TESTING HOMOGENEITY OF HIGH-DIMENSIONAL COVARIANCE MATRICES

[1] *Northeast Normal University,* [2] *The University of Texas MD Anderson Cancer Center and* [3] *The University of Hong Kong*

Supplementary Material

This Supplementary Material contains the proofs of the theorems and simulation results. In particular, Tables 1–4 summarize the simulation results for two samples, and Table 5 provides the simulation results for two samples under the ultra high-dimensional setting. Figure 1 presents the simulation results for three samples with Gaussian populations, and Figure 2 presents the simulation results for three samples with Gamma populations. For ease of exposition, the proof of Theorem 2 is given first and that of Theorem 1 is given later. We first provide two lemmas which are essential to the proofs of Theorems 1 and 2.

S1 Lemmas

Lemma 1. *Under Assumptions (A1)–(A2), we have*

$$\sigma_{A12}^{-1}\{\mathrm{tr}[(\mathbf{S}_1 - \mathbf{S}_2)^2] - \hat{\mu}_{21} - \mu_{A12}\} \to N(0,1)$$

where $\hat{\mu}_{21} = \sum_{k=1}^{2}(n_k^2 - n_k - 1)n_k^{-1}(n_k - 1)^{-2}(\mathrm{tr}\mathbf{S}_k)^2$,

$$\mu_{A12} = \mathrm{tr}[(\mathbf{\Sigma}_1 - \mathbf{\Sigma}_2)^2] + \sum_{k=1}^{2} \frac{n_k + 1}{(n_k - 1)^2}\mathrm{tr}(\mathbf{\Sigma}_k^2) + \sum_{k=1}^{2} \frac{\beta_k n_k}{(n_k - 1)^2} \sum_{\ell=1}^{p}(\mathbf{e}_\ell^T \mathbf{\Sigma}_k \mathbf{e}_\ell)^2$$

and

$$\begin{aligned}
\sigma_{A12}^2 \;=\;& 4[n_1^{-1}\mathrm{tr}(\mathbf{\Sigma}_1^2)]^2 + 4[n_2^{-1}\mathrm{tr}(\mathbf{\Sigma}_2^2)]^2 + 8n_1^{-1}n_2^{-1}[\mathrm{tr}(\mathbf{\Sigma}_1\mathbf{\Sigma}_2)]^2 \\
&+4[2n_1^{-1}\mathrm{tr}(\mathbf{\Sigma}_1^4) + \beta_1 n_1^{-1}\sum_{\ell=1}^{p}(\mathbf{e}_\ell^T \mathbf{\Sigma}_1^2 \mathbf{e}_\ell)^2] \\
&+4\{2n_1^{-1}\mathrm{tr}[(\mathbf{\Sigma}_1\mathbf{\Sigma}_2)^2] + \beta_1 n_1^{-1}\sum_{\ell=1}^{p}(\mathbf{e}_\ell^T \mathbf{\Sigma}_1^{1/2}\mathbf{\Sigma}_2\mathbf{\Sigma}_1^{1/2}\mathbf{e}_\ell)^2\} \\
&-8[2n_1^{-1}\mathrm{tr}(\mathbf{\Sigma}_1^3\mathbf{\Sigma}_2) + \beta_1 n_1^{-1}\sum_{\ell=1}^{p}\mathbf{e}_\ell^T \mathbf{\Sigma}_1^{1/2}\mathbf{\Sigma}_2\mathbf{\Sigma}_1^{1/2}\mathbf{e}_\ell\mathbf{e}_\ell^T\mathbf{\Sigma}_1^2\mathbf{e}_\ell] \\
&+4n_2^{-1}[2\mathrm{tr}(\mathbf{\Sigma}_2^4) + \beta_2\sum_{\ell=1}^{p}(\mathbf{e}_\ell^T \mathbf{\Sigma}_2^2 \mathbf{e}_\ell)^2] \\
&+4\{2n_2^{-1}\mathrm{tr}[(\mathbf{\Sigma}_2\mathbf{\Sigma}_1)^2] + \beta_2 n_2^{-1}\sum_{\ell=1}^{p}(\mathbf{e}_\ell^T \mathbf{\Sigma}_2^{1/2}\mathbf{\Sigma}_1\mathbf{\Sigma}_2^{1/2}\mathbf{e}_\ell)^2\} \\
&-8[2n_2^{-1}\mathrm{tr}(\mathbf{\Sigma}_2^3\mathbf{\Sigma}_1) + \beta_2 n_2^{-1}\sum_{\ell=1}^{p}\mathbf{e}_\ell^T \mathbf{\Sigma}_2^{1/2}\mathbf{\Sigma}_1\mathbf{\Sigma}_2^{1/2}\mathbf{e}_\ell\mathbf{e}_\ell^T\mathbf{\Sigma}_2^2\mathbf{e}_\ell].
\end{aligned}$$

Proof. Define $\mathbf{r}_{ki} = n_k^{-1/2}\mathbf{w}_{ki}$ and $\mathbf{w}_{ki} = (w_{k1i}, \ldots, w_{kpi})^T$ for $i = 1, \ldots, n_k$ and $k = 1, 2$. We have

$$(n_k-1)^2 n_k^{-2}\mathrm{tr}(\mathbf{S}_k^2) = \mathrm{tr}[(\sum_{i=1}^{n_k}\mathbf{\Sigma}_k^{1/2}\mathbf{r}_{ki}\mathbf{r}_{ki}^T\mathbf{\Sigma}_k^{1/2})^2] + n_k^2(\bar{\mathbf{r}}_k^T\mathbf{\Sigma}_k\bar{\mathbf{r}}_k)^2 - 2n_k\bar{\mathbf{r}}_k^T\mathbf{\Sigma}_k\sum_{i=1}^{n}\mathbf{r}_{ki}\mathbf{r}_{ki}^T\mathbf{\Sigma}_k\bar{\mathbf{r}}_k,$$

where $\bar{\mathbf{r}}_k = n_k^{-1}\sum_{i=1}^{n_k}\mathbf{r}_{ki}$ for $k = 1, \ldots, K$.

First, we will consider the bounded spectral norm case, that is, the maximum eigenvalue of $\mathbf{\Sigma}_k$ is bounded for $k = 1, 2$. When $\mathrm{tr}(\mathbf{\Sigma}_k^q) = O(p^q)$

holds for $q = 1, 2, 3, 4$ and at least one integer k in the index set $\{1, 2\}$, we only need to consider the CLT of $n^{-2/3}\mathrm{tr}[(\mathbf{S}_1 - \mathbf{S}_2)^2]$, and its proof is almost the same as the proof for the CLT of $\mathrm{tr}[(\mathbf{S}_1 - \mathbf{S}_2)^2]$ when $\mathbf{\Sigma}_1$ and $\mathbf{\Sigma}_2$ have the bounded spectral norms.

Specifically, for the bounded spectral norm case, the proof can be completed through the following steps.

Step 1. We will prove $n_k \bar{\mathbf{r}}_k^T \mathbf{\Sigma}_k \bar{\mathbf{r}}_k = n_k^{-1}\mathrm{tr}\mathbf{\Sigma}_k + o_p(1)$. We have $n_k \bar{\mathbf{r}}_k^T \mathbf{\Sigma}_k \bar{\mathbf{r}}_k = 2n_k^{-1}\sum_{i<j} \mathbf{r}_{ki}^T \mathbf{\Sigma}_k \mathbf{r}_{kj} + n_k^{-1}\sum_{i=1}^{n_k} \mathbf{r}_{ki}^T \mathbf{\Sigma}_k \mathbf{r}_{ki}$. As $\mathrm{E}(2n_k^{-1}\sum_{i<j} \mathbf{r}_{ki}^T \mathbf{\Sigma}_k \mathbf{r}_{kj}) = 0$ and

$$\mathrm{E}(2n_k^{-1}\sum_{i<j} \mathbf{r}_{ki}^T \mathbf{\Sigma}_k \mathbf{r}_{kj})^2 \leq 4\mathrm{E}(\mathbf{r}_{k1}^T \mathbf{\Sigma}_k \mathbf{r}_{k2} \mathbf{r}_{k2}^T \mathbf{\Sigma}_k \mathbf{r}_{k1}) = 4n_k^{-2}\mathrm{tr}(\mathbf{\Sigma}_k^2) \to 0,$$

we have $2n_k^{-1}\sum_{i<j} \mathbf{r}_{ki}^T \mathbf{\Sigma}_k \mathbf{r}_{kj} = o_p(1)$. Moreover, as $\mathrm{E}(n_k^{-1}\sum_{i=1}^{n_k} \mathbf{r}_{ki}^T \mathbf{\Sigma}_k \mathbf{r}_{ki}) = n_k^{-1}\mathrm{tr}\mathbf{\Sigma}_k$ and $\mathrm{Var}(n_k^{-1}\sum_{i=1}^{n_k} \mathbf{r}_{ki}^T \mathbf{\Sigma}_k \mathbf{r}_{ki}) = n_k^{-1}\mathrm{E}[(\mathbf{r}_{k1}^T \mathbf{\Sigma}_k \mathbf{r}_{k1} - n_k^{-1}\mathrm{tr}\mathbf{\Sigma}_k)^2] = n_k^{-2}[2\mathrm{tr}(\mathbf{\Sigma}_k^2) + \beta_k \sum_{j=1}^{p} (\mathbf{e}_j^T \mathbf{\Sigma}_k \mathbf{e}_j)^2] \to 0$ from (1.15) of Bai and Silverstein (2004), we have $n_k^{-1}\sum_{i=1}^{n_k} \mathbf{r}_{ki}^T \mathbf{\Sigma}_k \mathbf{r}_{ki} - n_k^{-1}\mathrm{tr}\mathbf{\Sigma}_k = o_p(1)$, and thus

$$n_k \bar{\mathbf{r}}_k^T \mathbf{\Sigma}_k \bar{\mathbf{r}}_k = n_k^{-1}\mathrm{tr}\mathbf{\Sigma}_k + o_p(1).$$

Step 2. We have

$$\begin{aligned} n_k \bar{\mathbf{r}}_k^T \mathbf{\Sigma}_k \sum_{i=1}^{n_k} \mathbf{r}_{ki} \mathbf{r}_{ki}^T \mathbf{\Sigma}_k \bar{\mathbf{r}}_k \quad &= \quad n_k^{-1} \sum_{i \neq j \neq \ell} \mathbf{r}_{ki}^T \mathbf{\Sigma}_k \mathbf{r}_{kj} \mathbf{r}_{kj}^T \mathbf{\Sigma}_k \mathbf{r}_{k\ell} + n_k^{-1} \sum_{i \neq j} \mathbf{r}_{ki}^T \mathbf{\Sigma}_k \mathbf{r}_{kj} \mathbf{r}_{kj}^T \mathbf{\Sigma}_k \mathbf{r}_{ki} \\ &\quad + 2n_k^{-1} \sum_{i \neq j} \mathbf{r}_{ki}^T \mathbf{\Sigma}_k \mathbf{r}_{ki} \mathbf{r}_{ki}^T \mathbf{\Sigma}_k \mathbf{r}_{kj} + n_k^{-1} \sum_{i=1}^{n_k} \mathbf{r}_{ki}^T \mathbf{\Sigma}_k \mathbf{r}_{ki} \mathbf{r}_{ki}^T \mathbf{\Sigma}_k \mathbf{r}_{ki}. \end{aligned}$$

Step 2.1. We have

$$\begin{aligned} n_k^{-1}\sum_{i=1}^{n_k}\mathbf{r}_{ki}^T\mathbf{\Sigma}_k\mathbf{r}_{ki}\mathbf{r}_{ki}^T\mathbf{\Sigma}_k\mathbf{r}_{ki} &= n_k^{-1}\sum_{i=1}^{n_k}(\mathbf{r}_{ki}^T\mathbf{\Sigma}_k\mathbf{r}_{ki}-n_k^{-1}\mathrm{tr}\mathbf{\Sigma}_k)^2+(n_k^{-1}\mathrm{tr}\mathbf{\Sigma}_k)^2 \\ &\quad +2n_k^{-1}\sum_{i=1}^{n_k}(n_k^{-1}\mathrm{tr}\mathbf{\Sigma}_k)(\mathbf{r}_{ki}^T\mathbf{\Sigma}_k\mathbf{r}_{ki}-n_k^{-1}\mathrm{tr}\mathbf{\Sigma}_k). \end{aligned}$$

As $n_k^{-1}\sum_{i=1}^{n_k}\mathrm{E}[(\mathbf{r}_{ki}^T\mathbf{\Sigma}_k\mathbf{r}_{ki}-n_k^{-1}\mathrm{tr}\mathbf{\Sigma}_k)^2]=n_k^{-2}[2\mathrm{tr}\mathbf{\Sigma}_k^2+\beta_k\sum_{j=1}^{p}(\mathbf{e}_j^T\mathbf{\Sigma}_k\mathbf{e}_j)^2]\to 0$, we have $n_k^{-1}\sum_{i=1}^{n_k}(\mathbf{r}_{ki}^T\mathbf{\Sigma}_k\mathbf{r}_{ki}-n_k^{-1}\mathrm{tr}\mathbf{\Sigma}_k)^2=o_p(1)$. As $n_k^{-1}\sum_{i=1}^{n_k}\mathrm{E}(\mathbf{r}_{ki}^T\mathbf{\Sigma}_k\mathbf{r}_{ki}-n_k^{-1}\mathrm{tr}\mathbf{\Sigma}_k)=0$ and

$$\mathrm{Var}[n_k^{-1}\sum_{i=1}^{n_k}(\mathbf{r}_{ki}^T\mathbf{\Sigma}_k\mathbf{r}_{ki}-n_k^{-1}\mathrm{tr}\mathbf{\Sigma}_k)]=n_k^{-3}[2\mathrm{tr}\mathbf{\Sigma}_k^2+\beta_k\sum_{j=1}^{p}(\mathbf{e}_j^T\mathbf{\Sigma}_k\mathbf{e}_j)^2]\to 0,$$

we have $n_k^{-1}\sum_{i=1}^{n_k}(\mathbf{r}_{ki}^T\mathbf{\Sigma}_k\mathbf{r}_{ki}-n_k^{-1}\mathrm{tr}\mathbf{\Sigma}_k)=o_p(1)$, and thus

$$n_k^{-1}\sum_{i=1}^{n_k}\mathbf{r}_{ki}^T\mathbf{\Sigma}_k\mathbf{r}_{ki}\mathbf{r}_{ki}^T\mathbf{\Sigma}_k\mathbf{r}_{ki}=(n_k^{-1}\mathrm{tr}\mathbf{\Sigma}_k)^2+o_p(1). \tag{S1.1}$$

Step 2.2. We have $n^{-1}\sum_{i\neq j\neq \ell}\mathrm{E}(\mathbf{r}_{ki}^T\mathbf{\Sigma}_k\mathbf{r}_{kj}\mathbf{r}_{kj}^T\mathbf{\Sigma}\mathbf{r}_{k\ell})=0$ and

$$\begin{aligned} & \mathrm{E}(n^{-1}\sum_{i\neq j\neq \ell}\mathbf{r}_i^T\mathbf{\Sigma}\mathbf{r}_j\mathbf{r}_j^T\mathbf{\Sigma}\mathbf{r}_\ell)^2 \\ \leq\ & 10n^{-2}\mathrm{tr}(\mathbf{\Sigma}^4)+(2n^{-3}+24n^{-4})[2\mathrm{tr}(\mathbf{\Sigma}^4)+\beta_k\sum_{j=1}^{p}(\mathbf{e}_j^T\mathbf{\Sigma}^2\mathbf{e}_j)^2+(\mathrm{tr}\mathbf{\Sigma}^2)^2] \\ & +24n^{-4}\beta_k\sum_{j=1}^{p}(\mathbf{e}_j^T\mathbf{\Sigma}^2\mathbf{e}_j)^2+8n^{-4}\beta_k^2\sum_{j=1}^{p}\sum_{\ell=1}^{p}(\mathbf{e}_j^T\mathbf{\Sigma}\mathbf{e}_\ell)^4\to 0. \end{aligned}$$

This leads to $n_k^{-1}\sum_{i\neq j\neq \ell}(\mathbf{r}_{ki}^T\mathbf{\Sigma}_k\mathbf{r}_{kj}\mathbf{r}_{kj}^T\mathbf{\Sigma}_k\mathbf{r}_{k\ell})=o_p(1)$.

