}\mathbf{e}_k)^2(\mathbf{e}_j^T\boldsymbol{\Gamma}\mathbf{e}_k)^2\Big] \\
& + \frac{(n-1)\mathrm{tr}(\mathbf{R}\mathbf{D}_2)}{n^3}\sum_{i=1}^{p}\sum_{j=1}^{p} \mathbf{e}_i^T\mathbf{D}_2\mathbf{e}_j\mathbf{e}_i^T\mathbf{R}\mathbf{e}_j\Big[2(\mathbf{e}_i^T\mathbf{R}\mathbf{e}_j)^2 + \beta_w\sum_{k=1}^{p}(\mathbf{e}_i^T\boldsymbol{\Gamma}\mathbf{e}_k)^2(\mathbf{e}_j^T\boldsymbol{\Gamma}\mathbf{e}_k)^2\Big] + o_p(1), \\
& \mathrm{tr}\big[\big(\mathrm{diag}(\widehat{\boldsymbol{\Sigma}}_n) - \mathbf{I}_p\big)\widehat{\boldsymbol{\Sigma}}_n\mathbf{D}_2\big(\mathrm{diag}(\widehat{\boldsymbol{\Sigma}}_n) - \mathbf{I}_p\big)\widehat{\boldsymbol{\Sigma}}_n\mathbf{D}_2\big] && \text{(S1.16)}\\
= \; & \frac{(n-1)(n-2)}{n^3}\sum_{i=1}^{p}\sum_{j=1}^{p} \mathbf{e}_i^T\mathbf{R}\mathbf{D}_2\mathbf{e}_j\mathbf{e}_j^T\mathbf{R}\mathbf{D}_2\mathbf{e}_i\Big[2(\mathbf{e}_i^T\mathbf{R}\mathbf{e}_j)^2 + \beta_w\sum_{k=1}^{p}(\mathbf{e}_i^T\boldsymbol{\Gamma}\mathbf{e}_k)^2(\mathbf{e}_j^T\boldsymbol{\Gamma}\mathbf{e}_k)^2\Big] + o_p(1),
\end{aligned}$$

where $\mathbf{C} = (c_{ij})_{i,j=1}^p$ is a $p \times p$ dimensional matrix whose (i,j)th element is given by

$$c_{ij} = 2r_{ij}^3 + \beta_w r_{ij}\sum_{\ell=1}^{p}(\mathbf{e}_i^T\boldsymbol{\Gamma}\mathbf{e}_\ell)^2(\mathbf{e}_j^T\boldsymbol{\Gamma}\mathbf{e}_\ell)^2,$$

and r_{ij} is the (i,j)th element of the population correlation matrix $\mathbf{R}$.

Combining (S1.9) and (S1.11)-(S1.12), we have

$$\mathrm{tr}(\widehat{\mathbf{R}}_n\mathbf{D}_1) = \mathrm{tr}\big(\widehat{\boldsymbol{\Sigma}}_n\mathbf{D}_1\big) - \mathrm{tr}\big[\big(\mathrm{diag}(\widehat{\boldsymbol{\Sigma}}_n) - \mathbf{I}_p\big)\widehat{\boldsymbol{\Sigma}}_n\mathbf{D}_1\big] + \frac{1}{4n}\mathrm{tr}(\mathbf{C}\mathbf{D}_1) \qquad \text{(S1.17)}$$

$$+\frac{3}{4n}\Big[2\mathrm{tr}(\mathbf{R}\mathbf{D}_1)+\beta_w\sum_{k=1}^{p}\mathbf{e}_k^T\mathbf{R}\mathbf{D}_1\mathbf{e}_k\sum_{\ell=1}^{p}(\mathbf{e}_k^T\boldsymbol{\Gamma}\mathbf{e}_\ell)^4\Big]+o_p(1).$$

Also combining (S1.10) and (S1.13)-(S1.16), we obtain

$$\begin{aligned}
&\mathrm{tr}\big(\widehat{\mathbf{R}}_n\mathbf{D}_2\widehat{\mathbf{R}}_n\mathbf{D}_2\big) \qquad\qquad (S1.18)\\
=\;&\mathrm{tr}\big(\widehat{\boldsymbol{\Sigma}}_n\mathbf{D}_2\widehat{\boldsymbol{\Sigma}}_n\mathbf{D}_2\big)-2\mathrm{tr}\big[\big(\mathrm{diag}(\widehat{\boldsymbol{\Sigma}}_n)-\mathbf{I}_p\big)\widehat{\boldsymbol{\Sigma}}_n\mathbf{D}_2\widehat{\boldsymbol{\Sigma}}_n\mathbf{D}_2\big]\\
&+\frac{3}{2n}\Big[2\mathrm{tr}(\mathbf{R}\mathbf{D}_2\mathbf{R}\mathbf{D}_2)+\beta_w\sum_{k=1}^{p}\mathbf{e}_k^T\mathbf{R}\mathbf{D}_2\mathbf{R}\mathbf{D}_2\mathbf{e}_k\sum_{\ell=1}^{p}(\mathbf{e}_k^T\boldsymbol{\Gamma}\mathbf{e}_\ell)^4\Big]\\
&+\frac{3\mathrm{tr}(\mathbf{R}\mathbf{D}_2)}{2n^2}\Big[2\mathrm{tr}(\mathbf{R}\mathbf{D}_2)+\beta_w\sum_{k=1}^{p}\mathbf{e}_k^T\mathbf{R}\mathbf{D}_2\mathbf{e}_k\sum_{\ell=1}^{p}(\mathbf{e}_k^T\boldsymbol{\Gamma}\mathbf{e}_\ell)^4\Big]\\
&+\frac{1}{2n}\sum_{i=1}^{p}\sum_{j=1}^{p}\mathbf{e}_i^T\mathbf{R}\mathbf{e}_j\mathbf{e}_i^T\mathbf{D}_2\mathbf{R}\mathbf{D}_2\mathbf{e}_j\Big[2(\mathbf{e}_i^T\mathbf{R}\mathbf{e}_j)^2+\beta_w\sum_{k=1}^{p}(\mathbf{e}_i^T\boldsymbol{\Gamma}\mathbf{e}_k)^2(\mathbf{e}_j^T\boldsymbol{\Gamma}\mathbf{e}_k)^2\Big]\\
&+\frac{1}{2n}\sum_{i=1}^{p}\sum_{j=1}^{p}\mathbf{e}_i^T\mathbf{D}_2\mathbf{e}_j\mathbf{e}_i^T\mathbf{R}\mathbf{D}_2\mathbf{R}\mathbf{e}_j\Big[2(\mathbf{e}_i^T\mathbf{R}\mathbf{e}_j)^2+\beta_w\sum_{k=1}^{p}(\mathbf{e}_i^T\boldsymbol{\Gamma}\mathbf{e}_k)^2(\mathbf{e}_j^T\boldsymbol{\Gamma}\mathbf{e}_k)^2\Big]\\
&+\frac{\mathrm{tr}(\mathbf{R}\mathbf{D}_2)}{2n^2}\sum_{i=1}^{p}\sum_{j=1}^{p}\mathbf{e}_i^T\mathbf{D}_2\mathbf{e}_j\mathbf{e}_i^T\mathbf{R}\mathbf{e}_j\Big[2(\mathbf{e}_i^T\mathbf{R}\mathbf{e}_j)^2+\beta_w\sum_{k=1}^{p}(\mathbf{e}_i^T\boldsymbol{\Gamma}\mathbf{e}_k)^2(\mathbf{e}_j^T\boldsymbol{\Gamma}\mathbf{e}_k)^2\Big]\\
&+\frac{1}{2n}\sum_{i=1}^{p}\sum_{j=1}^{p}\mathbf{e}_i^T\mathbf{R}\mathbf{D}_2\mathbf{e}_j\mathbf{e}_j^T\mathbf{R}\mathbf{D}_2\mathbf{e}_i\Big[2(\mathbf{e}_i^T\mathbf{R}\mathbf{e}_j)^2+\beta_w\sum_{k=1}^{p}(\mathbf{e}_i^T\boldsymbol{\Gamma}\mathbf{e}_k)^2(\mathbf{e}_j^T\boldsymbol{\Gamma}\mathbf{e}_k)^2\Big]+o_p(1).
\end{aligned}$$

S1.4 Analysis of the main terms

In this section, we derive the CLT for

$$\begin{pmatrix}\mathrm{tr}\big(\widehat{\boldsymbol{\Sigma}}_n\mathbf{D}_1\big)-\mathrm{tr}\big[\big(\mathrm{diag}(\widehat{\boldsymbol{\Sigma}}_n)-\mathbf{I}_p\big)\widehat{\boldsymbol{\Sigma}}_n\mathbf{D}_1\big]\\ \mathrm{tr}\big(\widehat{\boldsymbol{\Sigma}}_n\mathbf{D}_2\widehat{\boldsymbol{\Sigma}}_n\mathbf{D}_2\big)-2\mathrm{tr}\big[\big(\mathrm{diag}(\widehat{\boldsymbol{\Sigma}}_n)-\mathbf{I}_p\big)\widehat{\boldsymbol{\Sigma}}_n\mathbf{D}_2\widehat{\boldsymbol{\Sigma}}_n\mathbf{D}_2\big]\end{pmatrix}.$$

According to the expression of $\widehat{\boldsymbol{\Sigma}}_n$,

$$\widehat{\boldsymbol{\Sigma}}_n=\boldsymbol{\Xi}_n+\boldsymbol{\Delta}_n=\frac{n-1}{n}\sum_{j=1}^{n}\boldsymbol{\Gamma}\mathbf{r}_j\mathbf{r}_j^T\boldsymbol{\Gamma}^T-\frac{1}{n}\sum_{j\neq k}\boldsymbol{\Gamma}\mathbf{r}_j\mathbf{r}_k^T\boldsymbol{\Gamma}^T,$$

we have

$$\mathrm{tr}\big(\widehat{\boldsymbol{\Sigma}}_n\mathbf{D}_1\big)-\mathrm{Etr}\big(\widehat{\boldsymbol{\Sigma}}_n\mathbf{D}_1\big)=\mathrm{tr}(\boldsymbol{\Xi}_n\mathbf{D}_1)-\mathrm{Etr}(\boldsymbol{\Xi}_n\mathbf{D}_1)+o_p(1),$$

$$
\begin{aligned}
&\mathrm{tr}\big[\big(\mathrm{diag}(\widehat{\mathbf{\Sigma}}_n)-\mathbf{I}_p\big)\widehat{\mathbf{\Sigma}}_n\mathbf{D}_1\big]-\mathrm{Etr}\big[\big(\mathrm{diag}(\widehat{\mathbf{\Sigma}}_n)-\mathbf{I}_p\big)\widehat{\mathbf{\Sigma}}_n\mathbf{D}_1\big]\\
=&\mathrm{tr}[(\mathrm{diag}(\mathbf{\Xi}_n)-\mathbf{I}_p)\mathbf{\Xi}_n\mathbf{D}_1]-\mathrm{Etr}[(\mathrm{diag}(\mathbf{\Xi}_n)-\mathbf{I}_p)\mathbf{\Xi}_n\mathbf{D}_1]+o_p(1),\\
&\mathrm{tr}\big(\widehat{\mathbf{\Sigma}}_n\mathbf{D}_2\widehat{\mathbf{\Sigma}}_n\mathbf{D}_2\big)-\mathrm{Etr}\big(\widehat{\mathbf{\Sigma}}_n\mathbf{D}_2\widehat{\mathbf{\Sigma}}_n\mathbf{D}_2\big)=\mathrm{tr}(\mathbf{\Xi}_n\mathbf{D}_2\mathbf{\Xi}_n\mathbf{D}_2)-\mathrm{Etr}(\mathbf{\Xi}_n\mathbf{D}_2\mathbf{\Xi}_n\mathbf{D}_2)+o_p(1),\\
&\mathrm{tr}\big[\big(\mathrm{diag}(\widehat{\mathbf{\Sigma}}_n)-\mathbf{I}_p\big)\widehat{\mathbf{\Sigma}}_n\mathbf{D}_2\widehat{\mathbf{\Sigma}}_n\mathbf{D}_2\big]-\mathrm{Etr}\big[\big(\mathrm{diag}(\widehat{\mathbf{\Sigma}}_n)-\mathbf{I}_p\big)\widehat{\mathbf{\Sigma}}_n\mathbf{D}_2\widehat{\mathbf{\Sigma}}_n\mathbf{D}_2\big]\\
=&\mathrm{tr}[(\mathrm{diag}(\mathbf{\Xi}_n)-\mathbf{I}_p)\mathbf{\Xi}_n\mathbf{D}_2\mathbf{\Xi}_n\mathbf{D}_2]-\mathrm{Etr}[(\mathrm{diag}(\mathbf{\Xi}_n)-\mathbf{I}_p)\mathbf{\Xi}_n\mathbf{D}_2\mathbf{\Xi}_n\mathbf{D}_2]+o_p(1).
\end{aligned}
$$

Let $(\mathrm{E}_j-\mathrm{E}_{j-1})x=\mathrm{E}_j(x)-\mathrm{E}_{j-1}(x)$, where x is a random variable and $\mathrm{E}_j(\cdot)$ denotes the conditional expectation with respect to the σ-field generated by $\mathbf{r}_1,\ldots,\mathbf{r}_j$, then

$$
\begin{aligned}
&\mathrm{tr}(\mathbf{\Xi}_n\mathbf{D}_1)-\mathrm{Etr}(\mathbf{\Xi}_n\mathbf{D}_1)=\sum_{j=1}^n(\mathrm{E}_j-\mathrm{E}_{j-1})\mathrm{tr}(\mathbf{\Xi}_n\mathbf{D}_1),\\
&\mathrm{tr}[(\mathrm{diag}(\mathbf{\Xi}_n)-\mathbf{I}_p)\mathbf{\Xi}_n\mathbf{D}_1]-\mathrm{Etr}[(\mathrm{diag}(\mathbf{\Xi}_n)-\mathbf{I}_p)\mathbf{\Xi}_n\mathbf{D}_1]\\
=&\sum_{j=1}^n(\mathrm{E}_j-\mathrm{E}_{j-1})\mathrm{tr}[(\mathrm{diag}(\mathbf{\Xi}_n)-\mathbf{I}_p)\mathbf{\Xi}_n\mathbf{D}_1],\\
&\mathrm{tr}(\mathbf{\Xi}_n\mathbf{D}_2\mathbf{\Xi}_n\mathbf{D}_2)-\mathrm{Etr}(\mathbf{\Xi}_n\mathbf{D}_2\mathbf{\Xi}_n\mathbf{D}_2)=\sum_{j=1}^n(\mathrm{E}_j-\mathrm{E}_{j-1})\mathrm{tr}(\mathbf{\Xi}_n\mathbf{D}_2\mathbf{\Xi}_n\mathbf{D}_2),\\
&\mathrm{tr}[(\mathrm{diag}(\mathbf{\Xi}_n)-\mathbf{I}_p)\mathbf{\Xi}_n\mathbf{D}_2\mathbf{\Xi}_n\mathbf{D}_2]-\mathrm{Etr}[(\mathrm{diag}(\mathbf{\Xi}_n)-\mathbf{I}_p)\mathbf{\Xi}_n\mathbf{D}_2\mathbf{\Xi}_n\mathbf{D}_2]\\
=&\sum_{j=1}^n(\mathrm{E}_j-\mathrm{E}_{j-1})\mathrm{tr}[(\mathrm{diag}(\mathbf{\Xi}_n)-\mathbf{I}_p)\mathbf{\Xi}_n\mathbf{D}_2\mathbf{\Xi}_n\mathbf{D}_2].
\end{aligned}
$$

It is not difficult to prove that the martingale difference sequences

$$
\begin{aligned}
&\left\{(\mathrm{E}_j-\mathrm{E}_{j-1})\mathrm{tr}(\mathbf{\Xi}_n\mathbf{D}_1),j=1,...,n\right\},\\
&\left\{(\mathrm{E}_j-\mathrm{E}_{j-1})\mathrm{tr}[(\mathrm{diag}(\mathbf{\Xi}_n)-\mathbf{I}_p)\mathbf{\Xi}_n\mathbf{D}_1],j=1,...,n\right\},\\
&\left\{(\mathrm{E}_j-\mathrm{E}_{j-1})\mathrm{tr}(\mathbf{\Xi}_n\mathbf{D}_2\mathbf{\Xi}_n\mathbf{D}_2),j=1,...,n\right\},\\
&\left\{(\mathrm{E}_j-\mathrm{E}_{j-1})\mathrm{tr}[(\mathrm{diag}(\mathbf{\Xi}_n)-\mathbf{I}_p)\mathbf{\Xi}_n\mathbf{D}_2\mathbf{\Xi}_n\mathbf{D}_2],j=1,...,n\right\}
\end{aligned}
$$

satisfy the Lindeberg-type conditions, that is,

$$\sum_{j=1}^{n} \mathrm{E}[(\mathrm{E}_j - \mathrm{E}_{j-1})\mathrm{tr}(\mathbf{\Xi}_n \mathbf{D}_1)]^2 \delta_{\{|(\mathrm{E}_j-\mathrm{E}_{j-1})\mathrm{tr}(\mathbf{\Xi}_n\mathbf{D}_1)|>\varepsilon\}} \to 0,$$
$$\sum_{j=1}^{n} \mathrm{E}\{(\mathrm{E}_j - \mathrm{E}_{j-1})\mathrm{tr}[(\mathrm{diag}(\mathbf{\Xi}_n) - \mathbf{I}_p)\mathbf{\Xi}_n \mathbf{D}_1]\}^2 \delta_{\{|(\mathrm{E}_j-\mathrm{E}_{j-1})\mathrm{tr}[(\mathrm{diag}(\mathbf{\Xi}_n)-\mathbf{I}_p)\mathbf{\Xi}_n\mathbf{D}_1]|>\varepsilon\}} \to 0,$$
$$\sum_{j=1}^{n} \mathrm{E}[(\mathrm{E}_j - \mathrm{E}_{j-1})\mathrm{tr}(\mathbf{\Xi}_n \mathbf{D}_2 \mathbf{\Xi}_n \mathbf{D}_2)]^2 \delta_{\{|(\mathrm{E}_j-\mathrm{E}_{j-1})\mathrm{tr}(\mathbf{\Xi}_n\mathbf{D}_2\mathbf{\Xi}_n\mathbf{D}_2)|>\varepsilon\}} \to 0,$$
$$\sum_{j=1}^{n} \mathrm{E}\{(\mathrm{E}_j - \mathrm{E}_{j-1})\mathrm{tr}[(\mathrm{diag}(\mathbf{\Xi}_n) - \mathbf{I}_p)\mathbf{\Xi}_n \mathbf{D}_2 \mathbf{\Xi}_n \mathbf{D}_2]\}^2 \delta_{\{|(\mathrm{E}_j-\mathrm{E}_{j-1})\mathrm{tr}[(\mathrm{diag}(\mathbf{\Xi}_n)-\mathbf{I}_p)\mathbf{\Xi}_n\mathbf{D}_2\mathbf{\Xi}_n\mathbf{D}_2]|>\varepsilon\}} \to 0,$$

where $\delta_{\{\cdot\}}$ represents the indicator function, then the proof is omitted. We will calculate

$$\begin{aligned}
\sigma_{11}(\mathbf{D}_1) &= \sigma_{110}(\mathbf{D}_1) + \sigma_{220}(\mathbf{D}_1) - 2\sigma_{120}(\mathbf{D}_1), \\
\sigma_{22}(\mathbf{D}_2) &= \sigma_{330}(\mathbf{D}_2) + 4\sigma_{440}(\mathbf{D}_2) - 4\sigma_{340}(\mathbf{D}_2), \\
\sigma_{12}(\mathbf{D}_1, \mathbf{D}_2) &= \sigma_{130}(\mathbf{D}_1, \mathbf{D}_2) - 2\sigma_{140}(\mathbf{D}_1, \mathbf{D}_2) - \sigma_{230}(\mathbf{D}_1, \mathbf{D}_2) + 2\sigma_{240}(\mathbf{D}_1, \mathbf{D}_2),
\end{aligned}$$

