# Simultaneous testing of the mean vector and the covariance matrix for high-dimensional data

Masashi Hyodo*and Takahiro Nishiyama†

March 7, 2017

**Abstract**

In this article, we propose an $L^2$-norm-based test for simultaneous testing of the mean vector and the covariance matrix under high-dimensional non-normal populations. To construct this, we derive an asymptotic distribution of a test statistic based on both differences mean vectors and covariance matrices. We also investigate the asymptotic sizes and powers of the proposed tests using this result. Finally, we study the finite sample and dimension performance of this test via Monte Carlo simulations.

## 1 Introduction

Let $\boldsymbol{X}_{g1}, \boldsymbol{X}_{g2}, \ldots, \boldsymbol{X}_{gn_g}$ be $p$-dimensional random vectors from the $g$-th population $(g = 1, 2)$. We denote $g$-th population mean vector with $\boldsymbol{\mu}_g$, and $g$-th population covariance matrix with $\Sigma_g$. Assume that the vector $\boldsymbol{X}_{gi}$ have the following model:

$$\boldsymbol{X}_{gi} = \Sigma_g^{1/2}\boldsymbol{Z}_{gi} + \boldsymbol{\mu}_g \ \text{ for } i = 1, \ldots, n_g, \tag{1.1}$$

where $\boldsymbol{Z}_{gi}$ represents $p$-dimensional random vectors such that $\mathrm{E}[\boldsymbol{Z}_{gi}] = \boldsymbol{0}$, $\mathrm{Var}[\boldsymbol{Z}_g] = I_p$, a $p \times p$ identity matrix, and $\boldsymbol{Z}_{11}, \ldots, \boldsymbol{Z}_{1n_1}, \boldsymbol{Z}_{21}, \ldots, \boldsymbol{Z}_{2n_2}$ are mutually independent.

As part of the effort to discover significant differences between two high-dimensional distributions, we develop in this paper two-sample simultaneous test procedure for high-dimensional mean vectors and covariance matrices. Our primary interest is to test

$$H_0 : \boldsymbol{\mu}_1 = \boldsymbol{\mu}_2,\ \Sigma_1 = \Sigma_2 \text{ vs. } H_1 : \text{not } H_0. \tag{1.2}$$

Testing the above hypothesis when $p > \min\{n_1 - 1, n_2 - 1\}$ is a nontrivial statistical problem. When $p \leq \min\{n_1 - 1, n_2 - 1\}$, the likelihood ratio test (see Muirhead [8]) may be used for the above hypothesis. If we let

$$\overline{\boldsymbol{X}}_g = \frac{1}{n_g}\sum_{i=1}^{n_g}\boldsymbol{X}_{gi},\ W_g = \sum_{i=1}^{n_g}(\boldsymbol{X}_{gi} - \overline{\boldsymbol{X}}_g)(\boldsymbol{X}_{gi} - \overline{\boldsymbol{X}}_g)'$$

*Department of Mathematical Sciences, Graduate School of Engineering, Osaka Prefecture University, 1-1 Gakuen-cho, Naka-ku, Sakai, Osaka 599-8531, Japan. E-Mail:hyodo@ms.osakafu-u.ac.jp

†Department of Business Administration, Senshu University, 2-1-1, Higashimita, Tama-ku, Kawasaki-shi, Kanagawa 214-8580, Japan.

then the likelihood ratio for $H_0$ is

$$\lambda = \frac{\prod_{g=1}^{2} |W_g|^{n_g/2}}{|(n_1 n_2)/n(\overline{\boldsymbol{X}}_1 - \overline{\boldsymbol{X}}_2)(\overline{\boldsymbol{X}}_1 - \overline{\boldsymbol{X}}_2)' + \sum_{g=1}^{2} W_g|^{n/2}} \frac{n^{pn/2}}{\prod_{g=1}^{2} n_g^{pn_g/2}},$$

where $n = n_1 + n_2$. When $p > \min\{n_1 - 1, n_2 - 1\}$, at least one of the matrices $W_g$ is singular. This causes the likelihood ratio statistic $-2\log(\lambda)$ to be either infinite or undefined. Therefore, we need to consider other methods.

Note that test (1.2) is equivalent to the following test

$$H_0 : \|\boldsymbol{\delta}\|^2 = 0,\ \|\Delta\|_F^2 = 0 \text{ vs. } H_1 : \text{not } H_0,$$

where $\boldsymbol{\delta} = \boldsymbol{\mu}_1 - \boldsymbol{\mu}_2$ and $\Delta = \Sigma_1 - \Sigma_2$. Here, $\|\cdot\|$ denotes the Euclidean norm, and $\|\cdot\|_F$ denotes the Frobenius norm. The unbiased estimators of $\|\boldsymbol{\delta}\|^2$ and $\|\Delta\|_F^2$ are given as follows:

$$\begin{aligned} \widehat{\|\boldsymbol{\delta}\|^2} &= \|\overline{\boldsymbol{X}}_1 - \overline{\boldsymbol{X}}_2\|^2 - \frac{\mathrm{tr}S_1}{n_1} - \frac{\mathrm{tr}S_2}{n_2}, \\ \widehat{\|\Delta\|_F^2} &= \sum_{g=1}^{2} \widehat{\mathrm{tr}\Sigma_g^2} - 2\mathrm{tr}(S_1 S_2), \end{aligned}$$

where

$$\widehat{\mathrm{tr}\Sigma_g^2} = \frac{n_g - 1}{n_g(n_g-2)(n_g-3)}\{(n_g-1)(n_g-2)\mathrm{tr}S_g^2 + (\mathrm{tr}S_g)^2 - n_g K_g\}.$$

Here,

$$S_g = \frac{1}{n_g - 1} W_g,\ K_g = \frac{1}{n_g - 1}\sum_{i=1}^{n_g} \|\boldsymbol{X}_{gi} - \overline{\boldsymbol{X}}_g\|^4.$$

The unbiased estimator $\widehat{\mathrm{tr}\Sigma_g^2}$ is proposed in Himeno and Yamada [5] and Srivastava et al. [9]. These estimators are useful considering the following three points:

(i) These estimators are unbiased without the normality assumption.

(ii) It can be defined even when $p > \min\{n_1 - 1, n_2 - 1\}$.

(iii) Under appropriate assumptions, these estimators have asymptotic normality when $p$, $n_1$, and $n_2$ are large.

In high-dimensional settings, the unbiased estimators $\widehat{\|\boldsymbol{\delta}\|^2}$ and $\widehat{\|\Delta\|_F^2}$ are used in two-sample tests for mean vectors ($H_0 : \boldsymbol{\delta} = \mathbf{0}$ vs. $H_1 : \boldsymbol{\delta} \neq \mathbf{0}$) in Chen and Qin [1] and for covariance matrices ($H_0 : \Delta = O$ vs. $H_1 : \Delta \neq O$) in Li and Chen [7], respectively. In this paper, we consider a simultaneous test for these two nonexclusive hypotheses. In order to construct this test, we need to investigate the joint distribution of $\widehat{\|\boldsymbol{\delta}\|^2}$ and $\widehat{\|\Delta\|_F^2}$. We show the joint asymptotic normality of $\widehat{\|\boldsymbol{\delta}\|^2}$ and $\widehat{\|\Delta\|_F^2}$ under appropriate assumptions, and propose the approximate simultaneous test.

The rest of the paper is organized as follows. Section 2 presents the test procedure and its asymptotic size and power after establishing the asymptotic normality of the test statistic. In Section 3, the attained significance levels and powers of the suggested test are empirically analyzed. We also provide an example to apply our test. Finally, Section 4 concludes this paper. Some technical details are relegated to the appendix.

# 2 Simultaneous testing of the mean vector and the covariance matrix in high-dimensional data

## 2.1 Test statistic

Before formulating the test statistic for hypothesis $H_0$, we make several assumptions to investigate the asymptotic properties of the unbiased estimators $\widehat{\|\boldsymbol{\delta}\|^2}$ and $\widehat{\|\Delta\|_F^2}$.

We assume that the $j$-th element of $\boldsymbol{Z}_{gi}$ (denotes $Z_{gij}$) has a uniformly bounded 8th moment, and for any positive integers $r$ and $\alpha_\ell$ such that $\sum_{\ell=1}^r \alpha_\ell \leq 8$, $\mathrm{E}[\prod_{\ell=1}^r Z_{gij_\ell}^{\alpha_\ell}] = \prod_{\ell=1}^r \mathrm{E}[Z_{gij_\ell}^{\alpha_\ell}]$ whenever $j_1, j_2, \ldots, j_r$ are distinct indices. We also make the following assumptions:

(A1) Let $n_1$ and $n_2$ be functions of $p$. Then, $\min\{n_1, n_2\} \to \infty$ as $p \to \infty$.

(A2) Let $\mathrm{tr}(\Sigma_g\Sigma_h)$, and $\mathrm{tr}(\Sigma_g\Sigma_h)^2$ be functions of $p$. Then, for $g, h = 1, 2$,

$$\lim_{p\to\infty} \frac{\mathrm{tr}(\Sigma_g\Sigma_h)^2}{\{\mathrm{tr}(\Sigma_g\Sigma_h)\}^2} = 0.$$

(A3) Let $n_g$, $\boldsymbol{\delta}'\Sigma_g\boldsymbol{\delta}$, $\mathrm{tr}(\Sigma_g\Delta)^2$, and $\mathrm{tr}\Sigma_g^2$ be functions of $p$. Then, for $g = 1, 2$,

$$\limsup_{p\to\infty} \frac{n_g\boldsymbol{\delta}'\Sigma_g\boldsymbol{\delta}}{\mathrm{tr}\Sigma_g^2} < \infty \quad \text{and} \quad \limsup_{p\to\infty} \frac{n_g\mathrm{tr}(\Sigma_g\Delta)^2}{(\mathrm{tr}\Sigma_g^2)^2} < \infty.$$

(A4) Let $n_g$, $\boldsymbol{\delta}'\Sigma_g\boldsymbol{\delta}$, $\mathrm{tr}(\Sigma_g\Delta)^2$, and $\mathrm{tr}\Sigma_g^2$ be functions of $p$. Then, for $g = 1, 2$,

$$\lim_{p\to\infty} \frac{n_g\boldsymbol{\delta}'\Sigma_g\boldsymbol{\delta}}{\mathrm{tr}\Sigma_g^2} = \infty \quad \text{or} \quad \lim_{p\to\infty} \frac{n_g\mathrm{tr}(\Sigma_g\Delta)^2}{(\mathrm{tr}\Sigma_g^2)^2} = \infty.$$

Under assumption (A1), the leading variances of $\widehat{\|\boldsymbol{\delta}\|^2}$ and $\widehat{\|\Delta\|_F^2}$ and the leading covariances of $\widehat{\|\boldsymbol{\delta}\|^2}$ and $\widehat{\|\Delta\|_F^2}$ are

$$\begin{aligned}
\sigma_1^2 &= \sum_{g=1}^2 \frac{2\mathrm{tr}\Sigma_g^2}{n_g^2} + \frac{4\mathrm{tr}(\Sigma_1\Sigma_2)}{n_1n_2} + \sum_{g=1}^2 \frac{4\boldsymbol{\delta}'\Sigma_g\boldsymbol{\delta}}{n_g}, \\
\sigma_2^2 &= \sum_{g=1}^2 \frac{4(\mathrm{tr}\Sigma_g^2)^2}{n_g^2} + \frac{8\{\mathrm{tr}(\Sigma_1\Sigma_2)\}^2}{n_1n_2} \\
&\quad + \sum_{g=1}^2 \frac{8\mathrm{tr}(\Sigma_g\Delta)^2}{n_g} + \sum_{g=1}^2 \frac{4\kappa_{4g}\mathrm{tr}\{(\Sigma_g^{1/2}\Delta\Sigma_g^{1/2}) \odot (\Sigma_g^{1/2}\Delta\Sigma_g^{1/2})\}}{n_g}, \\
\sigma_{12} &= \sum_{g=1}^2 \frac{4\kappa_{3g}\sum_{i=1}^p \boldsymbol{\delta}'\Sigma_g^{1/2}\boldsymbol{e}_i\boldsymbol{e}_i'\Sigma_g^{1/2}\Delta\Sigma_g^{1/2}\boldsymbol{e}_i}{n_g},
\end{aligned}$$

where $\kappa_{3g} = \mathrm{E}[Z_{gij}^3]$, $\kappa_{4g} = \mathrm{E}[Z_{gij}^4] - 3$. Note that $\kappa_{4g} \geq -2$ since $0 \leq \mathrm{Var}(Z_{gij}^2) = \mathrm{E}[Z_{gij}^4] - (\mathrm{E}[Z_{gij}^2])^2 = \mathrm{E}[Z_{gij}^4] - 1$.

**Remark 1.** *We assume (A3). Then,* $-1 < \lim_{p\to\infty} \sigma_{12}/\sigma_1\sigma_2 < 1$*. See Section 2 in the supplemental material for full details. Also note that if* $\Delta = O$*,* $\boldsymbol{\delta} = \mathbf{0}$*, or* $\kappa_{3g} = 0$*, then* $\sigma_{12} = 0$*; if* $\boldsymbol{\delta} = \mathbf{0}$*, then*

$$\sigma_1^2 \;=\; \sum_{g=1}^{2} \frac{2\mathrm{tr}\Sigma_g^2}{n_g^2} + \frac{4\mathrm{tr}(\Sigma_1\Sigma_2)}{n_1 n_2} := \sigma_{10}^2;$$

*if* $\Delta = O$*, then*

$$\sigma_2^2 \;=\; \sum_{g=1}^{2} \frac{4(\mathrm{tr}\Sigma_g^2)^2}{n_g^2} + \frac{8\{\mathrm{tr}(\Sigma_1\Sigma_2)\}^2}{n_1 n_2} := \sigma_{20}^2.$$

The sum of the adjusted random variable and leading variance for each estimator is considered as 1:

$$\frac{\widehat{\|\boldsymbol{\delta}\|^2}}{\sigma_1} + \frac{\widehat{\|\Delta\|_F^2}}{\sigma_2}.$$

The variances $\sigma_1$ and $\sigma_2$ are unknown, but we can replace them by the unbiased estimators of $\sigma_1$ and $\sigma_2$ under $H_0$:

$$\widehat{\sigma}_{10}^2 \;=\; \sum_{g=1}^{2} \frac{2\widehat{\mathrm{tr}\Sigma_g^2}}{n_g^2} + \frac{4\mathrm{tr}(S_1S_2)}{n_1 n_2},\ \widehat{\sigma}_{20}^2 = \sum_{g=1}^{2} \frac{4(\widehat{\mathrm{tr}\Sigma_g^2})^2}{n_g^2} + \frac{8\{\mathrm{tr}(S_1S_2)\}^2}{n_1 n_2}.$$

Using these estimators, we propose the statistic

$$T = \frac{\widehat{\|\boldsymbol{\delta}\|^2}}{\widehat{\sigma}_{10}} + \frac{\widehat{\|\Delta\|_F^2}}{\widehat{\sigma}_{20}}.$$

From the following theorem, we derive the asymptotic distribution of $T$.

**Theorem 1.** *Let*

$$\widetilde{T}(c_1, c_2) = c_1\frac{\widehat{\|\boldsymbol{\delta}\|^2}}{\sigma_1} + c_2\frac{\widehat{\|\Delta\|_F^2}}{\sigma_2},$$

*where* $c_1$ *and* $c_2$ *are real constants that do not depend on* $p$*,* $n_1$*, and* $n_2$ *and* $(c_1, c_2)' \neq \mathbf{0}$*. Then, under assumptions (A1), (A2), and (A3),*

$$\frac{\widetilde{T}(c_1, c_2) - m(c_1, c_2)}{\sigma(c_1, c_2, \rho^*)} \xrightarrow{d} \mathcal{N}(0,1) \quad \text{as} \quad p \to \infty.$$

*Here,* $\xrightarrow{d}$ *denotes convergence in distribution*

$$m(c_1, c_2) = c_1\frac{\|\boldsymbol{\delta}\|^2}{\sigma_1} + c_2\frac{\|\Delta\|_F^2}{\sigma_2},\ \sigma(c_1, c_2, \rho^*) = \sqrt{c_1^2 + c_2^2 + 2c_1c_2\rho^*},$$

*where* $\rho^* = \lim_{p\to\infty} \sigma_{12}/(\sigma_1\sigma_2)$*.*

*Proof.* See Appendix B. □

Using Lemma A.1, we evaluate

$$\begin{aligned}
\mathrm{Var}[\widehat{\mathrm{tr}\Sigma_g^2}] &= O\left(\frac{\mathrm{tr}\Sigma_g^4}{n_g} + \frac{(\mathrm{tr}\Sigma_g^2)^2}{n_g^2}\right) = O\left(\frac{(\mathrm{tr}\Sigma_g^2)^2}{n_g}\right), \\
\mathrm{Var}[\mathrm{tr}(S_1S_2)] &= O\left(\frac{(n_1+n_2)\mathrm{tr}(\Sigma_1\Sigma_2)^2}{n_1n_2} + \frac{\{\mathrm{tr}(\Sigma_1\Sigma_2)\}^2}{n_1n_2}\right) \\
&= O\left(\frac{(n_1+n_2)\{\mathrm{tr}(\Sigma_1\Sigma_2)\}^2}{n_1n_2}\right)
\end{aligned}$$

under (A1). Thus, we obtain

$$\frac{\widehat{\mathrm{tr}\Sigma_g^2}}{\mathrm{tr}\Sigma_g^2} = 1 + o_p(1),\ \frac{\mathrm{tr}(S_1S_2)}{\mathrm{tr}(\Sigma_1\Sigma_2)} = 1 + o_p(1)$$

under (A1). Note that

$$\begin{aligned}
\frac{\widehat{\sigma}_{10}^2}{\sigma_{10}^2} &= \sum_{g=1,g\neq h}^{2} \frac{\widehat{\mathrm{tr}\Sigma_g^2}/\mathrm{tr}\Sigma_g^2 + n_g^2/\mathrm{tr}\Sigma_g^2\left\{\mathrm{tr}\Sigma_h^2/n_h^2 + 2\mathrm{tr}(\Sigma_1\Sigma_2)/(n_1n_2)\right\}}{1 + n_g^2/\mathrm{tr}\Sigma_g^2\left\{\mathrm{tr}\Sigma_h^2/n_h^2 + 2\mathrm{tr}(\Sigma_1\Sigma_2)/(n_1n_2)\right\}} \\
&\quad + \frac{\mathrm{tr}(S_1S_2)/\mathrm{tr}(\Sigma_1\Sigma_2) + n_1n_2/\{2\mathrm{tr}(\Sigma_1\Sigma_2)\}\sum_{g=1}^2 \mathrm{tr}\Sigma_g^2/n_g^2}{1 + n_1n_2/\{2\mathrm{tr}(\Sigma_1\Sigma_2)\}\sum_{g=1}^2 \mathrm{tr}\Sigma_g^2/n_g^2} - 2, \\
\frac{\widehat{\sigma}_{20}^2}{\sigma_{20}^2} &= \sum_{g=1,g\neq h}^{2} \frac{(\widehat{\mathrm{tr}\Sigma_g^2}/\mathrm{tr}\Sigma_g^2)^2 + n_g^2/(\mathrm{tr}\Sigma_g^2)^2\left[(\mathrm{tr}\Sigma_h^2/n_h)^2 + 2\{\mathrm{tr}(\Sigma_1\Sigma_2)\}^2/(n_1n_2)\right]}{1 + n_g^2/(\mathrm{tr}\Sigma_g^2)^2\left[(\mathrm{tr}\Sigma_h^2/n_h)^2 + 2\{\mathrm{tr}(\Sigma_1\Sigma_2)\}^2/(n_1n_2)\right]} \\
&\quad + \frac{\{\mathrm{tr}(S_1S_2)/\mathrm{tr}(\Sigma_1\Sigma_2)\}^2 + n_1n_2/[2\{\mathrm{tr}(\Sigma_1\Sigma_2)\}^2]\sum_{g=1}^2 \mathrm{tr}(\Sigma_g^2/n_g)^2}{1 + n_1n_2/[2\{\mathrm{tr}(\Sigma_1\Sigma_2)\}^2]\sum_{g=1}^2 \mathrm{tr}(\Sigma_g^2/n_g)^2} - 2.
\end{aligned}$$

Hence, $\widehat{\sigma}_{g0}^2/\sigma_{g0}^2 = 1 + o_p(1)$ under (A1). Moreover, under (A3), $\sigma_g/\sigma_{g0} \to r_g$ for some $r_g \in [1, \infty)$. Therefore,

$$T = \widetilde{T}(r_1, r_2) + o_p(1)$$

under (A1) and (A3). From this fact and Theorem 1, we obtain the following lemma.