Step 2.3. We have $n_k^{-1}\sum_{i\neq j}\mathrm{E}\mathbf{r}_{ki}^T\mathbf{\Sigma}_k\mathbf{r}_{kj}\mathbf{r}_{kj}^T\mathbf{\Sigma}_k\mathbf{r}_{ki}=(n_k-1)n_k^{-2}\mathrm{tr}(\mathbf{\Sigma}_k^2)$ and

$$\begin{aligned} & n_k^{-2}\mathrm{E}(\sum_{i\neq j}\mathbf{r}_{ki}^T\mathbf{\Sigma}_k\mathbf{r}_{kj}\mathbf{r}_{kj}^T\mathbf{\Sigma}_k\mathbf{r}_{ki})^2 \\ \leq\ & 6n_k^{-5}(n_k-1)\beta_k\sum_{j=1}^{p}(\mathbf{e}_j^T\mathbf{\Sigma}_k^2\mathbf{e}_j)^2+2n_k^{-5}(n_k-1)\beta_k^2\sum_{j=1}^{p}\sum_{\ell=1}^{p}(\mathbf{e}_j^T\mathbf{\Sigma}_k\mathbf{e}_\ell)^4 \\ & +n_k^{-5}(n_k-1)(n_k-2)(n_k-3)[\mathrm{tr}(\mathbf{\Sigma}_k^2)]^2 \\ & +2n_k^{-5}(n_k-1)(2n_k-1)[2\mathrm{tr}(\mathbf{\Sigma}_k^4)+\beta_k\sum_{j=1}^{p}(\mathbf{e}_j^T\mathbf{\Sigma}_k^2\mathbf{e}_j)^2+(\mathrm{tr}\mathbf{\Sigma}_k^2)^2]. \end{aligned}$$

As a result, we have $\mathrm{Var}(n_k^{-1}\sum_{i\neq j}\mathbf{r}_{ki}^T\mathbf{\Sigma}_k\mathbf{r}_{kj}\mathbf{r}_{kj}^T\mathbf{\Sigma}_k\mathbf{r}_{ki})\to 0$; that is,

$$n_k^{-1}\sum_{i\neq j}\mathbf{r}_{ki}^T\mathbf{\Sigma}_k\mathbf{r}_{kj}\mathbf{r}_{kj}^T\mathbf{\Sigma}_k\mathbf{r}_{ki}-(n_k-1)n_k^{-2}\mathrm{tr}(\mathbf{\Sigma}_k^2)=o_p(1).$$

Step 2.4. We have

$$\begin{aligned} & n_k^{-1}\sum_{i\neq j}\mathbf{r}_{ki}^T\mathbf{\Sigma}_k\mathbf{r}_{ki}\mathbf{r}_{ki}^T\mathbf{\Sigma}_k\mathbf{r}_{kj} \\ =\ & n_k^{-1}\sum_{i\neq j}(\mathbf{r}_{ki}^T\mathbf{\Sigma}_k\mathbf{r}_{ki}-n_k^{-1}\mathrm{tr}\mathbf{\Sigma}_k)\mathbf{r}_{ki}^T\mathbf{\Sigma}_k\mathbf{r}_{kj}+(n_k^{-2}\mathrm{tr}\mathbf{\Sigma}_k)\sum_{i\neq j}\mathbf{r}_{ki}^T\mathbf{\Sigma}_k\mathbf{r}_{kj}, \end{aligned}$$

and further,

$$\begin{aligned} & n_k^{-1}|\sum_{i\neq j}(\mathbf{r}_{ki}^T\mathbf{\Sigma}_k\mathbf{r}_{ki}-n_k^{-1}\mathrm{tr}\mathbf{\Sigma}_k)\mathbf{r}_{ki}^T\mathbf{\Sigma}_k\mathbf{r}_{kj}| \\ \leq\ & 2n_k^{-1}\sum_{i}(\mathbf{r}_{ki}^T\mathbf{\Sigma}_k\mathbf{r}_{ki}-n_k^{-1}\mathrm{tr}\mathbf{\Sigma}_k)^2+2n_k^{-1}\sum_{i\neq j}\mathbf{r}_{ki}^T\mathbf{\Sigma}_k\mathbf{r}_{kj}\mathbf{r}_{kj}^T\mathbf{\Sigma}_k\mathbf{r}_{ki}. \end{aligned}$$

Then, we have $n_k^{-1}\sum_{i\neq j}\mathbf{r}_{ki}^T\mathbf{\Sigma}_k\mathbf{r}_{ki}\mathbf{r}_{ki}^T\mathbf{\Sigma}_k\mathbf{r}_{kj}=o_p(1)$.

Thus, we have

$$\mathrm{tr}(\mathbf{S}_k^2)=\frac{n_k^2}{(n_k-1)^2}\mathrm{tr}[(\sum_{i=1}^{n_k}\mathbf{\Sigma}_k^{1/2}\mathbf{r}_{ki}\mathbf{r}_{ki}^T\mathbf{\Sigma}_k^{1/2})^2]-\frac{n_k+1}{n_k(n_k-1)^2}(\mathrm{tr}\mathbf{\Sigma}_k)^2-\frac{2}{n_k-1}\mathrm{tr}(\mathbf{\Sigma}_k^2)+o_p(1).$$

As $\mathrm{tr}\mathbf{S}_k = n_k(n_k-1)^{-1}(\sum_{i=1}^{n_k}\mathbf{r}_{ki}^T\mathbf{\Sigma}_k\mathbf{r}_{ki} - n_k\bar{\mathbf{r}}_k^T\mathbf{\Sigma}_k\bar{\mathbf{r}}_k)$, we have

$$\mathrm{tr}\mathbf{S}_k = n_k(n_k-1)^{-1}\sum_{i=1}^{n_k}\mathbf{r}_{ki}^T\mathbf{\Sigma}_k\mathbf{r}_{ki} - (n_k-1)^{-1}\mathrm{tr}\mathbf{\Sigma}_k + o_p(1).$$

As shown in Bai and Silverstein (2004, pp. 559–560),

$$\mathrm{tr}[(\sum_{i=1}^{n_k}\mathbf{\Sigma}_k^{1/2}\mathbf{r}_{ki}\mathbf{r}_{ki}^T\mathbf{\Sigma}_k^{1/2})^q] - \mathrm{tr}[(\sum_{i=1}^{n_k}\mathbf{\Sigma}_k^{1/2}\tilde{\mathbf{r}}_{ki}\tilde{\mathbf{r}}_{ki}^T\mathbf{\Sigma}_k^{1/2})^q] = o_p(1),\ \ q=1,2$$

where $\tilde{\mathbf{r}}_{ki} = n_k^{-1/2}\tilde{\mathbf{w}}_{ki}$, $\tilde{\mathbf{w}}_{ki} = (\tilde{w}_{k1i},\ldots,\tilde{w}_{kpi})^T$,

$$\tilde{w}_{k\ell i} = [\mathrm{Var}(w_{k\ell i}\delta_{\{|w_{k\ell i}|\le\sqrt{n_k}\eta_{n_k}\}})]^{-1/2}[w_{k\ell i}\delta_{\{|w_{k\ell i}|\le\sqrt{n_k}\eta_{n_k}\}} - \mathrm{E}(w_{k\ell i}\delta_{\{|w_{k\ell i}|\le\sqrt{n_k}\eta_{n_k}\}})],$$

$|\tilde{w}_{k\ell i}| \le c\sqrt{n_k}\eta_{n_k}$, $\mathrm{E}\tilde{w}_{k\ell i} = 0$, $\mathrm{E}(\tilde{w}_{k\ell i}^2) = 1$ and $\mathrm{E}(\tilde{w}_{k\ell i}^4) < \infty$ for $\ell = 1,\ldots,p$ and $i = 1,\ldots,n_k$ with $\eta_n \downarrow 0$, $\sqrt{n}\eta_n \to \infty$ and c being a positive constant. For simplicity, we rename the variable $\tilde{w}_{k\ell i}$ simply as $w_{k\ell i}$ and proceed by assuming that $|w_{k\ell i}| \le \sqrt{n_k}\eta_{n_k}$, $\mathrm{E}w_{k\ell i} = 0$, $\mathrm{E}(w_{k\ell i}^2) = 1$ and $\mathrm{E}(w_{k\ell i}^4) < \infty$ with $\eta_{n_k} \downarrow 0$ and $\sqrt{n_k}\eta_{n_k} \to \infty$. Let $\mathbf{B}_k = \sum_{i=1}^{n_k}\mathbf{\Sigma}_k^{1/2}\mathbf{r}_{ki}\mathbf{r}_{ki}^T\mathbf{\Sigma}_k^{1/2}$, then $\mathrm{tr}\mathbf{S}_k = n_k(n_k-1)^{-1}\mathrm{tr}\mathbf{B}_k - (n_k-1)^{-1}\mathrm{tr}\mathbf{\Sigma}_k + o_p(1)$ and

$$\mathrm{tr}(\mathbf{S}_k^2) = \frac{n_k^2}{(n_k-1)^2}\mathrm{tr}(\mathbf{B}_k^2) - \frac{n_k+1}{n_k(n_k-1)^2}(\mathrm{tr}\mathbf{\Sigma}_k)^2 - \frac{2}{n_k-1}\mathrm{tr}(\mathbf{\Sigma}_k^2) + o_p(1). \tag{S1.2}$$

Step 3. In this step, we show that

$$\begin{aligned} & \mathrm{tr}[(\mathbf{S}_1-\mathbf{S}_2)^2] - \mathrm{Etr}[(\mathbf{S}_1-\mathbf{S}_2)^2] \\ = \ & n_1^2(n_1-1)^{-2}[\mathrm{tr}(\mathbf{B}_1^2) - \mathrm{Etr}(\mathbf{B}_1^2)] + n_2^2(n_2-1)^{-2}[\mathrm{tr}(\mathbf{B}_2^2) - \mathrm{Etr}(\mathbf{B}_2^2)] \\ & -2n_1n_2(n_1-1)^{-1}(n_2-1)^{-1}[\mathrm{tr}(\mathbf{B}_1\mathbf{B}_2) - \mathrm{Etr}(\mathbf{\Sigma}_1\mathbf{\Sigma}_2)] + o_p(1) \end{aligned}$$

is asymptotically normal. When $\mathbf{\Sigma}_1 = \mathbf{\Sigma}_2 = \mathbf{\Sigma}$, we have

$$\begin{aligned}
& \mathrm{Etr}(\mathbf{S}_1^2) + \mathrm{Etr}(\mathbf{\Sigma}_2^2) - 2\mathrm{Etr}(\mathbf{S}_1\mathbf{S}_2) \\
= & \sum_{k=1}^{2} \frac{n_k+1}{(n_k-1)^2}\mathrm{tr}(\mathbf{\Sigma}^2) + \sum_{k=1}^{2} \frac{n_k^2-n_k-1}{n_k(n_k-1)^2}(\mathrm{tr}\mathbf{\Sigma})^2 + \sum_{k=1}^{2} \frac{\beta_k n_k}{(n_k-1)^2} \sum_{\ell=1}^{p} (\mathbf{e}_\ell^T \mathbf{\Sigma}_k \mathbf{e}_\ell)^2,
\end{aligned}$$

where

$$\begin{aligned}
& p^{-1}\mathrm{Etr}(\mathbf{S}_k^2) \\
= & \frac{n_k(n_k-1)+2}{p(n_k-1)^2}\mathrm{tr}(\mathbf{\Sigma}_k^2) + \frac{n_k^2-n_k-1}{pn_k(n_k-1)^2}(\mathrm{tr}\mathbf{\Sigma}_k)^2 + \frac{\beta_w n_k}{p(n_k-1)^2} \sum_{\ell=1}^{p} (\mathbf{e}_\ell^T \mathbf{\Sigma}_k \mathbf{e}_\ell)^2 \\
= & p^{-1}\mathrm{tr}(\mathbf{\Sigma}_k^2) + \frac{n_k+1}{p(n_k-1)^2}\mathrm{tr}(\mathbf{\Sigma}_k^2) + \frac{n_k^2-n_k-1}{pn_k(n_k-1)^2}(\mathrm{tr}\mathbf{\Sigma}_k)^2 + \frac{\beta_w n_k}{p(n_k-1)^2} \sum_{\ell=1}^{p} (\mathbf{e}_\ell^T \mathbf{\Sigma}_k \mathbf{e}_\ell)^2.
\end{aligned}$$

Since $\sum_{k=1}^2 (n_k^2 - n_k - 1)n_k^{-1}(n_k-1)^{-2}(\mathrm{tr}\mathbf{\Sigma})^2 \to \infty$, we need to establish the CLT of

$$\mathrm{tr}(\mathbf{S}_1 - \mathbf{S}_2)^2 - \hat{\mu}_{21},$$

where $\hat{\mu}_{21} = \sum_{k=1}^2 (n_k^2 - n_k - 1)n_k^{-1}(n_k-1)^{-2}(\mathrm{tr}\mathbf{S}_k)^2$. Let E_ℓ be the conditional expectation given $\{\mathbf{x}_{11}, \ldots, \mathbf{x}_{1\ell}\}$ and $\mathrm{E}_{\ell,\mathbf{S}_2}$ be the conditional expectation given $\{\mathbf{x}_{11}, \ldots, \mathbf{x}_{1\ell}, \mathbf{S}_2\}$. Based on the martingale difference central limit theory, it can be derived that conditional on $\mathbf{S}_2$,

$$\begin{aligned}
& \sigma_{11A}^{-1/2}[\mathrm{tr}(\mathbf{S}_1^2) - \mathrm{Etr}(\mathbf{S}_1^2) - 2\mathrm{tr}(\mathbf{S}_1\mathbf{S}_2) + 2\mathrm{tr}(\mathbf{\Sigma}_1\mathbf{S}_2) \\
& -\frac{n_1^2-n_1-1}{n_1(n_1-1)^2}\{(\mathrm{tr}\mathbf{S}_1)^2 - (\mathrm{tr}\mathbf{\Sigma}_1)^2\}] \xrightarrow{d} N(0,1), \qquad \text{(S1.3)}
\end{aligned}$$

where $\sigma_{11A} = \sigma_{110A} + 4\sigma_{220A} - 4\sigma_{120A} + 4n_1^{-2}(\mathrm{tr}\mathbf{\Sigma}_1)^2\sigma_{330A} - 4n_1^{-1}(\mathrm{tr}\mathbf{\Sigma}_1)\sigma_{130A} +$