where

$$\begin{aligned}
\sigma_{110}(\mathbf{D}_1) &= \sum_{j=1}^{n} \mathrm{E}_{j-1}\big[(\mathrm{E}_j - \mathrm{E}_{j-1})\mathrm{tr}(\mathbf{\Xi}_n \mathbf{D}_1)\big]^2, \\
\sigma_{220}(\mathbf{D}_1) &= \sum_{j=1}^{n} \mathrm{E}_{j-1}\big\{(\mathrm{E}_j - \mathrm{E}_{j-1})\mathrm{tr}[(\mathrm{diag}(\mathbf{\Xi}_n) - \mathbf{I}_p)\mathbf{\Xi}_n \mathbf{D}_1]\big\}^2, \\
\sigma_{120}(\mathbf{D}_1) &= \sum_{j=1}^{n} \mathrm{E}_{j-1}\big\{[(\mathrm{E}_j - \mathrm{E}_{j-1})\mathrm{tr}(\mathbf{\Xi}_n \mathbf{D}_1)][(\mathrm{E}_j - \mathrm{E}_{j-1})\mathrm{tr}[(\mathrm{diag}(\mathbf{\Xi}_n) - \mathbf{I}_p)\mathbf{\Xi}_n \mathbf{D}_1]]\big\}, \\
\sigma_{330}(\mathbf{D}_2) &= \sum_{j=1}^{n} \mathrm{E}_{j-1}\big[(\mathrm{E}_j - \mathrm{E}_{j-1})\mathrm{tr}(\mathbf{\Xi}_n \mathbf{D}_2 \mathbf{\Xi}_n \mathbf{D}_2)\big]^2 \\
\sigma_{440}(\mathbf{D}_2) &= \sum_{j=1}^{n} \mathrm{E}_{j-1}\big\{(\mathrm{E}_j - \mathrm{E}_{j-1})\mathrm{tr}[(\mathrm{diag}(\mathbf{\Xi}_n) - \mathbf{I}_p)\mathbf{\Xi}_n \mathbf{D}_2 \mathbf{\Xi}_n \mathbf{D}_2]\big\}^2 \\
\sigma_{340}(\mathbf{D}_2) &= \sum_{j=1}^{n} \mathrm{E}_{j-1}\big\{[(\mathrm{E}_j - \mathrm{E}_{j-1})\mathrm{tr}(\mathbf{\Xi}_n \mathbf{D}_2 \mathbf{\Xi}_n \mathbf{D}_2)] \\
&\qquad \times [(\mathrm{E}_j - \mathrm{E}_{j-1})\mathrm{tr}[(\mathrm{diag}(\mathbf{\Xi}_n) - \mathbf{I}_p)\mathbf{\Xi}_n \mathbf{D}_2 \mathbf{\Xi}_n \mathbf{D}_2]]\big\},
\end{aligned}$$

and

$$\sigma_{130}(\mathbf{D}_1,\mathbf{D}_2)=\sum_{j=1}^{n}\mathrm{E}_{j-1}\big\{[(\mathrm{E}_j-\mathrm{E}_{j-1})\mathrm{tr}(\mathbf{\Xi}_n\mathbf{D}_1)][(\mathrm{E}_j-\mathrm{E}_{j-1})\mathrm{tr}(\mathbf{\Xi}_n\mathbf{D}_2\mathbf{\Xi}_n\mathbf{D}_2)]\big\},$$

$$\sigma_{140}(\mathbf{D}_1,\mathbf{D}_2)=\sum_{j=1}^{n}\mathrm{E}_{j-1}\big\{[(\mathrm{E}_j-\mathrm{E}_{j-1})\mathrm{tr}(\mathbf{\Xi}_n\mathbf{D}_1)]$$
$$\times[(\mathrm{E}_j-\mathrm{E}_{j-1})\mathrm{tr}[(\mathrm{diag}(\mathbf{\Xi}_n)-\mathbf{I}_p)\mathbf{\Xi}_n\mathbf{D}_2\mathbf{\Xi}_n\mathbf{D}_2)]]\big\},$$

$$\sigma_{230}(\mathbf{D}_1,\mathbf{D}_2)=\sum_{j=1}^{n}\mathrm{E}_{j-1}\big\{[(\mathrm{E}_j-\mathrm{E}_{j-1})\mathrm{tr}[(\mathrm{diag}(\mathbf{\Xi}_n)-\mathbf{I}_p)\mathbf{\Xi}_n\mathbf{D}_1]]$$
$$\times[(\mathrm{E}_j-\mathrm{E}_{j-1})\mathrm{tr}(\mathbf{\Xi}_n\mathbf{D}_2\mathbf{\Xi}_n\mathbf{D}_2)]\big\},$$

$$\sigma_{240}(\mathbf{D}_1,\mathbf{D}_2)=\sum_{j=1}^{n}\mathrm{E}_{j-1}\big\{[(\mathrm{E}_j-\mathrm{E}_{j-1})\mathrm{tr}[(\mathrm{diag}(\mathbf{\Xi}_n)-\mathbf{I}_p)\mathbf{\Xi}_n\mathbf{D}_1]]$$
$$\times[(\mathrm{E}_j-\mathrm{E}_{j-1})\mathrm{tr}[(\mathrm{diag}(\mathbf{\Xi}_n)-\mathbf{I}_p)\mathbf{\Xi}_n\mathbf{D}_2\mathbf{\Xi}_n\mathbf{D}_2)]]\big\}.$$

Based on (1.15) of Bai and Silverstein (2004), after tedious calculations, we obtain

$$\begin{aligned}
\sigma_{110}(\mathbf{D}_1) &= \sum_{j=1}^{n}\mathrm{E}_{j-1}\big[(\mathrm{E}_j-\mathrm{E}_{j-1})\mathrm{tr}(\mathbf{\Xi}_n\mathbf{D}_1)\big]^2\\
&= \sum_{j=1}^{n}\mathrm{E}[\mathbf{r}_j^T\mathbf{\Gamma}^T\mathbf{D}_1\mathbf{\Gamma}\mathbf{r}_j-n^{-1}\mathrm{tr}(\mathbf{R}\mathbf{D}_1)]^2+o_p(1)\\
&= n^{-1}\Big[2\mathrm{tr}(\mathbf{R}\mathbf{D}_1)^2+\beta_w\sum_{\ell=1}^{p}(\mathbf{e}_\ell^T\mathbf{\Gamma}^T\mathbf{D}_1\mathbf{\Gamma}\mathbf{e}_\ell)^2\Big]+o_p(1),\\
\sigma_{220}(\mathbf{D}_1) &= \sum_{j=1}^{n}\mathrm{E}_{j-1}\big\{(\mathrm{E}_j-\mathrm{E}_{j-1})\mathrm{tr}[(\mathrm{diag}(\mathbf{\Xi}_n)-\mathbf{I}_p)\mathbf{\Xi}_n\mathbf{D}_1]\big\}^2\\
&= \sum_{j=1}^{n}\mathrm{E}_{j-1}\Big[\sum_{\ell=1}^{p}\sum_{i=1}^{j-1}\mathbf{r}_i^T\mathbf{\Gamma}^T\mathbf{D}_1\mathbf{e}_\ell\mathbf{e}_\ell^T\mathbf{\Gamma}\mathbf{r}_i(\mathbf{r}_j^T\mathbf{\Gamma}^T\mathbf{e}_\ell\mathbf{e}_\ell^T\mathbf{\Gamma}\mathbf{r}_j-n^{-1})\\
&\quad+\frac{n-j}{n}\sum_{\ell=1}^{p}\mathbf{e}_\ell^T\mathbf{R}\mathbf{D}_1\mathbf{e}_\ell(\mathbf{r}_j^T\mathbf{\Gamma}^T\mathbf{e}_\ell\mathbf{e}_\ell^T\mathbf{\Gamma}\mathbf{r}_j-n^{-1})\Big]^2+o_p(1)\\
&= n^{-1}\sum_{\ell_1=1}^{p}\sum_{\ell_2=1}^{p}\mathbf{e}_{\ell_1}^T\mathbf{R}\mathbf{D}_1\mathbf{e}_{\ell_1}\mathbf{e}_{\ell_2}^T\mathbf{R}\mathbf{D}_1\mathbf{e}_{\ell_2}\Big[2(\mathbf{e}_{\ell_1}^T\mathbf{R}\mathbf{e}_{\ell_2})^2+\beta_w\sum_{k=1}^{p}(\mathbf{e}_{\ell_1}^T\mathbf{\Gamma}\mathbf{e}_k)^2(\mathbf{e}_{\ell_2}^T\mathbf{\Gamma}\mathbf{e}_k)^2\Big]+o_p(1),\\
\sigma_{120}(\mathbf{D}_1) &= \sum_{j=1}^{n}\mathrm{E}_{j-1}\big\{[(\mathrm{E}_j-\mathrm{E}_{j-1})\mathrm{tr}(\mathbf{\Xi}_n\mathbf{D}_1)][(\mathrm{E}_j-\mathrm{E}_{j-1})\mathrm{tr}[(\mathrm{diag}(\mathbf{\Xi}_n)-\mathbf{I}_p)\mathbf{\Xi}_n\mathbf{D}_1]]\big\}
\end{aligned}$$

$$
\begin{aligned}
&= \sum_{j=1}^{n} \mathrm{E}_{j-1}\Big\{\Big[\sum_{\ell=1}^{p}\sum_{i=1}^{j-1} \mathbf{r}_i^T\boldsymbol{\Gamma}^T\mathbf{D}_1\mathbf{e}_\ell\mathbf{e}_\ell^T\boldsymbol{\Gamma}\mathbf{r}_i(\mathbf{r}_j^T\boldsymbol{\Gamma}^T\mathbf{e}_\ell\mathbf{e}_\ell^T\boldsymbol{\Gamma}\mathbf{r}_j - n^{-1}) \\
&\quad + \frac{n-j}{n}\sum_{\ell=1}^{p}\mathbf{e}_\ell^T\mathbf{R}\mathbf{D}_1\mathbf{e}_\ell(\mathbf{r}_j^T\boldsymbol{\Gamma}^T\mathbf{e}_\ell\mathbf{e}_\ell^T\boldsymbol{\Gamma}\mathbf{r}_j - n^{-1})\Big]\Big[\mathbf{r}_j^T\boldsymbol{\Gamma}^T\mathbf{D}_1\boldsymbol{\Gamma}\mathbf{r}_j - n^{-1}\mathrm{tr}(\mathbf{R}\mathbf{D}_1)\Big]\Big\} + o_p(1) \\
&= n^{-1}\sum_{\ell=1}^{p}\mathbf{e}_\ell^T\mathbf{R}\mathbf{D}_1\mathbf{e}_\ell\Big[2\mathbf{e}_\ell^T\mathbf{R}\mathbf{D}_1\mathbf{R}\mathbf{e}_\ell + \beta_w\sum_{k=1}^{p}(\mathbf{e}_\ell^T\boldsymbol{\Gamma}\mathbf{e}_k)^2\mathbf{e}_k^T\boldsymbol{\Gamma}^T\mathbf{D}_1\boldsymbol{\Gamma}\mathbf{e}_k\Big] + o_p(1).
\end{aligned}
$$

Therefore, we have

$$
\begin{aligned}
\sigma_{11}(\mathbf{D}_1) &= n^{-1}\Big[2\mathrm{tr}(\mathbf{R}\mathbf{D}_1)^2 + \beta_w\sum_{\ell=1}^{p}(\mathbf{e}_\ell^T\boldsymbol{\Gamma}^T\mathbf{D}_1\boldsymbol{\Gamma}\mathbf{e}_\ell)^2\Big] \\
&\quad + n^{-1}\sum_{\ell_1=1}^{p}\sum_{\ell_2=1}^{p}\mathbf{e}_{\ell_1}^T\mathbf{R}\mathbf{D}_1\mathbf{e}_{\ell_1}\mathbf{e}_{\ell_2}^T\mathbf{R}\mathbf{D}_1\mathbf{e}_{\ell_2}\Big[2(\mathbf{e}_{\ell_1}^T\mathbf{R}\mathbf{e}_{\ell_2})^2 + \beta_w\sum_{k=1}^{p}(\mathbf{e}_{\ell_1}^T\boldsymbol{\Gamma}\mathbf{e}_k)^2(\mathbf{e}_{\ell_2}^T\boldsymbol{\Gamma}\mathbf{e}_k)^2\Big] \\
&\quad - 2n^{-1}\sum_{\ell=1}^{p}\mathbf{e}_\ell^T\mathbf{R}\mathbf{D}_1\mathbf{e}_\ell\Big[2\mathbf{e}_\ell^T\mathbf{R}\mathbf{D}_1\mathbf{R}\mathbf{e}_\ell + \beta_w\sum_{k=1}^{p}(\mathbf{e}_\ell^T\boldsymbol{\Gamma}\mathbf{e}_k)^2\mathbf{e}_k^T\boldsymbol{\Gamma}^T\mathbf{D}_1\boldsymbol{\Gamma}\mathbf{e}_k\Big] + o_p(1).
\end{aligned} \tag{S1.19}
$$

Since the martingale differences of $\mathrm{tr}(\boldsymbol{\Xi}_n\mathbf{D}_2\boldsymbol{\Xi}_n\mathbf{D}_2)$ and $\mathrm{tr}[(\mathrm{diag}(\boldsymbol{\Xi}_n) - \mathbf{I}_p)\boldsymbol{\Xi}_n\mathbf{D}_2\boldsymbol{\Xi}_n\mathbf{D}_2]$ are

$$
\begin{aligned}
&(\mathrm{E}_j - \mathrm{E}_{j-1})\mathrm{tr}(\boldsymbol{\Xi}_n\mathbf{D}_2\boldsymbol{\Xi}_n\mathbf{D}_2) \\
&= \frac{2(n-j)}{n}\big(\mathbf{r}_j^T\boldsymbol{\Gamma}^T\mathbf{D}_2\mathbf{R}\mathbf{D}_2\boldsymbol{\Gamma}\mathbf{r}_j - n^{-1}\mathrm{tr}(\mathbf{R}\mathbf{D}_2\mathbf{R}\mathbf{D}_2)\big) \\
&\quad + 2n^{-1}\mathrm{tr}(\mathbf{R}\mathbf{D}_2)\big(\mathbf{r}_j^T\boldsymbol{\Gamma}^T\mathbf{D}_2\boldsymbol{\Gamma}\mathbf{r}_j - n^{-1}\mathrm{tr}(\mathbf{R}\mathbf{D}_2)\big) \\
&\quad + 2\sum_{k=1}^{j-1}\big(\mathbf{r}_j^T\boldsymbol{\Gamma}^T\mathbf{D}_2\boldsymbol{\Gamma}\mathbf{r}_k\mathbf{r}_k^T\boldsymbol{\Gamma}^T\mathbf{D}_2\boldsymbol{\Gamma}\mathbf{r}_j - n^{-1}\mathbf{r}_k^T\boldsymbol{\Gamma}^T\mathbf{D}_2\mathbf{R}\mathbf{D}_2\boldsymbol{\Gamma}\mathbf{r}_k\big) + o_p(1)
\end{aligned} \tag{S1.20}
$$

and

$$
\begin{aligned}
&(\mathrm{E}_j - \mathrm{E}_{j-1})\mathrm{tr}[(\mathrm{diag}(\boldsymbol{\Xi}_n) - \mathbf{I}_p)\boldsymbol{\Xi}_n\mathbf{D}_2\boldsymbol{\Xi}_n\mathbf{D}_2] \\
&= \sum_{\ell=1}^{p}\sum_{i=1}^{j-1}\sum_{h=1}^{j-1}\mathbf{e}_\ell^T\boldsymbol{\Gamma}\mathbf{r}_i\mathbf{r}_i^T\boldsymbol{\Gamma}^T\mathbf{D}_2\boldsymbol{\Gamma}\mathbf{r}_h\mathbf{r}_h^T\boldsymbol{\Gamma}^T\mathbf{D}_2\mathbf{e}_\ell(\mathbf{r}_j^T\boldsymbol{\Gamma}^T\mathbf{e}_\ell\mathbf{e}_\ell^T\boldsymbol{\Gamma}\mathbf{r}_j - n^{-1}) \\
&\quad + \frac{(n-j)(n-j-1)}{n^2}\sum_{\ell=1}^{p}\mathbf{e}_\ell^T\mathbf{R}\mathbf{D}_2\mathbf{R}\mathbf{D}_2\mathbf{e}_\ell(\mathbf{r}_j^T\boldsymbol{\Gamma}^T\mathbf{e}_\ell\mathbf{e}_\ell^T\boldsymbol{\Gamma}\mathbf{r}_j - n^{-1}) \\
&\quad + \frac{(n-j)\mathrm{tr}(\mathbf{R}\mathbf{D}_2)}{n^2}\sum_{\ell=1}^{p}\mathbf{e}_\ell^T\mathbf{R}\mathbf{D}_2\mathbf{e}_\ell(\mathbf{r}_j^T\boldsymbol{\Gamma}^T\mathbf{e}_\ell\mathbf{e}_\ell^T\boldsymbol{\Gamma}\mathbf{r}_j - n^{-1}) \\
&\quad + \frac{(n-j)}{n}\sum_{\ell=1}^{p}\sum_{h=1}^{j-1}\mathbf{r}_h^T\boldsymbol{\Gamma}^T\mathbf{D}_2\mathbf{R}\mathbf{D}_2\mathbf{e}_\ell\mathbf{e}_\ell^T\boldsymbol{\Gamma}\mathbf{r}_h(\mathbf{r}_j^T\boldsymbol{\Gamma}^T\mathbf{e}_\ell\mathbf{e}_\ell^T\boldsymbol{\Gamma}\mathbf{r}_j - n^{-1})
\end{aligned} \tag{S1.21}
$$

$$+\frac{(n-j)}{n}\sum_{\ell=1}^{p}\sum_{h=1}^{j-1}\mathbf{r}_h^T\boldsymbol{\Gamma}^T\mathbf{D}_2\mathbf{e}_\ell\mathbf{e}_\ell^T\mathbf{R}\mathbf{D}_2\boldsymbol{\Gamma}\mathbf{r}_h(\mathbf{r}_j^T\boldsymbol{\Gamma}^T\mathbf{e}_\ell\mathbf{e}_\ell^T\boldsymbol{\Gamma}\mathbf{r}_j-n^{-1})+o_p(1),$$