**Lemma 1.** *Under assumptions (A1), (A2), and (A3),*

$$\frac{T - (\sigma_{10}^{-1}\|\boldsymbol{\delta}\|^2 + \sigma_{20}^{-1}\|\Delta\|_F^2)}{\sigma(r_1, r_2, \rho^*)} \xrightarrow{d} \mathcal{N}(0,1) \quad \text{as} \quad p \to \infty.$$

**Remark 2.** $H_0 \Longrightarrow$ *(A3).*

## 2.2 Testing procedure and its asymptotic properties

Note that $\sigma(1, 1, 0) = 2$ under $H_0$. From Lemma 1 and Remark 2, under (A1), (A2), and $H_0$,

$$\frac{T}{\sqrt{2}} \xrightarrow{d} \mathcal{N}(0,1). \tag{2.1}$$

Based on the asymptotic normality of $T$ under $H_0$, we propose the following approximation test:

$$\text{rejecting } H_0 \Longleftrightarrow T \geq \sqrt{2} z_\alpha. \tag{2.2}$$

Here, $z_\alpha$ is the upper critical value of the standard normal distribution, $\mathcal{N}(0,1)$.

The size and power of test (2.2) are defined as

$$\begin{aligned} \textit{Size} &= \Pr(T \geq \sqrt{2} z_\alpha | \textit{sample generated by model (1.1) and } H_0), \\ \textit{Power} &= \Pr(T \geq \sqrt{2} z_\alpha | \textit{sample generated by model (1.1) and } H_1). \end{aligned}$$

**Remark 3.** *From (2.1), under (A1) and (A2),* $\textit{Size} = \alpha + o(1)$ *as* $p \to \infty$.

First, we investigate the power under (A3). Using Lemma 1, we obtain following proposition:

**Proposition 1.** *Under (A1), (A2), and (A3),*

$$\textit{Power} = \Phi\left( \frac{\sigma_{10}^{-1}\|\boldsymbol{\delta}\|^2 + \sigma_{20}^{-1}\|\Delta\|_F^2}{\sigma(r_1, r_2, \rho^*)} - \frac{\sqrt{2} z_\alpha}{\sigma(r_1, r_2, \rho^*)} \right) + o(1)$$

*as* $p \to \infty$.

Thus, if the difference between $\boldsymbol{\mu}_1$ ($\Sigma_1$) and $\boldsymbol{\mu}_2$ ($\Sigma_2$) is not so small that $\|\boldsymbol{\delta}\|^2$ ($\|\Delta\|_F^2$) has the same order as, or a higher order than, that $\sigma_{10}$ ($\sigma_{20}$), the test will be powerful. Conversely, if $\|\boldsymbol{\delta}\|^2$ and $\|\Delta\|_F^2$ are so small that $\|\boldsymbol{\delta}\|^2$ and $\|\Delta\|_F^2$ are of a lower order than $\sigma_{10}$ and $\sigma_{20}$, respectively, the test will not be powerful and cannot distinguish $H_0$ from $H_1$.

Next, we investigate the power under (A4). Let $n_g$, $\boldsymbol{\delta}'\Sigma_g\boldsymbol{\delta}$, $\mathrm{tr}(\Sigma_g\Delta)^2$, and $\mathrm{tr}\Sigma_g^2$ be functions of $p$. We consider

$$\begin{aligned} &\text{(A4-i)} \quad \lim_{p\to\infty} \frac{n_g \boldsymbol{\delta}'\Sigma_g\boldsymbol{\delta}}{\mathrm{tr}\Sigma_g^2} = \infty, \quad \limsup_{p\to\infty} \frac{n_g \mathrm{tr}(\Sigma_g\Delta)^2}{(\mathrm{tr}\Sigma_g^2)^2} < \infty, \\ &\text{(A4-ii)} \quad \limsup_{p\to\infty} \frac{n_g \boldsymbol{\delta}'\Sigma_g\boldsymbol{\delta}}{\mathrm{tr}\Sigma_g^2} < \infty, \quad \lim_{p\to\infty} \frac{n_g \mathrm{tr}(\Sigma_g\Delta)^2}{(\mathrm{tr}\Sigma_g^2)^2} = \infty, \\ &\text{(A4-iii)} \quad \lim_{p\to\infty} \frac{n_g \boldsymbol{\delta}'\Sigma_g\boldsymbol{\delta}}{\mathrm{tr}\Sigma_g^2} = \infty, \quad \lim_{p\to\infty} \frac{n_g \mathrm{tr}(\Sigma_g\Delta)^2}{(\mathrm{tr}\Sigma_g^2)^2} = \infty. \end{aligned}$$

Then,

$$\begin{aligned} T - \sqrt{2} z_\alpha &= \frac{\|\boldsymbol{\delta}\|^2}{\sigma_{10}} + \frac{\|\Delta\|_F^2}{\sigma_2} + o_p\left(\frac{\|\boldsymbol{\delta}\|^2}{\sigma_{10}}\right) \text{ under (A4-i)}, \\ T - \sqrt{2} z_\alpha &= \frac{\|\Delta\|_F^2}{\sigma_{20}} + \frac{\|\boldsymbol{\delta}\|^2}{\sigma_1} + o_p\left(\frac{\|\Delta\|_F^2}{\sigma_{20}}\right) \text{ under (A4-ii)}, \\ T - \sqrt{2} z_\alpha &= \frac{\|\boldsymbol{\delta}\|^2}{\sigma_{10}} + \frac{\|\Delta\|_F^2}{\sigma_{20}} + o_p\left(\frac{\|\boldsymbol{\delta}\|^2}{\sigma_{10}}\right) + o_p\left(\frac{\|\Delta\|_F^2}{\sigma_{20}}\right) \text{ under (A4-iii)}. \end{aligned}$$

See Section 3 in the supplemental material for full details. Thus, we obtain the following proposition.

**Proposition 2.** *Under (A4),* $\textit{Power} = 1 + o(1)$ as $p \to \infty$.

# 3 Simulation and example

## 3.1 Simulation

We generate Monte Carlo samples to evaluate the performance of the proposed test procedure, including its size and power. Notice that the data are generated from the model:

$$\boldsymbol{X}_{gi} = \Sigma_g^{1/2}\boldsymbol{Z}_{gi} + \boldsymbol{\mu}_g \quad \text{for } i = 1, \ldots, n_g. \tag{3.1}$$

We set the null hypothesis $H_0$ as follows:

$$\boldsymbol{\mu}_1 = \boldsymbol{\mu}_2 = \boldsymbol{0},\ \Sigma_1 = \Sigma_2 = B(0.3^{|i-j|})B, \tag{3.2}$$

where

$$B = \mathrm{diag}\left(\left(0.5 + \frac{1}{p+1}\right)^{1/2}, \left(0.5 + \frac{2}{p+1}\right)^{1/2}, \cdots, \left(0.5 + \frac{p}{p+1}\right)^{1/2}\right).$$

In addition, we set the alternative hypothesis $H_1$ as follows:

$$\begin{aligned} &\boldsymbol{\mu}_1 = \boldsymbol{0},\ \Sigma_1 = B(0.3^{|i-j|})B, \\ &\boldsymbol{\mu}_2 = (\mathrm{tr}\Sigma_1^2)^{1/4}(\sqrt{0.1}, \ldots, \sqrt{0.1})'\ ,\ \Sigma_2 = (1 - \sqrt{0.1})\Sigma_1. \end{aligned} \tag{3.3}$$

Note that $\|\boldsymbol{\delta}\|^2/(\mathrm{tr}\Sigma_1^2)^{1/2} = 0.1$ and $\|\Delta\|_F^2/\mathrm{tr}\Sigma_1^2 = 0.1$ under (3.3).

First, we compare the proposed test, denoted as HN, with the modified likelihood ratio test, designated MLR, and the proposed likelihood ratio based approximation test, described as HNLR. The MLR and HNLR tests are considered for multivariate normal populations where $p < \min\{n_1 - 1, n_2 - 1\}$. The MLR, proposed by Muirhead [8], is

$$\text{rejecting } H_0 \iff -2\rho \log \lambda \geq \chi^2_{p(p+3)/2}(\alpha),$$

where $\chi_a(\alpha)$ denotes the upper $100\alpha$ percentile of the chi-squared distribution with $a$ degrees of freedom,

$$\rho = 1 - \frac{(2p^2 + 9p + 11)(n_1^2 + n_2^2 + n_1 n_2)}{6(p+3)n n_1 n_2}.$$

Under a large sample framework, the size of this approximation test is $\alpha + O(n^{-2})$. Let $\tilde{\lambda} = \prod_{g=1}^2 n_g^{pn_g/2}/(n^{pn/2})\lambda$. By applying the same techniques as in Fujikoshi et al. [3] for the general moment of $\tilde{\lambda}$ derived by Muirhead [8], we can express the cumulant-generating function as the following infinite series:

$$K_{\tilde{\lambda}}(t) = \sum_{s=1}^{\infty} \frac{(\sqrt{-1}t)^s \kappa^{(s)}}{s!},$$

where

$$\kappa^{(s)} = (-1)^s \sum_{i=1}^{p} \left( \sum_{g=1}^{2} r_g^s \psi^{(s-1)}\left(\frac{n_g - i}{2}\right) - \psi^{(s-1)}\left(\frac{n-i}{2}\right)\right).$$

Here, $\psi^{(s)}(\cdot)$ denotes the polygamma function. Using the first- and second-order cumulants, we test the HNLR as follows:

$$\text{rejecting } H_0 \Longleftrightarrow \frac{\tilde{\lambda} - \kappa^{(1)}}{\sqrt{\kappa^{(2)}}} \geq z_\alpha.$$

We generated $n$ samples of $\boldsymbol{X}_{gi}$ $(i = 1, \ldots, n_g,\ g = 1, 2)$ from the $p$-variate normal distribution $\mathcal{N}_p(\boldsymbol{\mu}_g, \Sigma_g)$. We use the simulation settings (3.2) for size and (3.3) for power as $p = 10, 20, 40, 80$, $(n_1, n_2) = (p + 10, p + 10), (p + 10, p + 30), (p + 50, p + 50), (p + 20, p + 80)$ and conduct $10^6$ replications to calculate the size and power. The nominal significance level is $\alpha = 0.05$. Table 1 shows the empirical size and power of MLR, HNLR, and HN. The size of MLR greatly exceeds the significance level of 0.05 as the dimension increases. On the other hand, the size of both HN and HNLR is close to a significance level of 0.05 even in high-dimensional settings. The power of HN exceeds that of HNLR in all cases. These results indicate that the proposed HN test is superior to the MLR and HNLR tests in high-dimensional settings.

Next, we investigate the size and power of the proposed test under some distributions. For $j = 1, \ldots, n_g,\ g = 1, 2$, $\boldsymbol{Z}_{gj} = (Z_{gij})$ emerges in (3.1) for the following distributions:

(D1) $Z_{gij} \sim \mathcal{N}(0, 1)$,

(D2) $Z_{gij} = (U_{gij} - 10)/\sqrt{20}$ for $U_{gij} \sim \chi^2_{10}$,

(D3) $Z_{gij} = U_{gij}/\sqrt{5/4}$ for $U_{gij} \sim \mathcal{T}_{10}$,

(D4) $Z_{gij} = \left(1 - \frac{9}{5\pi}\right)^{-1/2} \left(U_{gij} + \frac{3}{\sqrt{5\pi}}\right)$ for $U_{gij} \sim \mathcal{SN}(-3)$.

Table 2 presents the third cumulant $\kappa_3$ and the fourth cumulant $\kappa_4$ for each distribution. We use the simulation settings (3.2) for size and (3.3) for power as $p = 32, 64, 128, 256$, $(n_1, n_2) = (20, 20), (30, 10), (50, 50), (70, 30), (80, 80), (120,$
$40), (100, 100), (150, 50)$. The nominal significance level is $\alpha = 0.05$. Tables 3-6 shows the empirical size and power of HN for each distribution. From Tables 3-6, we find that the power is monotonically increasing for $n$ and $p$, except for $n = 40$. The proposed test has the highest power in case (D4). We note that case (D4) has a third-order cumulant, the lowest among (D1)-(D4). Conversely, the proposed test has the lowest power in case (D2). We also note that case (D2) has third- and fourth-order cumulants, the highest among (D1)-(D4). In any case, the third- and fourth-order cumulants affect the power of the test in these simulations.

## 3.2 An example

**· Leukemia data set**

We apply our test to a microarray dataset analyzed by Dudoit et al. [2]. The dataset, obtained from affymetrix oligonucleotide microarrays, contains 72 cases of either acute lymphoblastic leukemia (ALL, $n_1 = 47$) or acute myeloid leukemia (AML, $n_2 = 25$). The dataset is publicly available at http://portals.broadinstitute.org/gpp/public/. We preprocess the dataset using the protocol written in Dudoit et al. [2]. The preprocessed dataset comprises $p = 3571$ variables. Using this dataset, we calculate

$$\widehat{\|\boldsymbol{\delta}\|^2} \approx 50.2,\ \widehat{\|\Delta\|^2} \approx 23758.9,\ \widehat{\sigma}_{10} \approx 11.7,\ \widehat{\sigma}_{20} \approx 2340.4.$$

From these values, we construct a closed testing procedure to simultaneously test the mean vector and the covariance matrix. Let $\mathcal{C} = \{H_{\delta,\Delta}, H_\delta, H_\Delta\}$. Here, $H_{\delta,\Delta} : \boldsymbol{\mu}_1 = \boldsymbol{\mu}_2,\ \Sigma_1 = \Sigma_2$, $H_\delta : \boldsymbol{\mu}_1 = \boldsymbol{\mu}_2$, $H_\Delta : \Sigma_1 = \Sigma_2$. Note that the family $\mathcal{C}$ is closed. Then, we test the hypotheses in $\mathcal{C}$ by using the following procedure:
**Step 1** We test hypothesis $H_{\delta,\Delta}$.
Case 1. If $T > \sqrt{2}z_\alpha$, we reject $H_{\delta,\Delta}$ and go to Step 2.
Case 2. If $T \le \sqrt{2}z_\alpha$, we retain all hypotheses, $H_{\delta,\Delta}, H_\delta,$ and $H_\Delta$.
**Step 2** We test hypotheses $H_\delta$ and $H_\Delta$.
If $\widehat{\|\boldsymbol{\delta}\|^2}/\widehat{\sigma}_{10} > z_\alpha$, we reject $H_\delta$. If $\widehat{\|\boldsymbol{\delta}\|^2}/\widehat{\sigma}_{10} \le z_\alpha$, we retain $H_\delta$.
If $\widehat{\|\Delta\|^2}/\widehat{\sigma}_{20} > z_\alpha$, we reject $H_\Delta$. If $\widehat{\|\Delta\|^2}/\widehat{\sigma}_{20} \le z_\alpha$, we retain $H_\Delta$.

We assume (A1) and (A2). Then, the following statements hold:

- If $H_{\delta,\Delta}$, $H_\delta$ and $H_\Delta$ are true, $\Pr\left(T > \sqrt{2}z_\alpha,\ \frac{\widehat{\|\Delta\|^2}}{\widehat{\sigma}_{20}} > z_\alpha,\ \frac{\widehat{\|\boldsymbol{\delta}\|^2}}{\widehat{\sigma}_{10}} > z_\alpha\right) \le \Pr\left(T > \sqrt{2}z_\alpha\right) = \alpha + o(1)$.
- If $H_\delta$ is true, $\Pr\left(T > \sqrt{2}z_\alpha,\ \frac{\widehat{\|\boldsymbol{\delta}\|^2}}{\widehat{\sigma}_{10}} > z_\alpha\right) \le \Pr\left(\frac{\widehat{\|\boldsymbol{\delta}\|^2}}{\widehat{\sigma}_{10}} > z_\alpha\right) = \alpha + o(1)$.
- If $H_\Delta$ is true, $\Pr\left(T > \sqrt{2}z_\alpha,\ \frac{\widehat{\|\Delta\|^2}}{\widehat{\sigma}_{20}} > z_\alpha\right) \le \Pr\left(\frac{\widehat{\|\Delta\|^2}}{\widehat{\sigma}_{20}} > z_\alpha\right) = \alpha + o(1)$.

From the above, the type-I familywise error rate of this procedure is not greater than $\alpha$.

We set a $\alpha = 0.05$ level of significance. In Step 1, the test statistic $T$ is calculated as $T \approx 48.6$, and the approximate critical values of $T$, at a $\alpha = 0.05$ level of significance, is obtained as $\sqrt{2}z_{0.05} \approx 2.326$. We reject $H_{\delta,\Delta}$ and go to Step 2. In Step 2, we calculate $\widehat{\|\boldsymbol{\delta}\|^2}/\widehat{\sigma}_{10} \approx 4.3$ and $\widehat{\|\Delta\|^2}/\widehat{\sigma}_{20} \approx 10.2$. Both values are greater than $z_{0.05} \approx 1.645$. Thus, we reject $H_\delta$ and reject $H_\Delta$.

**· Khan dataset**

We compare the population mean of four types of gene expression using training data. Khan et al. [6] study the expression of genes in four types of small round blue cell tumors of childhood (SRBCT). These were the Ewing family of tumors (EWS, 23 cases), Burkit lymphoma, a subset of non-Hodgkin lymphoma (BL, 8 cases), neuroblastoma (NB, 12 cases), and rhabdomyosarcoma (RMS, 21 cases). The data include the gene expression profiles obtained from both tumor biopsy and cell line samples. These were downloaded from a website containing a filtered dataset of 2308 gene expression profiles as described by Khan et al. [6]. This dataset is available from http://bioinf.ucd.ie/people/ae
din/R/. First, we construct a hypothesis test for all differences in the population mean vectors $\boldsymbol{\mu}_{\mathrm{EWS}}$, $\boldsymbol{\mu}_{\mathrm{BL}}$, $\boldsymbol{\mu}_{\mathrm{NB}}$, and $\boldsymbol{\mu}_{\mathrm{RMS}}$. By combining Bonferroni correction with the proposed test (2.1), we obtain pairwise comparison procedures for all possible pairs.
**Step 1** For $g, h = 1, \ldots, 4,\ g \ne h$, we define $H_0^{(g,h)}$ as $\boldsymbol{\mu}_g = \boldsymbol{\mu}_h,\ \Sigma_g = \Sigma_h$. We set the family of hypotheses:

$$\mathcal{F} = \left\{ H_0^{(g,h)}\ ;\ g < h,\ g, h = 1, \ldots, 4 \right\}.$$

**Step 2** We set the significance level at $\alpha \in (0, 1)$.
**Step 3** We choose $H_0^{(g,h)}$ in $\mathcal{F}$.
**Step 4** By using $(g, h)$ determined in Step 3, we calculate the statistic $t_{gh} = \widehat{\|\boldsymbol{\delta}_{gh}\|^2}/\widehat{\sigma}_{gh,10} + \widehat{\|\Delta_{gh}\|_F^2}/\widehat{\sigma}_{gh,20}$.
**Step 6** We reject $H_0^{(g,h)}$ if $t_{gh} > \sqrt{2} z_{\alpha/6}$.
We perform Steps 5 and 6 for all contrasts taken up in Step 4. The realized values of the test statistic $t_{gh}$ are summarized in Table 7.