$8n_1^{-1}(\mathrm{tr}\mathbf{\Sigma}_1)\sigma_{230A}$, with

$$\begin{aligned}
\sigma_{110A} &= \sum_{\ell=1}^{n_1}[(\mathrm{E}_\ell - \mathrm{E}_{\ell-1})\mathrm{tr}(\mathbf{B}_1^2)]^2,\\
\sigma_{220A} &= \sum_{\ell=1}^{n_1}[(\mathrm{E}_{\ell,\mathbf{B}_2} - \mathrm{E}_{\ell-1,\mathbf{B}_2})\mathrm{tr}(\mathbf{B}_1\mathbf{B}_2)]^2,\\
\sigma_{330A} &= \sum_{\ell=1}^{n_1}[(\mathrm{E}_{\ell,\mathbf{B}_2} - \mathrm{E}_{\ell-1,\mathbf{B}_2})\mathrm{tr}(\mathbf{B}_1)]^2,\\
\sigma_{120A} &= \sum_{\ell=1}^{n_1}[(\mathrm{E}_\ell - \mathrm{E}_{\ell-1})\mathrm{tr}(\mathbf{B}_1^2)][(\mathrm{E}_{\ell,\mathbf{B}_2} - \mathrm{E}_{\ell-1,\mathbf{B}_2})\mathrm{tr}(\mathbf{B}_1\mathbf{B}_2)],\\
\sigma_{130A} &= \sum_{\ell=1}^{n_1}[(\mathrm{E}_\ell - \mathrm{E}_{\ell-1})\mathrm{tr}(\mathbf{B}_1^2)][(\mathrm{E}_{\ell,\mathbf{B}_2} - \mathrm{E}_{\ell-1,\mathbf{B}_2})\mathrm{tr}\mathbf{B}_1],\\
\sigma_{230A} &= \sum_{\ell=1}^{n_1}[(\mathrm{E}_{\ell,\mathbf{B}_2} - \mathrm{E}_{\ell-1,\mathbf{B}_2})(\mathrm{tr}\mathbf{B}_1)][(\mathrm{E}_{\ell,\mathbf{B}_2} - \mathrm{E}_{\ell-1,\mathbf{B}_2})\mathrm{tr}(\mathbf{B}_1\mathbf{B}_2)].
\end{aligned}$$

Moreover, we have

$$\begin{aligned}
&\sigma_{22A}^{-1/2}[\mathrm{tr}(\mathbf{S}_2^2) - \mathrm{E}\mathrm{tr}(\mathbf{S}_2^2) - 2[\mathrm{tr}(\mathbf{\Sigma}_1\mathbf{S}_2) - \mathrm{tr}(\mathbf{\Sigma}_1\mathbf{\Sigma}_2)]\\
&-\frac{n_2^2 - n_2 - 1}{n_2(n_2-1)^2}\{(\mathrm{tr}\mathbf{S}_2)^2 - (\mathrm{tr}\mathbf{\Sigma}_2)^2\}] \xrightarrow{d} N(0,1), \qquad \text{(S1.4)}
\end{aligned}$$

where

$$\sigma_{22A} = \sigma_{440A} + 4\sigma_{550A} + 4n_2^{-2}(\mathrm{tr}\mathbf{\Sigma}_2)^2\sigma_{660A} - 4\sigma_{450A} - 4n_2^{-1}(\mathrm{tr}\mathbf{\Sigma}_2)\sigma_{460A} + 8n_2^{-1}(\mathrm{tr}\mathbf{\Sigma}_2)\sigma_{560A},$$

with

$$\begin{aligned}
\sigma_{440A} &= \sum_{\ell=1}^{n_2}[(\mathrm{E}_\ell - \mathrm{E}_{\ell-1})\mathrm{tr}(\mathbf{B}_2^2)]^2,\\
\sigma_{550A} &= \sum_{\ell=1}^{n_2}[(\mathrm{E}_\ell - \mathrm{E}_{\ell-1})\mathrm{tr}(\mathbf{B}_2\mathbf{\Sigma}_1)]^2,\\
\sigma_{660A} &= \sum_{\ell=1}^{n_2}[(\mathrm{E}_\ell - \mathrm{E}_{\ell-1})\mathrm{tr}\mathbf{B}_2]^2,
\end{aligned}$$

$$\begin{aligned}\sigma_{450A} &= \sum_{\ell=1}^{n_2}[(\mathrm{E}_\ell - \mathrm{E}_{\ell-1})\mathrm{tr}(\mathbf{B}_2^2)][(\mathrm{E}_\ell - \mathrm{E}_{\ell-1})\mathrm{tr}(\mathbf{B}_2\mathbf{\Sigma}_1)],\\ \sigma_{460A} &= \sum_{\ell=1}^{n_2}[(\mathrm{E}_\ell - \mathrm{E}_{\ell-1})\mathrm{tr}(\mathbf{B}_2^2)][(\mathrm{E}_\ell - \mathrm{E}_{\ell-1})\mathrm{tr}\mathbf{B}_2],\\ \sigma_{560A} &= \sum_{\ell=1}^{n_2}[(\mathrm{E}_\ell - \mathrm{E}_{\ell-1})\mathrm{tr}(\mathbf{B}_2\mathbf{\Sigma}_1)][(\mathrm{E}_\ell - \mathrm{E}_{\ell-1})\mathrm{tr}\mathbf{B}_2].\end{aligned}$$

Step 4. Next, we show that

$$\begin{aligned}\mu_{A2} &= \mathrm{Etr}(\mathbf{S}_2^2) + \mathrm{Etr}(\mathbf{S}_1^2) - 2\mathrm{tr}(\mathbf{\Sigma}_1\mathbf{\Sigma}_2) - \sum_{k=1}^{2}\frac{n_k^2 - n_k - 1}{n_k(n_k-1)^2}(\mathrm{tr}\mathbf{\Sigma}_k)^2\\ &= \mathrm{tr}[(\mathbf{\Sigma}_1 - \mathbf{\Sigma}_2)^2] + \frac{n_1+1}{(n_1-1)^2}\mathrm{tr}(\mathbf{\Sigma}_1^2) + \frac{n_2+1}{(n_2-1)^2}\mathrm{tr}(\mathbf{\Sigma}_2^2)\\ &\quad + \frac{\beta_1 n_1}{(n_1-1)^2}\sum_{k=1}^{p}(\mathbf{e}_k^T\mathbf{\Sigma}_1\mathbf{e}_k)^2 + \frac{\beta_2 n_2}{(n_2-1)^2}\sum_{k=1}^{p}(\mathbf{e}_k^T\mathbf{\Sigma}_2\mathbf{e}_k)^2.\end{aligned}$$

Moreover, we have

$$\begin{aligned}\sigma_{110A} &= \sum_{\ell=1}^{n_1}[(\mathrm{E}_\ell - \mathrm{E}_{\ell-1})\mathrm{tr}(\mathbf{B}_1^2)]^2\\ &= 4n_1^{-1}[2\mathrm{tr}(\mathbf{\Sigma}_1^4) + \beta_1\sum_{\ell=1}^{p}(\mathbf{e}_\ell^T\mathbf{\Sigma}_1^2\mathbf{e}_\ell)^2]\\ &\quad +2(n_1^{-1}\mathrm{tr}\mathbf{\Sigma}_1)^2 n_1^{-1}[2\mathrm{tr}(\mathbf{\Sigma}_1^2) + \beta_1\sum_{\ell}(\mathbf{e}_\ell^T\mathbf{\Sigma}_1\mathbf{e}_\ell)^2]\\ &\quad +4[n_1^{-1}\mathrm{tr}(\mathbf{\Sigma}_1^2)]^2 + 8(n_1^{-1}\mathrm{tr}\mathbf{\Sigma}_1)n_1^{-1}[2\mathrm{tr}(\mathbf{\Sigma}_1^3) + \beta_1\sum_{\ell=1}^{p}\mathbf{e}_\ell^T\mathbf{\Sigma}_1^2\mathbf{e}_\ell\mathbf{e}_\ell^T\mathbf{\Sigma}_1\mathbf{e}_\ell]\\ &\quad +2(n_1^{-1}\mathrm{tr}\mathbf{\Sigma}_1)^2 n_1^{-1}[2\mathrm{tr}(\mathbf{\Sigma}_1^2) + \beta_1\sum_{\ell=1}^{p}(\mathbf{e}_\ell^T\mathbf{\Sigma}_1\mathbf{e}_\ell)^2],\end{aligned}$$

$$
\begin{aligned}
\sigma_{220A} &= \sum_{\ell=1}^{n_1}[(\mathrm{E}_{\ell,\mathbf{B}_2}-\mathrm{E}_{\ell-1,\mathbf{B}_2})\mathrm{tr}(\mathbf{B}_1\mathbf{B}_2)]^2 \\
&= 2n_1^{-1}\mathrm{tr}[(\mathbf{\Sigma}_1\mathbf{B}_2)^2]+\beta_1 n_1^{-1}\sum_{\ell=1}^{p}(\mathbf{e}_\ell^T\mathbf{\Sigma}_1^{1/2}\mathbf{B}_2\mathbf{\Sigma}_1^{1/2}\mathbf{e}_\ell)^2 \\
&= \frac{2}{n_1n_2}[\mathrm{tr}(\mathbf{\Sigma}_1\mathbf{\Sigma}_2)]^2+[2n_1^{-1}\mathrm{tr}[(\mathbf{\Sigma}_1\mathbf{\Sigma}_2)^2]+\beta_1 n_1^{-1}\sum_{\ell=1}^{p}(\mathbf{e}_\ell^T\mathbf{\Sigma}_1^{1/2}\mathbf{\Sigma}_2\mathbf{\Sigma}_1^{1/2}\mathbf{e}_\ell)^2]+o_p(1), \\
\sigma_{330A} &= \sum_{\ell=1}^{n_1}[(\mathrm{E}_{\ell,\mathbf{B}_2}-\mathrm{E}_{\ell-1,\mathbf{B}_2})\mathrm{tr}\mathbf{B}_1]^2 \\
&= \sum_{\ell=1}^{n_1}\mathrm{E}(\mathbf{r}_\ell^T\mathbf{\Sigma}_1\mathbf{r}_\ell)^2 = 2n_1^{-1}\mathrm{tr}(\mathbf{\Sigma}_1^2)+\beta_1 n_1^{-1}\sum_{\ell=1}^{p}(\mathbf{e}_\ell^T\mathbf{\Sigma}_1\mathbf{e}_\ell)^2, \\
\sigma_{120A} &= \sum_{\ell=1}^{n_1}[(\mathrm{E}_\ell-\mathrm{E}_{\ell-1})\mathrm{tr}(\mathbf{B}_1^2)][(\mathrm{E}_{\ell,\mathbf{B}_2}-\mathrm{E}_{\ell-1,\mathbf{B}_2})\mathrm{tr}(\mathbf{B}_1\mathbf{B}_2)] \\
&= 2n_1^{-1}[2\mathrm{tr}(\mathbf{\Sigma}_1^3\mathbf{B}_2)+\beta_1\sum_{\ell=1}^{p}\mathbf{e}_\ell^T\mathbf{\Sigma}_1^{1/2}\mathbf{B}_2\mathbf{\Sigma}_1^{1/2}\mathbf{e}_\ell\mathbf{e}_\ell^T\mathbf{\Sigma}_1^2\mathbf{e}_\ell)] \\
&\quad +2(n_1^{-1}\mathrm{tr}\mathbf{\Sigma}_1)n_1^{-1}\left[2\mathrm{tr}(\mathbf{\Sigma}_1^2\mathbf{B}_2)+\beta_1\sum_{\ell=1}^{p}\mathbf{e}_\ell^T\mathbf{\Sigma}_1\mathbf{e}_\ell\mathbf{e}_\ell^T\mathbf{\Sigma}_1^{1/2}\mathbf{B}_2\mathbf{\Sigma}_1^{1/2}\mathbf{e}_\ell\right] \\
&= 2[2n_1^{-1}\mathrm{tr}(\mathbf{\Sigma}_1^3\mathbf{\Sigma}_2)+\beta_1 n_1^{-1}\sum_{\ell=1}^{p}\mathbf{e}_\ell^T\mathbf{\Sigma}_1^{1/2}\mathbf{\Sigma}_2\mathbf{\Sigma}_1^{1/2}\mathbf{e}_\ell\mathbf{e}_\ell^T\mathbf{\Sigma}_1^2\mathbf{e}_\ell] \\
&\quad +2(n_1^{-1}\mathrm{tr}\mathbf{\Sigma}_1)n_1^{-1}\left[2\mathrm{tr}(\mathbf{\Sigma}_1^2\mathbf{\Sigma}_2)+\beta_1\sum_{\ell=1}^{p}\mathbf{e}_\ell^T\mathbf{\Sigma}_1\mathbf{e}_\ell\mathbf{e}_\ell^T\mathbf{\Sigma}_1^{1/2}\mathbf{\Sigma}_2\mathbf{\Sigma}_1^{1/2}\mathbf{e}_\ell\right]+o_p(1), \\
\sigma_{130A} &= \sum_{\ell=1}^{n_1}[(\mathrm{E}_\ell-\mathrm{E}_{\ell-1})\mathrm{tr}(\mathbf{B}_1^2)][(\mathrm{E}_{\ell,\mathbf{B}_2}-\mathrm{E}_{\ell-1,\mathbf{B}_2})\mathrm{tr}(\mathrm{tr}\mathbf{B}_1)] \\
&= \sum_{\ell=1}^{n_1}[(\mathrm{E}_\ell-\mathrm{E}_{\ell-1})\mathrm{tr}(\mathbf{B}_1^2)][(\mathrm{E}_\ell-\mathrm{E}_{\ell-1})\mathrm{tr}\mathbf{B}_1] \\
&= 2(2n_1^{-1}\mathrm{tr}(\mathbf{\Sigma}_1^3)+\beta_1 n_1^{-1}\sum_{\ell=1}^{p}\mathbf{e}_\ell^T\mathbf{\Sigma}_1\mathbf{e}_\ell\mathbf{e}_\ell^T\mathbf{\Sigma}_1^2\mathbf{e}_\ell) \\
&\quad +2(n_1^{-1}\mathrm{tr}\mathbf{\Sigma}_1)[2n_1^{-1}\mathrm{tr}(\mathbf{\Sigma}_1^2)+\beta_1 n_1^{-1}\sum_{\ell=1}^{p}(\mathbf{e}_\ell^T\mathbf{\Sigma}_1\mathbf{e}_\ell)^2],
\end{aligned}
$$

