Multi-group profile analysis for high-dimensional elliptical populations

Masashi Hyodo†

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

Abstract

In this article, we deal with tests for the parallelism and flatness hypotheses in multi-group profile analysis for high-dimensional data. We extend procedures from the results for normal populations proposed by Harrar and Kong [S. W. Harrar, X. Kong, High-dimensional multivariate repeated measures analysis with unequal covariance matrices, J. Multivariate Anal. 145 (2016) 1–21] to results for elliptical populations. Specifically, for an elliptical population, we demonstrate the asymptotic normality of the statistics used in their study, and we propose a new approximate test by improving an estimator of the asymptotic variance. Using asymptotic normality, we show that the asymptotic size of the proposed test is equal to the nominal significance level, and we also derive the asymptotic power. Finally, we present simulation results and find that our results are superior to those found using the existing procedure.

AMS 2000 subject classification: Primary 62H15; secondary 62F05.

Key words: Profile analysis, statistical hypothesis testing, high dimension, elliptical population.

1. Introduction

We consider the multi-sample test problem for profile analysis for elliptical populations. For $g \in \{1, \ldots, a\}$, let $\boldsymbol{\mu}_g = (\mu_{g1}, \ldots, \mu_{gp})^\top$ be a p-dimensional real vector, Λ_g a $p \times p$ nonnegative definite matrix and $\xi_g(\cdot)$ a nonnegative function. The p-dimensional random vector $\boldsymbol{X}_g$ is said to have an elliptically contoured distribution if the characteristic function of $\boldsymbol{X}_g$ can be written as

$$\phi_g(\boldsymbol{t}) = e^{i\boldsymbol{t}^\top \boldsymbol{\mu}_g}\xi_g(\boldsymbol{t}^\top \Lambda_g \boldsymbol{t}).$$

This will be denoted by $\boldsymbol{X}_g \sim C_p(\xi_g, \boldsymbol{\mu}_g, \Lambda_g)$. Note that the expectation and covariance matrix of $\boldsymbol{X}_g$ are $\mathrm{E}(\boldsymbol{X}_g) = \boldsymbol{\mu}_g$ and $\mathrm{var}(\boldsymbol{X}_g) = -2\xi_g'(0)\Lambda_g := \Sigma_g$, respectively. Elliptical distributions include several special cases, for instance, the multivariate normal, multivariate t, and contaminated normal distributions referred to by Muirhead [9].

Let $\boldsymbol{X}_{gi}$ be n_g independent and identically distributed (i.i.d.) copies of $\boldsymbol{X}_g$ for $i \in \{1, \ldots, n_g\}$. We consider a procedure for testing the parallelism hypothesis

$$\mathcal{H}_{01} : \boldsymbol{\mu}_g - \boldsymbol{\mu}_a = \gamma_g \mathbf{1}_p \text{ for any } g \in \{1, \ldots, a-1\} \text{ vs. } \mathcal{A}_{01} : \text{not } \mathcal{H}_{01}. \tag{1.1}$$

Here, γ_g is a unknown real constant and $\mathbf{1}_p = (1, \ldots, 1)^\top$. We also consider a procedure for testing the flatness hypothesis

$$\mathcal{H}_{02} : \mu_{g1} = \cdots = \mu_{gp} \text{ for any } g \in \{1, \ldots, a-1\} \text{ vs. } \mathcal{A}_{02} : \text{not } \mathcal{H}_{02}, \tag{1.2}$$

and the level hypothesis

$$\mathcal{H}_{03} : \gamma_1 = \cdots = \gamma_{a-1} = 0 \text{ vs. } \mathcal{A}_{03} : \text{not } \mathcal{H}_{03}. \tag{1.3}$$

Harrar and Kong [4] showed other expressions that are equivalent to hypotheses (1.1)-(1.3). Expression (1.1) is equivalent to

$$\widetilde{\mathcal{H}}_{01} : \boldsymbol{\mu}^\top K_{01} \boldsymbol{\mu} = 0 \text{ vs. } \widetilde{\mathcal{A}}_{01} : \boldsymbol{\mu}^\top K_{01} \boldsymbol{\mu} > 0$$