respectively, then we have

$$\begin{aligned}
\sigma_{330}(\mathbf{D}_2) &= \sum_{j=1}^{n}\mathrm{E}_{j-1}\big[(\mathrm{E}_j-\mathrm{E}_{j-1})\mathrm{tr}(\boldsymbol{\Xi}_n\mathbf{D}_2\boldsymbol{\Xi}_n\mathbf{D}_2)\big]^2\\
&= 4n^{-1}\Big[2\mathrm{tr}(\mathbf{R}\mathbf{D}_2)^4+\beta_w\sum_{\ell=1}^{p}(\mathbf{e}_\ell^T\boldsymbol{\Gamma}^T\mathbf{D}_2\mathbf{R}\mathbf{D}_2\boldsymbol{\Gamma}\mathbf{e}_\ell)^2\Big]\\
&\quad +4[n^{-1}\mathrm{tr}(\mathbf{R}\mathbf{D}_2)]^2n^{-1}\Big[2\mathrm{tr}(\mathbf{R}\mathbf{D}_2)^2+\beta_w\sum_{\ell=1}^{p}(\mathbf{e}_\ell^T\boldsymbol{\Gamma}^T\mathbf{D}_2\boldsymbol{\Gamma}\mathbf{e}_\ell)^2\Big]\\
&\quad +4\big[n^{-1}\mathrm{tr}(\mathbf{R}\mathbf{D}_2)^2\big]^2+8n^{-1}\mathrm{tr}(\mathbf{R}\mathbf{D}_2)n^{-1}\Big[2\mathrm{tr}(\mathbf{R}\mathbf{D}_2)^3\\
&\quad +\beta_w\sum_{\ell=1}^{p}\mathbf{e}_\ell^T\boldsymbol{\Gamma}^T\mathbf{D}_2\mathbf{R}\mathbf{D}_2\boldsymbol{\Gamma}\mathbf{e}_\ell\mathbf{e}_\ell^T\boldsymbol{\Gamma}^T\mathbf{D}_2\boldsymbol{\Gamma}\mathbf{e}_\ell\Big]+o_p(1),\\
\sigma_{440}(\mathbf{D}_2) &= \sum_{j=1}^{n}\mathrm{E}_{j-1}\big\{(\mathrm{E}_j-\mathrm{E}_{j-1})\mathrm{tr}[(\mathrm{diag}(\boldsymbol{\Xi}_n)-\mathbf{I}_p)\boldsymbol{\Xi}_n\mathbf{D}_2\boldsymbol{\Xi}_n\mathbf{D}_2]\big\}^2\\
&= n^{-1}\sum_{\ell_1=1}^{p}\sum_{\ell_2=1}^{p}\mathbf{e}_{\ell_1}^T(\mathbf{R}\mathbf{D}_2)^2\mathbf{e}_{\ell_1}\mathbf{e}_{\ell_2}^T(\mathbf{R}\mathbf{D}_2)^2\mathbf{e}_{\ell_2}\\
&\qquad\times\Big[2(\mathbf{e}_{\ell_1}^T\mathbf{R}\mathbf{e}_{\ell_2})^2+\beta_w\sum_{k=1}^{p}(\mathbf{e}_{\ell_1}^T\boldsymbol{\Gamma}\mathbf{e}_k)^2(\mathbf{e}_{\ell_2}^T\boldsymbol{\Gamma}\mathbf{e}_k)^2\Big]\\
&\quad +2n^{-1}\mathrm{tr}(\mathbf{R}\mathbf{D}_2)n^{-1}\sum_{\ell_1=1}^{p}\sum_{\ell_2=1}^{p}\mathbf{e}_{\ell_1}^T\mathbf{R}\mathbf{D}_2\mathbf{e}_{\ell_1}\mathbf{e}_{\ell_2}^T(\mathbf{R}\mathbf{D}_2)^2\mathbf{e}_{\ell_2}\\
&\qquad\times\Big[2(\mathbf{e}_{\ell_1}^T\mathbf{R}\mathbf{e}_{\ell_2})^2+\beta_w\sum_{k=1}^{p}(\mathbf{e}_{\ell_1}^T\boldsymbol{\Gamma}\mathbf{e}_k)^2(\mathbf{e}_{\ell_2}^T\boldsymbol{\Gamma}\mathbf{e}_k)^2\Big]\\
&\quad +[n^{-1}\mathrm{tr}(\mathbf{R}\mathbf{D}_2)]^2n^{-1}\sum_{\ell_1=1}^{p}\sum_{\ell_2=1}^{p}\mathbf{e}_{\ell_1}^T\mathbf{R}\mathbf{D}_2\mathbf{e}_{\ell_1}\mathbf{e}_{\ell_2}^T\mathbf{R}\mathbf{D}_2\mathbf{e}_{\ell_2}\\
&\qquad\times\Big[2(\mathbf{e}_{\ell_1}^T\mathbf{R}\mathbf{e}_{\ell_2})^2+\beta_w\sum_{k=1}^{p}(\mathbf{e}_{\ell_1}^T\boldsymbol{\Gamma}\mathbf{e}_k)^2(\mathbf{e}_{\ell_2}^T\boldsymbol{\Gamma}\mathbf{e}_k)^2\Big]+o_p(1),\\
\sigma_{340}(\mathbf{D}_2) &= \sum_{j=1}^{n}\mathrm{E}_{j-1}\big\{[(\mathrm{E}_j-\mathrm{E}_{j-1})\mathrm{tr}(\boldsymbol{\Xi}_n\mathbf{D}_2\boldsymbol{\Xi}_n\mathbf{D}_2)]\\
&\qquad\times[(\mathrm{E}_j-\mathrm{E}_{j-1})\mathrm{tr}[(\mathrm{diag}(\boldsymbol{\Xi}_n)-\mathbf{I}_p)\boldsymbol{\Xi}_n\mathbf{D}_2\boldsymbol{\Xi}_n\mathbf{D}_2]]\big\}\\
&= 2n^{-1}\sum_{\ell=1}^{p}\mathbf{e}_\ell^T(\mathbf{R}\mathbf{D}_2)^2\mathbf{e}_\ell
\end{aligned}$$

$$
\begin{aligned}
&\times\Big[2\mathbf{e}_\ell^T(\mathbf{RD}_2)^2\mathbf{Re}_\ell+\beta_w\sum_{k=1}^p(\mathbf{e}_\ell^T\boldsymbol{\Gamma}\mathbf{e}_k)^2\mathbf{e}_k^T\boldsymbol{\Gamma}^T\mathbf{D}_2\mathbf{RD}_2\boldsymbol{\Gamma}\mathbf{e}_k\Big]\\
&+2n^{-1}\mathrm{tr}(\mathbf{RD}_2)n^{-1}\sum_{\ell=1}^p\mathbf{e}_\ell^T\mathbf{RD}_2\mathbf{e}_\ell\\
&\times\Big[2\mathbf{e}_\ell^T(\mathbf{RD}_2)^2\mathbf{Re}_\ell+\beta_w\sum_{k=1}^p(\mathbf{e}_\ell^T\boldsymbol{\Gamma}\mathbf{e}_k)^2\mathbf{e}_k^T\boldsymbol{\Gamma}^T\mathbf{D}_2\mathbf{RD}_2\boldsymbol{\Gamma}\mathbf{e}_k\Big]\\
&+2n^{-1}\mathrm{tr}(\mathbf{RD}_2)n^{-1}\sum_{\ell=1}^p\mathbf{e}_\ell^T(\mathbf{RD}_2)^2\mathbf{e}_\ell\\
&\times\Big[2\mathbf{e}_\ell^T\mathbf{RD}_2\mathbf{Re}_\ell+\beta_w\sum_{k=1}^p(\mathbf{e}_\ell^T\boldsymbol{\Gamma}\mathbf{e}_k)^2\mathbf{e}_k^T\boldsymbol{\Gamma}^T\mathbf{D}_2\boldsymbol{\Gamma}\mathbf{e}_k\Big]\\
&+2[n^{-1}\mathrm{tr}(\mathbf{RD}_2)]^2n^{-1}\sum_{\ell=1}^p\mathbf{e}_\ell^T\mathbf{RD}_2\mathbf{e}_\ell\\
&\times\Big[2\mathbf{e}_\ell^T\mathbf{RD}_2\mathbf{Re}_\ell+\beta_w\sum_{k=1}^p(\mathbf{e}_\ell^T\boldsymbol{\Gamma}\mathbf{e}_k)^2\mathbf{e}_k^T\boldsymbol{\Gamma}^T\mathbf{D}_2\boldsymbol{\Gamma}\mathbf{e}_k\Big]+o_p(1).
\end{aligned}
$$

Therefore, we have

$$
\begin{aligned}
\sigma_{22}(\mathbf{D}_2) \;=\; & 4n^{-1}\Big[2\mathrm{tr}(\mathbf{RD}_2)^4+\beta_w\sum_{\ell=1}^p(\mathbf{e}_\ell^T\boldsymbol{\Gamma}^T\mathbf{D}_2\mathbf{RD}_2\boldsymbol{\Gamma}\mathbf{e}_\ell)^2\Big] \qquad \text{(S1.22)}\\
&+4[n^{-1}\mathrm{tr}(\mathbf{RD}_2)]^2n^{-1}\Big[2\mathrm{tr}(\mathbf{RD}_2)^2+\beta_w\sum_{\ell=1}^p(\mathbf{e}_\ell^T\boldsymbol{\Gamma}^T\mathbf{D}_2\boldsymbol{\Gamma}\mathbf{e}_\ell)^2\Big]\\
&+4\big[n^{-1}\mathrm{tr}(\mathbf{RD}_2)^2\big]^2+8n^{-1}\mathrm{tr}(\mathbf{RD}_2)n^{-1}\Big[2\mathrm{tr}(\mathbf{RD}_2)^3\\
&+\beta_w\sum_{\ell=1}^p\mathbf{e}_\ell^T\boldsymbol{\Gamma}^T\mathbf{D}_2\mathbf{RD}_2\boldsymbol{\Gamma}\mathbf{e}_\ell\mathbf{e}_\ell^T\boldsymbol{\Gamma}^T\mathbf{D}_2\boldsymbol{\Gamma}\mathbf{e}_\ell\Big]\\
&+4n^{-1}\sum_{\ell_1=1}^p\sum_{\ell_2=1}^p\mathbf{e}_{\ell_1}^T(\mathbf{RD}_2)^2\mathbf{e}_{\ell_1}\mathbf{e}_{\ell_2}^T(\mathbf{RD}_2)^2\mathbf{e}_{\ell_2}\\
&\times\Big[2(\mathbf{e}_{\ell_1}^T\mathbf{Re}_{\ell_2})^2+\beta_w\sum_{k=1}^p(\mathbf{e}_{\ell_1}^T\boldsymbol{\Gamma}\mathbf{e}_k)^2(\mathbf{e}_{\ell_2}^T\boldsymbol{\Gamma}\mathbf{e}_k)^2\Big]\\
&+8n^{-1}\mathrm{tr}(\mathbf{RD}_2)n^{-1}\sum_{\ell_1=1}^p\sum_{\ell_2=1}^p\mathbf{e}_{\ell_1}^T\mathbf{RD}_2\mathbf{e}_{\ell_1}\mathbf{e}_{\ell_2}^T(\mathbf{RD}_2)^2\mathbf{e}_{\ell_2}\\
&\times\Big[2(\mathbf{e}_{\ell_1}^T\mathbf{Re}_{\ell_2})^2+\beta_w\sum_{k=1}^p(\mathbf{e}_{\ell_1}^T\boldsymbol{\Gamma}\mathbf{e}_k)^2(\mathbf{e}_{\ell_2}^T\boldsymbol{\Gamma}\mathbf{e}_k)^2\Big]\\
&+4[n^{-1}\mathrm{tr}(\mathbf{RD}_2)]^2n^{-1}\sum_{\ell_1=1}^p\sum_{\ell_2=1}^p\mathbf{e}_{\ell_1}^T\mathbf{RD}_2\mathbf{e}_{\ell_1}\mathbf{e}_{\ell_2}^T\mathbf{RD}_2\mathbf{e}_{\ell_2}
\end{aligned}
$$

$$
\begin{aligned}
&\times\Big[2(\mathbf{e}_{\ell_1}^T\mathbf{R}\mathbf{e}_{\ell_2})^2+\beta_w\sum_{k=1}^p(\mathbf{e}_{\ell_1}^T\boldsymbol{\Gamma}\mathbf{e}_k)^2(\mathbf{e}_{\ell_2}^T\boldsymbol{\Gamma}\mathbf{e}_k)^2\Big]\\
&-8n^{-1}\sum_{\ell=1}^p\mathbf{e}_\ell^T(\mathbf{R}\mathbf{D}_2)^2\mathbf{e}_\ell\\
&\quad\times\Big[2\mathbf{e}_\ell^T(\mathbf{R}\mathbf{D}_2)^2\mathbf{R}\mathbf{e}_\ell+\beta_w\sum_{k=1}^p(\mathbf{e}_\ell^T\boldsymbol{\Gamma}\mathbf{e}_k)^2\mathbf{e}_k^T\boldsymbol{\Gamma}^T\mathbf{D}_2\mathbf{R}\mathbf{D}_2\boldsymbol{\Gamma}\mathbf{e}_k\Big]\\
&-8n^{-1}\mathrm{tr}(\mathbf{R}\mathbf{D}_2)n^{-1}\sum_{\ell=1}^p\mathbf{e}_\ell^T\mathbf{R}\mathbf{D}_2\mathbf{e}_\ell\\
&\quad\times\Big[2\mathbf{e}_\ell^T(\mathbf{R}\mathbf{D}_2)^2\mathbf{R}\mathbf{e}_\ell+\beta_w\sum_{k=1}^p(\mathbf{e}_\ell^T\boldsymbol{\Gamma}\mathbf{e}_k)^2\mathbf{e}_k^T\boldsymbol{\Gamma}^T\mathbf{D}_2\mathbf{R}\mathbf{D}_2\boldsymbol{\Gamma}\mathbf{e}_k\Big]\\
&-8n^{-1}\mathrm{tr}(\mathbf{R}\mathbf{D}_2)n^{-1}\sum_{\ell=1}^p\mathbf{e}_\ell^T(\mathbf{R}\mathbf{D}_2)^2\mathbf{e}_\ell\\
&\quad\times\Big[2\mathbf{e}_\ell^T\mathbf{R}\mathbf{D}_2\mathbf{R}\mathbf{e}_\ell+\beta_w\sum_{k=1}^p(\mathbf{e}_\ell^T\boldsymbol{\Gamma}\mathbf{e}_k)^2\mathbf{e}_k^T\boldsymbol{\Gamma}^T\mathbf{D}_2\boldsymbol{\Gamma}\mathbf{e}_k\Big]\\
&-8[n^{-1}\mathrm{tr}(\mathbf{R}\mathbf{D}_2)]^2n^{-1}\sum_{\ell=1}^p\mathbf{e}_\ell^T\mathbf{R}\mathbf{D}_2\mathbf{e}_\ell\\
&\quad\times\Big[2\mathbf{e}_\ell^T\mathbf{R}\mathbf{D}_2\mathbf{R}\mathbf{e}_\ell+\beta_w\sum_{k=1}^p(\mathbf{e}_\ell^T\boldsymbol{\Gamma}\mathbf{e}_k)^2\mathbf{e}_k^T\boldsymbol{\Gamma}^T\mathbf{D}_2\boldsymbol{\Gamma}\mathbf{e}_k\Big]+o_p(1).
\end{aligned}
$$

Based on (S1.20) and (S1.21), after complicated calculation, we get

$$
\begin{aligned}
\sigma_{130}(\mathbf{D}_1,\mathbf{D}_2) &= \sum_{j=1}^n\mathrm{E}_{j-1}\big\{[(\mathrm{E}_j-\mathrm{E}_{j-1})\mathrm{tr}(\boldsymbol{\Xi}_n\mathbf{D}_1)][(\mathrm{E}_j-\mathrm{E}_{j-1})\mathrm{tr}(\boldsymbol{\Xi}_n\mathbf{D}_2\boldsymbol{\Xi}_n\mathbf{D}_2)]\big\}\\
&= 2n^{-1}\Big[2\mathrm{tr}(\mathbf{R}\mathbf{D}_1(\mathbf{R}\mathbf{D}_2)^2)+\beta_w\sum_{\ell=1}^p\mathbf{e}_\ell^T\boldsymbol{\Gamma}^T\mathbf{D}_1\boldsymbol{\Gamma}\mathbf{e}_\ell\mathbf{e}_\ell^T\boldsymbol{\Gamma}^T\mathbf{D}_2\mathbf{R}\mathbf{D}_2\boldsymbol{\Gamma}\mathbf{e}_\ell\Big]\\
&\quad+2n^{-1}\mathrm{tr}(\mathbf{R}\mathbf{D}_2)n^{-1}\Big[2\mathrm{tr}(\mathbf{R}\mathbf{D}_1\mathbf{R}\mathbf{D}_2)+\beta_w\sum_{\ell=1}^p\mathbf{e}_\ell^T\boldsymbol{\Gamma}^T\mathbf{D}_1\boldsymbol{\Gamma}\mathbf{e}_\ell\mathbf{e}_\ell^T\boldsymbol{\Gamma}^T\mathbf{D}_2\boldsymbol{\Gamma}\mathbf{e}_\ell\Big]+o_p(1),\\
\sigma_{140}(\mathbf{D}_1,\mathbf{D}_2) &= \sum_{j=1}^n\mathrm{E}_{j-1}\big\{[(\mathrm{E}_j-\mathrm{E}_{j-1})\mathrm{tr}(\boldsymbol{\Xi}_n\mathbf{D}_1)]\\
&\qquad\times[(\mathrm{E}_j-\mathrm{E}_{j-1})\mathrm{tr}[(\mathrm{diag}(\boldsymbol{\Xi}_n)-\mathbf{I}_p)\boldsymbol{\Xi}_n\mathbf{D}_2\boldsymbol{\Xi}_n\mathbf{D}_2)]]\big\}\\
&= n^{-1}\sum_{\ell=1}^p\mathbf{e}_\ell^T(\mathbf{R}\mathbf{D}_2)^2\mathbf{e}_\ell\Big[2\mathbf{e}_\ell^T\mathbf{R}\mathbf{D}_1\mathbf{R}\mathbf{e}_\ell+\beta_w\sum_{k=1}^p(\mathbf{e}_\ell^T\boldsymbol{\Gamma}\mathbf{e}_k)^2\mathbf{e}_k^T\boldsymbol{\Gamma}^T\mathbf{D}_1\boldsymbol{\Gamma}\mathbf{e}_k\Big]\\
&\quad+n^{-1}\mathrm{tr}(\mathbf{R}\mathbf{D}_2)n^{-1}\sum_{\ell=1}^p\mathbf{e}_\ell^T\mathbf{R}\mathbf{D}_2\mathbf{e}_\ell\Big[2\mathbf{e}_\ell^T\mathbf{R}\mathbf{D}_1\mathbf{R}\mathbf{e}_\ell+\beta_w\sum_{k=1}^p(\mathbf{e}_\ell^T\boldsymbol{\Gamma}\mathbf{e}_k)^2\mathbf{e}_k^T\boldsymbol{\Gamma}^T\mathbf{D}_1\boldsymbol{\Gamma}\mathbf{e}_k\Big]+o_p(1),
\end{aligned}
$$

$$\begin{aligned}
\sigma_{230}(\mathbf{D}_1,\mathbf{D}_2) &= \sum_{j=1}^{n} \mathrm{E}_{j-1}\big\{[(\mathrm{E}_j-\mathrm{E}_{j-1})\mathrm{tr}[(\mathrm{diag}(\boldsymbol{\Xi}_n)-\mathbf{I}_p)\boldsymbol{\Xi}_n\mathbf{D}_1]] \\
&\quad \times[(\mathrm{E}_j-\mathrm{E}_{j-1})\mathrm{tr}(\boldsymbol{\Xi}_n\mathbf{D}_2\boldsymbol{\Xi}_n\mathbf{D}_2)]\big\} \\
&= 2n^{-1}\sum_{\ell=1}^{p}\mathbf{e}_\ell^T\mathbf{R}\mathbf{D}_1\mathbf{e}_\ell\Big[2\mathbf{e}_\ell^T(\mathbf{R}\mathbf{D}_2)^2\mathbf{R}\mathbf{e}_\ell+\beta_w\sum_{k=1}^{p}(\mathbf{e}_\ell^T\boldsymbol{\Gamma}\mathbf{e}_k)^2\mathbf{e}_k^T\boldsymbol{\Gamma}^T\mathbf{D}_2\mathbf{R}\mathbf{D}_2\boldsymbol{\Gamma}\mathbf{e}_k\Big] \\
&\quad +2n^{-1}\mathrm{tr}(\mathbf{R}\mathbf{D}_2)n^{-1}\sum_{\ell=1}^{p}\mathbf{e}_\ell^T\mathbf{R}\mathbf{D}_1\mathbf{e}_\ell \\
&\quad \times\Big[2\mathbf{e}_\ell^T\mathbf{R}\mathbf{D}_2\mathbf{R}\mathbf{e}_\ell+\beta_w\sum_{k=1}^{p}(\mathbf{e}_\ell^T\boldsymbol{\Gamma}\mathbf{e}_k)^2\mathbf{e}_k^T\boldsymbol{\Gamma}^T\mathbf{D}_2\boldsymbol{\Gamma}\mathbf{e}_k\Big]+o_p(1), \\
\sigma_{240}(\mathbf{D}_1,\mathbf{D}_2) &= \sum_{j=1}^{n} \mathrm{E}_{j-1}\big\{[(\mathrm{E}_j-\mathrm{E}_{j-1})\mathrm{tr}[(\mathrm{diag}(\boldsymbol{\Xi}_n)-\mathbf{I}_p)\boldsymbol{\Xi}_n\mathbf{D}_1]] \\
&\quad \times[(\mathrm{E}_j-\mathrm{E}_{j-1})\mathrm{tr}[(\mathrm{diag}(\boldsymbol{\Xi}_n)-\mathbf{I}_p)\boldsymbol{\Xi}_n\mathbf{D}_2\boldsymbol{\Xi}_n\mathbf{D}_2)]]\big\} \\
&= n^{-1}\sum_{\ell_1=1}^{p}\sum_{\ell_2=1}^{p}\mathbf{e}_{\ell_1}^T\mathbf{R}\mathbf{D}_1\mathbf{e}_{\ell_1}\mathbf{e}_{\ell_2}^T(\mathbf{R}\mathbf{D}_2)^2\mathbf{e}_{\ell_2} \\
&\quad \times\Big[2(\mathbf{e}_{\ell_1}^T\mathbf{R}\mathbf{e}_{\ell_2})^2+\beta_w\sum_{k=1}^{p}(\mathbf{e}_{\ell_1}^T\boldsymbol{\Gamma}\mathbf{e}_k)^2(\mathbf{e}_{\ell_2}^T\boldsymbol{\Gamma}\mathbf{e}_k)^2\Big] \\
&\quad +n^{-1}\mathrm{tr}(\mathbf{R}\mathbf{D}_2)n^{-1}\sum_{\ell_1=1}^{p}\sum_{\ell_2=1}^{p}\mathbf{e}_{\ell_1}^T\mathbf{R}\mathbf{D}_1\mathbf{e}_{\ell_1}\mathbf{e}_{\ell_2}^T\mathbf{R}\mathbf{D}_2\mathbf{e}_{\ell_2} \\
&\quad \times\Big[2(\mathbf{e}_{\ell_1}^T\mathbf{R}\mathbf{e}_{\ell_2})^2+\beta_w\sum_{k=1}^{p}(\mathbf{e}_{\ell_1}^T\boldsymbol{\Gamma}\mathbf{e}_k)^2(\mathbf{e}_{\ell_2}^T\boldsymbol{\Gamma}\mathbf{e}_k)^2\Big]+o_p(1).
\end{aligned}$$