$$
\begin{aligned}
\sigma_{230A} &= \sum_{\ell=1}^{n_1}[(\mathrm{E}_{\ell,\mathbf{B}_2}-\mathrm{E}_{\ell-1,\mathbf{B}_2})(\mathrm{tr}\mathbf{B}_1)][(\mathrm{E}_{\ell,\mathbf{B}_2}-\mathrm{E}_{\ell-1,\mathbf{B}_2})\mathrm{tr}(\mathbf{B}_1\mathbf{B}_2)]\\
&= \sum_{\ell=1}^{n_1}[(\mathrm{E}_{\ell}-\mathrm{E}_{\ell-1})\mathrm{tr}\mathbf{B}_1][(\mathrm{E}_{\ell,\mathbf{B}_2}-\mathrm{E}_{\ell-1,\mathbf{B}_2})\mathrm{tr}(\mathbf{B}_1\mathbf{B}_2)]\\
&= \sum_{\ell=1}^{n_1}\mathrm{E}_{\ell-1,\mathbf{B}_2}(\mathbf{r}_\ell^T\mathbf{\Sigma}_1\mathbf{r}_\ell-n_1^{-1}\mathrm{tr}\mathbf{\Sigma}_1)[\mathbf{r}_\ell^T\mathbf{\Sigma}_1^{1/2}\mathbf{B}_2\mathbf{\Sigma}_1^{1/2}\mathbf{r}_\ell-n_1^{-1}\mathrm{tr}(\mathbf{\Sigma}_1\mathbf{B}_2)]\\
&= [2n_1^{-1}\mathrm{tr}(\mathbf{\Sigma}_1^2\mathbf{\Sigma}_2)+\beta_1 n_1^{-1}\sum_{\ell=1}^{p}(\mathbf{e}_\ell^T\mathbf{\Sigma}_1\mathbf{e}_\ell)(\mathbf{e}_\ell^T\mathbf{\Sigma}_1^{1/2}\mathbf{\Sigma}_2\mathbf{\Sigma}_1^{1/2}\mathbf{e}_\ell)],\\
\sigma_{440A} &= \sum_{\ell=1}^{n_2}[(\mathrm{E}_{\ell}-\mathrm{E}_{\ell-1})\mathrm{tr}(\mathbf{B}_2^2)]^2\\
&= 4n_2^{-1}[2\mathrm{tr}(\mathbf{\Sigma}_2^4)+\beta_2\sum_{\ell=1}^{p}(\mathbf{e}_\ell^T\mathbf{\Sigma}_2^2\mathbf{e}_\ell)^2]+2(n_2^{-1}\mathrm{tr}\mathbf{\Sigma}_2)^2 n_2^{-1}[2\mathrm{tr}(\mathbf{\Sigma}_2^2)+\beta_2\sum_{\ell}^{p}(\mathbf{e}_\ell^T\mathbf{\Sigma}_2\mathbf{e}_\ell)^2]\\
&\quad +4[n_2^{-1}\mathrm{tr}(\mathbf{\Sigma}_2^2)]^2+8(n_2^{-1}\mathrm{tr}\mathbf{\Sigma}_2)n_2^{-1}[2\mathrm{tr}(\mathbf{\Sigma}_2^3)+\beta_2\sum_{\ell=1}^{p}\mathbf{e}_\ell^T\mathbf{\Sigma}_2^2\mathbf{e}_\ell\mathbf{e}_\ell^T\mathbf{\Sigma}_2\mathbf{e}_\ell]\\
&\quad +2(n_2^{-1}\mathrm{tr}\mathbf{\Sigma}_2)^2 n_2^{-1}[2\mathrm{tr}(\mathbf{\Sigma}_2^2)+\beta_2\sum_{\ell=1}^{p}(\mathbf{e}_\ell^T\mathbf{\Sigma}_2\mathbf{e}_\ell)^2],\\
\sigma_{550A} &= \sum_{\ell=1}^{n_2}[(\mathrm{E}_{\ell}-\mathrm{E}_{\ell-1})\mathrm{tr}(\mathbf{B}_2\mathbf{\Sigma}_1)]^2\\
&= [2n_2^{-1}\mathrm{tr}[(\mathbf{\Sigma}_2\mathbf{\Sigma}_1)^2]+\beta_2 n_2^{-1}\sum_{\ell=1}^{p}(\mathbf{e}_\ell^T\mathbf{\Sigma}_2^{1/2}\mathbf{\Sigma}_1\mathbf{\Sigma}_2^{1/2}\mathbf{e}_\ell)^2],\\
\sigma_{660A} &= \sum_{\ell=1}^{n_2}[(\mathrm{E}_{\ell}-\mathrm{E}_{\ell-1})\mathrm{tr}\mathbf{B}_2]^2=[2n_2^{-1}\mathrm{tr}(\mathbf{\Sigma}_2^2)+\beta_2 n_2^{-1}\sum_{\ell=1}^{p}(\mathbf{e}_\ell^T\mathbf{\Sigma}_2\mathbf{e}_\ell)^2],\\
\sigma_{450A} &= \sum_{\ell=1}^{n_2}[(\mathrm{E}_{\ell}-\mathrm{E}_{\ell-1})\mathrm{tr}(\mathbf{B}_2^2)][(\mathrm{E}_{\ell}-\mathrm{E}_{\ell-1})\mathrm{tr}\mathbf{B}_2\mathbf{\Sigma}_1]\\
&= 2n_2^{-1}[2\mathrm{tr}(\mathbf{\Sigma}_2^3\mathbf{\Sigma}_1)+\beta_2\sum_{\ell=1}^{p}\mathbf{e}_\ell^T\mathbf{\Sigma}_2^{1/2}\mathbf{\Sigma}_1\mathbf{\Sigma}_2^{1/2}\mathbf{e}_\ell\mathbf{e}_\ell^T\mathbf{\Sigma}_2^2\mathbf{e}_\ell]\\
&\quad +2(n_2^{-1}\mathrm{tr}\mathbf{\Sigma}_2)n_2^{-1}\left[2\mathrm{tr}(\mathbf{\Sigma}_2^2\mathbf{\Sigma}_1)+\beta_2\sum_{\ell=1}^{p}\mathbf{e}_\ell^T\mathbf{\Sigma}_2\mathbf{e}_\ell\mathbf{e}_\ell^T\mathbf{\Sigma}_2^{1/2}\mathbf{\Sigma}_1\mathbf{\Sigma}_2^{1/2}\mathbf{e}_\ell\right],
\end{aligned}
$$

$$\begin{aligned}
\sigma_{460A} &= \sum_{\ell=1}^{n_2}[(\mathrm{E}_\ell - \mathrm{E}_{\ell-1})\mathrm{tr}(\mathbf{B}_2^2)][(\mathrm{E}_\ell - \mathrm{E}_{\ell-1})\mathrm{tr}\mathbf{B}_2] \\
&= 2[2n_2^{-1}\mathrm{tr}(\mathbf{\Sigma}_2^3) + \beta_2 n_2^{-1}\sum_{\ell=1}^{p}\mathbf{e}_\ell^T\mathbf{\Sigma}_2\mathbf{e}_\ell\mathbf{e}_\ell^T\mathbf{\Sigma}_2^2\mathbf{e}_\ell] \\
&\quad +2(n_2^{-1}\mathrm{tr}\mathbf{\Sigma}_2)\left[2n_2^{-1}\mathrm{tr}(\mathbf{\Sigma}_2^2) + \beta_2 n_2^{-1}\sum_{\ell=1}^{p}(\mathbf{e}_\ell^T\mathbf{\Sigma}_2\mathbf{e}_\ell)^2\right], \\
\sigma_{560A} &= \sum_{\ell=1}^{n_2}[(\mathrm{E}_\ell - \mathrm{E}_{\ell-1})\mathrm{tr}\mathbf{B}_2\mathbf{\Sigma}_1][(\mathrm{E}_\ell - \mathrm{E}_{\ell-1})\mathrm{tr}\mathbf{B}_2] \\
&= [2n_2^{-1}\mathrm{tr}[(\mathbf{\Sigma}_1\mathbf{\Sigma}_2^2)] + \beta_2 n_2^{-1}\sum_{\ell=1}^{p}(\mathbf{e}_\ell^T\mathbf{\Sigma}_2\mathbf{e}_\ell)(\mathbf{e}_\ell^T\mathbf{\Sigma}_2^{1/2}\mathbf{\Sigma}_1\mathbf{\Sigma}_2^{1/2}\mathbf{e}_\ell)].
\end{aligned}$$

Thus, we have

$$\begin{aligned}
\sigma_{11A} &= \sigma_{110A} + 4\sigma_{220A} + 4n_1^{-2}(\mathrm{tr}\mathbf{\Sigma}_1)^2\sigma_{330A} - 4\sigma_{120A} + 4n_1^{-1}(\mathrm{tr}\mathbf{\Sigma}_1)(-\sigma_{130A} + 2\sigma_{230A}) \\
&= 4[2n_1^{-1}\mathrm{tr}(\mathbf{\Sigma}_1^4) + \beta_1 n_1^{-1}\sum_{\ell=1}^{p}(\mathbf{e}_\ell^T\mathbf{\Sigma}_1^2\mathbf{e}_\ell)^2] + 4[n_1^{-1}\mathrm{tr}(\mathbf{\Sigma}_1^2)]^2 + \frac{8n_2}{n_1}(n_2^{-1}\mathrm{tr}\mathbf{\Sigma}_1\mathbf{\Sigma}_2)^2 \\
&\quad +4[2n_1^{-1}\mathrm{tr}[(\mathbf{\Sigma}_1\mathbf{\Sigma}_2)^2] + \beta_1 n_1^{-1}\sum_{\ell=1}^{p}(\mathbf{e}_\ell^T\mathbf{\Sigma}_1^{1/2}\mathbf{\Sigma}_2\mathbf{\Sigma}_1^{1/2}\mathbf{e}_\ell)^2] \\
&\quad -8[2n_1^{-1}\mathrm{tr}(\mathbf{\Sigma}_1^3\mathbf{\Sigma}_2) + \beta_1 n_1^{-1}\sum_{\ell=1}^{p}\mathbf{e}_\ell^T\mathbf{\Sigma}_1^{1/2}\mathbf{\Sigma}_2\mathbf{\Sigma}_1^{1/2}\mathbf{e}_\ell\mathbf{e}_\ell^T\mathbf{\Sigma}_1^2\mathbf{e}_\ell],
\end{aligned}$$

and

$$\begin{aligned}
\sigma_{22A} &= \sigma_{440A} + 4\sigma_{550A} + 4n_2^{-2}(\mathrm{tr}\mathbf{\Sigma}_2)^2\sigma_{660A} - 4\sigma_{450A} + 4n_2^{-1}(\mathrm{tr}\mathbf{\Sigma}_2)(-\sigma_{460A} + 2\sigma_{560A}) \\
&= 4n_2^{-1}[2\mathrm{tr}(\mathbf{\Sigma}_2^4) + \beta_2\sum_{\ell=1}^{p}(\mathbf{e}_\ell^T\mathbf{\Sigma}_2^2\mathbf{e}_\ell)^2] + 4[n_2^{-1}\mathrm{tr}(\mathbf{\Sigma}_2^2)]^2 \\
&\quad +4[2n_2^{-1}\mathrm{tr}[(\mathbf{\Sigma}_2\mathbf{\Sigma}_1)^2] + \beta_2 n_2^{-1}\sum_{\ell=1}^{p}(\mathbf{e}_\ell^T\mathbf{\Sigma}_2^{1/2}\mathbf{\Sigma}_1\mathbf{\Sigma}_2^{1/2}\mathbf{e}_\ell)^2] \\
&\quad -8(2n_2^{-1}\mathrm{tr}(\mathbf{\Sigma}_2^3\mathbf{\Sigma}_1) + \beta_2 n_2^{-1}\sum_{\ell=1}^{p}\mathbf{e}_\ell^T\mathbf{\Sigma}_2^{1/2}\mathbf{\Sigma}_1\mathbf{\Sigma}_2^{1/2}\mathbf{e}_\ell\mathbf{e}_\ell^T\mathbf{\Sigma}_2^2\mathbf{e}_\ell)].
\end{aligned}$$

Step 5. According to (S1.3) and (S1.4), we have

$$\sigma_{A2}^{-1}\{\mathrm{tr}[(\mathbf{S}_1-\mathbf{S}_2)^2]-\hat{\mu}_{21}-\mu_{A2}\}\to N(0,1),$$

where

$$\mu_{A2}=\mathrm{tr}[(\mathbf{\Sigma}_1-\mathbf{\Sigma}_2)^2]+\sum_{k=1}^{2}\frac{n_k+1}{(n_k-1)^2}\mathrm{tr}(\mathbf{\Sigma}_k^2)+\sum_{k=1}^{2}\frac{\beta_k n_k}{(n_k-1)^2}\sum_{\ell=1}^{p}(\mathbf{e}_\ell^T\mathbf{\Sigma}_k\mathbf{e}_\ell)^2$$

and

$$\begin{aligned}
\sigma_{A2}^2 &= \sigma_{11A}+\sigma_{22A}\\
&= 4[n_1^{-1}\mathrm{tr}(\mathbf{\Sigma}_1^2)]^2+4[n_2^{-1}\mathrm{tr}(\mathbf{\Sigma}_2^2)]^2+8n_1^{-1}n_2^{-1}[\mathrm{tr}(\mathbf{\Sigma}_1\mathbf{\Sigma}_2)]^2\\
&\quad +4\left[2n_1^{-1}\mathrm{tr}(\mathbf{\Sigma}_1^4)+\beta_1 n_1^{-1}\sum_{\ell=1}^{p}(\mathbf{e}_\ell^T\mathbf{\Sigma}_1^2\mathbf{e}_\ell)^2\right]\\
&\quad +4\{2n_1^{-1}\mathrm{tr}[(\mathbf{\Sigma}_1\mathbf{\Sigma}_2)^2]+\beta_1 n_1^{-1}\sum_{\ell=1}^{p}(\mathbf{e}_\ell^T\mathbf{\Sigma}_1^{1/2}\mathbf{\Sigma}_2\mathbf{\Sigma}_1^{1/2}\mathbf{e}_\ell)^2\}\\
&\quad -8[2n_1^{-1}\mathrm{tr}(\mathbf{\Sigma}_1^3\mathbf{\Sigma}_2)+\beta_1 n_1^{-1}\sum_{\ell=1}^{p}\mathbf{e}_\ell^T\mathbf{\Sigma}_1^{1/2}\mathbf{\Sigma}_2\mathbf{\Sigma}_1^{1/2}\mathbf{e}_\ell\mathbf{e}_\ell^T\mathbf{\Sigma}_1^2\mathbf{e}_\ell]\\
&\quad +4n_2^{-1}\left[2\mathrm{tr}(\mathbf{\Sigma}_2^4)+\beta_2\sum_{\ell=1}^{p}(\mathbf{e}_\ell^T\mathbf{\Sigma}_2^2\mathbf{e}_\ell)^2\right]\\
&\quad +4\{2n_2^{-1}\mathrm{tr}[(\mathbf{\Sigma}_2\mathbf{\Sigma}_1)^2]+\beta_2 n_2^{-1}\sum_{\ell=1}^{p}(\mathbf{e}_\ell^T\mathbf{\Sigma}_2^{1/2}\mathbf{\Sigma}_1\mathbf{\Sigma}_2^{1/2}\mathbf{e}_\ell)^2\}\\
&\quad -8[2n_2^{-1}\mathrm{tr}(\mathbf{\Sigma}_2^3\mathbf{\Sigma}_1)+\beta_2 n_2^{-1}\sum_{\ell=1}^{p}\mathbf{e}_\ell^T\mathbf{\Sigma}_2^{1/2}\mathbf{\Sigma}_1\mathbf{\Sigma}_2^{1/2}\mathbf{e}_\ell\mathbf{e}_\ell^T\mathbf{\Sigma}_2^2\mathbf{e}_\ell)].
\end{aligned}$$

As a result, the proof of Lemma 1 is completed.