# 4 Conclusion and Discussion

This paper treated two-sample simultaneous test for high-dimensional mean vectors and covariance matrices under non-normal populations. When $p \geq \min\{n_1 - 1, n_2 - 1\}$, the likelihood ratio test is not applicable for this problem. Therefore, we considered an $L^2$-norm-based test to simultaneously test the mean vector and the covariance matrix. The $L^2$-norm-based test for mean vectors ($H_0 : \boldsymbol{\delta} = \mathbf{0}$ vs. $H_1 : \boldsymbol{\delta} \neq \mathbf{0}$) was developed by Chen and Quin [1], and the $L^2$-norm based test for covariance matrices ($H_0 : \Delta = O$ vs. $H_1 : \Delta \neq O$) by Li and Chen [7]. Our test statistic is based on the sum of standardized estimators used in different tests, and we investigated the asymptotic size and power of the proposed test under a few asymptotic frameworks. We therefore infer that the power of this test depends on the distance between $\Sigma_1$ ($\boldsymbol{\mu}_1$) and $\Sigma_2$ ($\boldsymbol{\mu}_2$), a trace of the power of $\Sigma_g$, and the third- and fourth-order cumulants.

Further, we studied the finite sample and dimension performance of this test under a few non-normal settings via Monte Carlo simulations. Through the simulation results, we confirmed that the proposed test is satisfactory in most cases. The size of the proposed test is only slightly larger than the nominal significance level when the total sample size $n$ and the dimension $p$ is relatively large. The power monotonically increases with the total sample size $n$ and the dimension $p$. Even when the relatively dimension $p$ is relatively small and $p < \min\{n_1 - 1, n_2 - 1\}$, our test demonstrated better accuracy than the likelihood ratio test.

In conclusion, our test can be recommended for two-sample simultaneous testing of mean vectors and covariance matrices when the total sample size $n$ and the dimension $p$ are both large. However, our non-normal assumption does not include the family of elliptical distributions. A discussion on this distribution family is a task for the future.

**Acknowledgements**
The authors would like to thank Mr. Hayate Ogawa for his help. The second author was supported in part by a Grant-in-Aid for Young Scientists (B) (26730020) from the Japan Society for the Promotion of Science.

# References

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

[2] S. Dudoit, J. Fridlyand, P. T. Speed, Comparison of discrimination methods for the classification of tumors using gene expression data, *J. Amer. Statist. Assoc.*, **97** (2002) 77–87.

[3] Y. Fujikoshi, V. V. Ulyanov, R. Shimizu, *Multivariate Statistics:High-Dimensional and Large-Sample Approximations*, Wiley, Hobeken, N.J (2010).

[4] P. Hall, C. C. Heyde, *Martingale Limit Theory and Its Application*, Academic Press, New York (1980).

[5] T. Himeno, T. Yamada, Estimations for some functions of covariance matrix in high dimension under non-normality and its applications, *J. Multivariate Anal.*, **130** (2014) 27–44.

[6] J. Khan, J.S. Wei, M. Ringner, L.H. Saal, M. Ladanyi, F. Westermann, F. Berthold, M. Schwab, C.R. Antonescu, C. Peterson, Classification and diagnostic prediction of cancers using gene expression profiling and artificial neural networks, *Nat. Med.*, **7** (2001) 673–679.

[7] J. Li, S. X. Chen, Two sample tests for high-dimensional covariance matrices, *Ann. Statist.*, **40** (2012) 908–940.

[8] R. J. Muirhead, *Aspects of Multivariate Statistical Theory*, John Wiley & Sons, Inc., New York (1982).

[9] M. S. Srivastava, H. Yanagihara, T. Kubokawa, Tests for covariance matrices in high dimension with less sample size, *J. Multivariate Anal.*, **130** (2014) 289–309.

# Appendix

## A.Moment

**Lemma A. 1.** *Let $A$, $B$ be $p\times p$ non-random symmetric matrices and $\boldsymbol{b}$ be $p$-dimensional non-random vector. Assume that the $i$-th element of $p$-dimensional random vectors $\boldsymbol{Z}$ (denotes $Z_i$) and $\boldsymbol{W}$ (denotes $W_i$) has uniformly bounded 8th moment, and there exist finite constants $\kappa_3$, $\kappa_4$ such that $\mathrm{E}[Z_i^3]=\mathrm{E}[W_i^3]=\kappa_3$, $\mathrm{E}[Z_i^4]=\mathrm{E}[W_i^4]=\kappa_4+3$ $(\kappa_4\geq -2)$ and for any positive integers $r$ and $\alpha_\ell$ such that $\sum_{\ell=1}^r \alpha_\ell \leq 8$, $\mathrm{E}[\prod_{\ell=1}^r Z_{i_\ell}^{\alpha_\ell}]=\prod_{\ell=1}^r \mathrm{E}[Z_{i_\ell}^{\alpha_\ell}]$ and $\mathrm{E}[\prod_{\ell=1}^r W_{i_\ell}^{\alpha_\ell}]=\prod_{\ell=1}^r \mathrm{E}[W_{i_\ell}^{\alpha_\ell}]$ whenever $i_1,i_2,\ldots,i_r$ are distinct indices. Also, $\boldsymbol{Z}$ and $\boldsymbol{W}$ are independent. Then it holds that*

$$
\begin{aligned}
&\text{(i)} \quad \mathrm{E}[\boldsymbol{Z}'A\boldsymbol{Z}\boldsymbol{b}'\boldsymbol{Z}]=\kappa_3\sum_{i=1}^p \boldsymbol{e}_i'A\boldsymbol{e}_i\boldsymbol{b}'\boldsymbol{e}_i,\\
&\text{(ii)} \quad \mathrm{E}[\boldsymbol{Z}'A\boldsymbol{Z}\boldsymbol{Z}'B\boldsymbol{Z}]=\kappa_4\mathrm{tr}(A\odot B)+\mathrm{tr}A\mathrm{tr}B+2\mathrm{tr}(AB),\\
&\text{(iii)} \quad \mathrm{E}[(\boldsymbol{Z}'A\boldsymbol{W})^4]=\kappa_4^2\mathrm{tr}(A\odot A)^2+6\kappa_4\mathrm{tr}(A^2\odot A^2)+3(\mathrm{tr}A^2)^2+6\mathrm{tr}A^4,\\
&\text{(iv)} \quad \mathrm{E}[(\boldsymbol{Z}'A\boldsymbol{Z}-\mathrm{tr}A)^4]\leq const.(\mathrm{tr}A^2)^2.
\end{aligned}
$$

*Proof.* See Section 1 in supplemental material. □

## B.Proof of Theorem 1

Under assumptions (A1) and (A2),

$$
\widetilde{T}(c_1,c_2)=m(c_1,c_2)+\sum_{k=1}^n \varepsilon_k+o_p(1),
$$

where for $1 \le k \le n_1$,

$$\varepsilon_k = \frac{2c_1}{\sigma_1 n_1(n_1-1)}\boldsymbol{b}'_{1k}\boldsymbol{Y}_{1k} + \frac{2c_2}{\sigma_2 n_1(n_1-1)}\{\boldsymbol{Y}'_{1k}A_{1k}\boldsymbol{Y}_{1k} - \mathrm{tr}(A_{1k}\Sigma_1)\},$$

for $n_1+1 \le k \le n$,

$$\varepsilon_k = \frac{2c_1}{\sigma_1 n_2(n_2-1)}\boldsymbol{b}'_{2k}\boldsymbol{Y}_{2\ k-n_1} + \frac{2c_2}{\sigma_2 n_2(n_2-1)}\{\boldsymbol{Y}'_{2\ k-n_1}A_{2k}\boldsymbol{Y}_{2\ k-n_1} - \mathrm{tr}(A_{2k}\Sigma_2)\}.$$

Here,

$$\begin{aligned} A_{1k} &= \sum_{j=1}^{k-1}\mathcal{E}_j^{(1)} + (n_1-1)\Delta, \\ \boldsymbol{b}_{1k} &= \sum_{j=1}^{k-1}\boldsymbol{Y}_{1j} + (n_1-1)\boldsymbol{\delta}, \\ A_{2k} &= \sum_{j=1}^{k-n_1-1}\mathcal{E}_j^{(2)} - \frac{n_2-1}{n_1}\sum_{j=1}^{n_1}\mathcal{E}_j^{(1)} - (n_2-1)\Delta, \\ \boldsymbol{b}_{2k} &= \sum_{j=1}^{k-n_1-1}\boldsymbol{Y}_{2j} - \frac{n_2-1}{n_1}\sum_{j=1}^{n_1}\boldsymbol{Y}_{1j} - (n_2-1)\boldsymbol{\delta}, \end{aligned}$$

where

$$\mathcal{E}_j^{(1)} = \boldsymbol{Y}_{1j}\boldsymbol{Y}'_{1j} - \Sigma_1,\ \mathcal{E}_j^{(2)} = \boldsymbol{Y}_{2j}\boldsymbol{Y}'_{2j} - \Sigma_2.$$

Let $\mathcal{F}_0 = \{\emptyset, \Omega\}$, and $\mathcal{F}_k$ $(1 \le k)$ be the $\sigma$-algebra by the random variables $(\varepsilon_1, \dots, \varepsilon_k)$. Then we find that

$$\mathcal{F}_0 \subseteq \mathcal{F}_1 \subseteq \cdots \subseteq \mathcal{F}_\infty$$

and $\mathrm{E}\left[\varepsilon_k | \mathcal{F}_{k-1}\right] = 0$. We show asymptotic normality of $\sum_{k=1}^n \varepsilon_k$ by adapting martingale difference central limit theorem, see Hall and Heyde [4]. It is necessary to check the following two conditions to apply this theorem:

$$\begin{aligned} &\text{(I)} \quad \sum_{k=1}^n \mathrm{E}[\varepsilon_k^2|\mathcal{F}_{k-1}] = \sigma(c_1,c_2,\rho^*)^2 + o_p(1) \text{ as } p \to \infty, \\ &\text{(II)} \quad \sum_{k=1}^n \mathrm{E}[\varepsilon_k^4] = o(1) \text{ as } p \to \infty. \end{aligned}$$

For any $(c_1,c_2)' \in \mathcal{R}^2,\ (c_1,c_2)' \ne \boldsymbol{0}$, $\sigma(c_1,c_2,\rho^*)^2 > 0$ under (A3).

First we check (I). We define the following four random variables:

$$\begin{aligned}
C_k^{(g)} &= \frac{\boldsymbol{b}_{gk}'\Sigma_g\boldsymbol{b}_{gk}}{n_g^2(n_g-1)^2},\\
D_k^{(g)} &= \frac{\mathrm{tr}(\Sigma_g A_{gk})^2}{n_g^2(n_g-1)^2},\\
E_k^{(g)} &= \frac{\mathrm{tr}\{(\Sigma_g^{1/2}A_{gk}\Sigma_g^{1/2})\odot(\Sigma_g^{1/2}A_{gk}\Sigma_g^{1/2})\}}{n_g^2(n_g-1)^2},\\
F_k^{(g)} &= \frac{\sum_{i=1}^p(\boldsymbol{b}_{gk}'\Sigma_g^{1/2}\boldsymbol{e}_i\boldsymbol{e}_i'\Sigma_g^{1/2}A_{gk}\Sigma_g^{1/2}\boldsymbol{e}_i)}{n_g^2(n_g-1)^2}.
\end{aligned}$$

Then the sum of conditional mean $\sum_{k=1}^n \mathrm{E}[\varepsilon_k^2|\mathcal{F}_{k-1}]$ partially decompose as

$$\begin{aligned}
&4\sum_{k=1}^{n_1}\left\{\frac{c_1^2}{\sigma_1^2}C_k^{(1)}+\frac{c_2^2}{\sigma_2^2}(2D_k^{(1)}+\kappa_{41}E_k^{(1)})+\frac{2c_1c_2\kappa_{31}}{\sigma_1\sigma_2}F_k^{(1)}\right\}\\
&+4\sum_{k=n_1+1}^{n}\left\{\frac{c_1^2}{\sigma_1^2}C_k^{(2)}+\frac{c_2^2}{\sigma_2^2}(2D_k^{(2)}+\kappa_{42}E_k^{(2)})+\frac{2c_1c_2\kappa_{32}}{\sigma_1\sigma_2}F_k^{(2)}\right\}.
\end{aligned}$$

Under assumptions (A1) and (A2), it holds that following four statements:

$$\begin{aligned}
&\text{(A)}\quad \sigma_1^{-2}\left(\sum_{k=1}^{n_1}C_k^{(1)}+\sum_{k=n_1+1}^{n}C_k^{(2)}\right)-\frac{c_0}{\sigma_1^2}=o_p(1),\\
&\text{(B)}\quad \sigma_2^{-2}\left(\sum_{k=1}^{n_1}D_k^{(1)}+\sum_{k=n_1+1}^{n}D_k^{(2)}\right)-\frac{d_0}{\sigma_2^2}=o_p(1),\\
&\text{(C)}\quad \sigma_2^{-2}\left(\sum_{k=1}^{n_1}E_k^{(1)}+\sum_{k=n_1+1}^{n}E_k^{(2)}\right)-\frac{e_0}{\sigma_2^2}=o_p(1),\\
&\text{(D)}\quad (\sigma_1\sigma_2)^{-1}\left(\sum_{k=1}^{n_1}F_k^{(1)}+\sum_{k=n_1+1}^{n}F_k^{(2)}\right)-\frac{f_0}{\sigma_1\sigma_2}=o_p(1),
\end{aligned}$$

where

$$\begin{aligned}
c_0 &= \sum_{g=1}^2\frac{\mathrm{tr}\Sigma_g^2}{2n_g^2}+\frac{\mathrm{tr}(\Sigma_1\Sigma_2)}{n_1n_2}+\sum_{g=1}^2\frac{\boldsymbol{\delta}'\Sigma_g\boldsymbol{\delta}}{n_g},\\
d_0 &= \sum_{g=1}^2\frac{(\mathrm{tr}\Sigma_g^2)^2}{2n_g^2}+\frac{\{\mathrm{tr}(\Sigma_1\Sigma_2)\}^2}{n_1n_2}+\sum_{g=1}^2\frac{\mathrm{tr}(\Sigma_g\Delta)^2}{n_g},\\
e_0 &= \sum_{g=1}^2\frac{\mathrm{tr}\{(\Sigma_g^{1/2}\Delta\Sigma_g^{1/2})\odot(\Sigma_g^{1/2}\Delta\Sigma_g^{1/2})\}}{n_g},\\
f_0 &= \frac{\sum_{i=1}^p\boldsymbol{\delta}'\Sigma_g^{1/2}\boldsymbol{e}_i\boldsymbol{e}_i'\Sigma_g^{1/2}\Delta\Sigma_g^{1/2}\boldsymbol{e}_i}{n_g}.
\end{aligned}$$

These results are obtained by evaluating the moment of four random variables. See Section 4-7 in supplemental material for full details. Since (A)-(D) is satisfied, the statement (I) is true. Next we check (II). Under assumptions (A1) and (A2), it holds that following two statements:

$$\text{(E)} \quad \sum_{k=1}^{n_1} \mathrm{E}[\varepsilon_k^4] = o(1) \text{ as } p \to \infty,$$

$$\text{(F)} \quad \sum_{k=n_1+1}^{n} \mathrm{E}[\varepsilon_k^4] = o(1) \text{ as } p \to \infty.$$

See Section 8 and 9 in supplemental material for full details. Thus the asymptotic normality of $T(c_1, c_2)$ is established.

Table 1: Comparison of modified liklihood ratio test

| $p=10$ | $(n_1,n_2)$ | (20,20) | (20,40) | (60,60) | (30,90) |
|---|---|---|---|---|---|
| size | MLR | 0.068 | 0.066 | 0.048 | 0.055 |
| | HNLR | 0.056 | 0.057 | 0.054 | 0.056 |
| | HN | 0.074 | 0.072 | 0.068 | 0.070 |
| power | MLR | 0.146 | 0.232 | 0.594 | 0.490 |
| | HNLR | 0.120 | 0.209 | 0.620 | 0.495 |
| | HN | 0.302 | 0.350 | 0.796 | 0.572 |
| $p=20$ | $(n_1,n_2)$ | (30,30) | (30,50) | (70,70) | (40,100) |
| size | MLR | 0.156 | 0.130 | 0.055 | 0.081 |
| | HNLR | 0.057 | 0.054 | 0.056 | 0.052 |
| | HN | 0.066 | 0.065 | 0.063 | 0.064 |
| power | MLR | 0.307 | 0.381 | 0.653 | 0.607 |
| | HNLR | 0.128 | 0.210 | 0.656 | 0.528 |
| | HN | 0.458 | 0.532 | 0.903 | 0.753 |
| $p=40$ | $(n_1,n_2)$ | (50,50) | (50,70) | (90,90) | (60,120) |
| size | MLR | 0.665 | 0.484 | 0.093 | 0.233 |
| | HNLR | 0.055 | 0.054 | 0.053 | 0.055 |
| | HN | 0.060 | 0.060 | 0.061 | 0.059 |
| power | MLR | 0.862 | 0.822 | 0.785 | 0.863 |
| | HNLR | 0.155 | 0.236 | 0.682 | 0.570 |
| | HN | 0.765 | 0.833 | 0.985 | 0.946 |
| $p=80$ | $(n_1,n_2)$ | (90,90) | (90,110) | (130,130) | (100,160) |
| size | MLR | 1.000 | 0.998 | 0.451 | 0.897 |
| | HNLR | 0.051 | 0.050 | 0.052 | 0.051 |
| | HN | 0.057 | 0.057 | 0.056 | 0.057 |
| power | MLR | 1.000 | 1.000 | 1.000 | 1.000 |
| | HNLR | 0.226 | 0.326 | 0.756 | 0.668 |
| | HN | 0.990 | 0.995 | 0.987 | 0.999 |

Table 2: Third and fourth order cumulants

| | (D1) | (D2) | (D3) | (D4) |
|---|---|---|---|---|
| $\kappa_3$ | 0.00 | 0.89 | 0.00 | -0.67 |
| $\kappa_4$ | 0.00 | 1.20 | 1.00 | 0.51 |

Table 3: (D1) Normal distribution

| $(n_1, n_2)\backslash p$ | | 32 | 64 | 128 | 256 |
|---|---|---|---|---|---|
| (20,20) | size | 0.065 | 0.062 | 0.060 | 0.059 |
| | power | 0.302 | 0.301 | 0.297 | 0.297 |
| (30,10) | size | 0.062 | 0.058 | 0.056 | 0.054 |
| | power | 0.278 | 0.271 | 0.260 | 0.257 |
| (50,50) | size | 0.061 | 0.058 | 0.057 | 0.056 |
| | power | 0.757 | 0.782 | 0.799 | 0.814 |
| (70,30) | size | 0.061 | 0.058 | 0.055 | 0.054 |
| | power | 0.721 | 0.745 | 0.760 | 0.775 |
| (80,80) | size | 0.061 | 0.058 | 0.056 | 0.055 |
| | power | 0.958 | 0.974 | 0.982 | 0.987 |
| (120,40) | size | 0.060 | 0.057 | 0.055 | 0.054 |
| | power | 0.910 | 0.933 | 0.948 | 0.959 |
| (100,100) | size | 0.061 | 0.057 | 0.055 | 0.054 |
| | power | 0.991 | 0.996 | 0.998 | 0.999 |
| (150,50) | size | 0.060 | 0.056 | 0.054 | 0.053 |
| | power | 0.972 | 0.984 | 0.990 | 0.993 |