Thus, we have

$$\begin{aligned}
\sigma_{12}(\mathbf{D}_1,\mathbf{D}_2) &= 2n^{-1}\Big[2\mathrm{tr}(\mathbf{R}\mathbf{D}_1(\mathbf{R}\mathbf{D}_2)^2)+\beta_w\sum_{\ell=1}^{p}\mathbf{e}_\ell^T\boldsymbol{\Gamma}^T\mathbf{D}_1\boldsymbol{\Gamma}\mathbf{e}_\ell\mathbf{e}_\ell^T\boldsymbol{\Gamma}^T\mathbf{D}_2\mathbf{R}\mathbf{D}_2\boldsymbol{\Gamma}\mathbf{e}_\ell\Big] \qquad \text{(S1.23)} \\
&\quad +2n^{-1}\mathrm{tr}(\mathbf{R}\mathbf{D}_2)n^{-1}\Big[2\mathrm{tr}(\mathbf{R}\mathbf{D}_1\mathbf{R}\mathbf{D}_2)+\beta_w\sum_{\ell=1}^{p}\mathbf{e}_\ell^T\boldsymbol{\Gamma}^T\mathbf{D}_1\boldsymbol{\Gamma}\mathbf{e}_\ell\mathbf{e}_\ell^T\boldsymbol{\Gamma}^T\mathbf{D}_2\boldsymbol{\Gamma}\mathbf{e}_\ell\Big] \\
&\quad -2n^{-1}\sum_{\ell=1}^{p}\mathbf{e}_\ell^T(\mathbf{R}\mathbf{D}_2)^2\mathbf{e}_\ell\Big[2\mathbf{e}_\ell^T\mathbf{R}\mathbf{D}_1\mathbf{R}\mathbf{e}_\ell+\beta_w\sum_{k=1}^{p}(\mathbf{e}_\ell^T\boldsymbol{\Gamma}\mathbf{e}_k)^2\mathbf{e}_k^T\boldsymbol{\Gamma}^T\mathbf{D}_1\boldsymbol{\Gamma}\mathbf{e}_k\Big] \\
&\quad -2n^{-1}\sum_{\ell=1}^{p}\mathbf{e}_\ell^T\mathbf{R}\mathbf{D}_1\mathbf{e}_\ell\Big[2\mathbf{e}_\ell^T(\mathbf{R}\mathbf{D}_2)^2\mathbf{R}\mathbf{e}_\ell+\beta_w\sum_{k=1}^{p}(\mathbf{e}_\ell^T\boldsymbol{\Gamma}\mathbf{e}_k)^2\mathbf{e}_k^T\boldsymbol{\Gamma}^T\mathbf{D}_2\mathbf{R}\mathbf{D}_2\boldsymbol{\Gamma}\mathbf{e}_k\Big] \\
&\quad -2n^{-1}\mathrm{tr}(\mathbf{R}\mathbf{D}_2)n^{-1}\sum_{\ell=1}^{p}\mathbf{e}_\ell^T\mathbf{R}\mathbf{D}_2\mathbf{e}_\ell
\end{aligned}$$

$$
\begin{aligned}
&\times\Big[2\mathbf{e}_\ell^T\mathbf{R}\mathbf{D}_1\mathbf{R}\mathbf{e}_\ell+\beta_w\sum_{k=1}^p(\mathbf{e}_\ell^T\boldsymbol{\Gamma}\mathbf{e}_k)^2\mathbf{e}_k^T\boldsymbol{\Gamma}^T\mathbf{D}_1\boldsymbol{\Gamma}\mathbf{e}_k\Big]\\
&-2n^{-1}\mathrm{tr}(\mathbf{R}\mathbf{D}_2)n^{-1}\sum_{\ell=1}^p\mathbf{e}_\ell^T\mathbf{R}\mathbf{D}_1\mathbf{e}_\ell\\
&\times\Big[2\mathbf{e}_\ell^T\mathbf{R}\mathbf{D}_2\mathbf{R}\mathbf{e}_\ell+\beta_w\sum_{k=1}^p(\mathbf{e}_\ell^T\boldsymbol{\Gamma}\mathbf{e}_k)^2\mathbf{e}_k^T\boldsymbol{\Gamma}^T\mathbf{D}_2\boldsymbol{\Gamma}\mathbf{e}_k\Big]\\
&+2n^{-1}\sum_{\ell_1=1}^p\sum_{\ell_2=1}^p\mathbf{e}_{\ell_1}^T\mathbf{R}\mathbf{D}_1\mathbf{e}_{\ell_1}\mathbf{e}_{\ell_2}^T(\mathbf{R}\mathbf{D}_2)^2\mathbf{e}_{\ell_2}\\
&\times\Big[2(\mathbf{e}_{\ell_1}^T\mathbf{R}\mathbf{e}_{\ell_2})^2+\beta_w\sum_{k=1}^p(\mathbf{e}_{\ell_1}^T\boldsymbol{\Gamma}\mathbf{e}_k)^2(\mathbf{e}_{\ell_2}^T\boldsymbol{\Gamma}\mathbf{e}_k)^2\Big]\\
&+2n^{-1}\mathrm{tr}(\mathbf{R}\mathbf{D}_2)n^{-1}\sum_{\ell_1=1}^p\sum_{\ell_2=1}^p\mathbf{e}_{\ell_1}^T\mathbf{R}\mathbf{D}_1\mathbf{e}_{\ell_1}\mathbf{e}_{\ell_2}^T\mathbf{R}\mathbf{D}_2\mathbf{e}_{\ell_2}\\
&\times\Big[2(\mathbf{e}_{\ell_1}^T\mathbf{R}\mathbf{e}_{\ell_2})^2+\beta_w\sum_{k=1}^p(\mathbf{e}_{\ell_1}^T\boldsymbol{\Gamma}\mathbf{e}_k)^2(\mathbf{e}_{\ell_2}^T\boldsymbol{\Gamma}\mathbf{e}_k)^2\Big]+o_p(1).
\end{aligned}
$$

According to the CLT for martingale difference sequence, we obtain that

$$
\Lambda_n^{-1/2}\begin{pmatrix}\mathrm{tr}\big(\widehat{\boldsymbol{\Sigma}}_n\mathbf{D}_1\big)-\mathrm{tr}\big[\big(\mathrm{diag}(\widehat{\boldsymbol{\Sigma}}_n)-\mathbf{I}_p\big)\widehat{\boldsymbol{\Sigma}}_n\mathbf{D}_1\big]-\mu_{01}(\mathbf{D}_1)\\ \mathrm{tr}\big(\widehat{\boldsymbol{\Sigma}}_n\mathbf{D}_2\widehat{\boldsymbol{\Sigma}}_n\mathbf{D}_2\big)-2\mathrm{tr}\big[\big(\mathrm{diag}(\widehat{\boldsymbol{\Sigma}}_n)-\mathbf{I}_p\big)\widehat{\boldsymbol{\Sigma}}_n\mathbf{D}_2\widehat{\boldsymbol{\Sigma}}_n\mathbf{D}_2\big]-\mu_{02}(\mathbf{D}_2)\end{pmatrix} \tag{S1.24}
$$

is asymptotically distributed as the standard bivariate normal distribution with

$$
\Lambda_n=\begin{pmatrix}\sigma_{11}(\mathbf{D}_1) & \sigma_{12}(\mathbf{D}_1,\mathbf{D}_2)\\ \sigma_{21}(\mathbf{D}_1,\mathbf{D}_2) & \sigma_{22}(\mathbf{D}_2)\end{pmatrix},
$$

where $\sigma_{11}(\mathbf{D}_1),\sigma_{22}(\mathbf{D}_2),\sigma_{12}(\mathbf{D}_1,\mathbf{D}_2)$ are given in (S1.19)-(S1.22)-(S1.23) and

$$
\begin{aligned}
\mu_{01}(\mathbf{D}_1) &= \mathrm{Etr}\big(\widehat{\boldsymbol{\Sigma}}_n\mathbf{D}_1\big)-\mathrm{Etr}\big[\big(\mathrm{diag}(\widehat{\boldsymbol{\Sigma}}_n)-\mathbf{I}_p\big)\widehat{\boldsymbol{\Sigma}}_n\mathbf{D}_1\big],\\
\mu_{02}(\mathbf{D}_2) &= \mathrm{Etr}\big(\widehat{\boldsymbol{\Sigma}}_n\mathbf{D}_2\widehat{\boldsymbol{\Sigma}}_n\mathbf{D}_2\big)-2\mathrm{Etr}\big[\big(\mathrm{diag}(\widehat{\boldsymbol{\Sigma}}_n)-\mathbf{I}_p\big)\widehat{\boldsymbol{\Sigma}}_n\mathbf{D}_2\widehat{\boldsymbol{\Sigma}}_n\mathbf{D}_2\big].
\end{aligned}
$$

Because

$$
\begin{aligned}
&\mathrm{Etr}(\widehat{\boldsymbol{\Sigma}}_n\mathbf{D}_1)-\mathrm{Etr}[(\mathrm{diag}(\widehat{\boldsymbol{\Sigma}}_n)-\mathbf{I}_p)\widehat{\boldsymbol{\Sigma}}_n\mathbf{D}_1]\\
=\;&\mathrm{tr}(\mathbf{R}\mathbf{D}_1)-n^{-1}\Big[2\mathrm{tr}(\mathbf{R}\mathbf{D}_1)+\beta_w\sum_{k=1}^p\sum_{\ell=1}^p\mathbf{e}_k^T\mathbf{D}_1\boldsymbol{\Gamma}\mathbf{e}_\ell(\mathbf{e}_k^T\boldsymbol{\Gamma}\mathbf{e}_\ell)^3\Big]+o(1),
\end{aligned}
$$

$$
\begin{aligned}
&\mathrm{Etr}(\widehat{\boldsymbol{\Sigma}}_n\mathbf{D}_2\widehat{\boldsymbol{\Sigma}}_n\mathbf{D}_2)\\
=\;&\frac{(n-1)}{n^2}\mathrm{tr}^2(\mathbf{R}\mathbf{D}_2)+\frac{(n-1)}{n}\mathrm{tr}(\mathbf{R}\mathbf{D}_2)^2+\beta_w n^{-1}\sum_{k=1}^{p}(\mathbf{e}_k^T\boldsymbol{\Gamma}^T\mathbf{D}_2\boldsymbol{\Gamma}\mathbf{e}_k)^2+o(1),\\
&\mathrm{Etr}[(\mathrm{diag}(\widehat{\boldsymbol{\Sigma}}_n)-\mathbf{I}_p)\widehat{\boldsymbol{\Sigma}}_n\mathbf{D}_2\widehat{\boldsymbol{\Sigma}}_n\mathbf{D}_2]\\
=\;&n^{-2}\mathrm{tr}^2(\mathbf{R}\mathbf{D}_2)+n^{-1}\mathrm{tr}(\mathbf{R}\mathbf{D}_2)^2+2n^{-1}\sum_{k=1}^{p}\mathbf{e}_k^T\mathbf{R}\mathbf{D}_2\mathbf{e}_k\mathbf{e}_k^T\mathbf{R}\mathbf{D}_2\mathbf{R}\mathbf{e}_k\\
&+\beta_w n^{-1}\mathrm{tr}(\mathbf{R}\mathbf{D}_2)n^{-1}\sum_{k=1}^{p}\sum_{\ell=1}^{p}(\mathbf{e}_k^T\boldsymbol{\Gamma}\mathbf{e}_\ell)^3\mathbf{e}_k^T\mathbf{D}_2\boldsymbol{\Gamma}\mathbf{e}_\ell\\
&+\beta_w n^{-1}\sum_{k=1}^{p}\sum_{\ell=1}^{p}(\mathbf{e}_k^T\boldsymbol{\Gamma}\mathbf{e}_\ell)^3\mathbf{e}_k^T\mathbf{D}_2\mathbf{R}\mathbf{D}_2\boldsymbol{\Gamma}\mathbf{e}_\ell\\
&+\beta_w n^{-1}\sum_{k=1}^{p}\sum_{\ell=1}^{p}(\mathbf{e}_k^T\boldsymbol{\Gamma}\mathbf{e}_\ell)^2\mathbf{e}_k^T\mathbf{D}_2\boldsymbol{\Gamma}\mathbf{e}_\ell\mathbf{e}_k^T\mathbf{R}\mathbf{D}_2\boldsymbol{\Gamma}\mathbf{e}_\ell+o(1),
\end{aligned}
$$

then we have

$$
\begin{aligned}
\mu_{01}(\mathbf{D}_1)&=\mathrm{tr}(\mathbf{R}\mathbf{D}_1)-n^{-1}\Big[2\mathrm{tr}(\mathbf{R}\mathbf{D}_1)+\beta_w n^{-1}\sum_{k=1}^{p}\sum_{\ell=1}^{p}\mathbf{e}_k^T\mathbf{D}_1\boldsymbol{\Gamma}\mathbf{e}_\ell(\mathbf{e}_k^T\boldsymbol{\Gamma}\mathbf{e}_\ell)^3\Big]+o(1),\\
\mu_{02}(\mathbf{D}_2)&=\frac{n-3}{n^2}\mathrm{tr}^2(\mathbf{R}\mathbf{D}_2)-4n^{-1}\sum_{k=1}^{p}\mathbf{e}_k^T\mathbf{R}\mathbf{D}_2\mathbf{e}_k\mathbf{e}_k^T\mathbf{R}\mathbf{D}_2\mathbf{R}\mathbf{e}_k\\
&\quad+\frac{n-3}{n}\mathrm{tr}(\mathbf{R}\mathbf{D}_2)^2+\beta_w n^{-1}\sum_{k=1}^{p}(\mathbf{e}_k^T\boldsymbol{\Gamma}^T\mathbf{D}_2\boldsymbol{\Gamma}\mathbf{e}_k)^2\\
&\quad-2\beta_w n^{-1}\mathrm{tr}(\mathbf{R}\mathbf{D}_2)n^{-1}\sum_{k=1}^{p}\sum_{\ell=1}^{p}(\mathbf{e}_k^T\boldsymbol{\Gamma}\mathbf{e}_\ell)^3\mathbf{e}_k^T\mathbf{D}_2\boldsymbol{\Gamma}\mathbf{e}_\ell\\
&\quad-2\beta_w n^{-1}\sum_{k=1}^{p}\sum_{\ell=1}^{p}(\mathbf{e}_k^T\boldsymbol{\Gamma}\mathbf{e}_\ell)^3\mathbf{e}_k^T\mathbf{D}_2\mathbf{R}\mathbf{D}_2\boldsymbol{\Gamma}\mathbf{e}_\ell\\
&\quad-2\beta_w n^{-1}\sum_{k=1}^{p}\sum_{\ell=1}^{p}(\mathbf{e}_k^T\boldsymbol{\Gamma}\mathbf{e}_\ell)^2\mathbf{e}_k^T\mathbf{D}_2\boldsymbol{\Gamma}\mathbf{e}_\ell\mathbf{e}_k^T\mathbf{R}\mathbf{D}_2\boldsymbol{\Gamma}\mathbf{e}_\ell+o(1).
\end{aligned}
$$