Lemma 2. *Under Assumptions (A1)–(A2), we have*

$$\begin{pmatrix} \sigma^2_{Ak_1k_2} & \sigma_{Ak_1k_2k_2k_3} \\ \sigma_{Ak_1k_2k_2k_3} & \sigma^2_{Ak_2k_3j} \end{pmatrix}^{-1/2} \begin{pmatrix} \mathrm{tr}(\mathbf{S}_{k_1}-\mathbf{S}_{k_2})^2 - \hat{\mu}_{k_1k_2} - \mu_{Ak_1k_2} \\ \mathrm{tr}(\mathbf{S}_{k_2}-\mathbf{S}_{k_3})^2 - \hat{\mu}_{k_2k_3} - \mu_{Ak_2k_3} \end{pmatrix} \to N(\mathbf{0}_2, \mathbf{I}_2),$$

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

$$\begin{aligned}
\hat{\mu}_{ij} &= \sum_{k=i,j}(n_k^2 - n_k - 1)n_k^{-1}(n_k-1)^{-2}(\mathrm{tr}\mathbf{S}_k)^2, \\
\mu_{Aij} &= \mathrm{tr}[(\boldsymbol{\Sigma}_i - \boldsymbol{\Sigma}_j)^2] + \sum_{k=i,j}\frac{n_k+1}{(n_k-1)^2}\mathrm{tr}(\boldsymbol{\Sigma}_k^2) + \sum_{k=i,j}\frac{\beta_k n_k}{(n_k-1)^2}\sum_{\ell=1}^{p}(\mathbf{e}_\ell^T\boldsymbol{\Sigma}_k\mathbf{e}_\ell)^2, \\
\sigma^2_{Aij} &= 4[n_i^{-1}\mathrm{tr}(\boldsymbol{\Sigma}_i^2)]^2 + 4[n_j^{-1}\mathrm{tr}(\boldsymbol{\Sigma}_j^2)]^2 + 8n_i^{-1}n_j^{-1}[\mathrm{tr}(\boldsymbol{\Sigma}_i\boldsymbol{\Sigma}_j)]^2 \\
&\quad +4[2n_i^{-1}\mathrm{tr}(\boldsymbol{\Sigma}_i^4) + \beta_i n_i^{-1}\sum_{\ell=1}^{p}(\mathbf{e}_\ell^T\boldsymbol{\Sigma}_i^2\mathbf{e}_\ell)^2] \\
&\quad +4\{2n_i^{-1}\mathrm{tr}[(\boldsymbol{\Sigma}_i\boldsymbol{\Sigma}_j)^2] + \beta_i n_i^{-1}\sum_{\ell=1}^{p}(\mathbf{e}_\ell^T\boldsymbol{\Sigma}_i^{1/2}\boldsymbol{\Sigma}_j\boldsymbol{\Sigma}_i^{1/2}\mathbf{e}_\ell)^2\} \\
&\quad -8[2n_i^{-1}\mathrm{tr}(\boldsymbol{\Sigma}_i^3\boldsymbol{\Sigma}_j) + \beta_i n_i^{-1}\sum_{\ell=1}^{p}\mathbf{e}_\ell^T\boldsymbol{\Sigma}_i^{1/2}\boldsymbol{\Sigma}_j\boldsymbol{\Sigma}_i^{1/2}\mathbf{e}_\ell\mathbf{e}_\ell^T\boldsymbol{\Sigma}_i^2\mathbf{e}_\ell] \\
&\quad +4[2n_j^{-1}\mathrm{tr}(\boldsymbol{\Sigma}_j^4) + \beta_j n_j^{-1}\sum_{\ell=1}^{p}(\mathbf{e}_\ell^T\boldsymbol{\Sigma}_j^2\mathbf{e}_\ell)^2] \\
&\quad +4\{2n_j^{-1}\mathrm{tr}[(\boldsymbol{\Sigma}_j\boldsymbol{\Sigma}_i)^2] + \beta_j n_j^{-1}\sum_{\ell=1}^{p}(\mathbf{e}_\ell^T\boldsymbol{\Sigma}_j^{1/2}\boldsymbol{\Sigma}_i\boldsymbol{\Sigma}_j^{1/2}\mathbf{e}_\ell)^2\} \\
&\quad -8(2n_j^{-1}\mathrm{tr}(\boldsymbol{\Sigma}_j^3\boldsymbol{\Sigma}_i) + \beta_j n_j^{-1}\sum_{\ell=1}^{p}\mathbf{e}_\ell^T\boldsymbol{\Sigma}_j^{1/2}\boldsymbol{\Sigma}_i\boldsymbol{\Sigma}_j^{1/2}\mathbf{e}_\ell\mathbf{e}_\ell^T\boldsymbol{\Sigma}_j^2\mathbf{e}_\ell)],
\end{aligned}$$

$$
\begin{aligned}
&\sigma_{Ak_1k_2k_2k_3} \\
= \quad & 4[n_{k_2}^{-1}\mathrm{tr}(\mathbf{\Sigma}_{k_2}^2)]^2 + 4[2n_{k_2}^{-1}\mathrm{tr}(\mathbf{\Sigma}_{k_2}^4) + \beta_{k_2} n_{k_2}^{-1}\sum_{\ell=1}^{p}(\mathbf{e}_\ell^T\mathbf{\Sigma}_{k_2}^2\mathbf{e}_\ell)^2] \\
& -4[2n_{k_2}^{-1}\mathrm{tr}(\mathbf{\Sigma}_{k_2}^3\mathbf{\Sigma}_{k_3}) + \beta_{k_2} n_{k_2}^{-1}\sum_{\ell=1}^{p}\mathbf{e}_\ell^T\mathbf{\Sigma}_{k_2}^{1/2}\mathbf{\Sigma}_{k_3}\mathbf{\Sigma}_{k_2}^{1/2}\mathbf{e}_\ell\mathbf{e}_\ell^T\mathbf{\Sigma}_{k_2}^2\mathbf{e}_\ell] \\
& -4[2n_{k_2}^{-1}\mathrm{tr}(\mathbf{\Sigma}_{k_2}^3\mathbf{\Sigma}_{k_1}) + \beta_{k_2} n_{k_2}^{-1}\sum_{\ell=1}^{p}\mathbf{e}_\ell^T\mathbf{\Sigma}_{k_2}^{1/2}\mathbf{\Sigma}_{k_1}\mathbf{\Sigma}_{k_2}^{1/2}\mathbf{e}_\ell\mathbf{e}_\ell^T\mathbf{\Sigma}_{k_2}^2\mathbf{e}_\ell] \\
& +4[2n_{k_2}^{-1}\mathrm{tr}(\mathbf{\Sigma}_{k_1}\mathbf{\Sigma}_{k_2}\mathbf{\Sigma}_{k_3}\mathbf{\Sigma}_{k_2}) + \beta_{k_2} n_{k_2}^{-1}\sum_{\ell=1}^{p}\mathbf{e}_\ell^T\mathbf{\Sigma}_{k_2}^{1/2}\mathbf{\Sigma}_{k_1}\mathbf{\Sigma}_{k_2}^{1/2}\mathbf{e}_\ell\mathbf{e}_\ell^T\mathbf{\Sigma}_{k_2}^{1/2}\mathbf{\Sigma}_{k_3}\mathbf{\Sigma}_{k_2}^{1/2}\mathbf{e}_\ell].
\end{aligned}
$$

.

Proof. Similar to Lemma 1, we first consider $\mathbf{\Sigma}_k$ with the bounded spectral norm for all $k = k_1, k_2, k_3$. When $\mathrm{tr}(\mathbf{\Sigma}_k^q) = O(p^q)$ for $q = 1, 2, 3, 4$ and at least one k in the index set $\{k_1, k_2, k_3\}$, the proof mimics that in the bounded spectral norm case. Similar to the proof of Lemma 1, it can be shown that $\mathrm{tr}(\mathbf{S}_{k_1} - \mathbf{S}_{k_2})^2 - \hat{\mu}_{k_1k_2} - \mu_{Ak_1k_2}$ and $\mathrm{tr}(\mathbf{S}_{k_2} - \mathbf{S}_{k_3})^2 - \hat{\mu}_{k_2k_3} - \mu_{Ak_2k_3}$ are asymptotically distributed as a bivariate normal distribution with the asymptotic variances $\sigma^2_{Ak_1k_2}$, $\sigma^2_{Ak_2k_3}$ and asymptotic covariance $\sigma_{Ak_1k_2k_2k_3}$. For simplicity of presentation, we consider the computation of σ_{A1223}, which is given by

$$
\begin{aligned}
\sigma_{A1223} \quad = \quad & \tilde{\sigma}_{110A} - 2\tilde{\sigma}_{130A} - 2\tilde{\sigma}_{120A} + 4\tilde{\sigma}_{230A} + 4n_2^{-2}(\mathrm{tr}\mathbf{\Sigma}_2)^2\sigma_{660A} - 2(n_2^{-1}\mathrm{tr}\mathbf{\Sigma}_2)\sigma_{460A} \\
& +4(n_2^{-1}\mathrm{tr}\mathbf{\Sigma}_2)\sigma_{560A} - 2(n_2^{-1}\mathrm{tr}\mathbf{\Sigma}_2)\sigma_{460A} + 4(n_2^{-1}\mathrm{tr}\mathbf{\Sigma}_2)\sigma_{760A},
\end{aligned}
$$

where

$$
\begin{aligned}
\sigma_{760A} &= \sum_{\ell=1}^{n_2}[(\mathrm{E}_\ell - \mathrm{E}_{\ell-1})\mathrm{tr}(\mathbf{B}_2\mathbf{B}_3)][(\mathrm{E}_\ell - \mathrm{E}_{\ell-1})\mathrm{tr}(\mathbf{B}_2)], \\
\tilde{\sigma}_{110A} &= \sum_{\ell=1}^{n_2}[(\mathrm{E}_\ell - \mathrm{E}_{\ell-1})\mathrm{tr}(\mathbf{B}_2^2)]^2, \\
\tilde{\sigma}_{120A} &= \sum_{\ell=1}^{n_2}[(\mathrm{E}_\ell - \mathrm{E}_{\ell-1})\mathrm{tr}(\mathbf{B}_2^2)][(\mathrm{E}_\ell - \mathrm{E}_{\ell-1})\mathrm{tr}(\mathbf{B}_2\mathbf{B}_1)], \\
\tilde{\sigma}_{130A} &= \sum_{\ell=1}^{n_2}[(\mathrm{E}_\ell - \mathrm{E}_{\ell-1})\mathrm{tr}(\mathbf{B}_2^2)][(\mathrm{E}_\ell - \mathrm{E}_{\ell-1})\mathrm{tr}(\mathbf{B}_2\mathbf{B}_3)], \\
\tilde{\sigma}_{230A} &= \sum_{\ell=1}^{n_2}[(\mathrm{E}_\ell - \mathrm{E}_{\ell-1})\mathrm{tr}(\mathbf{B}_2\mathbf{B}_1)][(\mathrm{E}_{\ell,\mathbf{B}_3} - \mathrm{E}_{\ell-1,\mathbf{B}_3})\mathrm{tr}(\mathbf{B}_2\mathbf{B}_3)].
\end{aligned}
$$

Then, we have

$$
\begin{aligned}
\sigma_{760A} &= \sum_{\ell=1}^{n_2}[(\mathrm{E}_\ell - \mathrm{E}_{\ell-1})\mathrm{tr}(\mathbf{B}_2\mathbf{B}_3)][(\mathrm{E}_\ell - \mathrm{E}_{\ell-1})\mathrm{tr}(\mathbf{B}_2)] \\
&= [2n_2^{-1}\mathrm{tr}(\mathbf{\Sigma}_3\mathbf{\Sigma}_2^2) + \beta_2 n_2^{-1}\sum_{\ell=1}^{p}(\mathbf{e}_\ell^T\mathbf{\Sigma}_2\mathbf{e}_\ell)(\mathbf{e}_\ell^T\mathbf{\Sigma}_2^{1/2}\mathbf{\Sigma}_3\mathbf{\Sigma}_2^{1/2}\mathbf{e}_\ell)], \\
\tilde{\sigma}_{110A} &= \sum_{\ell=1}^{n_2}[(\mathrm{E}_\ell - \mathrm{E}_{\ell-1})\mathrm{tr}(\mathbf{B}_2^2)]^2 \\
&= 4n_2^{-1}[2\mathrm{tr}(\mathbf{\Sigma}_2^4) + \beta_2\sum_{\ell=1}^{p}(\mathbf{e}_\ell^T\mathbf{\Sigma}_2^2\mathbf{e}_\ell)^2] + 4(n_2^{-1}\mathrm{tr}\mathbf{\Sigma}_2)^2 n_2^{-1}[2\mathrm{tr}(\mathbf{\Sigma}_2^2) + \beta_2\sum_{\ell=1}^{p}(\mathbf{e}_\ell^T\mathbf{\Sigma}_2\mathbf{e}_\ell)^2] \\
&\quad +4[n_2^{-1}\mathrm{tr}(\mathbf{\Sigma}_2^2)]^2 + 8(n_2^{-1}\mathrm{tr}\mathbf{\Sigma}_2)n_2^{-1}[2\mathrm{tr}(\mathbf{\Sigma}_2^3) + \beta_2\sum_{\ell=1}^{p}\mathbf{e}_\ell^T\mathbf{\Sigma}_2^2\mathbf{e}_\ell\mathbf{e}_\ell^T\mathbf{\Sigma}_2\mathbf{e}_\ell],
\end{aligned}
$$