Table 4: (D2) Standardized chi-square distribution

| $(n_1, n_2)\backslash p$ | | 32 | 64 | 128 | 256 |
|---|---|---|---|---|---|
| (20,20) | size | 0.075 | 0.068 | 0.064 | 0.062 |
| | power | 0.178 | 0.173 | 0.167 | 0.165 |
| (30,10) | size | 0.071 | 0.064 | 0.060 | 0.057 |
| | power | 0.186 | 0.173 | 0.165 | 0.159 |
| (50,50) | size | 0.071 | 0.065 | 0.061 | 0.058 |
| | power | 0.443 | 0.447 | 0.450 | 0.455 |
| (70,30) | size | 0.070 | 0.064 | 0.059 | 0.057 |
| | power | 0.437 | 0.435 | 0.436 | 0.435 |
| (80,80) | size | 0.070 | 0.064 | 0.060 | 0.057 |
| | power | 0.716 | 0.739 | 0.758 | 0.770 |
| (120,40) | size | 0.070 | 0.063 | 0.059 | 0.057 |
| | power | 0.647 | 0.665 | 0.681 | 0.692 |
| (100,100) | size | 0.070 | 0.064 | 0.059 | 0.057 |
| | power | 0.849 | 0.873 | 0.891 | 0.904 |
| (150,50) | size | 0.069 | 0.063 | 0.058 | 0.056 |
| | power | 0.776 | 0.806 | 0.825 | 0.840 |

Table 5: (D3) Standardized $t$ distribution

| $(n_1, n_2)\backslash p$ | | 32 | 64 | 128 | 256 |
|---|---|---|---|---|---|
| (20,20) | size | 0.068 | 0.064 | 0.061 | 0.059 |
| | power | 0.285 | 0.282 | 0.280 | 0.275 |
| (30,10) | size | 0.064 | 0.059 | 0.057 | 0.055 |
| | power | 0.272 | 0.257 | 0.250 | 0.243 |
| (50,50) | size | 0.064 | 0.060 | 0.057 | 0.056 |
| | power | 0.720 | 0.743 | 0.762 | 0.777 |
| (70,30) | size | 0.063 | 0.059 | 0.056 | 0.055 |
| | power | 0.686 | 0.710 | 0.722 | 0.734 |
| (80,80) | size | 0.063 | 0.059 | 0.057 | 0.055 |
| | power | 0.942 | 0.962 | 0.971 | 0.978 |
| (120,40) | size | 0.062 | 0.058 | 0.056 | 0.054 |
| | power | 0.883 | 0.911 | 0.931 | 0.943 |
| (100,100) | size | 0.063 | 0.059 | 0.056 | 0.054 |
| | power | 0.985 | 0.990 | 0.996 | 0.997 |
| (150,50) | size | 0.062 | 0.058 | 0.055 | 0.053 |
| | power | 0.957 | 0.973 | 0.983 | 0.989 |

Table 6: (D4) Standardized skew normal distribution

| $(n_1, n_2)\backslash p$ | | 32 | 64 | 128 | 256 |
|---|---|---|---|---|---|
| (20,20) | size | 0.070 | 0.066 | 0.063 | 0.061 |
| | power | 0.397 | 0.403 | 0.404 | 0.402 |
| (30,10) | size | 0.067 | 0.061 | 0.057 | 0.056 |
| | power | 0.360 | 0.350 | 0.345 | 0.341 |
| (50,50) | size | 0.066 | 0.061 | 0.059 | 0.057 |
| | power | 0.877 | 0.902 | 0.919 | 0.934 |
| (70,30) | size | 0.066 | 0.061 | 0.058 | 0.055 |
| | power | 0.839 | 0.870 | 0.890 | 0.906 |
| (80,80) | size | 0.065 | 0.061 | 0.058 | 0.056 |
| | power | 0.992 | 0.996 | 0.998 | 0.999 |
| (120,40) | size | 0.065 | 0.060 | 0.056 | 0.055 |
| | power | 0.969 | 0.982 | 0.989 | 0.994 |
| (100,100) | size | 0.065 | 0.060 | 0.057 | 0.055 |
| | power | 0.999 | 1.000 | 1.000 | 1.000 |
| (150,50) | size | 0.065 | 0.060 | 0.057 | 0.055 |
| | power | 0.994 | 0.998 | 0.999 | 1.000 |

Table 7: The values of $t_{gh}$. The mark "$*$" represents 5 percent significance.

| | EWS | BL | NB | RMS |
|---|---|---|---|---|
| EWS | — | 25.51* | 28.40* | 15.90* |
| BL | — | — | 17.67* | 30.20* |
| NB | — | — | — | 24.35* |

# Supplemental Material: Simultaneous testing of the mean vector and the covariance matrix for high-dimensional data

Masashi Hyodo*and Takahiro Nishiyama†

March 6, 2017

## 1 Proof of Lemma A.1

Firstly, we show (i):

$$\mathrm{E}[\boldsymbol{Z}'A\boldsymbol{Z}\boldsymbol{b}'\boldsymbol{Z}] = \mathrm{E}\left[\sum_{i,j,\ell=1}^{p} Z_i a_{ij} Z_j b_\ell Z_\ell\right] = \sum_{i=1}^{p} a_{ii} b_i \mathrm{E}\left[Z_i^3\right] = \kappa_3 \sum_{i=1}^{p} a_{ii} b_i.$$

Secondly, we show (ii):

$$\begin{aligned}\mathrm{E}[\boldsymbol{Z}'A\boldsymbol{Z}\boldsymbol{Z}'B\boldsymbol{Z}] &= \mathrm{E}\left[\left(\sum_{i=1}^{p} a_{ii}Z_i^2 + \sum_{\substack{i,j=1\\ i\neq j}}^{m} a_{ij}Z_iZ_j\right)\left(\sum_{i=1}^{p} b_{ii}Z_i^2 + \sum_{\substack{i,j=1\\ i\neq j}}^{p} b_{ij}Z_iZ_j\right)\right]\\ &= \sum_{i=1}^{p} a_{ii}b_{ii}\mathrm{E}[Z_i^4] + \sum_{\substack{i,j=1\\ i\neq j}}^{p} a_{ii}b_{jj}\mathrm{E}[Z_i^2Z_j^2] + \sum_{\substack{i,j=1\\ i\neq j}}^{p}(a_{ij}b_{ij}+a_{ij}b_{ji})\mathrm{E}[Z_i^2Z_j^2]\\ &= \kappa_4\mathrm{tr}(A\odot B) + \mathrm{tr}A\mathrm{tr}B + 2\mathrm{tr}(AB).\end{aligned}$$

Thirdly, we show (iii):

$$\begin{aligned}\mathrm{E}[(\boldsymbol{Z}'A\boldsymbol{W})^4] &= \kappa_4\mathrm{E}[\mathrm{tr}\{(A\boldsymbol{W}\boldsymbol{W}'A)\odot(A\boldsymbol{W}\boldsymbol{W}'A)\}] + 3\mathrm{E}[(\boldsymbol{W}'A^2\boldsymbol{W})^2]\\ &= \kappa_4^2\sum_{i=1}^{p}\sum_{j=1}^{p}(\boldsymbol{e}_j'A\boldsymbol{e}_i\boldsymbol{e}_i'A\boldsymbol{e}_j)^2 + 3\kappa_4\sum_{i=1}^{p}(\boldsymbol{e}_i'A^2\boldsymbol{e}_i)^2\\ &\quad +3\kappa_4\mathrm{tr}(A^2\odot A^2) + 3(\mathrm{tr}A^2)^2 + 6\mathrm{tr}A^4.\end{aligned}$$

*Department of Mathematical Sciences, Graduate School of Engineering, Osaka Prefecture University, 1-1 Gakuen-cho, Naka-ku, Sakai, Osaka 599-8531, Japan. E-Mail:hyodo@ms.osakafu-u.ac.jp

†Department of Business Administration, Senshu University, 2-1-1, Higashimita, Tama-ku, Kawasaki-shi, Kanagawa 214-8580, Japan.

Finally, we show (iv). Using Hölder's inequality,

$$
\begin{aligned}
(\boldsymbol{Z}'A\boldsymbol{Z} - \mathrm{tr}A)^4 &= \left\{\sum_{i=1}^{p} a_{ii}(Z_i^2 - 1) + \sum_{i\neq j}^{p} a_{ij}Z_iZ_j\right\}^4 \\
&\leq 2^{4-1}\left[\left\{\sum_{i=1}^{p} a_{ii}(Z_i^2 - 1)\right\}^4 + \left(\sum_{i\neq j}^{p} a_{ij}Z_iZ_j\right)^4\right]
\end{aligned}
$$

Thus

$$
\mathrm{E}\left[(\boldsymbol{Z}'A\boldsymbol{Z} - \mathrm{tr}A)^4\right] \leq 8\mathrm{E}\left[\left\{\sum_{i=1}^{p} a_{ii}(Z_i^2 - 1)\right\}^4\right] + 8\mathrm{E}\left[\left(\sum_{i\neq j}^{p} a_{ij}Z_iZ_j\right)^4\right]. \tag{1.1}
$$

At first, we evaluate the first term on right hand side in (1.1):

$$
\begin{aligned}
\mathrm{E}\left[\left(\sum_{i=1}^{p} a_{ii}(Z_i^2 - 1)\right)^4\right] &= \mathrm{E}\left[\left\{\sum_{i=1}^{p} a_{ii}^2(Z_i^2 - 1)^2\right.\right. \\
&\qquad \left.\left. + \sum_{i\neq j}^{p} a_{ii}a_{jj}(Z_i^2 - 1)(Z_j^2 - 1)\right\}^2\right] \\
&= \sum_{i=1}^{p} a_{ii}^4 \mathrm{E}[(Z_i^2 - 1)^4] \\
&\quad +3\sum_{i\neq j} a_{ii}^2 a_{jj}^2 \mathrm{E}[(Z_i^2 - 1)^2]\mathrm{E}[(Z_j^2 - 1)^2] \\
&\leq (\kappa_8 + 24\kappa_6 + 56\kappa_3\kappa_5 + 35\kappa_4^2 + 156\kappa_4 \\
&\quad +240\kappa_3 + 60)\mathrm{tr}(A\odot A\odot A\odot A) \\
&\quad +3(\kappa_4 + 2)\{\mathrm{tr}(A\odot A)\}^2.
\end{aligned} \tag{1.2}
$$

Next we evaluate the second term on right hand side in (1.1). Using Hölder's inequality,

$$
\begin{aligned}
\mathrm{E}\left[\left(\sum_{i\neq j}^{p} a_{ij}Z_iZ_j\right)^4\right] &= \mathrm{E}\left[\left(2\sum_{i\neq j}^{p} a_{ij}^2 Z_i^2 Z_j^2 + 4\sum_{i\neq j\neq k\neq i}^{p} a_{ij}a_{ik}Z_i^2Z_jZ_k\right.\right. \\
&\qquad \left.\left. + \sum_{\substack{i\neq j\neq k\neq \ell \\ j\neq \ell\neq i\neq k}}^{p} a_{ij}a_{k\ell}Z_iZ_jZ_kZ_\ell\right)^2\right] \\
&\leq 12\mathrm{E}\left[\left(\sum_{i\neq j}^{p} a_{ij}^2 Z_i^2 Z_j^2\right)^2\right]
\end{aligned}
$$

$$+48\mathrm{E}\left[\left(\sum_{i\neq j\neq k\neq i}^{p} a_{ij}a_{ik}Z_i^2Z_jZ_k\right)^2\right]$$

$$+3\mathrm{E}\left[\left(\sum_{\substack{i\neq j\neq k\neq \ell\\ j\neq \ell\neq i\neq k}}^{p} a_{ij}a_{k\ell}Z_iZ_jZ_kZ_\ell\right)^2\right]. \tag{1.3}$$

The each term on right hand side in (1.3) is evaluated as

$$\begin{aligned}
\mathrm{E}\left[\left(\sum_{\substack{i\neq j\neq k\neq \ell\\ j\neq \ell\neq i\neq k}}^{p} a_{ij}a_{k\ell}Z_iZ_jZ_kZ_\ell\right)^2\right] &= \sum_{\substack{i\neq j\neq k\neq \ell\\ j\neq \ell\neq i\neq k}}^{p} 8(a_{ij}^2a_{k\ell}^2 + 2a_{ij}a_{k\ell}a_{ik}a_{j\ell})\mathrm{E}[Z_i^2Z_j^2Z_k^2Z_\ell^2] \\
&\leq 24\sum_{\substack{i\neq j\neq k\neq \ell\\ j\neq \ell\neq i\neq k}}^{p} a_{ij}^2a_{k\ell}^2 \\
&\leq 24\left(\sum_{i,j=1}^{p} a_{ij}^2\right)^2, \qquad (1.4)\\
\mathrm{E}\left[\left(\sum_{i\neq j\neq k\neq i}^{p} a_{ij}a_{ik}Z_i^2Z_jZ_k\right)^2\right] &= 2\sum_{i\neq j\neq k\neq i}^{p} a_{ij}^2a_{ik}^2\mathrm{E}[Z_i^4Z_j^2Z_k^2] \\
&\quad +4\sum_{i\neq j\neq k\neq i}^{p} a_{ij}a_{ik}a_{ji}a_{jk}\mathrm{E}[Z_i^3Z_j^3Z_k^2] \\
&\quad +2\sum_{\substack{i\neq j\neq k\neq \ell\\ j\neq \ell\neq i\neq k}}^{p} a_{ij}a_{ik}a_{\ell j}a_{\ell k}\mathrm{E}[Z_i^2Z_j^2Z_k^2Z_\ell^2] \\
&\leq 2(\kappa_4+3)\sum_{i\neq j\neq k\neq i}^{p} a_{ij}^2a_{ik}^2 \\
&\quad +4\kappa_3^2\sum_{i\neq j\neq k\neq i}^{p} a_{ij}a_{ik}a_{ji}a_{jk} \\
&\quad +2\sum_{\substack{i\neq j\neq k\neq \ell\\ j\neq \ell\neq i\neq k}}^{p} a_{ij}a_{ik}a_{\ell j}a_{\ell k} \\
&\leq 2(\kappa_4+2\kappa_3^2+4)\left(\sum_{i,j=1}^{p} a_{ij}^2\right)^2, \qquad (1.5)
\end{aligned}$$

and

$$\begin{aligned}
\mathrm{E}\left[\left(\sum_{i\neq j}^{p} a_{ij}^2 Z_i^2 Z_j^2\right)^2\right] &= 2\sum_{i\neq j}^{p} a_{ij}^4 \mathrm{E}[Z_i^4 Z_j^4] + 4\sum_{i\neq j\neq k\neq i}^{p} a_{ij}^2 a_{ik}^2 \mathrm{E}[Z_i^4 Z_j^2 Z_k^2] \\
&\quad + \sum_{\substack{i\neq j\neq k\neq \ell \\ j\neq \ell\neq i\neq k}}^{p} a_{ij}^2 a_{k\ell}^2 \mathrm{E}[Z_i^2 Z_j^2 Z_k^2 Z_\ell^2] \\
&\leq \left\{2(\kappa_4+3)^2 + 4(\kappa_4+3) + 1\right\} \left(\sum_{i,j=1}^{p} a_{ij}^2\right)^2. \qquad (1.6)
\end{aligned}$$

Combining (1.1)-(1.4), we obtain (iv). From (1.2), (1.5)-(1.6), we also note that "*Const.*" does not depend on the matrix $A$. $\square$

## 2 The asymptotic property of $\sigma_{12}/(\sigma_1\sigma_2)$

From Cauchy-Schwarz inequality,

$$|\kappa_{3g}| \le \mathrm{E}[|Z_i(Z_i^2-1)|] \le \sqrt{\mathrm{E}[Z_i^2]}\sqrt{\mathrm{E}[(Z_i^2-1)^2]} = \sqrt{\kappa_{4g}+2} \Rightarrow \kappa_{3g}^2 - \kappa_{4g} - 2 \le 0.$$

Note that

$$\begin{aligned}
0 \le \sigma_{12}^2 &= \left(\sum_{g=1}^{2} \frac{4\kappa_{3g}\sum_{i=1}^{p} \boldsymbol{\delta}'\Sigma_g^{1/2}\boldsymbol{e}_i\boldsymbol{e}_i'\Sigma_g^{1/2}\Delta\Sigma_g^{1/2}\boldsymbol{e}_i}{n_g}\right)^2 \\
&\le \left(\sum_{g=1}^{2}\sqrt{\frac{4\boldsymbol{\delta}'\Sigma_g\boldsymbol{\delta}}{n_g}}\sqrt{\frac{4\kappa_{3g}^2\mathrm{tr}\{(\Sigma_g^{1/2}\Delta\Sigma_g^{1/2})\odot(\Sigma_g^{1/2}\Delta\Sigma_g^{1/2})\}}{n_g}}\right)^2 \\
&\le \sum_{g=1}^{2}\frac{4\boldsymbol{\delta}'\Sigma_g\boldsymbol{\delta}}{n_g}\sum_{g=1}^{2}\frac{4\kappa_{3g}^2\mathrm{tr}\{(\Sigma_g^{1/2}\Delta\Sigma_g^{1/2})\odot(\Sigma_g^{1/2}\Delta\Sigma_g^{1/2})\}}{n_g} \\
&= (\sigma_1^2-\sigma_{10}^2)\left\{\sigma_2^2-\sigma_{20}^2+\sum_{g=1}^{2}\frac{4(\kappa_{3g}^2-\kappa_{4g}-2)\mathrm{tr}\{(\Sigma_g^{1/2}\Delta\Sigma_g^{1/2})\odot(\Sigma_g^{1/2}\Delta\Sigma_g^{1/2})\}}{n_g}\right. \\
&\quad \left.+\sum_{g=1}^{2}\frac{8[\mathrm{tr}\{(\Sigma_g^{1/2}\Delta\Sigma_g^{1/2})\odot(\Sigma_g^{1/2}\Delta\Sigma_g^{1/2})\}-\mathrm{tr}(\Delta\Sigma_g)^2]}{n_g}\right\} \\
&\le (\sigma_1^2-\sigma_{10}^2)(\sigma_2^2-\sigma_{20}^2).
\end{aligned}$$

Let $\rho = \sigma_{12}/(\sigma_1\sigma_2)$. Then we get

$$\rho^2 \le \left(1-\frac{\sigma_{10}^2}{\sigma_1^2}\right)\left(1-\frac{\sigma_{20}^2}{\sigma_2^2}\right). \tag{2.1}$$

Also note that

$$0 \le 1-\frac{\sigma_{10}^2}{\sigma_1^2} \le \sum_{g=1}^{2}\frac{2n_g\boldsymbol{\delta}'\Sigma_g\boldsymbol{\delta}}{\mathrm{tr}\Sigma_g^2}\left(1+\sum_{g=1}^{2}\frac{2n_g\boldsymbol{\delta}'\Sigma_g\boldsymbol{\delta}}{\mathrm{tr}\Sigma_g^2}\right)^{-1}, \tag{2.2}$$

$$0 \le 1-\frac{\sigma_{20}^2}{\sigma_2^2} \le \sum_{g=1}^{2}\frac{(2+\kappa_{4g})n_g\mathrm{tr}(\Delta\Sigma_g)^2}{(\mathrm{tr}\Sigma_g^2)^2}\left(1+\sum_{g=1}^{2}\frac{(2+\kappa_{4g})n_g\mathrm{tr}(\Delta\Sigma_g)^2}{(\mathrm{tr}\Sigma_g^2)^2}\right)^{-1}. \tag{2.3}$$

From (2.2) and (2.3), note that

$$(\mathrm{A3}) \Rightarrow \limsup_{p\to\infty}\frac{2n_g\boldsymbol{\delta}'\Sigma_g\boldsymbol{\delta}}{\mathrm{tr}\Sigma_g^2} < \infty \Rightarrow 0 \le \limsup_{p\to\infty}\left(1-\frac{\sigma_{10}^2}{\sigma_1^2}\right) < 1, \tag{2.4}$$

$$(\mathrm{A3}) \Rightarrow \limsup_{p\to\infty}\frac{n_g\mathrm{tr}(\Delta\Sigma_g)^2}{(\mathrm{tr}\Sigma_g^2)^2} < \infty \Rightarrow 0 \le \limsup_{p\to\infty}\left(1-\frac{\sigma_{20}^2}{\sigma_2^2}\right) < 1. \tag{2.5}$$

From (2.1), (2.4) and (2.5), $0 \le \lim_{p\to\infty}\rho^2 < 1$ under (A3).