By (S1.17)-(S1.18)-(S1.24), we obtain that

$$
\begin{pmatrix}\sigma_{11}(\mathbf{D}_1) & \sigma_{12}(\mathbf{D}_1,\mathbf{D}_2)\\ \sigma_{21}(\mathbf{D}_1,\mathbf{D}_2) & \sigma_{22}(\mathbf{D}_2)\end{pmatrix}^{-1/2}\begin{pmatrix}\mathrm{tr}\big(\widehat{\mathbf{R}}_n\mathbf{D}_1\big)-\nu_1(\mathbf{D}_1)\\ \mathrm{tr}\big(\widehat{\mathbf{R}}_n\mathbf{D}_2\widehat{\mathbf{R}}_n\mathbf{D}_2\big)-\nu_2(\mathbf{D}_2)\end{pmatrix}\xrightarrow{d}N\left(\mathbf{0}_2,\mathbf{I}_2\right),
$$

where $\mathbf{0}_2 = (0,0)^T$, $\mathbf{I}_2$ is the 2×2 dimensional identity matrix,

$$\begin{aligned}
\nu_1(\mathbf{D}_1) &= \mathrm{tr}(\mathbf{R}\mathbf{D}_1) - n^{-1}\Big[2\mathrm{tr}(\mathbf{R}\mathbf{D}_1) + \beta_w \sum_{k=1}^{p}\sum_{\ell=1}^{p} \mathbf{e}_k^T \mathbf{D}_1 \mathbf{\Gamma} \mathbf{e}_\ell (\mathbf{e}_k^T \mathbf{\Gamma} \mathbf{e}_\ell)^3\Big] \\
&\quad + \frac{1}{4n}\mathrm{tr}(\mathbf{C}\mathbf{D}_1) + \frac{3}{4n}\Big[2\mathrm{tr}(\mathbf{R}\mathbf{D}_1) + \beta_w \sum_{\ell=1}^{p} \mathbf{e}_\ell^T \mathbf{R}\mathbf{D}_1 \mathbf{e}_\ell \sum_{k=1}^{p} (\mathbf{e}_\ell^T \mathbf{\Gamma} \mathbf{e}_k)^4\Big], \\
\nu_2(\mathbf{D}_2) &= \frac{n-3}{n^2}\mathrm{tr}^2(\mathbf{R}\mathbf{D}_2) - 4n^{-1}\sum_{k=1}^{p} \mathbf{e}_k^T \mathbf{R}\mathbf{D}_2 \mathbf{e}_k \mathbf{e}_k^T \mathbf{R}\mathbf{D}_2 \mathbf{R} \mathbf{e}_k \\
&\quad + \frac{n-3}{n}\mathrm{tr}(\mathbf{R}\mathbf{D}_2)^2 + \beta_w n^{-1} \sum_{k=1}^{p} (\mathbf{e}_k^T \mathbf{\Gamma}^T \mathbf{D}_2 \mathbf{\Gamma} \mathbf{e}_k)^2 \\
&\quad - 2\beta_w n^{-1}\mathrm{tr}(\mathbf{R}\mathbf{D}_2) n^{-1} \sum_{k=1}^{p}\sum_{\ell=1}^{p} (\mathbf{e}_k^T \mathbf{\Gamma} \mathbf{e}_\ell)^3 \mathbf{e}_k^T \mathbf{D}_2 \mathbf{\Gamma} \mathbf{e}_\ell \\
&\quad - 2\beta_w n^{-1} \sum_{k=1}^{p}\sum_{\ell=1}^{p} (\mathbf{e}_k^T \mathbf{\Gamma} \mathbf{e}_\ell)^3 \mathbf{e}_k^T \mathbf{D}_2 \mathbf{R} \mathbf{D}_2 \mathbf{\Gamma} \mathbf{e}_\ell \\
&\quad - 2\beta_w n^{-1} \sum_{k=1}^{p}\sum_{\ell=1}^{p} (\mathbf{e}_k^T \mathbf{\Gamma} \mathbf{e}_\ell)^2 \mathbf{e}_k^T \mathbf{D}_2 \mathbf{\Gamma} \mathbf{e}_\ell \mathbf{e}_k^T \mathbf{R} \mathbf{D}_2 \mathbf{\Gamma} \mathbf{e}_\ell \\
&\quad + \frac{3}{2n}\Big[2\mathrm{tr}(\mathbf{R}\mathbf{D}_2\mathbf{R}\mathbf{D}_2) + \beta_w \sum_{k=1}^{p} \mathbf{e}_k^T \mathbf{R}\mathbf{D}_2\mathbf{R}\mathbf{D}_2 \mathbf{e}_k \sum_{\ell=1}^{p} (\mathbf{e}_k^T \mathbf{\Gamma} \mathbf{e}_\ell)^4\Big] \\
&\quad + \frac{3\mathrm{tr}(\mathbf{R}\mathbf{D}_2)}{2n^2}\Big[2\mathrm{tr}(\mathbf{R}\mathbf{D}_2) + \beta_w \sum_{k=1}^{p} \mathbf{e}_k^T \mathbf{R}\mathbf{D}_2 \mathbf{e}_k \sum_{\ell=1}^{p} (\mathbf{e}_k^T \mathbf{\Gamma} \mathbf{e}_\ell)^4\Big] \\
&\quad + \frac{1}{2n}\sum_{i=1}^{p}\sum_{j=1}^{p} \mathbf{e}_i^T \mathbf{R} \mathbf{e}_j \mathbf{e}_i^T \mathbf{D}_2 \mathbf{R} \mathbf{D}_2 \mathbf{e}_j \Big[2(\mathbf{e}_i^T \mathbf{R} \mathbf{e}_j)^2 + \beta_w \sum_{k=1}^{p} (\mathbf{e}_i^T \mathbf{\Gamma} \mathbf{e}_k)^2 (\mathbf{e}_j^T \mathbf{\Gamma} \mathbf{e}_k)^2\Big] \\
&\quad + \frac{1}{2n}\sum_{i=1}^{p}\sum_{j=1}^{p} \mathbf{e}_i^T \mathbf{D}_2 \mathbf{e}_j \mathbf{e}_i^T \mathbf{R} \mathbf{D}_2 \mathbf{R} \mathbf{e}_j \Big[2(\mathbf{e}_i^T \mathbf{R} \mathbf{e}_j)^2 + \beta_w \sum_{k=1}^{p} (\mathbf{e}_i^T \mathbf{\Gamma} \mathbf{e}_k)^2 (\mathbf{e}_j^T \mathbf{\Gamma} \mathbf{e}_k)^2\Big] \\
&\quad + \frac{\mathrm{tr}(\mathbf{R}\mathbf{D}_2)}{2n^2}\sum_{i=1}^{p}\sum_{j=1}^{p} \mathbf{e}_i^T \mathbf{D}_2 \mathbf{e}_j \mathbf{e}_i^T \mathbf{R} \mathbf{e}_j \Big[2(\mathbf{e}_i^T \mathbf{R} \mathbf{e}_j)^2 + \beta_w \sum_{k=1}^{p} (\mathbf{e}_i^T \mathbf{\Gamma} \mathbf{e}_k)^2 (\mathbf{e}_j^T \mathbf{\Gamma} \mathbf{e}_k)^2\Big] \\
&\quad + \frac{1}{2n}\sum_{i=1}^{p}\sum_{j=1}^{p} \mathbf{e}_i^T \mathbf{R} \mathbf{D}_2 \mathbf{e}_j \mathbf{e}_j^T \mathbf{R} \mathbf{D}_2 \mathbf{e}_i \Big[2(\mathbf{e}_i^T \mathbf{R} \mathbf{e}_j)^2 + \beta_w \sum_{k=1}^{p} (\mathbf{e}_i^T \mathbf{\Gamma} \mathbf{e}_k)^2 (\mathbf{e}_j^T \mathbf{\Gamma} \mathbf{e}_k)^2\Big].
\end{aligned}$$

S2 Proof of Theorem 1

Note that under H_0, $\boldsymbol{\theta} = \mathbf{A}^{-1}\mathbf{a}$ with $\mathbf{a} = (\mathrm{tr}[(\mathbf{R}-\mathbf{J}_0)\mathbf{J}_1], \ldots, \mathrm{tr}[(\mathbf{R}-\mathbf{J}_0)\mathbf{J}_K])^T$. By Lemma 1, for $k = 1, \ldots, K$, we have

$$\mathrm{tr}(\widehat{\mathbf{R}}_n\mathbf{J}_k) - \mathrm{tr}(\mathbf{R}\mathbf{J}_k) = O_p(1).$$

Since $\hat{a}_k = \mathrm{tr}[(\widehat{\mathbf{R}}_n - \mathbf{J}_0)\mathbf{J}_k]$ and $a_k = \mathrm{tr}[(\mathbf{R}-\mathbf{J}_0)\mathbf{J}_k]$, then we have $p^{-1}(\hat{a}_k - a_k) = o_p(1)$. It follows that

$$\hat{\boldsymbol{\theta}} - \boldsymbol{\theta} = \mathbf{A}^{-1}(\hat{\mathbf{a}} - \mathbf{a}) = \left(p^{-1}\mathbf{A}\right)^{-1}p^{-1}(\hat{\mathbf{a}} - \mathbf{a}) = O_p(n^{-1}) = o_p(1).$$

Then the proof of Theorem 1 is completed.

S3 Proof of Theorem 2

In this section, we derive the limiting distributions of T_{1n} and T_{2n}. Let

$$\mathbf{R}_P = \mathbf{J}_0 + \theta_{1P}\mathbf{J}_1 + \cdots + \theta_{KP}\mathbf{J}_K,$$

where $\boldsymbol{\theta}_P = (\theta_{1P}, ..., \theta_{KP})^T = \mathbf{A}^{-1}\mathbf{a}_P$ and $\mathbf{a}_P = (a_{1P}, ..., a_{KP})^T$ with $a_{kP} = \mathrm{tr}[(\mathbf{R}-\mathbf{J}_0)\mathbf{J}_k]$ for $k = 1, \ldots, K$. After simple calculation, we obtain

$$\begin{aligned}
\widehat{\mathbf{R}}_0^{-1} &= \mathbf{R}_P^{-1} - \sum_{k=1}^{K}(\hat{\theta}_k - \theta_{kP})\widehat{\mathbf{R}}_0^{-1}\mathbf{J}_k\mathbf{R}_P^{-1} \qquad \text{(S3.25)}\\
&= \mathbf{R}_P^{-1} - \sum_{k=1}^{K}(\hat{\theta}_k - \theta_{kP})(\widehat{\mathbf{R}}_0^{-1} - \mathbf{R}_P^{-1} + \mathbf{R}_P^{-1})\mathbf{J}_k\mathbf{R}_P^{-1}\\
&= \mathbf{R}_P^{-1} - \sum_{k=1}^{K}(\hat{\theta}_k - \theta_{kP})\mathbf{R}_P^{-1}\mathbf{J}_k\mathbf{R}_P^{-1} - \sum_{k=1}^{K}(\hat{\theta}_k - \theta_{kP})(\widehat{\mathbf{R}}_0^{-1} - \mathbf{R}_P^{-1})\mathbf{J}_k\mathbf{R}_P^{-1}\\
&= \mathbf{R}_P^{-1} - \sum_{k=1}^{K}(\hat{\theta}_k - \theta_{kP})\mathbf{R}_P^{-1}\mathbf{J}_k\mathbf{R}_P^{-1} + \sum_{k_1=1}^{K}\sum_{k_2=1}^{K}(\hat{\theta}_{k_1} - \theta_{k_1P})(\hat{\theta}_{k_2} - \theta_{k_2P})\widehat{\mathbf{R}}_0^{-1}\mathbf{J}_{k_1}\mathbf{R}_P^{-1}\mathbf{J}_{k_2}\mathbf{R}_P^{-1}.
\end{aligned}$$

By (S3.25) and $\hat{\theta}_k = \theta_{kP} + O_p(n^{-1})$ for $k = 1, \ldots, K$, we have

$$
\begin{aligned}
T_{1n} &= \operatorname{tr}[(\widehat{\mathbf{R}}_n \widehat{\mathbf{R}}_0^{-1} - \mathbf{I}_p)^2] \\
&= \operatorname{tr}\big\{[\widehat{\mathbf{R}}_n \mathbf{R}_P^{-1} - \mathbf{I}_p + \widehat{\mathbf{R}}_n(\widehat{\mathbf{R}}_0^{-1} - \mathbf{R}_P^{-1})]^2\big\} \\
&= \operatorname{tr}[(\widehat{\mathbf{R}}_n \mathbf{R}_P^{-1} - \mathbf{I}_p)^2] + 2\operatorname{tr}[(\widehat{\mathbf{R}}_n \mathbf{R}_P^{-1} - \mathbf{I}_p)\widehat{\mathbf{R}}_n(\widehat{\mathbf{R}}_0^{-1} - \mathbf{R}_P^{-1})] \\
&\quad + \operatorname{tr}[\widehat{\mathbf{R}}_n(\widehat{\mathbf{R}}_0^{-1} - \mathbf{R}_P^{-1})\widehat{\mathbf{R}}_n(\widehat{\mathbf{R}}_0^{-1} - \mathbf{R}_P^{-1})] \\
&= \operatorname{tr}[(\widehat{\mathbf{R}}_n \mathbf{R}_P^{-1} - \mathbf{I}_p)^2] + 2\sum_{k=1}^{K}(\hat{\theta}_k - \theta_{kP})\operatorname{tr}(\widehat{\mathbf{R}}_n \mathbf{R}_P^{-1}\mathbf{J}_k \mathbf{R}_P^{-1}) \\
&\quad -2\sum_{k=1}^{K}(\hat{\theta}_k - \theta_{kP})\operatorname{tr}(\widehat{\mathbf{R}}_n \mathbf{R}_P^{-1}\widehat{\mathbf{R}}_n \mathbf{R}_P^{-1}\mathbf{J}_k \mathbf{R}_P^{-1}) \\
&\quad +2\sum_{k_1=1}^{K}\sum_{k_2=1}^{K}(\hat{\theta}_{k_1} - \theta_{k_1P})(\hat{\theta}_{k_2} - \theta_{k_2P})\operatorname{tr}(\widehat{\mathbf{R}}_n \mathbf{R}_P^{-1}\widehat{\mathbf{R}}_n \widehat{\mathbf{R}}_0^{-1}\mathbf{J}_{k_1}\mathbf{R}_P^{-1}\mathbf{J}_{k_2}\mathbf{R}_P^{-1}) \\
&\quad -2\sum_{k_1=1}^{K}\sum_{k_2=1}^{K}(\hat{\theta}_{k_1} - \theta_{k_1P})(\hat{\theta}_{k_2} - \theta_{k_2P})\operatorname{tr}(\widehat{\mathbf{R}}_n \widehat{\mathbf{R}}_0^{-1}\mathbf{J}_{k_1}\mathbf{R}_P^{-1}\mathbf{J}_{k_2}\mathbf{R}_P^{-1}) \\
&\quad +\sum_{k_1=1}^{K}\sum_{k_2=1}^{K}(\hat{\theta}_{k_1} - \theta_{k_1P})(\hat{\theta}_{k_2} - \theta_{k_2P})\operatorname{tr}(\widehat{\mathbf{R}}_n \mathbf{R}_P^{-1}\mathbf{J}_{k_1}\mathbf{R}_P^{-1}\widehat{\mathbf{R}}_n \mathbf{R}_P^{-1}\mathbf{J}_{k_2}\mathbf{R}_P^{-1}) \\
&\quad -2\sum_{k_1,k_2,k_3=1}^{K}\prod_{\ell=1}^{3}(\hat{\theta}_{k_\ell} - \theta_{k_\ell P})\operatorname{tr}(\widehat{\mathbf{R}}_n \mathbf{R}_P^{-1}\mathbf{J}_{k_1}\mathbf{R}_P^{-1}\widehat{\mathbf{R}}_n \widehat{\mathbf{R}}_0^{-1}\mathbf{J}_{k_2}\mathbf{R}_P^{-1}\mathbf{J}_{k_3}\mathbf{R}_P^{-1}) \\
&\quad +\sum_{k_1,k_2,k_3,k_4=1}^{K}\prod_{\ell=1}^{4}(\hat{\theta}_{k_\ell} - \theta_{k_\ell P})\operatorname{tr}(\widehat{\mathbf{R}}_n \widehat{\mathbf{R}}_0^{-1}\mathbf{J}_{k_1}\mathbf{R}_P^{-1}\mathbf{J}_{k_2}\mathbf{R}_P^{-1}\widehat{\mathbf{R}}_n \widehat{\mathbf{R}}_0^{-1}\mathbf{J}_{k_3}\mathbf{R}_P^{-1}\mathbf{J}_{k_4}\mathbf{R}_P^{-1}) \\
&= \operatorname{tr}[(\widehat{\mathbf{R}}_n \mathbf{R}_P^{-1} - \mathbf{I}_p)^2] + 2\sum_{k=1}^{K} p(\hat{\theta}_k - \theta_{kP})p^{-1}\operatorname{tr}(\widehat{\mathbf{R}}_n \mathbf{R}_P^{-1}\mathbf{J}_k \mathbf{R}_P^{-1}) \\
&\quad -2\sum_{k=1}^{K} p(\hat{\theta}_k - \theta_{kP})p^{-1}\operatorname{tr}(\widehat{\mathbf{R}}_n \mathbf{R}_P^{-1}\widehat{\mathbf{R}}_n \mathbf{R}_P^{-1}\mathbf{J}_k \mathbf{R}_P^{-1}) + o_p(1).
\end{aligned}
$$

From the proof of Lemma 1, we get

$$p^{-1}\operatorname{tr}(\widehat{\mathbf{R}}_n \mathbf{M}_1) = p^{-1}\operatorname{tr}(\mathbf{R}\mathbf{M}_1) + o_p(1),$$

$$p^{-1}\operatorname{tr}(\widehat{\mathbf{R}}_n \mathbf{M}_1 \widehat{\mathbf{R}}_n \mathbf{M}_2) = y_n p^{-1}\operatorname{tr}(\mathbf{R}\mathbf{M}_1)p^{-1}\operatorname{tr}(\mathbf{R}\mathbf{M}_2) + p^{-1}\operatorname{tr}(\mathbf{R}\mathbf{M}_1\mathbf{R}\mathbf{M}_2) + o_p(1),$$

where $\mathbf{M}_1$ and $\mathbf{M}_2$ are $p \times p$ dimensional non-random symmetric matrices with uniformly bounded spectral norms. Hence, we have

$$
\begin{aligned}
T_{1n} &= \operatorname{tr}[(\widehat{\mathbf{R}}_n\mathbf{R}_P^{-1} - \mathbf{I}_p)^2] + 2\sum_{k=1}^{K}(\hat{\theta}_k - \theta_{kP})\operatorname{tr}(\mathbf{R}\mathbf{R}_P^{-1}\mathbf{J}_k\mathbf{R}_P^{-1}) \qquad \text{(S3.26)}\\
&\quad -2\sum_{k=1}^{K}(\hat{\theta}_k - \theta_{kP})\big[n^{-1}\operatorname{tr}(\mathbf{R}\mathbf{R}_P^{-1})\operatorname{tr}(\mathbf{R}\mathbf{R}_P^{-1}\mathbf{J}_k\mathbf{R}_P^{-1}) + \operatorname{tr}(\mathbf{R}\mathbf{R}_P^{-1}\mathbf{R}\mathbf{R}_P^{-1}\mathbf{J}_k\mathbf{R}_P^{-1})\big] + o_p(1)\\
&= \operatorname{tr}[(\widehat{\mathbf{R}}_n\mathbf{R}_P^{-1} - \mathbf{I}_p)^2] - 2n^{-1}\operatorname{tr}(\mathbf{R}\mathbf{R}_P^{-1})\operatorname{tr}(\widehat{\mathbf{R}}_0\mathbf{R}_P^{-1}\mathbf{R}\mathbf{R}_P^{-1}) + 2n^{-1}\operatorname{tr}^2(\mathbf{R}\mathbf{R}_P^{-1})\\
&\quad -2\operatorname{tr}[\widehat{\mathbf{R}}_0\mathbf{R}_P^{-1}\mathbf{R}\mathbf{R}_P^{-1}(\mathbf{R}\mathbf{R}_P^{-1} - \mathbf{I}_p)] + 2\operatorname{tr}[\mathbf{R}\mathbf{R}_P^{-1}(\mathbf{R}\mathbf{R}_P^{-1} - \mathbf{I}_p)] + o_p(1)\\
&= \operatorname{tr}[(\widehat{\mathbf{R}}_n\mathbf{R}_P^{-1} - \mathbf{I}_p)^2] - 2\operatorname{tr}(\widehat{\mathbf{R}}_0\mathbf{H}) + 2n^{-1}\operatorname{tr}^2(\mathbf{R}\mathbf{R}_P^{-1}) + 2\operatorname{tr}[\mathbf{R}\mathbf{R}_P^{-1}(\mathbf{R}\mathbf{R}_P^{-1} - \mathbf{I}_p)] + o_p(1)\\
&= \operatorname{tr}(\widehat{\mathbf{R}}_n\mathbf{R}_P^{-1})^2 - 2\operatorname{tr}(\widehat{\mathbf{R}}_n\mathbf{R}_P^{-1}) + p - 2\operatorname{tr}(\mathbf{J}_0\mathbf{H}) - 2\operatorname{tr}(\widehat{\mathbf{R}}_n\mathbf{B}_P)\\
&\quad +2\operatorname{tr}(\mathbf{J}_0\mathbf{B}_P) + 2n^{-1}\operatorname{tr}^2(\mathbf{R}\mathbf{R}_P^{-1}) + 2\operatorname{tr}[\mathbf{R}\mathbf{R}_P^{-1}(\mathbf{R}\mathbf{R}_P^{-1} - \mathbf{I}_p)] + o_p(1),
\end{aligned}
$$

where

$$
\begin{aligned}
&\mathbf{H} = n^{-1}\operatorname{tr}(\mathbf{R}\mathbf{R}_P^{-1})\mathbf{R}_P^{-1}\mathbf{R}\mathbf{R}_P^{-1} + \mathbf{R}_P^{-1}\mathbf{R}\mathbf{R}_P^{-1}(\mathbf{R}\mathbf{R}_P^{-1} - \mathbf{I}_p),\\
&\mathbf{B}_P = \sum_{k=1}^{K} h_{kP}\mathbf{J}_k, \quad h_{kP} = (\operatorname{tr}(\mathbf{J}_1\mathbf{H}), \ldots, \operatorname{tr}(\mathbf{J}_K\mathbf{H}))\mathbf{A}^{-1}\mathbf{e}_k,
\end{aligned}
$$

and $\mathbf{e}_k$ denotes the kth column of the identity matrix $\mathbf{I}_K$. Due to the fact that

$$
\begin{aligned}
\operatorname{tr}(\mathbf{J}_0\mathbf{H}) &= \operatorname{tr}(\mathbf{R}_P\mathbf{H}) - \operatorname{tr}(\mathbf{R}\mathbf{B}_P) + \operatorname{tr}(\mathbf{J}_0\mathbf{B}_P),\\
\operatorname{tr}(\mathbf{R}_P\mathbf{H}) &= n^{-1}\operatorname{tr}^2(\mathbf{R}\mathbf{R}_P^{-1}) + \operatorname{tr}[\mathbf{R}\mathbf{R}_P^{-1}(\mathbf{R}\mathbf{R}_P^{-1} - \mathbf{I}_p)],
\end{aligned}
$$

we obtain

$$
T_{1n} = \operatorname{tr}(\widehat{\mathbf{R}}_n\mathbf{R}_P^{-1})^2 - 2\operatorname{tr}[\widehat{\mathbf{R}}_n(\mathbf{R}_P^{-1} + \mathbf{B}_P)] + 2\operatorname{tr}(\mathbf{R}\mathbf{B}_P) + p + o_p(1).
$$