$$
\begin{aligned}
\tilde{\sigma}_{120A} &= \sum_{\ell=1}^{n_2}[(\mathrm{E}_\ell - \mathrm{E}_{\ell-1})\mathrm{tr}(\mathbf{B}_2^2)][(\mathrm{E}_\ell - \mathrm{E}_{\ell-1})\mathrm{tr}(\mathbf{B}_2\mathbf{B}_1)] \\
&= 2n_2^{-1}[2\mathrm{tr}(\boldsymbol{\Sigma}_2^3\mathbf{B}_1) + \beta_2\sum_{\ell=1}^{p}\mathbf{e}_\ell^T\boldsymbol{\Sigma}_2^{1/2}\mathbf{B}_1\boldsymbol{\Sigma}_2^{1/2}\mathbf{e}_\ell\mathbf{e}_\ell^T\boldsymbol{\Sigma}_2^2\mathbf{e}_\ell] \\
&\quad +2(n_2^{-1}\mathrm{tr}\boldsymbol{\Sigma}_2)n_2^{-1}[2\mathrm{tr}(\boldsymbol{\Sigma}_2^2\mathbf{B}_1) + \beta_2\sum_{\ell=1}^{p}\mathbf{e}_\ell^T\boldsymbol{\Sigma}_2\mathbf{e}_\ell\mathbf{e}_\ell^T\boldsymbol{\Sigma}_2^{1/2}\mathbf{B}_1\boldsymbol{\Sigma}_2^{1/2}\mathbf{e}_\ell] \\
&= 2n_2^{-1}[2\mathrm{tr}(\boldsymbol{\Sigma}_2^3\boldsymbol{\Sigma}_1) + \beta_2\sum_{\ell=1}^{p}\mathbf{e}_\ell^T\boldsymbol{\Sigma}_2^{1/2}\boldsymbol{\Sigma}_1\boldsymbol{\Sigma}_2^{1/2}\mathbf{e}_\ell\mathbf{e}_\ell^T\boldsymbol{\Sigma}_2^2\mathbf{e}_\ell] \\
&\quad +2(n_2^{-1}\mathrm{tr}\boldsymbol{\Sigma}_2)n_2^{-1}[2\mathrm{tr}(\boldsymbol{\Sigma}_2^2\boldsymbol{\Sigma}_1) + \beta_2\sum_{\ell=1}^{p}\mathbf{e}_\ell^T\boldsymbol{\Sigma}_2\mathbf{e}_\ell\mathbf{e}_\ell^T\boldsymbol{\Sigma}_2^{1/2}\boldsymbol{\Sigma}_1\boldsymbol{\Sigma}_2^{1/2}\mathbf{e}_\ell] + o_p(1), \\
\tilde{\sigma}_{130A} &= \sum_{\ell=1}^{n_2}[(\mathrm{E}_\ell - \mathrm{E}_{\ell-1})\mathrm{tr}(\mathbf{B}_2^2)][(\mathrm{E}_\ell - \mathrm{E}_{\ell-1})\mathrm{tr}(\mathbf{B}_2\mathbf{B}_3)] \\
&= 2n_2^{-1}[2\mathrm{tr}(\boldsymbol{\Sigma}_2^3\mathbf{B}_3) + \beta_2\sum_{\ell=1}^{p}\mathbf{e}_\ell^T\boldsymbol{\Sigma}_2^{1/2}\mathbf{B}_3\boldsymbol{\Sigma}_2^{1/2}\mathbf{e}_\ell\mathbf{e}_\ell^T\boldsymbol{\Sigma}_2^2\mathbf{e}_\ell] \\
&\quad +2(n_2^{-1}\mathrm{tr}\boldsymbol{\Sigma}_2)n_2^{-1}[2\mathrm{tr}(\boldsymbol{\Sigma}_2^2\mathbf{B}_3) + \beta_2\sum_{\ell=1}^{p}\mathbf{e}_\ell^T\boldsymbol{\Sigma}_2\mathbf{e}_\ell\mathbf{e}_\ell^T\boldsymbol{\Sigma}_2^{1/2}\mathbf{B}_3\boldsymbol{\Sigma}_2^{1/2}\mathbf{e}_\ell] \\
&= 2n_2^{-1}[2\mathrm{tr}(\boldsymbol{\Sigma}_2^3\boldsymbol{\Sigma}_3) + \beta_2\sum_{\ell=1}^{p}\mathbf{e}_\ell^T\boldsymbol{\Sigma}_2^{1/2}\boldsymbol{\Sigma}_3\boldsymbol{\Sigma}_2^{1/2}\mathbf{e}_\ell\mathbf{e}_\ell^T\boldsymbol{\Sigma}_2^2\mathbf{e}_\ell] \\
&\quad +2(n_2^{-1}\mathrm{tr}\boldsymbol{\Sigma}_2)n_2^{-1}[2\mathrm{tr}(\boldsymbol{\Sigma}_2^2\boldsymbol{\Sigma}_3) + \beta_2\sum_{\ell=1}^{p}\mathbf{e}_\ell^T\boldsymbol{\Sigma}_2\mathbf{e}_\ell\mathbf{e}_\ell^T\boldsymbol{\Sigma}_2^{1/2}\boldsymbol{\Sigma}_3\boldsymbol{\Sigma}_2^{1/2}\mathbf{e}_\ell] + o_p(1), \\
\tilde{\sigma}_{230A} &= \sum_{\ell=1}^{n_2}[(\mathrm{E}_\ell - \mathrm{E}_{\ell-1})\mathrm{tr}(\mathbf{B}_2\mathbf{B}_1)][(\mathrm{E}_{\ell,\mathbf{B}_3} - \mathrm{E}_{\ell-1,\mathbf{B}_3})\mathrm{tr}(\mathbf{B}_2\mathbf{B}_3)] \\
&= n_2^{-1}[2\mathrm{tr}(\mathbf{B}_1\boldsymbol{\Sigma}_2\mathbf{B}_3\boldsymbol{\Sigma}_2) + \beta_2\sum_{\ell=1}^{p}\mathbf{e}_\ell^T\boldsymbol{\Sigma}_2^{1/2}\mathbf{B}_1\boldsymbol{\Sigma}_2^{1/2}\mathbf{e}_\ell\mathbf{e}_\ell^T\boldsymbol{\Sigma}_2^{1/2}\mathbf{B}_3\boldsymbol{\Sigma}_2^{1/2}\mathbf{e}_\ell] \\
&= [2n_2^{-1}\mathrm{tr}(\boldsymbol{\Sigma}_1\boldsymbol{\Sigma}_2\boldsymbol{\Sigma}_3\boldsymbol{\Sigma}_2) + \beta_2 n_2^{-1}\sum_{\ell=1}^{p}\mathbf{e}_\ell^T\boldsymbol{\Sigma}_2^{1/2}\boldsymbol{\Sigma}_1\boldsymbol{\Sigma}_2^{1/2}\mathbf{e}_\ell\mathbf{e}_\ell^T\boldsymbol{\Sigma}_2^{1/2}\boldsymbol{\Sigma}_3\boldsymbol{\Sigma}_2^{1/2}\mathbf{e}_\ell] + o_p(1).
\end{aligned}
$$

Thus, we obtain

$$
\begin{aligned}
\sigma_{A1223} \;=\;& 4[n_2^{-1}\mathrm{tr}(\mathbf{\Sigma}_2^2)]^2+4n_2^{-1}[2\mathrm{tr}(\mathbf{\Sigma}_2^4)+\beta_2\sum_{\ell=1}^p(\mathbf{e}_\ell^T\mathbf{\Sigma}_2^2\mathbf{e}_\ell)^2] \\
&+4(n_2^{-1}\mathrm{tr}\mathbf{\Sigma}_2)^2n_2^{-1}[2\mathrm{tr}(\mathbf{\Sigma}_2^2)+\beta_2\sum_{\ell}^p(\mathbf{e}_\ell^T\mathbf{\Sigma}_2\mathbf{e}_\ell)^2] \\
&+8(n_2^{-1}\mathrm{tr}\mathbf{\Sigma}_2)n_2^{-1}[2\mathrm{tr}(\mathbf{\Sigma}_2^3)+\beta_2\sum_{\ell=1}^p\mathbf{e}_\ell^T\mathbf{\Sigma}_2^2\mathbf{e}_\ell\mathbf{e}_\ell^T\mathbf{\Sigma}_2\mathbf{e}_\ell] \\
&-4[2n_2^{-1}\mathrm{tr}(\mathbf{\Sigma}_2^3\mathbf{\Sigma}_3)+\beta_2n_2^{-1}\sum_{\ell=1}^p\mathbf{e}_\ell^T\mathbf{\Sigma}_2^{1/2}\mathbf{\Sigma}_3\mathbf{\Sigma}_2^{1/2}\mathbf{e}_\ell\mathbf{e}_\ell^T\mathbf{\Sigma}_2^2\mathbf{e}_\ell] \\
&-4(n_2^{-1}\mathrm{tr}\mathbf{\Sigma}_2)[2n_2^{-1}\mathrm{tr}(\mathbf{\Sigma}_2^2\mathbf{\Sigma}_3)+\beta_2n_2^{-1}\sum_{\ell=1}^p\mathbf{e}_\ell^T\mathbf{\Sigma}_2\mathbf{e}_\ell\mathbf{e}_\ell^T\mathbf{\Sigma}_2^{1/2}\mathbf{\Sigma}_3\mathbf{\Sigma}_2^{1/2}\mathbf{e}_\ell] \\
&-4[2n_2^{-1}\mathrm{tr}(\mathbf{\Sigma}_2^3\mathbf{\Sigma}_1)+\beta_2n_2^{-1}\sum_{\ell=1}^p\mathbf{e}_\ell^T\mathbf{\Sigma}_2^{1/2}\mathbf{\Sigma}_1\mathbf{\Sigma}_2^{1/2}\mathbf{e}_\ell\mathbf{e}_\ell^T\mathbf{\Sigma}_2^2\mathbf{e}_\ell] \\
&-4(n_2^{-1}\mathrm{tr}\mathbf{\Sigma}_2)[2n_2^{-1}\mathrm{tr}(\mathbf{\Sigma}_2^2\mathbf{\Sigma}_1)+\beta_2n_2^{-1}\sum_{\ell=1}^p\mathbf{e}_\ell^T\mathbf{\Sigma}_2\mathbf{e}_\ell\mathbf{e}_\ell^T\mathbf{\Sigma}_2^{1/2}\mathbf{\Sigma}_1\mathbf{\Sigma}_2^{1/2}\mathbf{e}_\ell] \\
&+4[2n_2^{-1}\mathrm{tr}(\mathbf{\Sigma}_1\mathbf{\Sigma}_2\mathbf{\Sigma}_3\mathbf{\Sigma}_2)+\beta_2n_2^{-1}\sum_{\ell=1}^p\mathbf{e}_\ell^T\mathbf{\Sigma}_2^{1/2}\mathbf{\Sigma}_1\mathbf{\Sigma}_2^{1/2}\mathbf{e}_\ell\mathbf{e}_\ell^T\mathbf{\Sigma}_2^{1/2}\mathbf{\Sigma}_3\mathbf{\Sigma}_2^{1/2}\mathbf{e}_\ell] \\
&+4(n_2^{-1}\mathrm{tr}\mathbf{\Sigma}_2)^2[2n_2^{-1}\mathrm{tr}(\mathbf{\Sigma}_2^2)+\beta_2n_2^{-1}\sum_{\ell=1}^p(\mathbf{e}_\ell^T\mathbf{\Sigma}_2\mathbf{e}_\ell)^2] \\
&-4(n_2^{-1}\mathrm{tr}\mathbf{\Sigma}_2)[2n_2^{-1}\mathrm{tr}(\mathbf{\Sigma}_2^3)+\beta_2n_2^{-1}\sum_{\ell=1}^p\mathbf{e}_\ell^T\mathbf{\Sigma}_2\mathbf{e}_\ell\mathbf{e}_\ell^T\mathbf{\Sigma}_2^2\mathbf{e}_\ell] \\
&-4(n_2^{-1}\mathrm{tr}\mathbf{\Sigma}_2)(n_2^{-1}\mathrm{tr}\mathbf{\Sigma}_2)[2n_2^{-1}\mathrm{tr}(\mathbf{\Sigma}_2^2)+\beta_2n_2^{-1}\sum_{\ell=1}^p(\mathbf{e}_\ell^T\mathbf{\Sigma}_2\mathbf{e}_\ell)^2] \\
&+4(n_2^{-1}\mathrm{tr}\mathbf{\Sigma}_2)[2n_2^{-1}\mathrm{tr}(\mathbf{\Sigma}_1\mathbf{\Sigma}_2^2)+\beta_2n_2^{-1}\sum_{\ell=1}^p(\mathbf{e}_\ell^T\mathbf{\Sigma}_2\mathbf{e}_\ell)(\mathbf{e}_\ell^T\mathbf{\Sigma}_2^{1/2}\mathbf{\Sigma}_1\mathbf{\Sigma}_2^{1/2}\mathbf{e}_\ell)] \\
&-4(n_2^{-1}\mathrm{tr}\mathbf{\Sigma}_2)[2n_2^{-1}\mathrm{tr}(\mathbf{\Sigma}_2^3)+\beta_2n_2^{-1}\sum_{\ell=1}^p\mathbf{e}_\ell^T\mathbf{\Sigma}_2\mathbf{e}_\ell\mathbf{e}_\ell^T\mathbf{\Sigma}_2^2\mathbf{e}_\ell] \\
&-4(n_2^{-1}\mathrm{tr}\mathbf{\Sigma}_2)(n_2^{-1}\mathrm{tr}\mathbf{\Sigma}_2)[2n_2^{-1}\mathrm{tr}(\mathbf{\Sigma}_2^2)+\beta_2n_2^{-1}\sum_{\ell=1}^p(\mathbf{e}_\ell^T\mathbf{\Sigma}_2\mathbf{e}_\ell)^2] \\
&+4(n_2^{-1}\mathrm{tr}\mathbf{\Sigma}_2)[2n_2^{-1}\mathrm{tr}(\mathbf{\Sigma}_3\mathbf{\Sigma}_2^2)+\beta_2n_2^{-1}\sum_{\ell=1}^p(\mathbf{e}_\ell^T\mathbf{\Sigma}_2\mathbf{e}_\ell)(\mathbf{e}_\ell^T\mathbf{\Sigma}_2^{1/2}\mathbf{\Sigma}_3\mathbf{\Sigma}_2^{1/2}\mathbf{e}_\ell)]
\end{aligned}
$$

$$
\begin{aligned}
&= 4[n_2^{-1}\mathrm{tr}(\mathbf{\Sigma}_2^2)]^2 + 4[2n_2^{-1}\mathrm{tr}(\mathbf{\Sigma}_2^4) + \beta_2 n_2^{-1}\sum_{\ell=1}^{p}(\mathbf{e}_\ell^T\mathbf{\Sigma}_2^2\mathbf{e}_\ell)^2] \\
&\quad -4[2n_2^{-1}\mathrm{tr}(\mathbf{\Sigma}_2^3\mathbf{\Sigma}_3) + \beta_2 n_2^{-1}\sum_{\ell=1}^{p}\mathbf{e}_\ell^T\mathbf{\Sigma}_2^{1/2}\mathbf{\Sigma}_3\mathbf{\Sigma}_2^{1/2}\mathbf{e}_\ell\mathbf{e}_\ell^T\mathbf{\Sigma}_2^2\mathbf{e}_\ell] \\
&\quad -4[2n_2^{-1}\mathrm{tr}(\mathbf{\Sigma}_2^3\mathbf{\Sigma}_1) + \beta_2 n_2^{-1}\sum_{\ell=1}^{p}\mathbf{e}_\ell^T\mathbf{\Sigma}_2^{1/2}\mathbf{\Sigma}_1\mathbf{\Sigma}_2^{1/2}\mathbf{e}_\ell\mathbf{e}_\ell^T\mathbf{\Sigma}_2^2\mathbf{e}_\ell] \\
&\quad +4[2n_2^{-1}\mathrm{tr}(\mathbf{\Sigma}_1\mathbf{\Sigma}_2\mathbf{\Sigma}_3\mathbf{\Sigma}_2) + \beta_2 n_2^{-1}\sum_{\ell=1}^{p}\mathbf{e}_\ell^T\mathbf{\Sigma}_2^{1/2}\mathbf{\Sigma}_1\mathbf{\Sigma}_2^{1/2}\mathbf{e}_\ell\mathbf{e}_\ell^T\mathbf{\Sigma}_2^{1/2}\mathbf{\Sigma}_3\mathbf{\Sigma}_2^{1/2}\mathbf{e}_\ell].
\end{aligned}
$$