## 3 The asymptotic property of test statistic under (A4)

The following inequality holds:

$$\begin{aligned}\frac{\sigma_1^2}{\|\boldsymbol{\delta}\|^4} &= \frac{1}{(\boldsymbol{\delta}'\boldsymbol{\delta})^2}\left\{\sum_{g=1}^{2}\frac{2\mathrm{tr}\Sigma_g^2}{n_g^2}+\frac{4\mathrm{tr}(\Sigma_1\Sigma_2)}{n_1n_2}+\sum_{g=1}^{2}\frac{4\boldsymbol{\delta}'\Sigma_g\boldsymbol{\delta}}{n_g}\right\}\\ &\leq \sum_{g=1}^{2}2\left(\frac{\mathrm{tr}\Sigma_g^2}{n_g\boldsymbol{\delta}'\Sigma_g\boldsymbol{\delta}}\right)^2+\prod_{g=1}^{2}\frac{4\mathrm{tr}\Sigma_g^2}{n_g\boldsymbol{\delta}'\Sigma_g\boldsymbol{\delta}}+\sum_{g=1}^{2}\frac{4\mathrm{tr}\Sigma_g^2}{n_g\boldsymbol{\delta}'\Sigma_g\boldsymbol{\delta}},\end{aligned}$$

since

$$\begin{aligned}\frac{\mathrm{tr}\Sigma_g^2}{(\boldsymbol{\delta}'\boldsymbol{\delta})^2n_g^2} &= \frac{(\boldsymbol{\delta}'\Sigma_g\boldsymbol{\delta})^2}{(\boldsymbol{\delta}'\boldsymbol{\delta})^2\mathrm{tr}\Sigma_g^2}\left(\frac{\mathrm{tr}\Sigma_g^2}{n_g\boldsymbol{\delta}'\Sigma_g\boldsymbol{\delta}}\right)^2\\ &\leq \frac{(\boldsymbol{\delta}'\boldsymbol{\delta})^2\mathrm{tr}\Sigma_g^2}{(\boldsymbol{\delta}'\boldsymbol{\delta})^2\mathrm{tr}\Sigma_g^2}\left(\frac{\mathrm{tr}\Sigma_g^2}{n_g\boldsymbol{\delta}'\Sigma_g\boldsymbol{\delta}}\right)^2=\left(\frac{\mathrm{tr}\Sigma_g^2}{n_g\boldsymbol{\delta}'\Sigma_g\boldsymbol{\delta}}\right)^2,\\ \frac{\mathrm{tr}(\Sigma_1\Sigma_2)}{(\boldsymbol{\delta}'\boldsymbol{\delta})^2n_1n_2} &\leq \prod_{g=1}^{2}\frac{\sqrt{\mathrm{tr}\Sigma_g^2}}{\boldsymbol{\delta}'\boldsymbol{\delta}n_g}\leq\prod_{g=1}^{2}\frac{\mathrm{tr}\Sigma_g^2}{n_g\boldsymbol{\delta}'\Sigma_g\boldsymbol{\delta}},\\ \frac{\boldsymbol{\delta}'\Sigma_g\boldsymbol{\delta}}{n_g(\boldsymbol{\delta}'\boldsymbol{\delta})^2} &= \frac{(\boldsymbol{\delta}'\Sigma_g\boldsymbol{\delta})^2}{n_g\boldsymbol{\delta}'\Sigma_g\boldsymbol{\delta}(\boldsymbol{\delta}'\boldsymbol{\delta})^2}\leq\frac{\mathrm{tr}\Sigma_g^2}{n_g\boldsymbol{\delta}'\Sigma_g\boldsymbol{\delta}}.\end{aligned}$$

The following inequality holds:

$$\begin{aligned}\frac{\sigma_2^2}{\|\Delta\|^4} &\leq \frac{1}{(\mathrm{tr}\Delta^2)^2}\left\{\sum_{g=1}^{2}\frac{4(\mathrm{tr}\Sigma_g^2)^2}{n_g^2}+\frac{8\mathrm{tr}\Sigma_1^2\mathrm{tr}\Sigma_2^2}{n_1n_2}+\sum_{g=1}^{2}\frac{4(2+\kappa_{4g})\mathrm{tr}(\Sigma_g\Delta)^2}{n_g}\right\}\\ &\leq \sum_{g=1}^{2}4\left\{\frac{(\mathrm{tr}\Sigma_g^2)^2}{n_g\mathrm{tr}(\Sigma_g\Delta)^2}\right\}^2+8\prod_{g=1}^{2}\frac{(\mathrm{tr}\Sigma_g^2)^2}{n_g\mathrm{tr}(\Sigma_g\Delta)^2}+\sum_{g=1}^{2}\frac{4(2+\kappa_{4g})(\mathrm{tr}\Sigma_g^2)^2}{n_g\mathrm{tr}(\Sigma_g\Delta)^2},\end{aligned}$$

since

$$\begin{aligned}\frac{(\mathrm{tr}\Sigma_g^2)^2}{(\mathrm{tr}\Delta^2)^2n_g^2} &= \frac{\{\mathrm{tr}(\Sigma_g\Delta)^2\}^2}{(\mathrm{tr}\Delta^2)^2(\mathrm{tr}\Sigma_g^2)^2}\left\{\frac{(\mathrm{tr}\Sigma_g^2)^2}{n_g\mathrm{tr}(\Sigma_g\Delta)^2}\right\}^2\leq\left\{\frac{(\mathrm{tr}\Sigma_g^2)^2}{n_g\mathrm{tr}(\Sigma_g\Delta)^2}\right\}^2,\\ \frac{\{\mathrm{tr}(\Sigma_1\Sigma_2)\}^2}{(\mathrm{tr}\Delta^2)^2n_1n_2} &\leq \prod_{g=1}^{2}\frac{\mathrm{tr}\Sigma_g^2}{\mathrm{tr}\Delta^2n_g}\leq\prod_{g=1}^{2}\frac{(\mathrm{tr}\Sigma_g^2)^2}{n_g\mathrm{tr}(\Sigma_g\Delta)^2},\\ \frac{\mathrm{tr}(\Sigma_g\Delta)^2}{n_g(\mathrm{tr}\Delta^2)^2} &= \frac{\{\mathrm{tr}(\Sigma_g\Delta)^2\}^2}{(\mathrm{tr}\Delta^2)^2}\frac{1}{n_g\mathrm{tr}(\Sigma_g\Delta)^2}\leq\frac{(\mathrm{tr}\Sigma_g^2)^2}{n_g\mathrm{tr}(\Sigma_g\Delta)^2}.\end{aligned}$$

Thus,

$$\begin{aligned}&\frac{\sigma_{10}^2}{\|\boldsymbol{\delta}\|^4}\to 0,\ \frac{\sigma_1^2}{\|\boldsymbol{\delta}\|^4}\to 0\ \text{ if }\ \frac{n_g\boldsymbol{\delta}'\Sigma_g\boldsymbol{\delta}}{\mathrm{tr}\Sigma_g^2}\to\infty,\\ &\frac{\sigma_{20}^2}{\|\Delta\|^4}\to 0,\ \frac{\sigma_2^2}{\|\Delta\|^4}\to 0\ \text{ if }\ \frac{n_g\mathrm{tr}(\Sigma_g\Delta)^2}{(\mathrm{tr}\Sigma_g^2)^2}\to\infty.\end{aligned}$$

Hence,

$$\text{(i)} \quad \frac{\widehat{\|\boldsymbol{\delta}\|^2}}{\|\boldsymbol{\delta}\|^2} = 1 + o_p(1) \ \text{ if } \ \frac{n_g \boldsymbol{\delta}' \Sigma_g \boldsymbol{\delta}}{\text{tr}\Sigma_g^2} \to \infty. \tag{3.1}$$

$$\text{(ii)} \quad \frac{\widehat{\|\Delta\|_F^2}}{\|\Delta\|_F^2} = 1 + o_p(1) \ \text{ if } \ \frac{n_g \text{tr}(\Sigma_g \Delta)^2}{(\text{tr}\Sigma_g^2)^2} \to \infty. \tag{3.2}$$

From (3.1),

$$\begin{aligned}
\frac{\widehat{\|\boldsymbol{\delta}\|^2}}{\widehat{\sigma}_{10}} &= \frac{\|\boldsymbol{\delta}\|^2}{\sigma_{10}} \times \frac{\sigma_{10}}{\widehat{\sigma}_{10}} \times \frac{\widehat{\|\boldsymbol{\delta}\|^2}}{\|\boldsymbol{\delta}\|^2} \\
&= \frac{\|\boldsymbol{\delta}\|^2}{\sigma_{10}} \times (1 + o_p(1)) \times (1 + o_p(1)) \\
&= \frac{\|\boldsymbol{\delta}\|^2}{\sigma_{10}} (1 + o_p(1))
\end{aligned}$$

under $n_g \boldsymbol{\delta}' \Sigma_g \boldsymbol{\delta} / \text{tr}\Sigma_g^2 \to \infty$. Similarly, from (3.2),

$$\frac{\widehat{\|\Delta\|^2}}{\widehat{\sigma}_{20}} = \frac{\|\Delta\|^2}{\sigma_{20}} (1 + o_p(1))$$

under $n_g \text{tr}(\Sigma_g \Delta)^2 / (\text{tr}\Sigma_g^2)^2 \to \infty$.

Let $n_g$, $\boldsymbol{\delta}' \Sigma_g \boldsymbol{\delta}$, $\text{tr}(\Sigma_g \Delta)^2$, and $\text{tr}\Sigma_g^2$ be functions of $p$. We consider

$$\begin{aligned}
&\text{(A4-i)} \quad \lim_{p\to\infty} \frac{n_g \boldsymbol{\delta}' \Sigma_g \boldsymbol{\delta}}{\text{tr}\Sigma_g^2} = \infty, \quad \limsup_{p\to\infty} \frac{n_g \text{tr}(\Sigma_g \Delta)^2}{(\text{tr}\Sigma_g^2)^2} < \infty, \\
&\text{(A4-ii)} \quad \limsup_{p\to\infty} \frac{n_g \boldsymbol{\delta}' \Sigma_g \boldsymbol{\delta}}{\text{tr}\Sigma_g^2} < \infty, \quad \lim_{p\to\infty} \frac{n_g \text{tr}(\Sigma_g \Delta)^2}{(\text{tr}\Sigma_g^2)^2} = \infty, \\
&\text{(A4-iii)} \quad \lim_{p\to\infty} \frac{n_g \boldsymbol{\delta}' \Sigma_g \boldsymbol{\delta}}{\text{tr}\Sigma_g^2} = \infty, \quad \lim_{p\to\infty} \frac{n_g \text{tr}(\Sigma_g \Delta)^2}{(\text{tr}\Sigma_g^2)^2} = \infty.
\end{aligned}$$

Under (A4-i), it holds that

$$\begin{aligned}
T - \sqrt{2} z_\alpha &= \frac{\|\boldsymbol{\delta}\|^2}{\sigma_{10}} (1 + o_p(1)) + \frac{\|\Delta\|_F^2}{\sigma_2} + O_p(1) - \sqrt{2} z_\alpha \\
&= \frac{\|\boldsymbol{\delta}\|^2}{\sigma_{10}} \left( 1 + o_p(1) + \frac{\sigma_{10} \|\Delta\|_F^2}{\sigma_2 \|\boldsymbol{\delta}\|^2} + O_p \left( \frac{\sigma_{10}}{\|\boldsymbol{\delta}\|^2} \right) \right) \\
&= \frac{\|\boldsymbol{\delta}\|^2}{\sigma_{10}} \left( 1 + \frac{\sigma_{10} \|\Delta\|_F^2}{\sigma_2 \|\boldsymbol{\delta}\|^2} + o_p(1) \right).
\end{aligned}$$

Under (A4-ii), it holds that

$$\begin{aligned}
T - \sqrt{2} z_\alpha &= \frac{\|\boldsymbol{\delta}\|^2}{\sigma_1} + O_p(1) + \frac{\|\Delta\|_F^2}{\sigma_{20}} (1 + o_p(1)) - \sqrt{2} z_\alpha \\
&= \frac{\|\Delta\|_F^2}{\sigma_{20}} \left( \frac{\sigma_{20} \|\boldsymbol{\delta}\|^2}{\sigma_1 \|\Delta\|_F^2} + O_p \left( \frac{\sigma_{20}}{\|\Delta\|_F^2} \right) + 1 + o_p(1) \right) \\
&= \frac{\|\Delta\|_F^2}{\sigma_{20}} \left( 1 + \frac{\sigma_{20} \|\boldsymbol{\delta}\|^2}{\sigma_1 \|\Delta\|_F^2} + o_p(1) \right).
\end{aligned}$$

Under (A4-iii), it holds that

$$\begin{aligned} T-\sqrt{2}z_{\alpha} &= \frac{\|\boldsymbol{\delta}\|^2}{\sigma_{10}}(1+o_p(1))+\frac{\|\Delta\|_F^2}{\sigma_{20}}(1+o_p(1))-\sqrt{2}z_{\alpha} \\ &= \frac{\|\boldsymbol{\delta}\|^2}{\sigma_{10}}(1+o_p(1))+\frac{\|\Delta\|_F^2}{\sigma_{20}}(1+o_p(1)). \end{aligned}$$

# 4 Derivation of (A)

At first, we evaluate $\sum_{k=1}^{n_1} \mathrm{E}[C_k^{(1)}]$. We rewrite

$$\sum_{k=1}^{n_1} C_k^{(1)} = \frac{\boldsymbol{\delta}'\Sigma_1\boldsymbol{\delta}}{n_1} + \sum_{i=1}^{3} \widetilde{C}_i^{(1)},$$

where

$$\begin{aligned}
\widetilde{C}_1^{(1)} &= \sum_{k=1}^{n_1} \frac{n_1 - k}{n_1^2(n_1-1)^2} \boldsymbol{Y}'_{1k}\Sigma_1\boldsymbol{Y}_{1k}, \\
\widetilde{C}_2^{(1)} &= \sum_{k=2}^{n_1}\sum_{\ell=1}^{k-1} \frac{2(n_1 - k)}{n_1^2(n_1-1)^2} \boldsymbol{Y}'_{1k}\Sigma_1\boldsymbol{Y}_{1\ell}, \\
\widetilde{C}_3^{(1)} &= \sum_{k=1}^{n_1} \frac{2(n_1 - k)}{n_1^2(n_1-1)} \boldsymbol{\delta}'\Sigma_1\boldsymbol{Y}_{1k}.
\end{aligned}$$

By using Lemma A.1, we have

$$\mathrm{E}\left[\widetilde{C}_1^{(1)}\right] = \frac{1}{2n_1(n_1-1)}\mathrm{tr}\Sigma_1^2,\ \mathrm{E}\left[\widetilde{C}_2^{(1)}\right] = \mathrm{E}\left[\widetilde{C}_3^{(1)}\right] = 0 \tag{4.1}$$

$$\mathrm{Var}\left[\widetilde{C}_1^{(1)}\right] = O\left(\frac{\mathrm{tr}\Sigma_1^4}{n_1^5}\right) = o(\sigma_1^4), \tag{4.2}$$

$$\mathrm{Var}\left[\widetilde{C}_2^{(1)}\right] = O\left(\frac{\mathrm{tr}\Sigma_1^4}{n_1^4}\right) = o(\sigma_1^4), \tag{4.3}$$

$$\mathrm{Var}\left[\widetilde{C}_3^{(1)}\right] = O\left(\frac{\boldsymbol{\delta}'\Sigma_1\boldsymbol{\delta}}{n_1}\frac{\sqrt{\mathrm{tr}\Sigma_1^4}}{n_1^2}\right) = o(\sigma_1^4). \tag{4.4}$$

Using Chebyshev's inequality with (4.1)-(4.4), we obtain

$$\frac{\widetilde{C}_1^{(1)}}{\mathrm{tr}\Sigma_1^2/(2n_1^2)} = 1 + o_p(1),\ \frac{\widetilde{C}_k^{(1)}}{\sigma_1^2} = o_p(1)$$

for $k = 2, 3$. Thus

$$\frac{\sum_{k=1}^{n_1} C_k^{(1)}}{\sigma_1^2} = \frac{\mathrm{tr}\Sigma_1^2/(2n_1^2) + \boldsymbol{\delta}'\Sigma_1\boldsymbol{\delta}/n_1}{\sigma_1^2} + o_p(1). \tag{4.5}$$

Next, we evaluate $\sum_{k=n_1+1}^{n} \mathrm{E}[C_k^{(2)}]$. We rewrite

$$\sum_{k=n_1+1}^{n} C_k^{(2)} = \frac{\boldsymbol{\delta}'\Sigma_2\boldsymbol{\delta}}{n_2} + \sum_{i=1}^{5} \widetilde{C}_i^{(2)},$$

where

$$\begin{aligned}
\widetilde{C}_1^{(2)} &= \sum_{k=1}^{n_2} \frac{n_2-k}{n_2^2(n_2-1)^2} \boldsymbol{Y}'_{2k}\Sigma_2\boldsymbol{Y}_{2k} + 2\sum_{k=2}^{n_2}\sum_{\ell=1}^{k-1} \frac{n_2-k}{n_2^2(n_2-1)^2} \boldsymbol{Y}'_{2k}\Sigma_2\boldsymbol{Y}_{2\ell}, \\
\widetilde{C}_2^{(2)} &= \frac{1}{n_1^2 n_2}\left(\sum_{k=1}^{n_1} \boldsymbol{Y}'_{1k}\Sigma_2\boldsymbol{Y}_{1k} + \sum_{k\neq\ell}^{n_1} \boldsymbol{Y}'_{1k}\Sigma_2\boldsymbol{Y}_{1\ell}\right), \\
\widetilde{C}_3^{(2)} &= -\frac{2}{n_1 n_2^2(n_2-1)} \sum_{k=1}^{n_2}\sum_{\ell=1}^{n_1} (n_2-k)\boldsymbol{Y}'_{2k}\Sigma_2\boldsymbol{Y}_{1\ell}, \\
\widetilde{C}_4^{(2)} &= -\sum_{k=1}^{n_2} \frac{2(n_2-k)}{n_2^2(n_2-1)} \boldsymbol{Y}'_{2k}\Sigma_2\boldsymbol{\delta}, \widetilde{C}_5^{(2)} = \sum_{k=1}^{n_1} \frac{2}{n_1 n_2} \boldsymbol{Y}'_{1k}\Sigma_2\boldsymbol{\delta}.
\end{aligned}$$