July 5, 2017

with

$$K_{01} = P_a \otimes P_p,$$

where $\boldsymbol{\mu} = (\boldsymbol{\mu}_1^\top, \ldots, \boldsymbol{\mu}_a^\top)^\top$ and $P_k = I_k - k^{-1}\mathbf{1}_k\mathbf{1}_k^\top$ for $k \in \{a, p\}$. Other expressions equivalent to hypotheses (1.2) and (1.3) are also obtained:

$$\widetilde{\mathcal{H}}_{0x} : \boldsymbol{\mu}^\top K_{0x}\boldsymbol{\mu} = 0 \text{ vs. } \widetilde{\mathcal{A}}_{0x} : \boldsymbol{\mu}^\top K_{0x}\boldsymbol{\mu} > 0$$

for $x \in \{2, 3\}$ with

$$K_{02} = (a^{-1}\mathbf{1}_a\mathbf{1}_a^\top) \otimes P_p \text{ and } K_{03} = D_a \otimes (p^{-1}\mathbf{1}_p\mathbf{1}_p^\top),$$

where $D_a = \mathrm{diag}(n_1, \ldots, n_a) - n_{(a)}^{-1}\boldsymbol{n}\boldsymbol{n}^\top$. Here, $\boldsymbol{n} = (n_1, \ldots, n_a)^\top$ and $n_{(a)} = \sum_{g=1}^{a} n_g$.

Srivastava [12] derived the likelihood ratio test for hypotheses (1.1), (1.2), and (1.3) for two normal populations. However, the likelihood ratio test for (1.1) and (1.2) cannot be applied to data sets, such as microarray data, for $n_{(a)} \ll p$, even for normal populations with covariance homogeneity.

In profile analysis, Takahashi and Shutoh [13] considered approximation tests for hypotheses (1.1) and (1.2) for two normal populations with equal covariance matrices. Harrar and Kong [4] extended these tests for multi-group normal populations without assuming equal covariance matrices. They also obtained the approximation test for hypothesis (1.3) based on matching moments.

On the other hand, some previous studies of profile analysis have investigated the effects of non-normality in profile analysis. Okamoto et al. [10] used a perturbation method to obtain the asymptotic expansions of the distributions of test statistics for elliptical populations. Maruyama [7] extended the results under more general conditions using a different method introduced by Kano [6]. Note that these results are derived for a large asymptotic $n_{(a)}$.

In this paper, we propose new approximation tests for (1.1) and (1.2) for high-dimensional elliptical populations without assuming equal covariance matrices. We note that the rank of K_{03} is at most $a - 1$. That is, it does not grow with p and hence, it does not make sense to consider a large asymptotic $(n_{(a)}, p)$. Thus, our primary interest is to test (1.1) and (1.2). To propose these approximation tests, for high-dimensional elliptical populations, we show the asymptotic normality of the test statistics proposed by Harrar and Kong [4]. As a result, asymptotic normality is also established for a high-dimensional elliptical population, but that is not a trivial result. Furthermore, improving the estimator of the asymptotic variance of these test statistics enables us to propose a new approximate test for (1.1) and (1.2) for a high-dimensional elliptical population.

The remainder of this paper is organized as follows. The preliminary results for approximation tests are presented in Section 2. Using the asymptotic results in Section 2, we construct approximate tests for (1.1) and (1.2) and derive the asymptotic sizes and powers of these tests for elliptical populations in Section 3. In Section 4, the numerical accuracy of the proposed tests is investigated. The application of the results is illustrated with a real data example in Section 5.

2. Preliminary asymptotic results

We define a non-random $a \times a$ matrix

$$(R_a)_{ij} = \begin{cases} d_i \text{ if } i = j \\ \psi\delta_i\delta_j \text{ if } i \neq j. \end{cases}$$

with $d_i, \delta_i, \psi \in \mathbb{R}$ for $i \in \{1, \ldots, a\}$. Then we consider following random variable:

$$T = \overline{\boldsymbol{X}}^\top (R_a \otimes P_p)\overline{\boldsymbol{X}} - \sum_{g=1}^{a} \frac{d_g \mathrm{tr}(P_p S_g)}{n_g},$$

where

$$\overline{\boldsymbol{X}} = (\overline{\boldsymbol{X}}_1^\top, \ldots, \overline{\boldsymbol{X}}_a^\top)^\top, \quad S_g = \frac{1}{n_g - 1}\sum_{i=1}^{n_g}(\boldsymbol{X}_{gi} - \overline{\boldsymbol{X}}_g)(\boldsymbol{X}_{gi} - \overline{\boldsymbol{X}}_g)^\top.$$

Here, $\overline{\boldsymbol{X}}_g = n_g^{-1}\sum_{i=1}^{n_g} \boldsymbol{X}_{gi}$.