Generally, $\sigma_{Ak_1k_2k_2k_3}$ is obtained by replacing n_i, $\mathbf{\Sigma}_i$ and β_i by n_{k_i}, $\mathbf{\Sigma}_{k_i}$ and β_{k_i} in σ_{A1223}. That is,

$$
\begin{aligned}
\sigma_{Ak_1k_2k_2k_3} &= 4[n_{k_2}^{-1}\mathrm{tr}(\mathbf{\Sigma}_{k_2}^2)]^2 + 4[2n_{k_2}^{-1}\mathrm{tr}(\mathbf{\Sigma}_{k_2}^4) + \beta_{k_2} n_{k_2}^{-1}\sum_{\ell=1}^{p}(\mathbf{e}_\ell^T\mathbf{\Sigma}_{k_2}^2\mathbf{e}_\ell)^2] \\
&\quad -4[2n_{k_2}^{-1}\mathrm{tr}(\mathbf{\Sigma}_{k_2}^3\mathbf{\Sigma}_{k_3}) + \beta_{k_2} n_{k_2}^{-1}\sum_{\ell=1}^{p}\mathbf{e}_\ell^T\mathbf{\Sigma}_{k_2}^{1/2}\mathbf{\Sigma}_{k_3}\mathbf{\Sigma}_{k_2}^{1/2}\mathbf{e}_\ell\mathbf{e}_\ell^T\mathbf{\Sigma}_{k_2}^2\mathbf{e}_\ell] \\
&\quad -4[2n_{k_2}^{-1}\mathrm{tr}(\mathbf{\Sigma}_{k_2}^3\mathbf{\Sigma}_{k_1}) + \beta_{k_2} n_{k_2}^{-1}\sum_{\ell=1}^{p}\mathbf{e}_\ell^T\mathbf{\Sigma}_{k_2}^{1/2}\mathbf{\Sigma}_{k_1}\mathbf{\Sigma}_{k_2}^{1/2}\mathbf{e}_\ell\mathbf{e}_\ell^T\mathbf{\Sigma}_{k_2}^2\mathbf{e}_\ell] \\
&\quad +4[2n_{k_2}^{-1}\mathrm{tr}(\mathbf{\Sigma}_{k_1}\mathbf{\Sigma}_{k_2}\mathbf{\Sigma}_{k_3}\mathbf{\Sigma}_{k_2}) + \beta_{k_2} n_{k_2}^{-1}\sum_{\ell=1}^{p}\mathbf{e}_\ell^T\mathbf{\Sigma}_{k_2}^{1/2}\mathbf{\Sigma}_{k_1}\mathbf{\Sigma}_{k_2}^{1/2}\mathbf{e}_\ell\mathbf{e}_\ell^T\mathbf{\Sigma}_{k_2}^{1/2}\mathbf{\Sigma}_{k_3}\mathbf{\Sigma}_{k_2}^{1/2}\mathbf{e}_\ell].
\end{aligned}
$$

The proof of Lemma 2 is completed.

S2 Proofs of Theorem 2 and Theorem 1

By Lemma 2 and the delta method, under the conditions of Lemma 2, we have

$$\sigma_{AK}^{-1}(T_{K1} - \hat{\mu}_{K1} - \mu_{AK}) \xrightarrow{d} N(0,1),$$

where

$$\hat{\mu}_{K1} = \sum_{1\le k_1<k_2\le K} \omega_{k_1k_2}\hat{\mu}_{k_1k_2} = \sum_{1\le k_1<k_2\le K} \omega_{k_1k_2}\Big[\sum_{k=k_1,k_2}(n_k^2 - n_k - 1)n_k^{-1}(n_k - 1)^{-2}(\mathrm{tr}\mathbf{S}_k)^2\Big],$$

and $\mu_{AK} = \sum_{1\le i<j\le K}\omega_{ij}\mu_{Aij}$ and

$$\begin{aligned}\sigma_{AK}^2 &= \sum_{1\le i<j\le K}\omega_{ij}^2\sigma_{Aij}^2 + 2\sum_{i<j<k}\omega_{ij}\omega_{jk}\sigma_{Aijjk} \\ &\quad +2\sum_{i<k<j}\omega_{ij}\omega_{jk}\sigma_{Aijjk} + 2\sum_{j<i<k}\omega_{ij}\omega_{jk}\sigma_{Aijjk},\end{aligned}$$

with the weights $\{\omega_{ij}, 1\le i,j\le K\}$ and $\omega_{ij} = \omega_{ji}$. By (32) and (33) in Cai, Liu and Xia (2013), we have

$$\begin{aligned}&\max_{1\le \ell_1\le \ell_2\le p}\delta_{k_1k_2\ell_1\ell_2} - s(n_{k_1}, n_{k_2}, p) \\ \overset{a.s.}{\ge}\ & 0.5\max_{1\le i\le j\le p}\frac{(\sigma_{k_1ij}-\sigma_{k_2ij})^2}{\hat{\theta}_{k_1ij}/n_{k_1}+\hat{\theta}_{k_2ij}/n_{k_2}} - s(n_{k_1}, n_{k_2}, p) - 4\log p + 0.5\log\log p \\ \overset{a.s.}{\ge}\ & 0.5\max_{1\le i\le j\le p}\frac{(\sigma_{k_1ij}-\sigma_{k_2ij})^2}{\theta_{k_1ij}/n_{k_1}+\theta_{k_2ij}/n_{k_2}} - s(n_{k_1}, n_{k_2}, p) - 4\log p + 0.5\log\log p\end{aligned}$$

Then, there exists a pair of $1\le k_1, k_2\le K$ satisfying

$$0.5\max_{1\le i\le j\le p}\frac{(\sigma_{k_1ij}-\sigma_{k_2ij})^2}{\theta_{k_1ij}/n_{k_1}+\theta_{k_2ij}/n_{k_2}} \ge s(n_1, n_2, p) + 4\log p.$$

Then we have

$$\max_{1\le \ell_1\le \ell_2\le p}\delta_{k_1k_2\ell_1\ell_2} - s(n_{k_1}, n_{k_2}, p) \overset{a.s.}{\ge} 0.5\log\log p$$

and

$$\max_{1\le \ell_1\le \ell_2\le p}\delta_{k_1k_2\ell_1\ell_2} - s(n_{k_1}, n_{k_2}, p) \overset{a.s.}{>} 0$$

Therefore, we have

$$\max_{1\le i<j\le K}\{I\{\max_{1\le \ell_1\le \ell_2\le p}\delta_{ij\ell_1\ell_2} > s(n_i,n_j,p)\}\} \stackrel{a.s.}{=} 1.$$

That is, $T_{K2} \stackrel{a.s.}{=} K_0$. Then, we have

$$\sigma_{AK}^{-1}(T_K - K_0 - \hat{\mu}_{K1} - \mu_{AK}) \xrightarrow{d} N(0,1).$$

We first focus on the proof of the case with $K=2$, and that of $K \ge 3$ can be shown in a similar way. When $K=2$, the power function is

$$\begin{aligned} g_2(\mathbf{\Sigma}_1,\mathbf{\Sigma}_2) &= P_{H_{A2}}(T_2 - \hat{\mu}_{21} > \hat{\mu}_2 + z_{1-\alpha}\hat{\sigma}_2) \\ &= P_{H_{A2}}\left(\frac{T_2-\hat{\mu}_{21}-\mu_{A2}}{\sigma_{A2}} > \frac{\hat{\mu}_2-\mu_{A2}+z_{1-\alpha}\hat{\sigma}_2}{\sigma_{A2}}\right) \\ &> P_{H_{A2}}\left(\frac{T_{21}-\hat{\mu}_{21}-\mu_{A2}}{\sigma_{A2}} > \frac{z_{1-\alpha}\hat{\sigma}_2}{\sigma_{A2}}\right). \end{aligned}$$

Step 1. When $p^{-1}\mathrm{tr}\mathbf{A}_{12}^2 \to 0$ and $\mathrm{tr}\mathbf{A}_{12}^2 > c_0$ for some positive constant c_0, $\sigma_{A2}/\hat{\sigma}_2 \to 1$. That is, $(z_{1-\alpha}\hat{\sigma}_2)/\sigma_{A2} - z_{1-\alpha} = o_p(1)$. Then, when n_1, n_2 are large enough, we have $g_2(\mathbf{\Sigma}_1,\mathbf{\Sigma}_2) > \alpha$.

Step 2. We have $g_2(\mathbf{\Sigma}_1,\mathbf{\Sigma}_2) > P_{H_{A2}}((T_{21} - \hat{\mu}_{21} - \mu_{A2})\sigma_{A2}^{-1} > (\hat{\mu}_2 - \mu_{A2} + z_{1-\alpha}\hat{\sigma}_2)\sigma_{A2}^{-1})$. When $\mathrm{tr}(\mathbf{A}_{12}^2) \to \infty$, we have $(\hat{\mu}_2 - \mu_{A2})\sigma_{A2}^{-1} \to -\infty$ in probability. Then, the power function satisfies $g_2(\mathbf{\Sigma}_1,\mathbf{\Sigma}_2) \to 1$.

Step 3. We have

$$
\begin{aligned}
g_2(\mathbf{\Sigma}_1, \mathbf{\Sigma}_2) &= P_{H_{A2}}(T_2 - \hat{\mu}_{21} > \hat{\mu}_2 + z_{1-\alpha}\hat{\sigma}_2) \\
&= P_{H_{A2}}\left(\frac{T_2 - \hat{\mu}_{21} - \mu_{A2}}{\sigma_{A2}} > \frac{\hat{\mu}_2 - \mu_{A2} + z_{1-\alpha}\hat{\sigma}_2}{\sigma_{A2}}\right) \\
&> P_{H_{A2}}\left(\frac{T_2 - \hat{\mu}_{21} - \mu_{A2}}{\sigma_{A2}} > \frac{z_{1-\alpha}\hat{\sigma}_2}{\sigma_{A2}}\right) \\
&> P_{H_{A2}}\left(\frac{T_{21} - \hat{\mu}_{21} - \mu_{A2}}{\sigma_{A2}} + \frac{T_{K2}}{\sigma_{A2}} > \frac{z_{1-\alpha}\hat{\sigma}_2}{\sigma_{A2}}\right).
\end{aligned}
$$

Because $\frac{T_{21} - \hat{\mu}_{21} - \mu_{A2}}{\sigma_{A2}}$ is asymptotically normal under H_1 and $T_{K2} \stackrel{a.s.}{=} K_0$, if

$$
0.5 \max_{1 \le i \le j \le p} \frac{(\sigma_{k_1 ij} - \sigma_{k_2 ij})^2}{\theta_{k_1 ij}/n_{k_1} + \theta_{k_2 ij}/n_{k_2}} \ge s(n_1, n_2, p) + 4\log p,
$$

then the power function $g_2(\mathbf{\Sigma}_1, \mathbf{\Sigma}_2)$ tends to one.

The proof of Theorem 2 is completed. Moreover, Theorem 1 is a special case of Theorem 2.

Bibliography

Bai, Z. D. and Silverstein, J. W. (2004). CLT for linear spectral statistics of large dimensional sample covariance matrices. *Ann. Probab. 32*, pp. 553-605.

Cai, T. T., Liu, W. D. and Xia, Y. (2013). Two-sample covariance matrix testing and support recovery in high dimensional and sparse settings. *J. Am. Stat. Assoc. 108*, pp. 265-277.

Li, J. and Chen, S. X. (2012). Two-sample tests for high dimensional covariance matrices. *Ann. Stat. 40*, pp. 908-940.

Schott, J. R. (2007). A test for the equality of covariance matrices when the dimension is large relative to the sample sizes. *Comput. Stat. Data Anal. 51*, pp. 6535-6542.

Srivastava, M. S. and Yanagihara, H. (2010). Testing the equality of several covariance matrices with fewer observations than the dimension. *J. Multivar. Anal. 101*, pp. 1319-1329.

Yang Q. and Pan G. (2017). Weighted statistic in detecting faint and sparse alternatives for high-dimensional covariance matrices. *J. Am. Stat. Assoc. 517*, pp. 188–200.

Table 1: Comparison of empirical sizes and power (in percentage) of the proposed two-sample test T_2 with four existing methods under Scenario 1.

			Gaussian					Gamma			
p	n_1	n_2	T_2	CLX	LC	SC	YP	T_2	CLX	LC	YP
Size (%)											
100	60	60	6.1	5.3	4.9	5.2	7.1	5.8	3.3	6.3	21.4
	100	200	5.8	5.5	4.2	5.0	—	5.4	3.6	6.6	—
	200	200	6.3	4.9	3.8	5.5	5.0	5.7	3.4	4.2	13.5
	300	300	5.9	4.6	5.4	5.4	6.0	5.9	3.8	5.2	11.9
300	60	60	5.9	6.0	4.2	4.9	10.5	5.0	3.4	4.6	22.5
	100	200	5.2	4.5	7.8	5.2	—	4.9	3.2	5.2	—
	200	200	5.6	4.4	4.2	5.2	4.1	5.4	2.8	6.0	13.9
	300	300	5.6	4.2	5.6	5.4	6.7	5.4	3.2	4.0	13.4
Power (%)											
100	60	60	63.8	6.9	62.2	66.2	94.6	63.3	4.7	62.8	95.6
	100	200	68.7	8.7	68.8	70.2	—	67.1	6.2	66.4	—
	200	200	32.1	6.2	30.8	32.0	36.8	31.9	4.4	29.8	44.0
	300	300	26.1	5.6	27.2	26.1	24.2	26.1	4.7	26.8	36.0
300	60	60	99.9	7.5	100.0	100.0	99.8	99.9	4.9	100.0	97.5
	100	200	100.0	9.9	100.0	100.0	—	100.0	7.4	100.0	—
	200	200	96.4	5.6	95.8	96.6	100.0	95.8	3.5	96.0	100.0
	300	300	87.4	5.8	88.6	87.7	99.9	87.6	3.8	86.8	100.0

Note: Four existing tests include Yang and Pan (2017) (YP), Li and Chen (2012) (LC), Cai, Liu and Xia (2013) (CLX) and Schott (2007) (SC), and "—" denotes "not applicable".