By Lemma 1 and the Delta method, we have

$$
\sigma_{1n}^{-1}(T_{1n} - \mu_{1n}) \xrightarrow{d} N(0,1),
$$

where

$$
\begin{aligned}
\mu_{1n} &= \nu_2(\mathbf{R}_P^{-1}) - 2\nu_1(\mathbf{R}_P^{-1} + \mathbf{B}_P) + 2\mathrm{tr}(\mathbf{R}\mathbf{B}_P) + p, \\
\sigma_{1n}^2 &= \sigma_{22}(\mathbf{R}_P^{-1}) + 4\sigma_{11}(\mathbf{R}_P^{-1} + \mathbf{B}_P) - 4\sigma_{12}(\mathbf{R}_P^{-1} + \mathbf{B}_P, \mathbf{R}_P^{-1}).
\end{aligned}
$$

Note that $\widehat{\mathbf{R}}_0 - \mathbf{R}_P = \sum_{k=1}^{K} (\hat{\theta}_k - \theta_{kP})\mathbf{J}_k$. From $\hat{\theta}_k - \theta_{kP} = O_p(n^{-1})$, we get

$$
\begin{aligned}
T_{2n} &= \mathrm{tr}[(\widehat{\mathbf{R}}_n - \widehat{\mathbf{R}}_0)^2] \\
&= \mathrm{tr}[(\widehat{\mathbf{R}}_n - \mathbf{R}_P)^2] - 2\mathrm{tr}[(\widehat{\mathbf{R}}_n - \mathbf{R}_P)(\widehat{\mathbf{R}}_0 - \mathbf{R}_P)] + \mathrm{tr}[(\widehat{\mathbf{R}}_0 - \mathbf{R}_P)^2] \\
&= \mathrm{tr}[(\widehat{\mathbf{R}}_n - \mathbf{R}_P)^2] - 2\sum_{k=1}^{K} (\hat{\theta}_k - \theta_{kP})\mathrm{tr}[(\widehat{\mathbf{R}}_n - \mathbf{R}_P)\mathbf{J}_k] + o_p(1) \\
&= \mathrm{tr}[(\widehat{\mathbf{R}}_n - \mathbf{R}_P)^2] - 2\sum_{k=1}^{K} (\hat{\theta}_k - \theta_{kP})\mathrm{tr}[(\mathbf{R} - \mathbf{R}_P)\mathbf{J}_k] + o_p(1) \\
&= \mathrm{tr}[(\widehat{\mathbf{R}}_n - \mathbf{R}_P)^2] - 2\mathrm{tr}(\widehat{\mathbf{R}}_n\tilde{\mathbf{B}}_P) + 2\mathrm{tr}[(\mathbf{R} - \mathbf{R}_P)\mathbf{R}_P] \\
&\quad -2\mathrm{tr}[(\mathbf{R} - \mathbf{R}_P)\mathbf{J}_0] + 2\mathrm{tr}(\mathbf{J}_0\tilde{\mathbf{B}}_P) + o_p(1),
\end{aligned}
$$

where $\tilde{\mathbf{B}}_P = \sum_{k=1}^{K} \tilde{h}_{kP}\mathbf{J}_k$ and $\tilde{h}_{kP} = \big(\mathrm{tr}[(\mathbf{R} - \mathbf{R}_P)\mathbf{J}_1], \ldots, \mathrm{tr}[(\mathbf{R} - \mathbf{R}_P)\mathbf{J}_K]\big)\mathbf{A}^{-1}\mathbf{e}_k$. Due to the fact that $\mathrm{tr}[(\mathbf{R} - \mathbf{R}_P)(\mathbf{R}_P - \mathbf{J}_0)] = \mathrm{tr}(\mathbf{R}\tilde{\mathbf{B}}_P) - \mathrm{tr}(\mathbf{J}_0\tilde{\mathbf{B}}_P)$, we have

$$
\begin{aligned}
T_{2n} &= \mathrm{tr}[(\widehat{\mathbf{R}}_n - \mathbf{R}_P)^2] - 2\mathrm{tr}(\widehat{\mathbf{R}}_n\tilde{\mathbf{B}}_P) + 2\mathrm{tr}(\mathbf{R}\tilde{\mathbf{B}}_P) + o_p(1) \qquad \text{(S3.27)} \\
&= \mathrm{tr}(\widehat{\mathbf{R}}_n^2) - 2\mathrm{tr}[\widehat{\mathbf{R}}_n(\mathbf{R}_P + \tilde{\mathbf{B}}_P)] + \mathrm{tr}(\mathbf{R}_P^2) + 2\mathrm{tr}(\mathbf{R}\tilde{\mathbf{B}}_P) + o_p(1).
\end{aligned}
$$

By Lemma 1 and the Delta method, we have

$$
\sigma_{2n}^{-1}(T_{2n} - \mu_{2n}) \xrightarrow{d} N(0, 1),
$$

where

$$
\begin{aligned}
\mu_{2n} &= \nu_2(\mathbf{I}_p) - 2\nu_1(\mathbf{R}_P + \tilde{\mathbf{B}}_P) + \mathrm{tr}(\mathbf{R}_P^2) + 2\mathrm{tr}(\mathbf{R}\tilde{\mathbf{B}}_P), \\
\sigma_{2n}^2 &= \sigma_{22}(\mathbf{I}_p) + 4\sigma_{11}(\mathbf{R}_P + \tilde{\mathbf{B}}_P) - 4\sigma_{12}(\mathbf{R}_P + \tilde{\mathbf{B}}_P, \mathbf{I}_p).
\end{aligned}
$$

S4 Proof of Theorem 3

Note that under the null hypothesis H_0, $\mathbf{R}_P$ degenerates to $\mathbf{R}_0$, in this case, $\mathbf{H} = y_n\mathbf{R}_0^{-1}$ and $\mathbf{B}_P = y_n\mathbf{B}$, where $\mathbf{B} = \sum_{k=1}^{K} h_k\mathbf{J}_k$ and $h_k = (\mathrm{tr}(\mathbf{J}_1\mathbf{R}_0^{-1}), \dots, \mathrm{tr}(\mathbf{J}_K\mathbf{R}_0^{-1}))\mathbf{A}^{-1}\mathbf{e}_k$. Based on the conclusion (a) in Theorem 2, we have $\sigma_{10}^{-1}(T_{1n} - \mu_{10}) \xrightarrow{d} N(0,1)$, where

$$
\begin{aligned}
\mu_{10} &= \nu_2(\mathbf{R}_0^{-1}) - 2\nu_1(\mathbf{C}_1) + 2y_n\mathrm{tr}(\mathbf{R}_0\mathbf{B}) + p, \\
\sigma_{10}^2 &= \sigma_{22}(\mathbf{R}_0^{-1}) + 4\sigma_{11}(\mathbf{C}_1) - 4\sigma_{12}(\mathbf{C}_1, \mathbf{R}_0^{-1}),
\end{aligned}
$$

$\mathbf{C}_1 = \mathbf{R}_0^{-1} + y_n\mathbf{B}$ and $\nu_1(\cdot), \nu_2(\cdot), \sigma_{11}(\cdot), \sigma_{12}(\cdot,\cdot), \sigma_{22}(\cdot)$ are defined in Lemma 1. After calculation and simplification, we get

$$
\begin{aligned}
\mu_{10} &= py_n - 3y_n^2 - 7y_n + \beta_w y_n - 0.5n^{-1}\mathrm{tr}(\mathbf{C}_0\mathbf{C}_1) \\
&\quad -(2y_n+4)\beta_w n^{-1}\sum_{k=1}^{p}\sum_{\ell=1}^{p}(\mathbf{e}_k^T\mathbf{\Gamma}\mathbf{e}_\ell)^3\mathbf{e}_k^T\mathbf{R}_0^{-1}\mathbf{\Gamma}\mathbf{e}_\ell \\
&\quad +\Big(1.5y_n+2\Big)n^{-1}\Big[2p + \beta_w\sum_{k=1}^{p}\sum_{\ell=1}^{p}(\mathbf{e}_k^T\mathbf{\Gamma}\mathbf{e}_\ell)^4\Big] \\
&\quad +\Big(0.5y_n+1\Big)n^{-1}\sum_{i=1}^{p}\sum_{j=1}^{p}\mathbf{e}_i^T\mathbf{R}_0\mathbf{e}_j\mathbf{e}_i^T\mathbf{R}_0^{-1}\mathbf{e}_j \\
&\qquad \times\Big[2(\mathbf{e}_i^T\mathbf{R}_0\mathbf{e}_j)^2 + \beta_w\sum_{k=1}^{p}(\mathbf{e}_i^T\mathbf{\Gamma}\mathbf{e}_k)^2(\mathbf{e}_j^T\mathbf{\Gamma}\mathbf{e}_k)^2\Big] \\
&\quad +2n^{-1}\Big[2\mathrm{tr}(\mathbf{R}_0\mathbf{C}_1) + \beta_w\sum_{k=1}^{p}\sum_{\ell=1}^{p}(\mathbf{e}_k^T\mathbf{\Gamma}\mathbf{e}_\ell)^3\mathbf{e}_k^T\mathbf{C}_1\mathbf{\Gamma}\mathbf{e}_\ell\Big] \\
&\quad -1.5n^{-1}\Big[2\mathrm{tr}(\mathbf{R}_0\mathbf{C}_1) + \beta_w\sum_{k=1}^{p}\mathbf{e}_k^T\mathbf{R}_0\mathbf{C}_1\mathbf{e}_k\sum_{\ell=1}^{p}(\mathbf{e}_k^T\mathbf{\Gamma}\mathbf{e}_\ell)^4\Big],
\end{aligned}
$$

and

$$
\begin{aligned}
\sigma_{10}^2 &= 4y_n^2 - 4(2+\beta_w)y_n(1+y_n)^2 \\
&\quad +4(1+y_n)^2n^{-1}\Big[2\mathrm{tr}(\mathbf{R}_0^2) + \beta_w\sum_{k=1}^{p}(\mathbf{e}_k^T\mathbf{\Gamma}^T\mathbf{\Gamma}\mathbf{e}_k)^2\Big]
\end{aligned}
$$

$$+4n^{-1}\Big[2\mathrm{tr}(\mathbf{R}_0\mathbf{C}_1)^2+\beta_w\sum_{k=1}^{p}(\mathbf{e}_k^T\mathbf{\Gamma}^T\mathbf{C}_1\mathbf{\Gamma}\mathbf{e}_k)^2\Big]$$
$$+4n^{-1}\sum_{i=1}^{p}\sum_{j=1}^{p}\mathbf{e}_i^T\mathbf{R}_0\mathbf{C}_1\mathbf{e}_i\mathbf{e}_j^T\mathbf{R}_0\mathbf{C}_1\mathbf{e}_j\Big[2(\mathbf{e}_i^T\mathbf{R}_0\mathbf{e}_j)^2+\beta_w\sum_{k=1}^{p}(\mathbf{e}_i^T\mathbf{\Gamma}\mathbf{e}_k)^2(\mathbf{e}_j^T\mathbf{\Gamma}\mathbf{e}_k)^2\Big]$$
$$-8n^{-1}\sum_{i=1}^{p}\mathbf{e}_i^T\mathbf{R}_0\mathbf{C}_1\mathbf{e}_i\Big[2\mathbf{e}_i^T\mathbf{R}_0\mathbf{C}_1\mathbf{R}_0\mathbf{e}_i+\beta_w\sum_{k=1}^{p}(\mathbf{e}_i^T\mathbf{\Gamma}\mathbf{e}_k)^2\mathbf{e}_k^T\mathbf{\Gamma}^T\mathbf{C}_1\mathbf{\Gamma}\mathbf{e}_k\Big]$$
$$+8(1+y_n)n^{-1}\Big[2\mathrm{tr}(\mathbf{R}_0^2\mathbf{C}_1)+\beta_w\sum_{k=1}^{p}\mathbf{e}_k^T\mathbf{\Gamma}^T\mathbf{\Gamma}\mathbf{e}_k\mathbf{e}_k^T\mathbf{\Gamma}^T\mathbf{C}_1\mathbf{\Gamma}\mathbf{e}_k\Big]$$
$$-8(1+y_n)n^{-1}\sum_{i=1}^{p}\mathbf{e}_i^T\mathbf{R}_0\mathbf{C}_1\mathbf{e}_i\Big[2\mathbf{e}_i^T\mathbf{R}_0^2\mathbf{e}_i+\beta_w\sum_{k=1}^{p}(\mathbf{e}_i^T\mathbf{\Gamma}\mathbf{e}_k)^2\mathbf{e}_k^T\mathbf{\Gamma}^T\mathbf{\Gamma}\mathbf{e}_k\Big],$$

where $\mathbf{C}_0$ is a $p\times p$ matrix with (i,j)th element being $c_{0ij}=2r_{0ij}^3+\beta_w r_{0ij}\sum_{\ell=1}^{p}(\mathbf{e}_i^T\mathbf{\Gamma}\mathbf{e}_\ell)^2(\mathbf{e}_j^T\mathbf{\Gamma}\mathbf{e}_\ell)^2$.

Similarly, under the null hypothesis H_0, we have $\tilde{\mathbf{B}}_P=\mathbf{0}_{p\times p}$. By the conclusion (b) in Theorem 2, we get $\sigma_{20}^{-1}(T_{2n}-\mu_{20})\xrightarrow{d}N(0,1)$, where

$$\begin{aligned}\mu_{20} &= \nu_2(\mathbf{I}_p)-2\nu_1(\mathbf{R}_0)+\mathrm{tr}(\mathbf{R}_0^2),\\ \sigma_{20}^2 &= \sigma_{22}(\mathbf{I}_p)+4\sigma_{11}(\mathbf{R}_0)-4\sigma_{12}(\mathbf{R}_0,\mathbf{I}_p).\end{aligned}$$

After calculation and simplification, we obtain that

$$\begin{aligned}\mu_{20} &= py_n+y_n^2+n^{-1}\mathrm{tr}(\mathbf{R}_0^2)+\beta_w n^{-1}\sum_{k=1}^{p}(\mathbf{e}_k^T\mathbf{\Gamma}^T\mathbf{\Gamma}\mathbf{e}_k)^2-0.5n^{-1}\mathrm{tr}(\mathbf{C}_0\mathbf{R}_0)\\ &\quad -2n^{-1}\Big[2\mathrm{tr}(\mathbf{R}_0^2)+\beta_w\sum_{k=1}^{p}\sum_{\ell=1}^{p}(\mathbf{e}_k^T\mathbf{\Gamma}\mathbf{e}_\ell)^3\mathbf{e}_k^T\mathbf{R}_0\mathbf{\Gamma}\mathbf{e}_\ell\Big]\\ &\quad +0.5n^{-1}\Big[2\mathrm{tr}(\mathbf{R}_0^2)+\beta_w\sum_{k=1}^{p}\mathbf{e}_k^T\mathbf{R}_0^2\mathbf{e}_k\sum_{\ell=1}^{p}(\mathbf{e}_k^T\mathbf{\Gamma}\mathbf{e}_\ell)^4\Big]\\ &\quad +n^{-1}\sum_{i=1}^{p}\sum_{j=1}^{p}(\mathbf{e}_i^T\mathbf{R}_0\mathbf{e}_j)^2\Big[2(\mathbf{e}_i^T\mathbf{R}_0\mathbf{e}_j)^2+\beta_w\sum_{k=1}^{p}(\mathbf{e}_i^T\mathbf{\Gamma}\mathbf{e}_k)^2(\mathbf{e}_j^T\mathbf{\Gamma}\mathbf{e}_k)^2\Big],\end{aligned}$$

and $\sigma_{20}^2=4[n^{-1}\mathrm{tr}(\mathbf{R}_0^2)]^2$.

S5 Proof of Theorem 5

The proof of Theorem 5 is similar to those of Theorem 1 in Cai et al. (2013) and Theorem 2.1 in Yang (2020). Throughout this section, we denote the constants that do not depend on p, n by C and may vary from one expression to the next. Before presenting the proof, we list some technical lemmas that will be used.

S5.1 Technical lemmas

The first lemma is on the large deviations for $\hat{\eta}_{ij}$.

Lemma 2 (Lemma 4.3 of Yang (2020)). *Under assumptions F or F*, there exists some constant $C > 0$ such that for any $M > 0$ and $\epsilon > 0$,*

$$P\left(\max_{1\leq i<j\leq p} |\hat{\eta}_{ij} - \eta_{ij}| \geq C\frac{\varepsilon_n}{\log p}\right) = O(p^{-M} + n^{-\epsilon/8}),$$

where $\varepsilon_n = (\log p)^{3/2}/n^{1/2}$ if assumption F holds and $\varepsilon_n = (\log p)^{-1}$ if assumption F holds.*

The second lemma is an extension of the Bernstein inequality.

Lemma 3 (Lemma 8 of Xiao and Wu (2013)). *Let $X, X_1, \ldots, X_n$ be i.i.d. random variables with mean zero and unit variance. Assume that for some $0 < \alpha < 1$,*

$$\mathrm{E}\left(|X|^{3(1-\alpha)} e^{t|X|^\alpha}\right) \leq A, \quad \textit{for all } \ 0 \leq t < T.$$

Let $S_n = X_1 + \cdots + X_n$. If $x^{1-\alpha} \geq 2A/T^2$, then we have

$$P\left(S_n \geq x\right) \leq \exp\left\{-\frac{x^2}{2(n + x^{2-\alpha}/T)}\right\} + nP(X \geq x).$$

The next lemma is the classical Bonferroni's inequality.