Remark 2.1. *If $\psi = -1$, $d_i = 1 - 1/a$, and $\delta_i = 1/\sqrt{a}$ for $i \in \{1, \dots, a\}$, then T is the test statistic for $\mathcal{H}_{01}$. If $\psi = d_i = a^{-1}$, $\delta_i = 1$ then T is the test statistic for $\mathcal{H}_{02}$.*

Here, T is an unbiased estimator of $\boldsymbol{\mu}^\top(R_a \otimes P_p)\boldsymbol{\mu}$, i.e.,

$$\mathrm{E}(T) = \boldsymbol{\mu}^\top(R_a \otimes P_p)\boldsymbol{\mu}. \tag{2.1}$$

In addition, the variance of T is obtained as follows:

$$\sigma^2 = \sum_{g=1}^{a} \frac{2d_g^2 \mathrm{tr}\{(P_p\Sigma_g)^2\}}{n_g(n_g-1)} + \sum_{g=2}^{a}\sum_{h=1}^{g-1} \frac{4\psi^2\delta_g^2\delta_h^2 \mathrm{tr}(P_p\Sigma_g P_p\Sigma_h)}{n_g n_h} + 4\boldsymbol{\mu}^\top(R_a \otimes P_p)\left\{\sum_{g=1}^{a}\frac{1}{n_g}(\boldsymbol{e}_g\boldsymbol{e}_g^\top)\otimes\Sigma_g\right\}(R_a \otimes P_p)\boldsymbol{\mu}, \tag{2.2}$$

where $\boldsymbol{e}_i$ denotes i-th basis vector.

Remark 2.2. *If R_a is an idempotent matrix, then $R_a \otimes P_p$ is also an idempotent matrix, and $\boldsymbol{\mu}^\top(R_a \otimes P_p)\boldsymbol{\mu} = 0$ is equivalent to $(R_a \otimes P_p)\boldsymbol{\mu} = \mathbf{0}$. Thus, if R_a is an idempotent matrix and $\boldsymbol{\mu}^\top(R_a \otimes P_p)\boldsymbol{\mu} = 0$, then*

$$\begin{aligned} \mathrm{E}(T) &= 0, \\ \sigma^2 &= \sum_{g=1}^{a} \frac{2d_g^2 \mathrm{tr}\{(P_p\Sigma_g)^2\}}{n_g(n_g-1)} + \sum_{g=2}^{a}\sum_{h=1}^{g-1} \frac{4\psi^2\delta_g^2\delta_h^2 \mathrm{tr}(P_p\Sigma_g P_p\Sigma_h)}{n_g n_h}. \end{aligned}$$

We investigate the asymptotic distribution of T for elliptical populations. Our primary objective in this section is to derive the asymptotic distribution of T under some assumptions.

Let n_g for $g \in \{1, \dots, a\}$ be function of p, i.e., $n_g = n_g(p)$, and

$$\mathrm{tr}\{(P_p\Sigma_g P_p\Sigma_h)^i\} \text{ for any } g, h \in \{1, \dots a\},\ i \in \{1, 2\}$$

be function of p. Then we assume following conditions:

(A1) $\lim_{p\to\infty} n_g(p) = \infty,\ 0 < \lim_{p\to\infty} \frac{n_g(p)}{n_h(p)} < \infty$ for any $g, h \in \{1, \dots a\}$,

(A2) $\kappa_g < \infty$ for any $g \in \{1, \dots a\}$,

(A3) $\dfrac{\mathrm{tr}\{(P_p\Sigma_g P_p\Sigma_h)^2\}}{\{\mathrm{tr}(P_p\Sigma_g P_p\Sigma_h)\}^2} = o(1)$ for any $g, h \in \{1, \dots a\}$,

where

$$\kappa_g = \frac{\mathrm{E}[\{(\boldsymbol{X}_g - \boldsymbol{\mu}_g)^\top\Sigma_g^{-1}(\boldsymbol{X}_g - \boldsymbol{\mu}_g)\}^2]}{p(p+2)} - 1.$$

The parameter κ_g is called a kurtosis parameter.