Table 2: Comparison of empirical sizes and power (in percentage) of the proposed two-sample test $T_2$with four existing methods under Scenario 2.

			Gaussian					Gamma			
p	n_1	n_2	T_2	CLX	LC	SC	YP	T_2	CLX	LC	YP
			Size (%)								
100	60	60	6.6	5.7	5.0	5.7	100.0	6.0	3.8	6.1	100.0
	100	200	5.5	5.9	4.8	4.4	—	5.8	4.4	5.6	—
	200	200	6.3	4.9	4.0	5.7	99.8	6.1	3.7	6.6	97.7
	300	300	5.6	4.7	2.6	5.3	70.6	5.4	3.3	5.4	96.3
300	60	60	5.8	5.4	5.6	4.9	100.0	5.1	3.4	4.9	100.0
	100	200	5.8	4.9	5.6	5.1	—	6.5	3.8	4.4	—
	200	200	6.1	4.3	4.4	6.0	100.0	5.8	3.2	4.8	100.0
	300	300	5.4	4.6	6.8	5.1	53.5	5.0	3.7	4.4	99.9
			Power (%)								
100	60	60	37.9	8.4	38.5	38.9	100.0	38.0	6.5	40.4	100.0
	100	200	90.0	56.3	89.8	88.7	—	88.8	53.5	86.2	—
	200	200	99.3	53.7	99.6	99.3	100.0	99.1	42.8	99.2	99.9
	300	300	100.0	91.8	100.0	100.0	99.8	100.0	81.0	100.0	95.3
300	60	60	36.2	6.4	37.8	39.0	100.0	36.8	4.8	38.4	100.0
	100	200	91.5	53.7	85.0	90.5	—	91.4	47.9	89.4	—
	200	200	99.7	44.0	99.8	99.7	100.0	99.7	31.4	99.2	100.0
	300	300	100.0	90.3	100.0	100.0	100.0	100.0	75.7	100.0	99.7

Table 3: Comparison of empirical sizes and power (in percentage) of the proposed two-sample test T_2 with four existing methods under Scenario 3.

			Gaussian					Gamma			
p	n_1	n_2	T_2	CLX	LC	SC	YP	T_2	CLX	LC	YP
						Size (%)					
100	60	60	6.3	5.2	5.6	5.2	100.0	5.5	3.3	4.5	100.0
	100	200	5.9	4.5	5.0	5.3	—	5.7	3.5	7.0	—
	200	200	5.8	4.6	6.8	5.6	50.0	5.8	3.1	5.4	99.5
	300	300	5.6	4.4	5.4	5.1	9.4	5.5	4.0	3.6	30.8
300	60	60	5.4	5.4	5.1	4.8	99.7	4.6	3.1	5.2	100.0
	100	200	5.6	5.4	6.2	5.0	—	5.0	3.5	5.8	—
	200	200	5.5	4.4	4.4	4.9	77.3	5.2	2.6	6.4	9.5
	300	300	5.2	4.7	5.6	5.0	23.9	5.5	3.5	4.6	80.6
							Power (%)				
100	60	60	43.5	53.8	7.9	11.5	100.0	37.2	42.9	16.0	100.0
	100	200	99.0	99.6	13.4	24.5	—	96.1	98.3	29.8	—
	200	200	100.0	100.0	43.2	41.2	100.0	99.8	99.9	36.6	100.0
	300	300	100.0	100.0	85.4	66.2	100.0	100.0	100.0	88.6	100.0
300	60	60	24.2	33.3	5.9	6.5	100.0	20.0	25.7	6.6	100.0
	100	200	95.8	98.5	13.6	8.9	—	89.9	95.4	11.6	—
	200	200	99.9	100.0	6.8	12.0	99.9	99.1	99.7	11.0	100.0
	300	300	100.0	100.0	27.0	18.2	100.0	100.0	100.0	12.6	100.0

Table 4: Comparison of empirical sizes and power (in percentage) of the proposed two-sample test T_2 with four existing methods under Scenario 4.

			Gaussian					Gamma			
p	n_1	n_2	T_2	CLX	LC	SC	YP	T_2	CLX	LC	YP
						Size (%)					
100	60	60	6.0	5.4	4.5	5.3	8.0	5.0	3.8	5.5	17.8
	100	200	5.5	4.5	5.0	5.2	—	6.2	3.9	6.0	—
	200	200	5.8	4.2	7.2	5.2	6.5	6.1	3.3	6.0	11.6
	300	300	5.1	4.7	4.4	4.8	5.5	5.5	3.8	5.0	12.8
300	60	60	6.0	5.4	4.7	4.9	7.7	4.9	3.0	6.6	20.1
	100	200	5.2	5.5	4.4	4.6	—	5.3	3.5	4.6	—
	200	200	4.8	5.1	4.8	4.6	5.8	5.4	2.7	5.2	11.3
	300	300	5.3	4.4	5.4	5.0	5.7	5.4	3.3	4.8	13.7
						Power (%)					
100	60	60	36.2	17.6	31.0	33.8	86.7	32.9	8.7	34.4	84.9
	100	200	96.1	98.1	47.2	48.7	—	71.7	69.4	48.3	—
	200	200	99.2	99.8	67.2	70.3	99.7	85.5	79.6	65.2	97.5
	300	300	100.0	100.0	87.2	89.5	99.9	98.1	97.3	83.0	99.2
300	60	60	65.1	16.1	62.0	67.2	96.7	65.0	7.7	65.4	79.3
	100	200	99.3	99.7	57.0	55.5	—	81.9	76.5	80.3	—
	200	200	99.8	99.9	68.0	68.0	100.0	89.2	83.8	67.2	99.8
	300	300	100.0	100.0	78.5	79.7	100.0	98.6	98.1	75.6	100.0

Table 5: Empirical sizes and empirical power (in percentage) of the proposed two-sample test T_2 for ultra high-dimensional cases under Scenarios 1–4.

			Size (%)				Power (%)			
p	n_1	n_2	1	2	3	4	1	2	3	4
500	60	60	5.5	5.8	5.9	5.5	100.0	35.5	18.8	84.7
	100	200	5.0	5.2	5.5	5.6	100.0	91.8	93.7	99.7
	200	200	4.9	5.5	5.4	5.0	100.0	99.8	99.8	99.9
	300	300	5.5	5.2	5.4	5.1	100.0	100.0	100.0	100.0
1000	60	60	4.8	5.2	5.3	4.5	100.0	34.5	14.1	98.1
	100	200	5.1	5.3	4.9	5.3	100.0	92.4	90.0	100.0
	200	200	5.7	5.4	4.3	4.7	100.0	99.5	99.4	100.0
	300	300	4.9	5.2	5.4	4.9	100.0	100.0	100.0	100.0

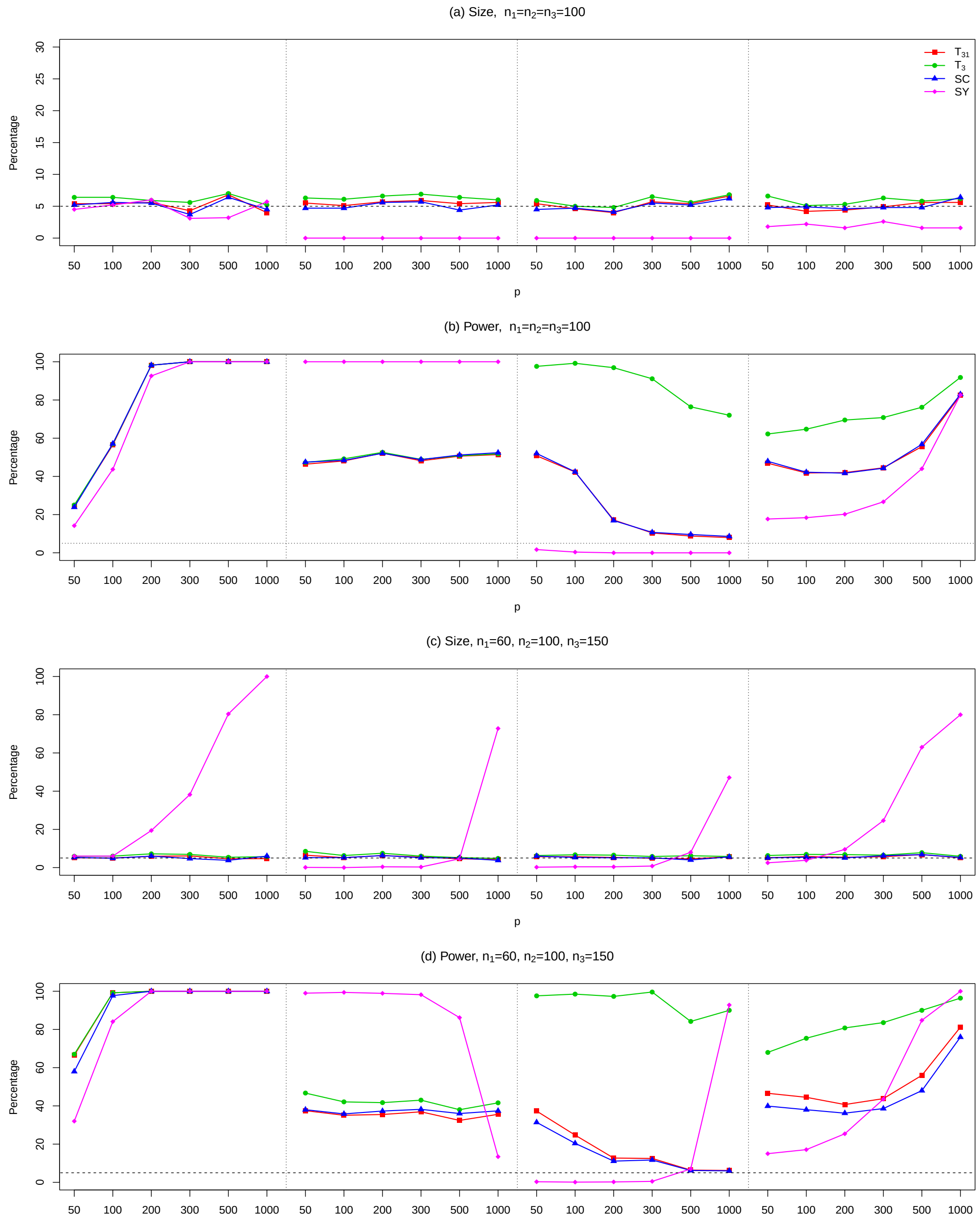

Figure 1: Simulation results for testing the equality of three covariance matrices with Gaussian populations under Scenarios 1–4 in comparison with two existing tests of Schott (2007) (SC) and Srivastava and Yanagihara (2010) (SY).

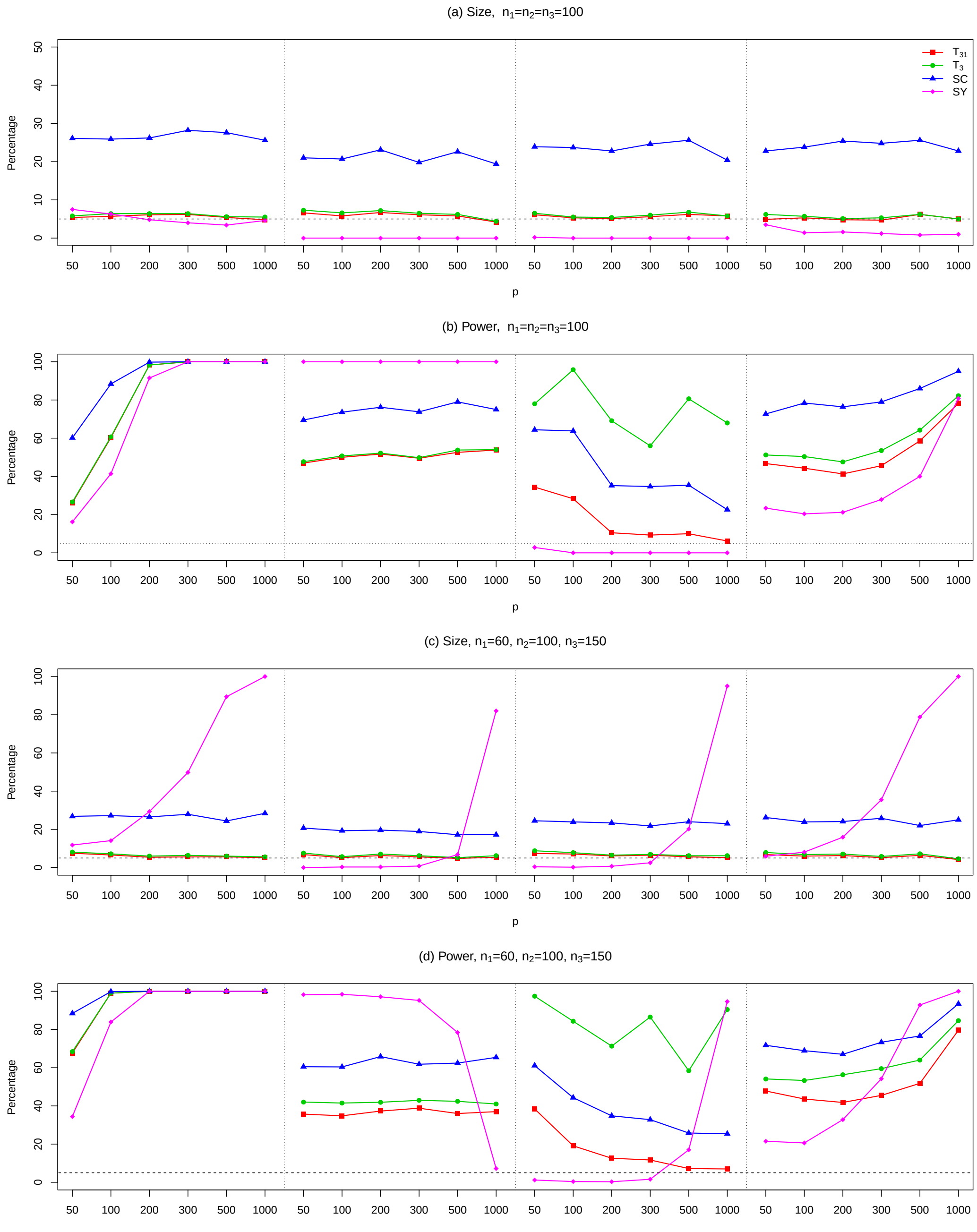

Figure 2: Simulation results for testing the equality of three covariance matrices with Gamma populations under Scenarios 1–4 in comparison with two existing tests of Schott (2007) (SC) and Srivastava and Yanagihara (2010) (SY).