By using Lemma A.1, we have

$$\mathrm{E}\left[\widetilde{C}_1^{(2)}\right] = \frac{\mathrm{tr}\Sigma_2^2}{2n_2(n_2-1)}, \mathrm{E}\left[\widetilde{C}_2^{(2)}\right] = \frac{\mathrm{tr}(\Sigma_1\Sigma_2)}{n_1 n_2},$$

$$\mathrm{E}\left[\widetilde{C}_k^{(2)}\right] = 0 \ (k=3,4,5), \tag{4.6}$$

$$\mathrm{Var}\left[\widetilde{C}_1^{(2)}\right] = O\left(\frac{\mathrm{tr}\Sigma_2^4}{n_2^4}\right) = o(\sigma_1^4), \tag{4.7}$$

$$\mathrm{Var}\left[\widetilde{C}_2^{(2)}\right] = O\left(\frac{\mathrm{tr}(\Sigma_1\Sigma_2)^2}{n_1^2 n_2^2}\right) = o(\sigma_1^4), \tag{4.8}$$

$$\mathrm{Var}\left[\widetilde{C}_3^{(2)}\right] = O\left(\frac{\sqrt{\mathrm{tr}(\Sigma_1\Sigma_2)^2\mathrm{tr}\Sigma_2^4}}{n_1 n_2^3}\right) = o(\sigma_1^4), \tag{4.9}$$

$$\mathrm{Var}\left[\widetilde{C}_4^{(2)}\right] = O\left(\frac{\boldsymbol{\delta}'\Sigma_2\boldsymbol{\delta}}{n_2}\frac{\sqrt{\mathrm{tr}\Sigma_2^4}}{n_2^2}\right) = o(\sigma_1^4), \tag{4.10}$$

$$\mathrm{Var}\left[\widetilde{C}_5^{(2)}\right] = O\left(\frac{\boldsymbol{\delta}'\Sigma_2\boldsymbol{\delta}}{n_2}\frac{\sqrt{\mathrm{tr}(\Sigma_1\Sigma_2)^2}}{n_1 n_2}\right) = o(\sigma_1^4). \tag{4.11}$$

Using Chebyshev's inequality with (4.6)-(4.11), we obtain

$$\frac{\widetilde{C}_1^{(2)}}{\mathrm{tr}\Sigma_2^2/(2n_1^2)} = 1 + o_p(1), \ \frac{\widetilde{C}_2^{(2)}}{\mathrm{tr}(\Sigma_1\Sigma_2)/(n_1 n_2)} = 1 + o_p(1),$$

and

$$\frac{\widetilde{C}_k^{(2)}}{\sigma_1^2} = o_p(1)$$

for $k = 3, 4, 5$. Thus

$$\frac{\sum_{k=n_1+1}^{n} C_k^{(2)}}{\sigma_1^2} = \frac{\mathrm{tr}\Sigma_2^2/(2n_2^2) + \mathrm{tr}(\Sigma_1\Sigma_2)/(n_1 n_2) + \boldsymbol{\delta}'\Sigma_2\boldsymbol{\delta}/n_2}{\sigma_1^2} + o_p(1). \tag{4.12}$$

From (4.5) and (4.12), statement (A) follows. □

# 5 Derivation of (B)

We assume (A1) and (A2). At first, we evaluate $\sum_{k=1}^{n_1} D_k^{(1)}$. We rewrite

$$D_k^{(1)} = \sum_{j=1}^{k-1} \frac{\mathrm{tr}(\Sigma_1 \mathcal{E}_j^{(1)})^2}{n_1^2(n_1-1)^2} + \sum_{j\neq\ell}^{k-1} \frac{\mathrm{tr}(\Sigma_1 \mathcal{E}_j^{(1)} \Sigma_1 \mathcal{E}_\ell^{(1)})}{n_1^2(n_1-1)^2} + \frac{2\sum_{j=1}^{k-1} \mathrm{tr}(\Sigma_1 \Delta \Sigma_1 \mathcal{E}_j^{(1)})}{n_1^2(n_1-1)} + \frac{\mathrm{tr}(\Sigma_1\Delta)^2}{n_1^2}.$$

By using Lemma A.1, we have

$$\begin{aligned}
\mathrm{E}[\mathrm{tr}(\Sigma_1 \mathcal{E}_1^{(1)})^2] &= \mathrm{E}[(\boldsymbol{Z}_{11}' \Sigma_1^2 \boldsymbol{Z}_{11})^2 - 2(\boldsymbol{Z}_{11}' \Sigma_1^4 \boldsymbol{Z}_{11}) + \mathrm{tr}\Sigma_1^4] \\
&= \kappa_{41}\mathrm{tr}(\Sigma_1^2 \odot \Sigma_1^2) + (\mathrm{tr}\Sigma_1^2)^2 + \mathrm{tr}\Sigma_1^4, && (5.1)\\
\mathrm{E}[\mathrm{tr}(\Sigma_1 \mathcal{E}_1^{(1)} \Sigma_1 \mathcal{E}_2^{(1)})] &= 0, \mathrm{E}[\mathrm{tr}(\Delta\Sigma_1 \mathcal{E}_1^{(1)})] = 0. && (5.2)
\end{aligned}$$

From (5.1)-(5.2),

$$\mathrm{E}\left[\sum_{k=1}^{n_1} D_k^{(1)}\right] = \frac{\kappa_{41}\mathrm{tr}(\Sigma_1^2 \odot \Sigma_1^2) + (\mathrm{tr}\Sigma_1^2)^2 + \mathrm{tr}\Sigma_1^4}{2n_1(n_1-1)} + \frac{\mathrm{tr}(\Sigma_1\Delta)^2}{n_1}. \quad (5.3)$$

We decompose the centered random variable as following:

$$\begin{aligned}
\sum_{k=1}^{n_1} D_k^{(1)} - \mathrm{E}\left[\sum_{k=1}^{n_1} D_k^{(1)}\right] &= \sum_{k=1}^{n_1} \frac{(n_1-k)\{\mathrm{tr}(\Sigma_1 \mathcal{E}_k^{(1)})^2 - \mathrm{E}[\mathrm{tr}(\Sigma_1 \mathcal{E}_k^{(1)})^2]\}}{n_1^2(n_1-1)^2} \\
&\quad + \sum_{k=2}^{n_1}\sum_{\ell=1}^{k-1} \frac{2(n_1-k)\mathrm{tr}(\Sigma_1 \mathcal{E}_k^{(1)} \Sigma_1 \mathcal{E}_\ell^{(1)})}{n_1^2(n_1-1)^2} \\
&\quad + \sum_{k=1}^{n_1} \frac{2(n_1-k)\mathrm{tr}(\Sigma_1\Delta\Sigma_1 \mathcal{E}_k^{(1)})}{n_1^2(n_1-1)}.
\end{aligned}$$

By using Holder's inequality, we obtain the following

$$\begin{aligned}
\mathrm{Var}\left[\sum_{k=1}^{n_1} D_k^{(1)}\right] &\leq 3\left(\frac{2n_1-1}{6n_1^3(n_1-1)^3}\mathrm{Var}[\mathrm{tr}(\Sigma_1 \mathcal{E}_1^{(1)})^2]\right. \\
&\quad + \frac{2(2n_1-1)}{3n_1^3(n_1-1)^2}\mathrm{E}[\{\mathrm{tr}(\Sigma_1 \mathcal{E}_1^{(1)} \Sigma_1 \mathcal{E}_2^{(1)})\}^2] \\
&\quad \left. + \frac{2(2n_1-1)}{3n_1^3(n_1-1)}\mathrm{E}[\{\mathrm{tr}(\Sigma_1\Delta\Sigma_1 \mathcal{E}_1^{(1)})\}^2]\right). && (5.4)
\end{aligned}$$

By using Lemma A.1, we have

$$\begin{aligned}
\mathrm{Var}[\mathrm{tr}(\Sigma_1 \mathcal{E}_1^{(1)})^2] &= O(\mathrm{tr}\Sigma_1^4(\mathrm{tr}\Sigma_1^2)^2), && (5.5)\\
\mathrm{E}[\{\mathrm{tr}(\Sigma_1 \mathcal{E}_1^{(1)} \Sigma_1 \mathcal{E}_2^{(1)})\}^2] &= O((\mathrm{tr}\Sigma_1^4)^2), && (5.6)\\
\mathrm{E}[\{\mathrm{tr}(\Sigma_1\Delta\Sigma_1 \mathcal{E}_1^{(1)})\}^2] &= O(\mathrm{tr}\Sigma_1^4\mathrm{tr}(\Delta\Sigma_1)^2). && (5.7)
\end{aligned}$$

Substituting (5.5)-(5.7) to (5.4),

$$\mathrm{Var}\left[\sum_{k=1}^{n_1} D_k^{(1)}\right] = O\left(\frac{1}{n_1}\frac{\mathrm{tr}\Sigma_1^4}{n_1^2}\frac{(\mathrm{tr}\Sigma_1^2)^2}{n_1^2}\right) + O\left(\frac{(\mathrm{tr}\Sigma_1^4)^2}{n_1^4}\right) + O\left(\frac{\mathrm{tr}(\Sigma_1\Delta)^2}{n_1}\frac{\mathrm{tr}\Sigma_1^4}{n_1^2}\right). \quad (5.8)$$

From (5.3) and (5.8),

$$\frac{\sum_{k=1}^{n_1} D_k^{(1)}}{\sigma_2^2} = \frac{(\mathrm{tr}\Sigma_1^2)^2/(2n_1^2) + \mathrm{tr}(\Sigma_1\Delta)^2/n_1}{\sigma_2^2} + o_p(1). \tag{5.9}$$

Next, we evaluate $\sum_{k=n_1+1}^{n} D_k^{(2)}$. We rewrite

$$\begin{aligned} D_k^{(2)} &= \sum_{j=1}^{k-n_1-1} \frac{\mathrm{tr}(\Sigma_2\mathcal{E}_j^{(2)})^2}{n_2^2(n_2-1)^2} + \sum_{j\neq\ell}^{k-n_1-1} \frac{\mathrm{tr}(\Sigma_2\mathcal{E}_j^{(2)}\Sigma_2\mathcal{E}_\ell^{(2)})}{n_2^2(n_2-1)^2} \\ &\quad -\frac{2}{n_1n_2^2(n_2-1)} \sum_{j=1}^{k-n_1-1}\sum_{\ell=1}^{n_1} \mathrm{tr}(\Sigma_2\mathcal{E}_j^{(2)}\Sigma_2\mathcal{E}_\ell^{(1)}) \\ &\quad -\frac{2}{n_2^2(n_2-1)} \sum_{j=1}^{k-n_1-1} \mathrm{tr}(\Sigma_2\Delta\Sigma_2\mathcal{E}_j^{(2)}) \\ &\quad +\frac{1}{n_1^2n_2^2}\sum_{j=1}^{n_1}\mathrm{tr}(\Sigma_2\mathcal{E}_j^{(1)})^2 + \frac{1}{n_1^2n_2^2}\sum_{j\neq\ell}^{n_1}\mathrm{tr}(\Sigma_2\mathcal{E}_j^{(1)}\Sigma_2\mathcal{E}_\ell^{(1)}) \\ &\quad +\frac{2}{n_1n_2^2}\sum_{j=1}^{n_1}\mathrm{tr}(\Sigma_2\mathcal{E}_j^{(1)}\Sigma_2\Delta) + \frac{1}{n_2^2}\mathrm{tr}(\Sigma_2\Delta)^2. \end{aligned}$$

By using Lemma A.1, we have

$$\mathrm{E}[\mathrm{tr}(\Sigma_2\mathcal{E}_1^{(2)})^2] = \kappa_{42}\mathrm{tr}(\Sigma_2^2\odot\Sigma_2^2) + (\mathrm{tr}\Sigma_2^2)^2 + \mathrm{tr}\Sigma_2^4, \tag{5.10}$$

$$\mathrm{E}[\mathrm{tr}(\Sigma_2\mathcal{E}_1^{(2)})(\Sigma_2\mathcal{E}_2^{(2)})] = 0,\ \mathrm{E}[\mathrm{tr}(\Sigma_2\mathcal{E}_1^{(2)}\Sigma_2\mathcal{E}_1^{(1)})] = 0, \tag{5.11}$$

$$\mathrm{E}[\mathrm{tr}(\Sigma_2\Delta\Sigma_2\mathcal{E}_1^{(2)})] = 0, \tag{5.12}$$

$$\begin{aligned} \mathrm{E}[\mathrm{tr}(\Sigma_2\mathcal{E}_1^{(1)})^2] &= \kappa_{41}\mathrm{tr}\{(\Sigma_1^{1/2}\Sigma_2\Sigma_1^{1/2})\odot(\Sigma_1^{1/2}\Sigma_2\Sigma_1^{1/2})\} \\ &\quad +\{\mathrm{tr}(\Sigma_1\Sigma_2)\}^2 + \mathrm{tr}(\Sigma_1\Sigma_2)^2, \end{aligned} \tag{5.13}$$

$$\mathrm{E}[\mathrm{tr}(\Sigma_2\mathcal{E}_1^{(1)})(\Sigma_2\mathcal{E}_2^{(1)})] = 0, \mathrm{E}[\mathrm{tr}(\Sigma_2\mathcal{E}_1^{(1)}\Sigma_2\Delta)] = 0. \tag{5.14}$$

From (5.10)-(5.14),

$$\begin{aligned} \mathrm{E}\left[\sum_{k=n_1+1}^{n} D_k^{(2)}\right] &= \frac{\kappa_{42}\mathrm{tr}(\Sigma_2^2\odot\Sigma_2^2) + (\mathrm{tr}\Sigma_2^2)^2 + \mathrm{tr}\Sigma_2^4}{2n_2(n_2-1)} \\ &\quad +\frac{1}{n_1n_2}\Big[\kappa_{41}\mathrm{tr}\{(\Sigma_1^{1/2}\Sigma_2\Sigma_1^{1/2})\odot(\Sigma_1^{1/2}\Sigma_2\Sigma_1^{1/2})\} \\ &\quad +\{\mathrm{tr}(\Sigma_1\Sigma_2)\}^2 + \mathrm{tr}(\Sigma_1\Sigma_2)^2\Big] + \frac{1}{n_2}\mathrm{tr}(\Sigma_2\Delta)^2. \end{aligned} \tag{5.15}$$

We decompose the centered random variable as following:

$$
\begin{aligned}
\sum_{k=n_1+1}^{n} D_k^{(2)} - \mathrm{E}\left[\sum_{k=n_1+1}^{n} D_k^{(2)}\right] &= \sum_{k=1}^{n_2} \frac{(n_2-k)\{\mathrm{tr}(\Sigma_2\mathcal{E}_k^{(2)})^2 - \mathrm{E}[\mathrm{tr}(\Sigma_2\mathcal{E}_k^{(2)})^2]\}}{n_2^2(n_2-1)^2} \\
&\quad + \sum_{k=2}^{n_2}\sum_{\ell=1}^{k-1} \frac{2(n_2-k)\mathrm{tr}(\Sigma_2\mathcal{E}_k^{(2)}\Sigma_2\mathcal{E}_\ell^{(2)})}{n_2^2(n_2-1)^2} \\
&\quad - \sum_{\ell=1}^{n_1}\sum_{k=1}^{n_2} \frac{2(n_2-k)\mathrm{tr}(\Sigma_2\mathcal{E}_k^{(2)}\Sigma_2\mathcal{E}_\ell^{(1)})}{n_1 n_2^2(n_2-1)} \\
&\quad - \sum_{k=1}^{n_2} \frac{2(n_2-k)\mathrm{tr}(\Sigma_2\Delta\Sigma_2\mathcal{E}_k^{(2)})}{n_2^2(n_2-1)} \\
&\quad + \sum_{k=1}^{n_1} \frac{\mathrm{tr}(\Sigma_2\mathcal{E}_k^{(1)})^2 - \mathrm{E}[\mathrm{tr}(\Sigma_2\mathcal{E}_k^{(1)})^2]}{n_1^2 n_2} \\
&\quad + \sum_{k=2}^{n_1}\sum_{\ell=1}^{k-1} \frac{2\mathrm{tr}(\Sigma_2\mathcal{E}_k^{(1)}\Sigma_2\mathcal{E}_\ell^{(1)})}{n_1^2 n_2} \\
&\quad + \sum_{k=1}^{n_1} \frac{2\mathrm{tr}(\Sigma_2\mathcal{E}_j^{(1)}\Sigma_2\Delta)}{n_1 n_2}.
\end{aligned}
$$

By Holder's inequality, we obtain the following

$$
\begin{aligned}
\mathrm{Var}\left[\sum_{k=n_1+1}^{n} D_k^{(2)}\right] &\le 7\Bigg(\frac{2n_2-1}{6n_2^3(n_2-1)^3}\mathrm{Var}[\mathrm{tr}(\Sigma_2\mathcal{E}_1^{(2)})^2] \\
&\quad + \frac{n_2-2}{3n_2^3(n_2-1)^2}\mathrm{E}[\{\mathrm{tr}(\Sigma_2\mathcal{E}_1^{(2)}\Sigma_2\mathcal{E}_2^{(2)})\}^2] \\
&\quad + \frac{2(2n_2-1)}{3n_1 n_2^3(n_2-1)}\mathrm{E}[\{\mathrm{tr}(\Sigma_2\mathcal{E}_1^{(2)}\Sigma_2\mathcal{E}_1^{(1)})\}^2] \\
&\quad + \frac{2(2n_2-1)}{3n_2^3(n_2-1)}\mathrm{E}[\{\mathrm{tr}(\Sigma_2\Delta\Sigma_2\mathcal{E}_1^{(2)})\}^2] \\
&\quad + \frac{1}{n_1^3 n_2^2}\mathrm{Var}[\mathrm{tr}(\Sigma_2\mathcal{E}_1^{(1)})^2] \\
&\quad + \frac{2(n_1-1)}{n_1^3 n_2^2}\mathrm{E}[\{\mathrm{tr}(\Sigma_2\mathcal{E}_1^{(1)}\Sigma_2\mathcal{E}_2^{(1)})\}^2] \\
&\quad + \frac{4}{n_1 n_2^2}\mathrm{E}[\{\mathrm{tr}(\Sigma_2\mathcal{E}_1^{(1)}\Sigma_2\Delta)\}^2]\Bigg). \qquad (5.16)
\end{aligned}
$$

By using Lemma A.1, we have

$$\begin{aligned}
\mathrm{Var}[\mathrm{tr}(\Sigma_2\mathcal{E}_1^{(2)})^2] &= O((\mathrm{tr}\Sigma_2^4)(\mathrm{tr}\Sigma_2^2)^2), &(5.17)\\
\mathrm{E}[\{\mathrm{tr}(\Sigma_2\mathcal{E}_1^{(2)}\Sigma_2\mathcal{E}_2^{(2)})\}^2] &= O((\mathrm{tr}\Sigma_2^4)^2), &(5.18)\\
\mathrm{E}[\{\mathrm{tr}(\Sigma_2\mathcal{E}_1^{(2)}\Sigma_2\mathcal{E}_1^{(1)})\}^2] &= O(\mathrm{tr}(\Sigma_1\Sigma_2)^2\mathrm{tr}\Sigma_2^4), &(5.19)\\
\mathrm{E}[\{\mathrm{tr}(\Sigma_2\Delta\Sigma_2\mathcal{E}_1^{(2)})\}^2] &= O(\mathrm{tr}\Sigma_2^4\mathrm{tr}(\Sigma_2\Delta)^2), &(5.20)\\
\mathrm{Var}[\mathrm{tr}(\Sigma_2\mathcal{E}_1^{(1)})^2] &= O(\mathrm{tr}(\Sigma_1\Sigma_2)^2\{\mathrm{tr}(\Sigma_1\Sigma_2)\}^2), &(5.21)\\
\mathrm{E}[\{\mathrm{tr}(\Sigma_2\mathcal{E}_1^{(1)}\Sigma_2\mathcal{E}_2^{(1)})\}^2] &= O(\{\mathrm{tr}(\Sigma_1\Sigma_2)^2\}^2), &(5.22)\\
\mathrm{E}[\{\mathrm{tr}(\Sigma_2\mathcal{E}_1^{(1)}\Sigma_2\Delta)\}^2] &= O(\mathrm{tr}(\Sigma_1\Sigma_2)^2\mathrm{tr}(\Delta\Sigma_2)^2). &(5.23)
\end{aligned}$$