Lemma 4 (Lemma 1 of Cai et al. (2013)). *Let* $B = \bigcup_{t=1}^{p} B_t$. *For any* $k < [p/2]$, *we have*

$$\sum_{t=1}^{2k}(-1)^{t-1}E_t \leq P(B) \leq \sum_{t=1}^{2k-1}(-1)^{t-1}E_t,$$

where $E_t = \sum_{1\leq i_1<\cdots<i_t\leq p} P(B_{i_1} \cap \cdots \cap B_{i_t})$.

Suppose that $(N_i)_{i\in\mathcal{I}_n}$ is a Gaussian random vector whose entries have zero mean and unit variance, where $\mathcal{I}_n$ is an index set with cardinality $|\mathcal{I}_n| = s_n$. Let $\mathbf{\Sigma}_n = (\sigma_{ij})_{i,j\in\mathcal{I}_n}$ be the covariance matrix of $(N_i)_{i\in\mathcal{I}_n}$. Assume that $s_n \to \infty$ as $n \to \infty$.

Consider either of the following two conditions.

(C1) For any sequence $\{b_n\}$ such that $b_n \to \infty$, $\gamma(n, b_n) = o(1/\log b_n)$;

and $\limsup_{n\to\infty} \gamma_n < 1$.

(C2) For any sequence $\{b_n\}$ such that $b_n \to \infty$, $\gamma(n, b_n) = o(1)$;

$\sum_{i\neq j\in\mathcal{I}_n} \sigma_{ij}^2 = O\left(s_n^{2-\delta}\right)$ for some $\delta > 0$; and $\limsup_{n\to\infty} \gamma_n < 1$,

where

$$\gamma(n, b_n) := \sup_{i\in\mathcal{I}_n} \sup_{\mathcal{A}\subset\mathcal{I}_n,|\mathcal{A}|=b_n} \inf_{j\in\mathcal{A}} |\sigma_{ij}| \quad \text{and} \quad \gamma_n := \sup_{i,j\in\mathcal{I}_n;i\neq j} |\sigma_{ij}|.$$

Lemma 5 (Lemma 7 of Xiao and Wu (2013)). *Assume that either (C1) or (C2) holds. For a fixed* $z \in \mathbb{R}$ *and a sequence* $\{z_n\}$ *satisfying* $z_n^2 = 2\log s_n - \log\log s_n - \log\pi + 2z + o(1)$, *denote*

$$A_i' = \{|N_i| > z_n\} \quad \text{and} \quad Q_d' = \sum_{\mathcal{A}\subset\mathcal{I}_n,|\mathcal{A}|=d} P\left(\bigcap_{i\in\mathcal{A}} A_i'\right),$$

then for all $d \geq 1$, *it holds that*

$$\lim_{n\to\infty} Q_d' = \frac{e^{-dz}}{d!}.$$

S5.2 The detailed proof

Without loss of generality, we assume that $\boldsymbol{\mu} = \mathbf{0}$ and $\sigma_{ii} = 1$ for $i = 1, \ldots, p$. We divide the proof of Theorem 5 into two steps:

Step 1: Effects of estimated variances and means. Denote

$$\widetilde{M}_{n,1} = \max_{1 \le i < j \le p} \frac{n(\hat{r}_{ij} - r_{ij})^2}{\eta_{ij}}.$$

Under the event $\{|\hat{\eta}_{ij}/\eta_{ij} - 1| \le C\varepsilon_n/\log p\}$, we have

$$|\widetilde{M}_n - \widetilde{M}_{n,1}| \le C\widetilde{M}_n \frac{\varepsilon_n}{\log p}.$$

Thus, by Lemma 2, it suffices to show that for any $t \in \mathbb{R}$,

$$P\left(\widetilde{M}_{n,1} - 4\log p + \log\log p \le t\right) \to \exp\left(-\frac{1}{\sqrt{8\pi}} \exp\left(-\frac{t}{2}\right)\right).$$

Let $\widetilde{\mathbf{\Sigma}}_n$ be the non-centralized sample covariance matrix, that is, $\widetilde{\mathbf{\Sigma}}_n = (\tilde{\sigma}_{ij})_{i,j=1}^p = n^{-1}\sum\limits_{k=1}^n \mathbf{x}_k\mathbf{x}_k^T$, and $\widetilde{\mathbf{R}}_n = (\tilde{r}_{ij})_{i,j=1}^p$ be the sample correlation matrix corresponding to $\widetilde{\mathbf{\Sigma}}_n$. Denote

$$\widetilde{M}_{n,2} = \max_{1 \le i < j \le p} \frac{n(\tilde{r}_{ij} - r_{ij})^2}{\eta_{ij}}.$$

Using Lemma 3 and the Bernstein inequality, we can show that

$$\max_{1 \le i \le p} |\bar{x}_i| = O_p\left(\sqrt{\log p/n}\right), \tag{S5.28}$$

$$\max_{1 \le i \le j \le p} |\tilde{\sigma}_{ij} - \sigma_{ij}| = O_p\left(\sqrt{\log p/n}\right). \tag{S5.29}$$

From (S5.28), (S5.29), and the first order Taylor expansion of the 3 variate function $x(yz)^{-1/2}$ for $x \in \mathbb{R}$ and $y, z > 0$ (see equation (5) in Cai and Zhang (2016)),

$$\frac{\hat{x}}{(\hat{y}\hat{z})^{1/2}} = \frac{x}{(yz)^{1/2}} + \frac{\hat{x} - x}{(yz)^{1/2}} - \frac{x}{(yz)^{1/2}}\left(\frac{\hat{y} - y}{2y} + \frac{\hat{z} - z}{2z}\right) + o(\hat{x} - x) + o(\hat{y} - y) + o(\hat{z} - z),$$

we get

$$\tilde{r}_{ij} - r_{ij} = \frac{1}{n}\sum_{k=1}^{n}\left[x_{ik}x_{jk} - \frac{r_{ij}}{2}\left(x_{ik}^2 + x_{jk}^2\right)\right] + o_p(n^{-1/2}), \tag{S5.30}$$

$$\hat{r}_{ij} - r_{ij} = \frac{1}{n}\sum_{k=1}^{n}\left[(x_{ik} - \bar{x}_i)(x_{jk} - \bar{x}_j) - \frac{r_{ij}}{2}\left((x_{ik} - \bar{x}_i)^2 + (x_{jk} - \bar{x}_j)^2\right)\right] + o_p(n^{-1/2}).$$

It follows that

$$\hat{r}_{ij} - \tilde{r}_{ij} = \bar{x}_i\bar{x}_j - \frac{r_{ij}}{2}\left(\bar{x}_i^2 + \bar{x}_j^2\right) + o_p(n^{-1/2}).$$

Thus, from (S5.29), we get

$$\left|\widetilde{M}_{n,1} - \widetilde{M}_{n,2}\right| \leq Cn \max_{1\leq i\leq p} |\bar{x}_i|^4 + Cn^{1/2}\widetilde{M}_{n,2}^{1/2} \max_{1\leq i\leq p} |\bar{x}_i|^2 + o_p(1),$$

which together with (S5.28) implies that we only need to prove that for any $t \in \mathbb{R}$,

$$P\left(\widetilde{M}_{n,2} - 4\log p + \log\log p \leq t\right) \to \exp\left(-\frac{1}{\sqrt{8\pi}}\exp\left(-\frac{t}{2}\right)\right).$$

Step 2: Truncation. From (S5.29) and (S5.30), we get $\widetilde{M}_{n,2} = \max_{\alpha\in\mathcal{I}_n} Q_\alpha^2 + o_p(1)$, where

$$Q_\alpha = \frac{1}{\sqrt{n\eta_{ij}}}\sum_{k=1}^{n} V_{k\alpha},$$

and

$$V_{k\alpha} = x_{ik}x_{jk} - \frac{r_{ij}}{2}\left(x_{ik}^2 + x_{jk}^2\right).$$

Let $\hat{V}_{k\alpha} = V_{k\alpha}I\{|V_{k\alpha}| \leq \tau_n\} - \mathrm{E}V_{k\alpha}I\{|V_{k\alpha}| \leq \tau_n\}$, where $\tau_n = \eta^{-1}8\log(p+n)$ if assumption F holds and $\tau_n = \sqrt{n}/(\log p)^8$ if assumption F* holds. Denote

$$\hat{Q}_\alpha = \frac{1}{\sqrt{n\eta_{ij}}}\sum_{k=1}^{n} \hat{V}_{k\alpha}.$$

Note that $|V_{k\alpha}| \leq x_{ik}^2 + x_{jk}^2$ for $k = 1, \ldots, n$. If assumption F holds, then we have

$$\max_{\alpha\in\mathcal{I}_n} \frac{1}{\sqrt{n\eta_{ij}}}\sum_{k=1}^{n} \mathrm{E}|V_{k\alpha}|I\{|V_{k\alpha}| \geq \tau_n\} \leq C\sqrt{n}\max_{\alpha\in\mathcal{I}_n} \mathrm{E}|V_{1\alpha}|I\{|V_{1\alpha}| \geq \tau_n\}$$

$$\leq C\sqrt{n}(p+n)^{-2}\max_{\alpha\in\mathcal{I}_n} \mathrm{E}|V_{1\alpha}|\exp(\eta|V_{1\alpha}|/4) \leq C\sqrt{n}(p+n)^{-2},$$

and if assumption F* holds, then we have

$$\max_{\alpha\in\mathcal{I}_n}\frac{1}{\sqrt{n\eta_{ij}}}\sum_{k=1}^{n}\mathrm{E}|V_{k\alpha}|I\{|V_{k\alpha}|\geq\tau_n\}\leq C\sqrt{n}\max_{\alpha\in\mathcal{I}_n}\mathrm{E}|V_{1\alpha}|I\{|V_{1\alpha}|\geq\tau_n\}$$

$$\leq C\sqrt{n}\max_{\alpha\in\mathcal{I}_n}\mathrm{E}|V_{1\alpha}|^{2+2\gamma_0+\epsilon/2}/\tau_n^{1+2\gamma_0+\epsilon/2}\leq Cn^{-\gamma_0-\epsilon/8}.$$

Thus, we get

$$\begin{aligned}
&P\left(\max_{\alpha\in\mathcal{I}_n}\left|\hat{Q}_\alpha-\hat{Q}_\alpha\right|\geq(\log p)^{-1}\right)\\
\leq&P\left(\max_{\alpha\in\mathcal{I}_n}\frac{1}{\sqrt{n\eta_{ij}}}\sum_{k=1}^{n}|V_{k\alpha}I\{|V_{k\alpha}|\geq\tau_n\}-\mathrm{E}V_{k\alpha}I\{|V_{k\alpha}|\geq\tau_n\}|\geq(\log p)^{-1}\right)\\
\leq&P\left(\max_{\alpha\in\mathcal{I}_n}\max_{1\leq k\leq n}|V_{k\alpha}|\geq\tau_n\right)\leq\sum_{k=1}^{n}P\left(\max_{\alpha\in\mathcal{I}_n}|V_{k\alpha}|\geq\tau_n\right)\\
=&nP\left(\max_{\alpha\in\mathcal{I}_n}|V_{1\alpha}|\geq\tau_n\right)\leq np\max_{1\leq i\leq p}P\left(x_{i1}^2\geq\tau_n/2\right)\\
=&O(p^{-1}+n^{-\epsilon/8}).
\end{aligned}\tag{S5.31}$$

Note that

$$\left|\max_{\alpha\in\mathcal{I}_n}Q_\alpha^2-\max_{\alpha\in\mathcal{I}_n}\hat{Q}_\alpha^2\right|\leq2\max_{\alpha\in\mathcal{I}_n}\left|\hat{Q}_\alpha\right|\max_{\alpha\in\mathcal{I}_n}\left|Q_\alpha-\hat{Q}_\alpha\right|+\max_{\alpha\in\mathcal{I}_n}\left|Q_\alpha-\hat{Q}_\alpha\right|^2.\tag{S5.32}$$

Based on (S5.31) and (S5.32), it suffices to prove that for any $t\in\mathbb{R}$,

$$P\left(\max_{\alpha\in\mathcal{I}_n}\hat{Q}_\alpha^2-4\log p+\log\log p\leq t\right)\to\exp\left(-\frac{1}{\sqrt{8\pi}}\exp\left(-\frac{t}{2}\right)\right).\tag{S5.33}$$

Let q be the cardinality of $\mathcal{I}_n$ and $y_p=t+4\log p-\log\log p$. We arrange the two dimensional indices $\{(i,j):1\leq i<j\leq p\}$ in any ordering and set them as $\{(i_m,j_m):1\leq m\leq q\}$. Denote $V_{km}=x_{i_mk}x_{j_mk}-\frac{r_{i_mj_m}}{2}\left(x_{i_mk}^2+x_{j_mk}^2\right)$, $\eta_m=\eta_{i_mj_m}$, and $Q_m=\frac{1}{\sqrt{n\eta_m}}\sum_{k=1}^{n}V_{km}$. By Lemma 4, for any integer s with $0<s<q/2$, we have

$$\begin{aligned}
&\sum_{d=1}^{2s}(-1)^{d-1}\sum_{1\leq m_1<\cdots<m_d\leq q}P\left(\bigcap_{j=1}^{d}E_{m_j}\right)\leq P\left(\max_{1\leq m\leq q}\hat{Q}_m^2\geq y_p\right)\\
&\leq\sum_{d=1}^{2s-1}(-1)^{d-1}\sum_{1\leq m_1<\cdots<m_d\leq q}P\left(\bigcap_{j=1}^{d}E_{m_j}\right),
\end{aligned}\tag{S5.34}$$

where $E_{m_j} = \{\hat{Q}^2_{m_j} \geq y_p\}$. Let $\tilde{V}_{km} = \hat{V}_{km}/(\eta_m)^{1/2}$ for $1 \leq m \leq q$ and $\mathbf{V}_k = (\tilde{V}_{km_1}, \ldots, \tilde{V}_{km_d})$ for $1 \leq k \leq n$. Denote $|\mathbf{a}|_{\min} = \min_{1\leq i\leq d} |a_i|$ for any vector $\mathbf{a} \in \mathbb{R}^d$. Then we have

$$P\left(\bigcap_{j=1}^{d} E_{m_j}\right) = P\left(\left|n^{-1/2}\sum_{k=1}^{n}\mathbf{V}_k\right|_{\min} \geq y_p^{1/2}\right).$$

By Lemma 9 in Xiao and Wu (2013), we get

$$\begin{aligned} P\left(\left|n^{-1/2}\sum_{k=1}^{n}\mathbf{V}_k\right|_{\min} \geq y_p^{1/2}\right) \leq & P\left(|\mathbf{N}_d|_{\min} \geq y_p^{1/2} - \epsilon_n(\log p)^{-1/2}\right) \\ & + c_{1,d}\exp\left(-\frac{n^{1/2}\epsilon_n}{c_{2,d}\tau_n(\log p)^{1/2}}\right), \end{aligned} \tag{S5.35}$$

where $c_{j,d} = c_j d^{5/2},\ j = 1, 2,\ c_j > 0$ are constants, $\mathbf{N}_d = (N_{m_1}, \ldots, N_{m_d})$ is a d-dimensional Gaussian vector with zero mean and $\mathrm{Cov}(\mathbf{N}_d) = \mathrm{Cov}(\mathbf{V}_1)$. Let

$$\epsilon_n = \begin{cases} (\log p)^{3/2}/n^{3/10}, & \text{if assumption F holds,} \\ (\log p)^{-1/2}, & \text{if assumption F* holds,} \end{cases}$$

then we have $\epsilon_n \to 0$ and for any $M > 0$,

$$c_{1,d}\exp\left(-\frac{n^{1/2}\epsilon_n}{c_{2,d}\tau_n(\log p)^{1/2}}\right) = O(p^{-M}). \tag{S5.36}$$

It follows from (S5.34), (S5.35), and (S5.36) that

$$P\left(\max_{1\leq m\leq q}\hat{Q}^2_m \geq y_p\right) \leq \sum_{d=1}^{2s-1}(-1)^{d-1}\sum_{1\leq m_1<\cdots<m_d\leq q} P\left(|\mathbf{N}_d|_{\min} \geq y_p^{1/2} - \epsilon_n(\log p)^{-1/2}\right) + o(1). \tag{S5.37}$$

Similarly, we also have

$$P\left(\max_{1\leq m\leq q}\hat{Q}^2_m \geq y_p\right) \geq \sum_{d=1}^{2s}(-1)^{d-1}\sum_{1\leq m_1<\cdots<m_d\leq q} P\left(|\mathbf{N}_d|_{\min} \geq y_p^{1/2} + \epsilon_n(\log p)^{-1/2}\right) - o(1). \tag{S5.38}$$

For $0 < k, \ell < q$, denote $C_{k\ell} = 1/(\eta_{m_k}\eta_{m_\ell})^{1/2}$. By some elementary calculations, we get if assumptions F holds,

$$\max_{k,\ell}|\mathrm{Cov}(N_{m_k}, N_{m_\ell}) - C_{k\ell}\mathrm{Cov}(V_{1m_k}, V_{1m_\ell})| \leq C(p+n)^{-1}, \tag{S5.39}$$

and if assumption F* holds,

$$\max_{k,\ell} |\mathrm{Cov}(N_{m_k}, N_{m_\ell}) - C_{k\ell}\mathrm{Cov}(V_{1m_k}, V_{1m_\ell})| \le C\tau_n^{-2\gamma_0-\epsilon/2}. \tag{S5.40}$$

Based on assumption E (or E*), (S5.39), and (S5.40), we know that the covariance matrix of $(N_1, \cdots, N_q)$ satisfies assumption E (or E*). Thus, by Lemma 5, we get

$$\lim_{n\to\infty} \sum_{1\le m_1<\cdots<m_d\le q} P\left(|\mathbf{N}_d|_{\min} \ge y_p^{1/2} \pm \epsilon_n(\log p)^{-1/2}\right) = \frac{1}{d!}\left(\frac{1}{\sqrt{8\pi}}\exp\left(-\frac{t}{2}\right)\right)^d. \tag{S5.41}$$

Submitting (S5.41) into (S5.37) and (S5.38), we get

$$\limsup_{n\to\infty} P\left(\max_{1\le m\le q} \hat{Q}_m^2 \ge y_p\right) \le \sum_{d=1}^{2s-1}(-1)^{d-1}\frac{1}{d!}\left(\frac{1}{\sqrt{8\pi}}\exp\left(-\frac{t}{2}\right)\right)^d$$

and

$$\liminf_{n\to\infty} P\left(\max_{1\le m\le q} \hat{Q}_m^2 \ge y_p\right) \ge \sum_{d=1}^{2s}(-1)^{d-1}\frac{1}{d!}\left(\frac{1}{\sqrt{8\pi}}\exp\left(-\frac{t}{2}\right)\right)^d$$

for any positive integer s. Letting $s \to \infty$, we prove that (S5.33) holds.