To discuss examples satisfying the assumptions (A2) and (A3), we consider the following three density function of $\boldsymbol{Z} = \Sigma_g^{-1/2}(\boldsymbol{X}_g - \boldsymbol{\mu}_g)$.

a) The multivariate normal distribution with density function

$$f(z) = \frac{1}{(2\pi)^{p/2}}\exp\left(-\frac{z^\top z}{2}\right).$$

b) The ϵ-contaminated normal distribution with density function

$$f(z) = \frac{1-\epsilon}{(2\pi)^{p/2}}\exp\left(-\frac{z^\top z}{2}\right) + \frac{\epsilon}{(2\pi\eta^2)^{p/2}}\exp\left(-\frac{z^\top z}{2\eta^2}\right), \tag{2.3}$$

for $\varepsilon \in [0, 1]$ and $\eta \in (0, \infty)$.

c) The multivariate t distribution with k degrees of freedom with the density function

$$f(z) = \frac{\Gamma[(k+p)/2]}{\Gamma[k/2](k\pi)^{p/2}} \left(1 + \frac{z^\top z}{k}\right)^{-(k+p)/2}$$

for $k \in \mathbb{N}$. Here, $\Gamma[\cdot]$ denotes gamma function.

These distributions satisfy assumption (A2). Actually, the kurtosis parameter of a) is 0, the kurtosis parameter of b) is

$$\frac{1 + \epsilon(\eta^4 - 1)}{\{1 + \epsilon(\eta^2 - 1)\}^2} - 1,$$

and the kurtosis parameter of c) is $2/(k-4)$ for $k > 4$. Examples of covariance matrices that satisfy (A3) are those with compound symmetry. Actually, if $\Sigma_g = (1-\rho_g)I_p + \rho_g(\mathbf{1}_p\mathbf{1}_p^\top)$, for $g \in \{1, \dots, a\}$ and $\rho_g \in (-1/(p-1), 1)$, then $\mathrm{tr}\{(P_p\Sigma_g P_p\Sigma_h)^2\}/\{\mathrm{tr}(P_p\Sigma_g P_p\Sigma_h)\}^2 = 1/(p-1)$.

The following lemma provides the asymptotic normality of T under assumptions (A1), (A2), and (A3). The lemma assures us that the asymptotic normality of the statistic T is maintained for an elliptical population.

Lemma 2.1. *Under assumptions (A1), (A2), and (A3),*

$$\frac{T - \boldsymbol{\mu}^\top(R_a \otimes P_p)\boldsymbol{\mu}}{\sigma} \rightsquigarrow \mathcal{N}(0,1) \text{ as } p \to \infty.$$

Proof. See, Appendix A.2. □

3. Principal results

3.1. Proposed test

In this subsection, we propose approximation tests using a normal approximation based on Lemma 2.1. The test statistics for (1.1) and (1.2) are given by

$$\begin{aligned} T_{01} &= \overline{\boldsymbol{X}}^\top K_{01}\overline{\boldsymbol{X}} - \sum_{g=1}^{a}\left(1 - \frac{1}{a}\right)\frac{\mathrm{tr}(P_pS_g)}{n_g}, \\ T_{02} &= \overline{\boldsymbol{X}}^\top K_{02}\overline{\boldsymbol{X}} - \sum_{g=1}^{a}\frac{\mathrm{tr}(P_pS_g)}{an_g}, \end{aligned}$$

respectively. These statistics are also used in Harrar and Kong [4]. From (2.1), (2.2), and Remark 2.2, their expectation and variance for elliptical populations are summarized by the following equations. For $x \in \{1,2\}$,

$$\begin{aligned} \mathrm{E}(T_{0x}) &= \begin{cases} 0 & \text{under } \mathcal{H}_{0x}, \\ \boldsymbol{\mu}^\top K_{0x}\boldsymbol{\mu}\ (>0) & \text{under } \mathcal{A}_{0x}. \end{cases} \\ \mathrm{var}(T_{0x}) &= \begin{cases} \sigma^2_{\mathcal{H}_{0x}} & \text{under } \mathcal{H}_{0x}, \\ \sigma^2_{\mathcal{A}_{0x}} & \text{under } \mathcal{A}_{0x}. \end{cases} \end{aligned}$$