Substituting (5.17)-(5.23) to (5.16),

$$\begin{aligned}
\mathrm{Var}\left[\sum_{k=n_1+1}^{n} D_k^{(2)}\right] &= O\left(\frac{1}{n_2}\frac{\mathrm{tr}\Sigma_2^4}{n_2^2}\frac{(\mathrm{tr}\Sigma_2^2)^2}{n_2^2}\right) + O\left(\frac{(\mathrm{tr}\Sigma_2^4)^2}{n_2^4}\right)\\
&\quad + O\left(\frac{\mathrm{tr}(\Sigma_1\Sigma_2)^2}{n_1n_2}\frac{\mathrm{tr}\Sigma_2^4}{n_2^2}\right) + O\left(\frac{\mathrm{tr}(\Sigma_2\Delta)^2}{n_2}\frac{\mathrm{tr}\Sigma_2^4}{n_2^2}\right)\\
&\quad + O\left(\frac{1}{n_1}\frac{\mathrm{tr}(\Sigma_1\Sigma_2)^2}{n_1n_2}\frac{\{\mathrm{tr}(\Sigma_1\Sigma_2)\}^2}{n_1n_2}\right)\\
&\quad + O\left(\left\{\frac{\mathrm{tr}(\Sigma_1\Sigma_2)^2}{n_1n_2}\right\}^2\right) + O\left(\frac{\mathrm{tr}(\Sigma_1\Sigma_2)^2}{n_1n_2}\frac{\mathrm{tr}(\Delta\Sigma_2)^2}{n_2}\right).
\end{aligned} \tag{5.24}$$

From (5.15) and (5.24),

$$\frac{\sum_{k=n_1+1}^{n} D_k^{(2)}}{\sigma_2^2} = \frac{(\mathrm{tr}\Sigma_2^2)^2/(2n_2^2) + \{\mathrm{tr}(\Sigma_1\Sigma_2)\}^2/(n_1n_2) + \mathrm{tr}(\Delta\Sigma_2)^2/n_2}{\sigma_2^2} + o_p(1). \tag{5.25}$$

From (5.9) and (5.25), statement (B) follows. □

## 6 Derivation of (C)

At first, we evaluate $\sum_{k=1}^{n_1} E_k^{(1)}$. We rewrite

$$\sum_{k=1}^{n_1} E_k^{(1)} = \frac{1}{n_1}\mathrm{tr}\{(\Sigma_1^{1/2}\Delta\Sigma_1^{1/2}) \odot (\Sigma_1^{1/2}\Delta\Sigma_1^{1/2})\} + \sum_{i=1}^{3} \widetilde{E}_i^{(k)},$$

where

$$\begin{aligned}
\widetilde{E}_1^{(1)} &= \sum_{k=1}^{n_1} \frac{(n_1-k)\mathrm{tr}\{(\Sigma_1^{1/2}\mathcal{E}_k^{(1)}\Sigma_1^{1/2}) \odot (\Sigma_1^{1/2}\mathcal{E}_k^{(1)}\Sigma_1^{1/2})\}}{n_1^2(n_1-1)^2} \\
\widetilde{E}_2^{(1)} &= \sum_{k=2}^{n_1}\sum_{\ell=1}^{k-1} \frac{2(n_1-k)\mathrm{tr}\{(\Sigma_1^{1/2}\mathcal{E}_k^{(1)}\Sigma_1^{1/2}) \odot (\Sigma_1^{1/2}\mathcal{E}_\ell^{(1)}\Sigma_1^{1/2})\}}{n_1^2(n_1-1)^2} \\
\widetilde{E}_3^{(1)} &= \sum_{j=1}^{k-1} \frac{2(n_1-k)\mathrm{tr}\{(\Sigma_1^{1/2}\Delta\Sigma_1^{1/2}) \odot (\Sigma_1^{1/2}\mathcal{E}_k^{(1)}\Sigma_1^{1/2})\}}{n_1^2(n_1-1)}.
\end{aligned}$$

We show that $\widetilde{E}_1$, $\widetilde{E}_2$ and $\widetilde{E}_3$ are negligible. The expectations of $|\widetilde{E}_1^{(1)}|$, $\widetilde{E}_2^{(1)^2}$, and $\widetilde{E}_3^{(1)^2}$ are evaluated as

$$\mathrm{E}[|\widetilde{E}_1^{(1)}|] = O\left(\frac{\mathrm{tr}\Sigma_1^4}{n_1^2}\right) = o(\sigma_2^2), \tag{6.1}$$

$$\mathrm{E}[\widetilde{E}_2^{(1)^2}] = O\left(\left(\frac{\mathrm{tr}\Sigma_1^4}{n_1^2}\right)^2\right) = o(\sigma_2^4), \tag{6.2}$$

$$\mathrm{E}[\widetilde{E}_3^{(1)^2}] = O\left(\frac{\mathrm{tr}(\Sigma_1\Delta)^2}{n_1}\frac{\mathrm{tr}\Sigma_1^4}{n_1^2}\right) = o(\sigma_2^4). \tag{6.3}$$

Using Markov's inequality with (6.1) and Chebyshev's inequality with (6.2), (6.3), we obtain

$$\frac{\sum_{k=1}^{n_1} E_k^{(1)}}{\sigma_2^2} = \frac{\mathrm{tr}\{(\Sigma_1^{1/2}\Delta\Sigma_1^{1/2}) \odot (\Sigma_1^{1/2}\Delta\Sigma_1^{1/2})\}/n_1}{\sigma_2^2} + o_p(1). \tag{6.4}$$

Next, we evaluate $\sum_{k=n_1+1}^{n} E_k^{(2)}$. We decompose the centered random variable as following:

$$\sum_{k=n_1+1}^{n} E_k^{(2)} = \frac{\mathrm{tr}\{(\Sigma_2^{1/2}\Delta\Sigma_2^{1/2}) \odot (\Sigma_2^{1/2}\Delta\Sigma_2^{1/2})\}}{n_2} + \sum_{i=1}^{7} \widetilde{E}_i^{(2)},$$

where

$$
\begin{aligned}
\widetilde{E}_1^{(2)} &= \sum_{k=1}^{n_2} \frac{(n_2-k)\mathrm{tr}\{(\Sigma_2^{1/2}\mathcal{E}_k^{(2)}\Sigma_2^{1/2}) \odot (\Sigma_2^{1/2}\mathcal{E}_k^{(2)}\Sigma_2^{1/2})\}}{n_2^2(n_2-1)^2}, \\
\widetilde{E}_2^{(2)} &= \sum_{k=2}^{n_2}\sum_{\ell=1}^{k-1} \frac{2(n_2-k)\mathrm{tr}\{(\Sigma_2^{1/2}\mathcal{E}_k^{(2)}\Sigma_2^{1/2}) \odot (\Sigma_2^{1/2}\mathcal{E}_\ell^{(2)}\Sigma_2^{1/2})\}}{n_2^2(n_2-1)^2}, \\
\widetilde{E}_3^{(2)} &= -\sum_{\ell=1}^{n_1}\sum_{k=1}^{n_2} \frac{2(n_2-k)\mathrm{tr}\{(\Sigma_2^{1/2}\mathcal{E}_k^{(2)}\Sigma_2^{1/2}) \odot (\Sigma_2^{1/2}\mathcal{E}_\ell^{(1)}\Sigma_2^{1/2})\}}{n_1 n_2^2(n_2-1)}, \\
\widetilde{E}_4^{(2)} &= -\sum_{k=1}^{n_2} \frac{2(n_2-k)\mathrm{tr}\{(\Sigma_2^{1/2}\Delta\Sigma_2^{1/2}) \odot (\Sigma_2^{1/2}\mathcal{E}_k^{(2)}\Sigma_2^{1/2})\}}{n_2^2(n_2-1)}, \\
\widetilde{E}_5^{(2)} &= \sum_{k=1}^{n_1} \frac{\mathrm{tr}\{(\Sigma_2^{1/2}\mathcal{E}_k^{(1)}\Sigma_2^{1/2}) \odot (\Sigma_2^{1/2}\mathcal{E}_k^{(1)}\Sigma_2^{1/2})\}}{n_1^2 n_2}, \\
\widetilde{E}_6^{(2)} &= \sum_{k=2}^{n_1}\sum_{\ell=1}^{k-1} \frac{2\mathrm{tr}\{(\Sigma_2^{1/2}\mathcal{E}_k^{(1)}\Sigma_2^{1/2}) \odot (\Sigma_2^{1/2}\mathcal{E}_\ell^{(1)}\Sigma_2^{1/2})\}}{n_1^2 n_2}, \\
\widetilde{E}_7^{(2)} &= \sum_{k=1}^{n_1} \frac{2\mathrm{tr}\{(\Sigma_2^{1/2}\mathcal{E}_k^{(1)}\Sigma_2^{1/2}) \odot (\Sigma_2^{1/2}\Delta\Sigma_2^{1/2})\}}{n_1 n_2}.
\end{aligned}
$$

The expectations of $|\widetilde{E}_1^{(2)}|$, $\widetilde{E}_2^{(2)^2}$, and $\widetilde{E}_3^{(2)^2}$, $\widetilde{E}_4^{(2)^2}$, $|\widetilde{E}_5^{(2)}|$, $\widetilde{E}_6^{(2)^2}$ and $\widetilde{E}_7^{(2)^2}$ are evaluated as

$$
\begin{aligned}
\mathrm{E}[|\widetilde{E}_1^{(2)}|] &= O\left(\frac{\mathrm{tr}\Sigma_2^4}{n_2^2}\right) = o(\sigma_2^2), &(6.5)\\
\mathrm{E}[\widetilde{E}_2^{(2)^2}] &= O\left(\left(\frac{\mathrm{tr}\Sigma_2^4}{n_2^2}\right)^2\right) = o(\sigma_2^4), &(6.6)\\
\mathrm{E}[\widetilde{E}_3^{(2)^2}] &= O\left(\frac{\mathrm{tr}\Sigma_2^4}{n_2^2}\frac{\mathrm{tr}(\Sigma_1\Sigma_2)^2}{n_1 n_2}\right) = o(\sigma_2^4), &(6.7)\\
\mathrm{E}[\widetilde{E}_4^{(2)^2}] &= O\left(\frac{\mathrm{tr}(\Sigma_2\Delta)^2}{n_2}\frac{\mathrm{tr}\Sigma_2^4}{n_2^2}\right) = o(\sigma_2^4), &(6.8)\\
\mathrm{E}[|\widetilde{E}_5^{(2)}|] &= O\left(\frac{\mathrm{tr}(\Sigma_1\Sigma_2)^2}{n_1 n_2}\right) = o(\sigma_2^2), &(6.9)\\
\mathrm{E}[\widetilde{E}_6^{(2)^2}] &= O\left(\left\{\frac{\mathrm{tr}(\Sigma_1\Sigma_2)^2}{n_1 n_2}\right\}^2\right) = o(\sigma_2^4), &(6.10)\\
\mathrm{E}[\widetilde{E}_7^{(2)^2}] &= O\left(\frac{\mathrm{tr}(\Sigma_2\Delta)^2}{n_2}\frac{\mathrm{tr}(\Sigma_1\Sigma_2)^2}{n_1 n_2}\right) = o(\sigma_2^4). &(6.11)
\end{aligned}
$$

Using Markov's inequality with (6.6) and (6.9) and Chebyshev's inequality with (6.6)-(6.8), (6.10) and (6.11), we obtain

$$
\frac{\sum_{k=n_1+1}^{n} E_k^{(2)}}{\sigma_2^2} = \frac{\mathrm{tr}\{(\Sigma_2^{1/2}\Delta\Sigma_2^{1/2}) \odot (\Sigma_2^{1/2}\Delta\Sigma_2^{1/2})\}/n_2}{\sigma_2^2} + o_p(1). \qquad (6.12)
$$

From (6.4) and (6.12), statement (C) follows. □

# 7 Derivation of (D)

We assume (A1) and (A2). At first, we evaluate $\sum_{k=1}^{n_1} F_k^{(1)}$. We rewrite

$$\sum_{k=1}^{n_1} F_k^{(1)} = \frac{\sum_{i=1}^p \boldsymbol{\delta}'\Sigma_1^{1/2}\boldsymbol{e}_i\boldsymbol{e}_i'\Sigma_1^{1/2}\Delta\Sigma_1^{1/2}\boldsymbol{e}_i}{n_1} + \sum_{i=1}^4 \widetilde{F}_i^{(1)},$$

where

$$\begin{aligned}
\widetilde{F}_1^{(1)} &= \sum_{k=1}^{n_1} \frac{(n_1-k)\sum_{i=1}^p \boldsymbol{Y}_{1k}'\Sigma_1^{1/2}\boldsymbol{e}_i\boldsymbol{e}_i'\Sigma_1^{1/2}\mathcal{E}_k^{(1)}\Sigma_1^{1/2}\boldsymbol{e}_i}{n_1^2(n_1-1)^2}, \\
\widetilde{F}_2^{(1)} &= \sum_{k=2}^{n_1}\sum_{\ell=1}^{k-1}(n_1-k)\left(\frac{\sum_{i=1}^p \boldsymbol{Y}_{1k}'\Sigma_1^{1/2}\boldsymbol{e}_i\boldsymbol{e}_i'\Sigma_1^{1/2}\mathcal{E}_\ell^{(1)}\Sigma_1^{1/2}\boldsymbol{e}_i}{n_1^2(n_1-1)^2}\right. \\
&\quad \left. + \frac{\sum_{i=1}^p \boldsymbol{Y}_{1\ell}'\Sigma_1^{1/2}\boldsymbol{e}_i\boldsymbol{e}_i'\Sigma_1^{1/2}\mathcal{E}_k^{(1)}\Sigma_1^{1/2}\boldsymbol{e}_i}{n_1^2(n_1-1)^2}\right), \\
\widetilde{F}_3^{(1)} &= \sum_{k=1}^{n_1} \frac{(n_1-k)\sum_{i=1}^p \boldsymbol{Y}_{1k}'\Sigma_1^{1/2}\boldsymbol{e}_i\boldsymbol{e}_i'\Sigma_1^{1/2}\Delta\Sigma_1^{1/2}\boldsymbol{e}_i}{n_1^2(n_1-1)}, \\
\widetilde{F}_4^{(1)} &= \sum_{k=1}^{n_1} \frac{(n_1-k)\sum_{i=1}^p \boldsymbol{\delta}'\Sigma_1^{1/2}\boldsymbol{e}_i\boldsymbol{e}_i'\Sigma_1^{1/2}\mathcal{E}_k^{(1)}\Sigma_1^{1/2}\boldsymbol{e}_i}{n_1^2(n_1-1)}.
\end{aligned}$$

We show that $\widetilde{F}_1$, $\widetilde{F}_2$ and $\widetilde{F}_3$ are negligible. The expectations of $|\widetilde{F}_1^{(1)}|$, $\widetilde{F}_2^{(1)^2}$, and $\widetilde{F}_3^{(1)^2}$ are evaluated as

$$\mathrm{E}[|\widetilde{F}_1^{(1)}|] = O\left(\sqrt{\frac{\mathrm{tr}\Sigma_1^2}{n_1^2}\frac{\mathrm{tr}\Sigma_1^4}{n_1^2}}\right) = o(\sigma_1\sigma_2), \tag{7.1}$$

$$\mathrm{E}[\widetilde{F}_2^{(1)^2}] = O\left(\frac{\mathrm{tr}\Sigma_1^2}{n_1^2}\frac{\mathrm{tr}\Sigma_1^4}{n_1^2}\right) = o(\sigma_1^2\sigma_2^2), \tag{7.2}$$

$$\mathrm{E}[\widetilde{F}_3^{(1)^2}] = O\left(\frac{\sqrt{\mathrm{tr}\Sigma_1^4}}{n_1^2}\frac{\mathrm{tr}(\Sigma_1\Delta)^2}{n_1}\right) = o(\sigma_1^2\sigma_2^2), \tag{7.3}$$

$$\mathrm{E}[\widetilde{F}_4^{(1)^2}] = O\left(\frac{\mathrm{tr}\Sigma_1^4}{n_1^2}\frac{\boldsymbol{\delta}'\Sigma_1\boldsymbol{\delta}}{n_1}\right) = o(\sigma_1^2\sigma_2^2) \tag{7.4}$$

Using Markov's inequality with (7.1) and Chebyshev's inequality with (7.2)-(7.4), we obtain

$$\frac{\sum_{k=1}^{n_1} F_k^{(1)}}{\sigma_1\sigma_2} = \frac{\sum_{i=1}^p \boldsymbol{\delta}'\Sigma_1^{1/2}\boldsymbol{e}_i\boldsymbol{e}_i'\Sigma_1^{1/2}\Delta\Sigma_1^{1/2}\boldsymbol{e}_i}{n_1\sigma_1\sigma_2} + o_p(1). \tag{7.5}$$

Next, we evaluate $\sum_{k=n_1+1}^{n} F_k^{(2)}$. We decompose the centered random variable as following:

$$\sum_{k=n_1+1}^{n} F_k^{(2)} = \frac{\sum_{i=1}^p (\boldsymbol{\delta}'\Sigma_2^{1/2}\boldsymbol{e}_i)(\boldsymbol{e}_i'\Sigma_2^{1/2}\Delta\Sigma_2^{1/2}\boldsymbol{e}_i)}{n_2} + \sum_{i=1}^{10} \widetilde{F}_i^{(2)},$$

where

$$
\begin{aligned}
\widetilde{F}_1^{(2)} &= \sum_{k=1}^{n_2} \frac{(n_2-k)\sum_{i=1}^p \boldsymbol{Y}'_{2k}\Sigma_2^{1/2}\boldsymbol{e}_i\boldsymbol{e}_i'\Sigma_2^{1/2}\mathcal{E}_k^{(2)}\Sigma_2^{1/2}\boldsymbol{e}_i}{n_2^2(n_2-1)^2}, \\
\widetilde{F}_2^{(2)} &= \sum_{k=2}^{n_2}\sum_{\ell=1}^{k-1}(n_2-k)\left(\frac{\sum_{i=1}^p \boldsymbol{Y}'_{2k}\Sigma_2^{1/2}\boldsymbol{e}_i\boldsymbol{e}_i'\Sigma_2^{1/2}\mathcal{E}_\ell^{(2)}\Sigma_2^{1/2}\boldsymbol{e}_i}{n_2^2(n_2-1)^2}\right. \\
&\quad \left. + \frac{\sum_{i=1}^p \boldsymbol{Y}'_{2\ell}\Sigma_2^{1/2}\boldsymbol{e}_i\boldsymbol{e}_i'\Sigma_2^{1/2}\mathcal{E}_k^{(2)}\Sigma_2^{1/2}\boldsymbol{e}_i}{n_2^2(n_2-1)^2}\right), \\
\widetilde{F}_3^{(2)} &= -\sum_{\ell=1}^{n_1}\sum_{k=1}^{n_2} \frac{(n_2-k)\sum_{i=1}^p \boldsymbol{Y}'_{2k}\Sigma_2^{1/2}\boldsymbol{e}_i\boldsymbol{e}_i'\Sigma_2^{1/2}\mathcal{E}_\ell^{(1)}\Sigma_2^{1/2}\boldsymbol{e}_i}{n_1n_2^2(n_2-1)}, \\
\widetilde{F}_4^{(2)} &= -\sum_{k=1}^{n_2} \frac{(n_2-k)\sum_{i=1}^p \boldsymbol{Y}'_{2k}\Sigma_2^{1/2}\boldsymbol{e}_i\boldsymbol{e}_i'\Sigma_2^{1/2}\Delta\Sigma_2^{1/2}\boldsymbol{e}_i}{n_2^2(n_2-1)}, \\
\widetilde{F}_5^{(2)} &= -\sum_{k=1}^{n_2}\sum_{\ell=1}^{n_1} \frac{(n_2-k)\sum_{i=1}^p \boldsymbol{Y}'_{1\ell}\Sigma_2^{1/2}\boldsymbol{e}_i\boldsymbol{e}_i'\Sigma_2^{1/2}\mathcal{E}_k^{(2)}\Sigma_2^{1/2}\boldsymbol{e}_i}{n_1n_2^2(n_2-1)}, \\
\widetilde{F}_6^{(2)} &= \sum_{k=1}^{n_1} \frac{\sum_{i=1}^p \boldsymbol{Y}_{1k}\Sigma_2^{1/2}\boldsymbol{e}_i\boldsymbol{e}_i'\Sigma_2^{1/2}\mathcal{E}_k^{(1)}\Sigma_2^{1/2}\boldsymbol{e}_i}{n_1^2n_2}, \\
\widetilde{F}_7^{(2)} &= \sum_{k=2}^{n_1}\sum_{\ell=1}^{k-1}\left(\frac{\sum_{i=1}^p \boldsymbol{Y}'_{1k}\Sigma_2^{1/2}\boldsymbol{e}_i\boldsymbol{e}_i'\Sigma_2^{1/2}\mathcal{E}_\ell^{(1)}\Sigma_2^{1/2}\boldsymbol{e}_i}{n_1^2n_2}\right. \\
&\quad \left. + \frac{\sum_{i=1}^p \boldsymbol{Y}'_{1\ell}\Sigma_2^{1/2}\boldsymbol{e}_i\boldsymbol{e}_i'\Sigma_2^{1/2}\mathcal{E}_k^{(1)}\Sigma_2^{1/2}\boldsymbol{e}_i}{n_1^2n_2}\right), \\
\widetilde{F}_8^{(2)} &= \sum_{k=1}^{n_1} \frac{\sum_{i=1}^p \boldsymbol{Y}'_{1k}\Sigma_2^{1/2}\boldsymbol{e}_i\boldsymbol{e}_i'\Sigma_2^{1/2}\Delta\Sigma_2^{1/2}\boldsymbol{e}_i}{n_1n_2}, \\
\widetilde{F}_9^{(2)} &= -\sum_{k=1}^{n_2} \frac{(n_2-k)\sum_{i=1}^p \boldsymbol{\delta}'\Sigma_2^{1/2}\boldsymbol{e}_i\boldsymbol{e}_i'\Sigma_2^{1/2}\mathcal{E}_k^{(2)}\Sigma_2^{1/2}\boldsymbol{e}_i}{n_2^2(n_2-1)}, \\
\widetilde{F}_{10}^{(2)} &= \sum_{k=1}^{n_1} \frac{\sum_{i=1}^p \boldsymbol{\delta}'\Sigma_2^{1/2}\boldsymbol{e}_i\boldsymbol{e}_i'\Sigma_2^{1/2}\mathcal{E}_k^{(1)}\Sigma_2^{1/2}\boldsymbol{e}_i}{n_1n_2}.
\end{aligned}
$$