S6 Proof of Corollary 3

Note that $\mathbf{R}_0 = (r_{0ij})_{i,j=1}^p$ represents the structured population correlation matrix under the null hypothesis H_0. Let

$$M_{n,1} = \max_{1\le i<j\le p} \frac{n(\hat{r}_{ij} - r_{0ij})^2}{\hat{\eta}_{ij}}.$$

By Theorem 5, we get for any $t \in \mathbb{R}$,

$$P\left(M_{n,1} - 4\log p + \log\log p \le t\right) \to \exp\left(-\frac{1}{\sqrt{8\pi}}\exp\left(-\frac{t}{2}\right)\right).$$

Under the null hypothesis H_0, we have

$$\eta_{ij} = \mathrm{Var}\left(\frac{(x_{i1}-\mu_i)(x_{j1}-\mu_j)}{(\sigma_{ii}\sigma_{jj})^{1/2}} - \frac{r_{0ij}}{2}\left(\frac{(x_{i1}-\mu_i)^2}{\sigma_{ii}} + \frac{(x_{j1}-\mu_j)^2}{\sigma_{jj}}\right)\right),$$

and

$$n^{-1}\sum_{k=1}^{n}\frac{(x_{ik}-\bar{x}_i)(x_{jk}-\bar{x}_j)}{(\hat{\sigma}_{ii}\hat{\sigma}_{jj})^{1/2}}-\frac{\hat{r}_{0ij}}{2}\left(\frac{(\hat{x}_{ik}-\bar{x}_i)^2}{\hat{\sigma}_{ii}}+\frac{(x_{jk}-\bar{x}_j)^2}{\hat{\sigma}_{jj}}\right)=\hat{r}_{ij}-\hat{r}_{0ij},$$

then η_{ij} can be estimated by

$$\hat{\eta}_{0ij}=n^{-1}\sum_{k=1}^{n}\left[\frac{(x_{ik}-\bar{x}_i)(x_{jk}-\bar{x}_j)}{(\hat{\sigma}_{ii}\hat{\sigma}_{jj})^{1/2}}-\frac{\hat{r}_{0ij}}{2}\left(\frac{(x_{ik}-\bar{x}_i)^2}{\hat{\sigma}_{ii}}+\frac{(x_{jk}-\bar{x}_j)^2}{\hat{\sigma}_{jj}}\right)-(\hat{r}_{ij}-\hat{r}_{0ij})\right]^2.$$

Next, we prove that $\max_{1\le i<j\le p}|\hat{\eta}_{0ij}-\hat{\eta}_{ij}|=o_p(1)$. After some elementary calculations, we get

$$\hat{\eta}_{0ij}=n^{-1}\sum_{k=1}^{n}\left[T_{kij}+(\hat{r}_{ij}-\hat{r}_{0ij})R_{kij}\right]^2,$$

where

$$\begin{aligned}T_{kij}&=\frac{(x_{ik}-\bar{x}_i)(x_{jk}-\bar{x}_j)}{(\hat{\sigma}_{ii}\hat{\sigma}_{jj})^{1/2}}-\frac{\hat{r}_{ij}}{2}\left(\frac{(x_{ik}-\bar{x}_i)^2}{\hat{\sigma}_{ii}}+\frac{(x_{jk}-\bar{x}_j)^2}{\hat{\sigma}_{jj}}\right),\\ R_{kij}&=\frac{(x_{ik}-\bar{x}_i)^2}{2\hat{\sigma}_{ii}}+\frac{(x_{jk}-\bar{x}_j)^2}{2\hat{\sigma}_{jj}}-1.\end{aligned}$$

Thus, we have

$$\hat{\eta}_{0ij}-\hat{\eta}_{ij}=2(\hat{r}_{ij}-\hat{r}_{0ij})n^{-1}\sum_{k=1}^{n}T_{kij}R_{kij}+(\hat{r}_{ij}-\hat{r}_{0ij})^2n^{-1}\sum_{k=1}^{n}R_{kij}^2. \tag{S6.42}$$

Note that $n^{-1}\sum_{k=1}^{n}R_{kij}^2\le\max_{1\le i\le p}\hat{\vartheta}_{ii}/\hat{\sigma}_{ii}^2$, where $\hat{\vartheta}_{ii}=n^{-1}\sum_{k=1}^{n}\left((x_{ik}-\bar{x}_i)^2-\hat{\sigma}_{ii}\right)^2$. By Lemma 3 of Cai et al. (2013), we have

$$P\left(\max_i|\hat{\vartheta}_{ii}-\vartheta_{ii}|/\sigma_{ii}^2\ge C\frac{\varepsilon_n}{\log p}\right)=O(p^{-1}+n^{-\epsilon/8}),$$

where $\varepsilon_n=\max\left((\log p)^{1/6}/n^{1/2},(\log p)^{-1}\right)$. Therefore, we get

$$\max_{1\le i<j\le p}n^{-1}\sum_{k=1}^{n}R_{kij}^2=O_p(1). \tag{S6.43}$$

Due to the conclusion that $\hat{\theta}_k=\theta_k+O_p(n^{-1})$ in the proof of Theorem 2, we have

$$\max_{1\le i<j\le p}|\hat{r}_{0ij}-r_{0ij}|\le\|\widehat{\mathbf{R}}_0-\mathbf{R}_0\|=o_p(n^{-1/2}). \tag{S6.44}$$

From (S5.28) and (S5.29), we get $\max_{1\le i<j\le p} |\hat{r}_{ij} - r_{0ij}| = o_p(1)$. It follows that

$$\max_{1\le i<j\le p} |\hat{r}_{ij} - \hat{r}_{0ij}| = o_p(1). \tag{S6.45}$$

Denote

$$M_{n,2} = \max_{1\le i<j\le p} \frac{n(\hat{r}_{ij} - r_{0ij})^2}{\hat{\eta}_{0ij}}.$$

Combining (S6.42), (S6.43), and (S6.45), we have $\max_{1\le i<j\le p} |\hat{\eta}_{0ij} - \hat{\eta}_{ij}| = o_p(1)$. It follows that

$$|M_{n,1} - M_{n,2}| \le M_{n,1} \max_{1\le i<j\le p} |\hat{\eta}_{ij}/\hat{\eta}_{0ij} - 1| = o_p(1).$$

From (S6.44), we get

$$|M_{n,2} - M_n| \le Cn \max_{1\le i<j\le p} |\hat{r}_{0ij} - r_{0ij}|^2 + Cn^{1/2} M_{n,2}^{1/2} \max_{1\le i<j\le p} |\hat{r}_{0ij} - r_{0ij}| = o_p(1).$$

Based on the Slutsky's theorem, we derive that M_n converges to the Type I extreme value distribution under the null hypothesis H_0. Thus, the proof of Corollary 3 is completed.

Table S1: Key notations used in the paper

Notation	Definition
p	the dimension
n	the sample size
y_n	the dimension-versus-sample-size ratio, defined as p/n
$\mathbf{x}_1, \cdots, \mathbf{x}_n$	the *i.i.d.* sample
$\boldsymbol{\Sigma} = (\sigma_{ij})_{i,j=1}^{p}$	the population covariance matrix
$\mathbf{R} = (r_{ij})_{i,j=1}^{p}$	the population correlation matrix
$\mathrm{diag}(\boldsymbol{\Sigma})$	the diagonal matrix formed by the diagonal elements of $\boldsymbol{\Sigma}$
$\boldsymbol{\Gamma}$	a $p \times p$ matrix, defined as $[\mathrm{diag}(\boldsymbol{\Sigma})]^{-1/2}\boldsymbol{\Sigma}^{1/2}$
$\widehat{\boldsymbol{\Sigma}}_n = (\hat{\sigma}_{ij})_{i,j=1}^{p}$	the sample covariance matrix
$\widehat{\mathbf{R}}_n = (\hat{r}_{ij})_{i,j=1}^{p}$	the sample correlation matrix
β_w	the kurtosis, defined in Assumption A
$\mathbf{J}_0, \mathbf{J}_1, \cdots, \mathbf{J}_K$	the basis matrices involved in H_0
$\mathbf{R}_0 = (r_{0ij})_{i,j=1}^{p}$	the population correlation matrix specified under H_0
$\widehat{\mathbf{R}}_0 = (\hat{r}_{0ij})_{i,j=1}^{p}$	the structured estimator of $\mathbf{R}$ under H_0
$\mathbf{A}$	a $K \times K$ matrix with the (i,j)th entry being $\mathrm{tr}(\mathbf{J}_i\mathbf{J}_j)$
$\mathbf{I}_p$	the $p \times p$ identity matrix
$\mathbf{e}_i$	the ith column of the identity matrix, and its dimension is determined by the matrix it multiplies
$\mathbf{C}_0 = (c_{0ij})_{i,j=1}^{p}$	a matrix with $c_{0ij} = 2r_{0ij}^3 + \beta_w r_{0ij} \sum_{k=1}^{p} (\mathbf{e}_i^T\boldsymbol{\Gamma}\mathbf{e}_k)^2(\mathbf{e}_j^T\boldsymbol{\Gamma}\mathbf{e}_k)^2$
h_k	a quantity, defined as $(\mathrm{tr}(\mathbf{J}_1\mathbf{R}_0^{-1}), \ldots, \mathrm{tr}(\mathbf{J}_K\mathbf{R}_0^{-1}))\mathbf{A}^{-1}\mathbf{e}_k$
$\mathbf{B}$	a $p \times p$ matrix, defined as $\sum_{k=1}^{K} h_k\mathbf{J}_k$
$\mathbf{C}_1$	a $p \times p$ matrix, defined as $\mathbf{R}_0^{-1} + y_n\mathbf{B}$

Table S2: Empirical size and power for the test statistics based on infinite norm (M_n), ratio-based quadratic norm (T_{1n}) and distance-based quadratic norm (T_{2n}) under scenarios 1–4, where n observations with dimension p are generated from the Gamma population.

		Scenario 1			Scenario 2			Scenario 3			Scenario 4		
n	p	M_n	T_{1n}	T_{2n}	M_n	T_{1n}	T_{2n}	M_n	T_{1n}	T_{2n}	M_n	T_{1n}	T_{2n}
		Empirical size (%)											
	50	4.1	5.1	3.8	3.6	4.9	3.7	3.9	4.5	4.3	3.7	4.6	4.7
	100	3.4	5.4	4.4	3.3	5.1	4.5	3.1	5.0	4.7	3.1	4.6	4.8
100	300	3.2	5.7	4.6	3.1	5.1	4.6	2.6	5.3	4.8	2.2	4.6	4.6
	500	2.6	6.0	5.1	2.7	5.3	4.7	2.9	5.4	5.1	2.7	4.9	4.8
	1000	2.6	6.3	4.6	2.9	5.9	5.0	2.5	5.8	4.5	2.4	4.5	4.7
	50	3.6	5.5	4.2	3.7	5.4	3.8	3.8	5.2	4.4	3.7	4.9	4.8
	100	3.5	5.3	4.7	3.5	5.2	4.3	3.3	5.2	4.6	3.6	4.8	5.0
300	300	3.3	5.2	4.7	3.3	5.2	4.6	3.3	4.9	4.5	3.2	4.7	4.5
	500	3.1	5.3	4.4	2.8	4.8	4.7	2.9	5.0	4.4	3.1	4.7	4.6
	1000	3.0	5.7	5.1	2.7	5.7	5.0	2.7	5.2	4.9	2.4	5.2	5.2
		Empirical power (%)											
	50	5.6	100.0	11.6	6.8	19.3	54.4	73.6	4.9	9.1	5.0	22.4	14.1
	100	4.5	100.0	11.5	7.0	62.6	89.3	58.7	4.8	6.8	4.5	28.9	15.0
100	300	3.2	100.0	11.5	7.4	99.9	100.0	35.0	5.9	4.9	2.7	57.2	15.3
	500	3.1	100.0	11.9	7.8	100.0	100.0	26.7	5.9	5.3	3.0	80.3	16.4
	1000	2.7	100.0	11.8	8.3	100.0	100.0	17.4	5.9	4.9	2.5	99.4	16.3
	50	11.4	100.0	30.5	10.2	8.0	90.6	99.7	15.3	21.9	9.9	57.8	42.1
	100	8.4	100.0	31.1	11.2	35.1	99.8	99.3	6.2	11.9	7.4	67.0	43.5
300	300	5.2	100.0	32.0	12.7	98.4	100.0	96.6	5.0	5.5	5.0	88.7	46.5
	500	4.7	100.0	31.5	12.6	100.0	100.0	93.9	5.3	5.1	4.3	97.1	46.5
	1000	3.9	100.0	32.5	12.8	100.0	100.0	89.0	5.9	5.1	3.2	100.0	47.9

Table S3: Empirical size and power for the test statistics based on infinite norm (M_n), ratio-based quadratic norm (T_{1n}) and distance-based quadratic norm (T_{2n}) under scenarios 5–8, where n observations with dimension p are generated from the Gamma population.

		Scenario 5			Scenario 6			Scenario 7			Scenario 8		
n	p	M_n	T_{1n}	T_{2n}	M_n	T_{1n}	T_{2n}	M_n	T_{1n}	T_{2n}	M_n	T_{1n}	T_{2n}
		Empirical size (%)											
100	50	3.7	5.1	3.9	3.8	4.3	4.3	3.3	5.2	4.0	2.9	6.3	2.3
	100	3.2	5.0	4.6	3.2	4.5	4.4	3.6	5.3	4.4	2.9	5.6	3.1
	300	2.7	4.9	4.6	2.6	4.8	4.8	2.9	4.7	4.4	2.7	5.0	3.4
	500	2.9	5.3	4.9	2.9	4.8	4.9	2.9	5.3	5.1	2.7	5.3	3.3
	1000	2.3	4.9	4.7	2.6	4.5	4.6	2.8	4.9	5.2	2.7	5.1	2.7
300	50	3.5	5.3	4.0	3.7	4.9	4.9	3.6	5.5	4.0	2.8	6.1	2.6
	100	3.4	5.0	4.4	3.7	4.8	4.8	3.2	4.7	3.9	3.0	5.5	3.3
	300	3.2	4.8	4.4	3.2	5.0	4.9	3.0	5.1	4.7	2.9	5.1	3.7
	500	2.9	5.1	4.7	3.0	4.9	4.7	2.7	5.0	4.8	2.6	5.1	3.6
	1000	2.5	5.1	4.9	2.6	4.9	4.8	2.6	4.9	5.3	2.4	5.0	3.0
		Empirical power (%)											
100	50	87.1	35.1	50.0	57.1	46.0	40.5	3.7	26.6	4.4	45.3	26.4	2.9
	100	75.3	46.7	75.8	56.6	42.8	36.3	3.6	44.9	4.7	34.2	23.4	3.2
	300	50.9	95.5	99.8	54.3	44.8	33.2	3.0	95.2	4.6	26.9	40.6	3.1
	500	39.8	99.8	100.0	52.7	48.0	32.2	3.0	100.0	5.4	25.7	60.6	3.0
	1000	27.1	100.0	100.0	48.8	59.7	32.0	2.7	100.0	5.5	24.8	91.2	2.1
300	50	99.9	83.9	94.1	98.4	94.5	91.8	4.0	45.9	4.9	97.5	61.2	5.6
	100	99.8	91.8	99.4	98.5	93.1	88.8	3.5	68.9	4.9	95.4	51.2	3.6
	300	99.4	100.0	100.0	98.0	92.4	85.4	3.1	99.3	5.9	93.1	65.3	2.7
	500	98.7	100.0	100.0	97.5	93.4	84.0	2.9	100.0	5.8	92.8	83.0	2.6
	1000	97.4	100.0	100.0	97.5	96.0	83.7	2.7	100.0	6.2	93.2	98.6	2.1

Table S4: Empirical size and power for the Tippett's minimum p-value test T_{tn} and the Cauchy combination test T_{cn} under scenarios 1–8, where n observations with dimension p are generated from the Gamma population.

Scenario		1		2		3		4		5		6		7		8	
n	p	T_{tn}	T_{cn}	T_{tn}	T_{cn}	T_{tn}	T_{cn}	T_{tn}	T_{cn}	T_{tn}	T_{cn}	T_{tn}	T_{cn}	T_{tn}	T_{cn}	T_{tn}	T_{cn}
		Empirical size (%)															
	50	4.0	4.3	3.4	3.9	3.5	4.3	3.1	4.3	3.8	4.2	2.6	4.0	3.7	4.1	3.8	3.8
	100	4.1	4.3	3.9	4.1	3.8	4.3	2.7	4.1	3.8	4.3	2.4	3.7	4.0	4.4	3.4	3.6
100	300	4.3	4.4	4.3	4.3	3.7	4.1	2.6	3.6	3.8	4.1	2.5	3.8	3.7	3.8	3.8	3.6
	500	4.4	4.3	4.1	4.1	4.0	4.3	3.1	4.1	3.8	4.0	2.9	4.0	3.9	4.1	3.7	3.7
	1000	4.5	4.5	4.6	4.5	4.4	4.2	2.6	3.6	3.9	3.9	2.5	3.8	4.1	4.1	3.1	3.0
	50	4.0	4.4	3.8	4.1	3.7	4.3	3.2	4.3	3.7	4.2	3.0	4.2	3.8	4.2	3.8	3.9
	100	4.0	4.6	4.3	4.5	3.8	4.4	3.2	4.3	3.8	4.2	2.9	4.3	3.3	3.7	3.7	3.7
300	300	3.9	4.1	4.0	4.0	3.6	4.0	2.9	4.0	3.6	4.0	2.7	4.0	3.8	4.2	3.7	3.4
	500	3.9	4.1	4.0	4.1	3.7	4.0	3.0	3.9	3.6	3.8	2.5	3.9	3.5	3.8	3.4	3.2
	1000	4.4	4.5	4.5	4.2	4.0	4.1	2.9	3.9	3.9	4.1	2.6	4.0	4.0	4.2	3.3	3.2
		Empirical power (%)															
	50	99.9	99.9	51.7	53.7	61.4	61.7	14.2	17.2	84.4	86.8	59.9	66.2	18.3	18.7	41.2	42.4
	100	100.0	100.0	94.3	95.0	46.7	47.0	18.0	20.5	86.4	88.4	59.0	64.9	33.6	33.2	33.2	34.2
100	300	100.0	100.0	100.0	100.0	26.6	26.9	40.9	41.4	99.7	99.8	59.0	64.0	90.7	90.2	39.9	40.7
	500	100.0	100.0	100.0	100.0	20.4	21.0	67.6	66.4	100.0	100.0	60.1	65.1	99.8	99.8	55.5	56.3
	1000	100.0	100.0	100.0	100.0	13.9	14.0	98.7	98.4	100.0	100.0	65.5	68.6	100.0	100.0	86.4	86.7
	50	100.0	100.0	87.7	88.4	99.4	99.4	44.4	49.7	100.0	100.0	99.6	99.8	34.3	34.0	96.8	97.1
	100	100.0	100.0	99.9	100.0	98.2	98.2	52.2	56.7	100.0	100.0	99.6	99.8	55.5	55.1	94.0	94.4
300	300	100.0	100.0	100.0	100.0	93.8	93.6	79.4	80.0	100.0	100.0	99.5	99.6	98.4	98.3	94.5	95.3
	500	100.0	100.0	100.0	100.0	90.2	90.1	93.3	92.9	100.0	100.0	99.5	99.7	100.0	100.0	96.7	97.2
	1000	100.0	100.0	100.0	100.0	83.7	83.6	99.9	99.9	100.0	100.0	99.6	99.8	100.0	100.0	99.7	99.8

References

Bai, Z. D. and J. W. Silverstein (2004). CLT for linear spectral statistics of large-dimensional sample covariance matrices. *The Annals of Probability 32*(1A), 553–605.

Bai, Z. D. and J. W. Silverstein (2010). *Spectral Analysis of Large Dimensional Random Matrices*. 2nd Edition. Science Press.

Cai, T., W. Liu, and Y. Xia (2013). Two-sample covariance matrix testing and support recovery in high-dimensional and sparse settings. *Journal of the American Statistical Association 108*(501), 265–277.

Cai, T. T. and A. Zhang (2016). Inference for high-dimensional differential correlation matrices. *Journal of Multivariate Analysis 143*, 107–126.

Noureddine, E. K. (2009). Concentration of measure and spectra of random matrices: with applications to correlation matrices, elliptical distributions and beyond. *The Annals of Applied Probability 19*(6), 2362–2405.

Xiao, H. and W. Wu (2013). Asymptotic theory for maximum deviations of sample covariance matrix estimates. *Stochastic Processes and their Applications 123*, 2899–2920.

Yang, X. (2020). Test on the equality of two high-dimensional correlation matrices with sparse settings. Master Thesis, Southern University of Science and Technology.

Yin, Y. and Y. Ma (2022). Properties of eigenvalues and eigenvectors of large-dimensional sample correlation matrices. *The Annals of Applied Probability 32*(6), 4763–4802.

Yin, Y., S. Zheng, and T. Zou (2023). Supplement on "central limit theorem of linear spectral statistics of high-dimensional sample correlation matrices".