The moments of above random variables are evaluated as

$$
\mathrm{E}[|\widetilde{F}_1^{(2)}|] = O\left(\sqrt{\frac{\mathrm{tr}\Sigma_2^2}{n_2^2}\frac{\mathrm{tr}\Sigma_2^4}{n_2^2}}\right) = o(\sigma_1\sigma_2), \tag{7.6}
$$

$$
\mathrm{E}[\widetilde{F}_2^{(2)^2}] = O\left(\frac{\mathrm{tr}\Sigma_2^2}{n_2^2}\frac{\mathrm{tr}\Sigma_2^4}{n_2^2}\right) = o(\sigma_1^2\sigma_2^2), \tag{7.7}
$$

$$\mathrm{E}[\widetilde{F}_3^{(2)^2}] = O\left(\frac{\mathrm{tr}\Sigma_2^2}{n_2^2}\frac{\mathrm{tr}(\Sigma_1\Sigma_2)^2}{n_1n_2}\right) = o(\sigma_1^2\sigma_2^2), \tag{7.8}$$

$$\mathrm{E}[\widetilde{F}_4^{(2)^2}] = O\left(\frac{\sqrt{\mathrm{tr}\Sigma_2^4}}{n_2^2}\frac{\mathrm{tr}(\Sigma_2\Delta)^2}{n_2}\right) = o(\sigma_1^2\sigma_2^2), \tag{7.9}$$

$$\mathrm{E}[\widetilde{F}_5^{(2)^2}] = O\left(\frac{\mathrm{tr}(\Sigma_1\Sigma_2)}{n_1n_2}\frac{\mathrm{tr}\Sigma_2^4}{n_2^2}\right) = o(\sigma_1^2\sigma_2^2), \tag{7.10}$$

$$\mathrm{E}[|\widetilde{F}_6^{(2)}|] = O\left(\sqrt{\frac{\mathrm{tr}(\Sigma_1\Sigma_2)}{n_1n_2}\frac{\mathrm{tr}(\Sigma_1\Sigma_2)^2}{n_1n_2}}\right) = o(\sigma_1\sigma_2), \tag{7.11}$$

$$\mathrm{E}[\widetilde{F}_7^{(2)^2}] = O\left(\frac{\mathrm{tr}(\Sigma_1\Sigma_2)}{n_1n_2}\frac{\mathrm{tr}(\Sigma_1\Sigma_2)^2}{n_1n_2}\right) = o(\sigma_1^2\sigma_2^2), \tag{7.12}$$

$$\mathrm{E}[\widetilde{F}_8^{(2)^2}] = O\left(\frac{\sqrt{\mathrm{tr}(\Sigma_1\Sigma_2)^2}}{n_1n_2}\frac{\mathrm{tr}(\Sigma_2\Delta)^2}{n_2}\right) = o(\sigma_1^2\sigma_2^2), \tag{7.13}$$

$$\mathrm{E}[\widetilde{F}_9^{(2)^2}] = O\left(\frac{\boldsymbol{\delta}'\Sigma_2\boldsymbol{\delta}}{n_2}\frac{\mathrm{tr}\Sigma_2^4}{n_2^2}\right) = o(\sigma_1^2\sigma_2^2), \tag{7.14}$$

$$\mathrm{E}[\widetilde{F}_{10}^{(2)^2}] = O\left(\frac{\boldsymbol{\delta}'\Sigma_2\boldsymbol{\delta}}{n_2}\frac{\mathrm{tr}(\Sigma_1\Sigma_2)^2}{n_1n_2}\right) = o(\sigma_1^2\sigma_2^2). \tag{7.15}$$

Using Markov's inequality with (7.6) and (7.11) and Chebyshev's inequality with (7.7)-(7.10), (7.12)-(7.15), we obtain

$$\frac{\sum_{k=n_1+1}^{n} F_k^{(2)}}{\sigma_1\sigma_2} = \frac{\sum_{i=1}^{p}\boldsymbol{\delta}'\Sigma_2^{1/2}\boldsymbol{e}_i\boldsymbol{e}_i'\Sigma_2^{1/2}\Delta\Sigma_2^{1/2}\boldsymbol{e}_i}{n_2\sigma_1\sigma_2} + o_p(1). \tag{7.16}$$

From (7.5) and (7.16), statement (D) follows. □

## 8 Derivation of (E)

Let

$$\begin{aligned}
\varepsilon_{1k} &= \frac{2c_1}{\sigma_1 n_1(n_1-1)}\sum_{j=1}^{k-1}\boldsymbol{Y}'_{1k}\boldsymbol{Y}_{1j},\\
\varepsilon_{2k} &= \frac{2c_1}{\sigma_1 n_1}\boldsymbol{\delta}'\boldsymbol{Y}_{1k},\\
\varepsilon_{3k} &= \frac{2c_2}{\sigma_2 n_1(n_1-1)}\sum_{j=1}^{k-1}\{\boldsymbol{Y}'_{1k}\mathcal{E}_j^{(1)}\boldsymbol{Y}_{1k}-\mathrm{tr}(\Sigma_1\mathcal{E}_j^{(1)})\},\\
\varepsilon_{4k} &= \frac{2c_2}{\sigma_2 n_1}\{\boldsymbol{Y}_{1k}\Delta\boldsymbol{Y}_{1k}-\mathrm{tr}(\Sigma_1\Delta)\}.
\end{aligned}$$

Note that $\varepsilon_k=\sum_{g=1}^4\varepsilon_{gk}$ for $1\le k\le n_1$. Then it is sufficient to show that $\sum_{k=1}^{n_1}\mathrm{E}[\varepsilon_{gk}^4]=o(1)$. Firstly, we evaluate $\sum_{k=1}^{n_1}\mathrm{E}[\varepsilon_{1k}^4]$:

$$\begin{aligned}
\sum_{k=1}^{n_1}\mathrm{E}[\varepsilon_{1k}^4] &\le \frac{16c_1^4(\kappa_{41}+2)}{\sigma_1^4 n_1^4(n_1-1)^4}\sum_{k=1}^{n_1}\mathrm{E}\left[\mathrm{tr}\left(\sum_{j,\ell=1}^{k-1}\Sigma_1^{1/2}\boldsymbol{Y}_{1j}\boldsymbol{Y}'_{1\ell}\Sigma_1^{1/2}\right)^2\right]\\
&\quad+\frac{16c_1^4}{\sigma_1^4 n_1^4(n_1-1)^4}\sum_{k=1}^{n_1}\mathrm{E}\left[\left\{\mathrm{tr}\left(\sum_{j,\ell=1}^{k-1}\Sigma_1^{1/2}\boldsymbol{Y}_{1j}\boldsymbol{Y}'_{1\ell}\Sigma_1^{1/2}\right)\right\}^2\right]\\
&= \frac{8c_1^4(\kappa_{41}+3)}{\sigma_1^4 n_1^3(n_1-1)^3}\mathrm{E}[(\boldsymbol{Z}'_{11}\Sigma_1^2\boldsymbol{Z}_{11})^2]+\frac{16c_1^4(n_1-2)(\kappa_{41}+3)}{3\sigma_1^4 n_1^3(n_1-1)^3}(\mathrm{E}[\boldsymbol{Z}'_{11}\Sigma_1^2\boldsymbol{Z}_{11}])^2\\
&\quad+\frac{16c_1^4(n_1-2)(\kappa_{41}+3)}{3\sigma_1^4 n_1^3(n_1-1)^3}\mathrm{E}[(\boldsymbol{Z}'_{11}\Sigma_1^2\boldsymbol{Z}_{12})^2]\\
&= O\left(\frac{1}{n_1}\right)=o(1). \qquad (8.1)
\end{aligned}$$

Secondly, we evaluate $\sum_{k=1}^{n_1}\mathrm{E}[\varepsilon_{2k}^4]$:

$$\begin{aligned}
\sum_{k=1}^{n_1}\mathrm{E}[\varepsilon_{2k}^4] &\le \frac{16c_1^4}{\sigma_1^4 n_1^3}\mathrm{E}\left[(\boldsymbol{Y}'_{1j}\boldsymbol{\delta})^4\right]\\
&\le \frac{16c_1^4(\kappa_{41}+3)}{n_1}\frac{(\boldsymbol{\delta}'\Sigma_1\boldsymbol{\delta})^2}{\sigma_1^4 n_1^2}=O\left(\frac{1}{n_1}\right)=o(1). \qquad (8.2)
\end{aligned}$$

Thirdly, we evaluate $\sum_{k=1}^{n_1}\mathrm{E}[\varepsilon_{3k}^4]$:

$$\begin{aligned}
\sum_{k=1}^{n_1}\mathrm{E}[\varepsilon_{3k}^4] &\le \frac{16c_2^4\gamma_1}{\sigma_2^4 n_1^4(n_1-1)^4}\sum_{k=1}^{n_1}\mathrm{E}\left[\left\{\mathrm{tr}\left(\sum_{j=1}^{k-1}\Sigma_1^{1/2}\mathcal{E}_j^{(1)}\Sigma_1^{1/2}\right)^2\right\}^2\right]\\
&= \frac{8c_2^4\gamma_1}{\sigma_2^4 n_1^3(n_1-1)^3}\mathrm{E}\left[\left\{\mathrm{tr}\left(\Sigma_1^{1/2}\mathcal{E}_1^{(1)}\Sigma_1^{1/2}\right)^2\right\}^2\right]\\
&\quad+\frac{16c_2^4\gamma_1(n_1-2)}{3\sigma_2^4 n_1^3(n_1-1)^3}(\mathrm{E}[\mathrm{tr}(\Sigma_1^{1/2}\mathcal{E}_1^{(1)}\Sigma_1^{1/2})^2])^2\\
&\quad+\frac{32c_2^4\gamma_1(n_1-2)}{3\sigma_2^4 n_1^3(n_1-1)^3}\mathrm{E}[\{\mathrm{tr}(\Sigma_1^{1/2}\mathcal{E}_1\Sigma_1^{1/2})(\Sigma_1^{1/2}\mathcal{E}_2\Sigma_1^{1/2})\}^2].
\end{aligned}$$

Here, $\gamma_1$ is some constant which are not depend on $n_g$, $p$, $\Sigma_g$, and $\boldsymbol{\mu}_g$. By using Lemma A.1,

$$\mathrm{E}\left[\left\{\mathrm{tr}\left(\Sigma_1^{1/2}\mathcal{E}_1^{(1)}\Sigma_1^{1/2}\right)^2\right\}^2\right] = O((\mathrm{tr}\Sigma_1^2)^4), \tag{8.3}$$

$$\mathrm{E}[\mathrm{tr}(\Sigma_1^{1/2}\mathcal{E}_1\Sigma_1^{1/2})^2] = O((\mathrm{tr}\Sigma_1^4)^2). \tag{8.4}$$

Plug in (8.3)-(8.4) for (8.3),

$$\sum_{k=1}^{n_1}\mathrm{E}[\varepsilon_{3k}^4] = O\left(\frac{1}{n_1}\right) = o(1). \tag{8.5}$$

Finally, we evaluate $\sum_{k=1}^{n_1}\mathrm{E}[\varepsilon_{4k}^4]$:

$$\begin{aligned}\sum_{k=1}^{n_1}\mathrm{E}[\varepsilon_{4k}^4] &= \frac{16c_2^4}{\sigma_2^4 n_1^4}\sum_{k=1}^{n_1}\mathrm{E}\left[(\boldsymbol{Y}'_{1k}\Delta\boldsymbol{Y}_{1k} - \mathrm{tr}(\Sigma_1\Delta))^4\right] \\ &\leq \frac{16c_2^4\gamma_1}{\sigma_2^4 n_1^3}\{\mathrm{tr}(\Sigma_1\Delta)^2\}^2 = O\left(\frac{1}{n_1}\right) = o(1). \end{aligned} \tag{8.6}$$

From (8.1), (8.2), (8.5) and (8.6), statement (E) follows. □

# 9 Derivation of (F)

Let

$$
\begin{aligned}
\varepsilon_{1k} &= \frac{2c_1}{\sigma_2 n_2(n_2-1)} \sum_{j=1}^{k-n_1-1} \boldsymbol{Y}'_{2j} \boldsymbol{Y}_{2k-n_1}, \\
\varepsilon_{2k} &= -\frac{2c_1}{\sigma_1 n_2} \boldsymbol{\delta}' \boldsymbol{Y}_{2k-n_1}, \\
\varepsilon_{3k} &= -\frac{2c_1}{\sigma_1 n_2} \overline{\boldsymbol{Y}}'_1 \boldsymbol{Y}_{2k-n_1}, \\
\varepsilon_{4k} &= \frac{2c_2}{\sigma_2 n_2(n_2-1)} \sum_{j=1}^{k-n_1-1} \{\boldsymbol{Y}'_{2k-n_1} \mathcal{E}_j^{(2)} \boldsymbol{Y}_{2k-n_1} - \mathrm{tr}(\Sigma_2 \mathcal{E}_j^{(2)})\}, \\
\varepsilon_{5k} &= -\frac{2c_2}{\sigma_2 n_1 n_2} \sum_{j=1}^{n_1} \{\boldsymbol{Y}'_{2k-n_1} \mathcal{E}_j^{(1)} \boldsymbol{Y}_{2k-n_1} - \mathrm{tr}(\Sigma_2 \mathcal{E}_j^{(1)})\}, \\
\varepsilon_{6k} &= -\frac{2c_2}{\sigma_2 n_2} \{\boldsymbol{Y}_{2k-n_1} \Delta \boldsymbol{Y}_{2k-n_1} - \mathrm{tr}(\Sigma_2 \Delta)\}.
\end{aligned}
$$

Note that $\varepsilon_k = \sum_{g=1}^{6} \varepsilon_{gk}$ for $n_1+1 \le k \le n_1+n_2$. Then it is sufficient to show that $\sum_{k=n_1+1}^{n_1+n_2} \mathrm{E}[\varepsilon_{gk}^4] = o(1)$. Firstly, we evaluate $\sum_{k=n_1+1}^{n_1+n_2} \mathrm{E}[\varepsilon_{1k}^4]$:

$$
\sum_{k=n_1+1}^{n_1+n_2} \mathrm{E}[\varepsilon_{1k}^4] = O\left(\frac{1}{n_2}\right) = o(1). \tag{9.1}
$$

Secondly, we evaluate $\sum_{k=n_1+1}^{n_1+n_2} \mathrm{E}[\varepsilon_{2k}^4]$:

$$
\sum_{k=n_1+1}^{n_1+n_2} \mathrm{E}[\varepsilon_{2k}^4] = O\left(\frac{1}{n_2}\right) = o(1). \tag{9.2}
$$

Thirdly, we evaluate $\sum_{k=n_1+1}^{n_1+n_2} \mathrm{E}[\varepsilon_{3k}^4]$:

$$
\begin{aligned}
\sum_{k=n_1+1}^{n_1+n_2} \mathrm{E}[\varepsilon_{3k}^4] &\le \frac{16c_1^4(\kappa_{42}+3)}{\sigma_1^4 n_1^3 n_2^3} \mathrm{E}[(\boldsymbol{Z}'_{11} \Sigma_1^{1/2} \Sigma_2 \Sigma_1^{1/2} \boldsymbol{Z}_{11})^2] \\
&\quad + \frac{16c_1^4(n_1-1)(\kappa_{42}+3)}{\sigma_1^4 n_1^3 n_2^3} (\mathrm{E}[\boldsymbol{Z}'_{11} \Sigma_1^{1/2} \Sigma_2 \Sigma_1^{1/2} \boldsymbol{Z}_{11}])^2 \\
&\quad + \frac{16c_1^4(n_1-1)(\kappa_{42}+3)}{\sigma_1^4 n_1^3 n_2^3} \mathrm{E}[(\boldsymbol{Z}'_{11} \Sigma_1^{1/2} \Sigma_2 \Sigma_1^{1/2} \boldsymbol{Z}_{12})^2] \\
&= O\left(\frac{1}{n_2}\right) = o(1).
\end{aligned} \tag{9.3}
$$

Fourthly, we evaluate $\sum_{k=n_1+1}^{n_1+n_2} \mathrm{E}[\varepsilon_{4k}^4]$:

$$
\sum_{k=n_1+1}^{n_1+n_2} \mathrm{E}[\varepsilon_{4k}^4] = O\left(\frac{1}{n_2}\right) = o(1). \tag{9.4}
$$

Fifthly, we evaluate $\sum_{k=n_1+1}^{n_1+n_2} \mathrm{E}[\varepsilon_{5k}^4]$:

$$\sum_{k=n_1+1}^{n_1+n_2} \mathrm{E}[\varepsilon_{5k}^4] \quad = \quad O\left(\frac{1}{n_2}\right) = o(1). \tag{9.5}$$

Finally, we evaluate $\sum_{k=n_1+1}^{n_1+n_2} \mathrm{E}[\varepsilon_{6k}^4]$:

$$\sum_{k=n_1+1}^{n_1+n_2} \mathrm{E}[\varepsilon_{6k}^4] = O\left(\frac{1}{n_2}\right) = o(1). \tag{9.6}$$

From (9.1)-(9.6), statement (F) follows. □