Here, for $x \in \{1,2\}$ with

$$\sigma^2_{\mathcal{A}_{0x}} = \sigma^2_{\mathcal{H}_{0x}} + 4\boldsymbol{\mu}^\top K_{0x}\left\{\sum_{g=1}^{a}\frac{1}{n_g}(\boldsymbol{e}_g\boldsymbol{e}_g^\top) \otimes \Sigma_g\right\}K_{0x}\boldsymbol{\mu},$$

where

$$
\begin{aligned}
\sigma^2_{\mathcal{H}_{01}} &= \sum_{g=1}^{a}\left(1-\frac{1}{a}\right)^2 \frac{2\mathrm{tr}\{(P_p\Sigma_g)^2\}}{n_g(n_g-1)} + \sum_{g=2}^{a}\sum_{h=1}^{g-1}\frac{4\mathrm{tr}(P_p\Sigma_g P_p\Sigma_h)}{a^2 n_g n_h}, \\
\sigma^2_{\mathcal{H}_{02}} &= \sum_{g=1}^{a}\frac{2\mathrm{tr}\{(P_p\Sigma_g)^2\}}{a^2 n_g(n_g-1)} + \sum_{g=2}^{a}\sum_{h=1}^{g-1}\frac{4\mathrm{tr}(P_p\Sigma_g P_p\Sigma_h)}{a^2 n_g n_h}.
\end{aligned}
$$

In practical use, it is necessary to estimate the asymptotic variance $\sigma^2_{\mathcal{H}_{0x}}$. Harrar and Kong [4] used the following estimator:

$$
\begin{aligned}
\mathrm{tr}\{\widetilde{(P_p\Sigma_g)^2}\} &= \frac{(n_g-1)^2}{(n_g+1)(n_g-2)}\left[\mathrm{tr}\{(P_pS_g)^2\} - \frac{\{\mathrm{tr}(P_pS_g)\}^2}{n_g-1}\right], &(3.1)\\
\mathrm{tr}(\widehat{P_p\Sigma_g P_p}\Sigma_h) &= \mathrm{tr}(P_pS_gP_pS_h). &(3.2)
\end{aligned}
$$

Under elliptical populations, the expectations of them are calculated as following:

$$
\begin{aligned}
\mathrm{E}[\mathrm{tr}\{\widetilde{(P_p\Sigma_g)^2}\}] &= \mathrm{tr}\{(P_p\Sigma_g)^2\} + \frac{\kappa_g(n_g-1)}{n_g(n_g+1)}[\{\mathrm{tr}(P_p\Sigma_g)\}^2 + 2\mathrm{tr}\{(P_p\Sigma_g)^2\}], \\
\mathrm{E}\{\mathrm{tr}(\widehat{P_p\Sigma_g P_p}\Sigma_h)\} &= \mathrm{tr}(P_p\Sigma_g P_p\Sigma_h).
\end{aligned}
$$

Thus, the estimator (3.1) has a bias for elliptical populations except when $\kappa_g = 0$.

We use the same estimator of $\mathrm{tr}(P_p\Sigma_g P_p\Sigma_h)$ as Harrar and Kong [4], but different estimators of $\mathrm{tr}\{(P_p\Sigma_g)^2\}$, which is defined as follows:

$$
\mathrm{tr}\{\widehat{(P_p\Sigma_g)^2}\} = \frac{n_g-1}{n_g(n_g-2)(n_g-3)}\left[(n_g-1)(n_g-2)\mathrm{tr}\{(P_pS_g)^2\} + \{\mathrm{tr}(P_pS_g)\}^2 - n_gM_g\right], \quad (3.3)
$$

where

$$
M_g = \frac{1}{n_g-1}\sum_{i=1}^{n_g}\{(\boldsymbol{X}_{gi}-\overline{\boldsymbol{X}}_g)^\top P_p(\boldsymbol{X}_{gi}-\overline{\boldsymbol{X}}_g)\}^2.
$$

Some properties of the estimators (3.2) and (3.3) are summarized in the following lemma.

Lemma 3.1. *The estimators* $\mathrm{tr}\{\widehat{(P_p\Sigma_g)^2}\}$ *and* $\mathrm{tr}(\widehat{P_p\Sigma_g P_p}\Sigma_h)$ *are unbiased, and rate consistent estimator, i.e. under (A1) and (A2),*

$$
\frac{\mathrm{tr}\{\widehat{(P_p\Sigma_g)^2}\}}{\mathrm{tr}\{(P_p\Sigma_g)^2\}} = 1 + o_p(1), \quad \frac{\mathrm{tr}(\widehat{P_p\Sigma_g P_p}\Sigma_h)}{\mathrm{tr}(P_p\Sigma_g P_p\Sigma_h)} = 1 + o_p(1) \text{ as } p \to \infty.
$$

Proof. See, Appendix A.3. □

Remark 3.1. *If* P_p *is replaced by* I_p, $\mathrm{tr}\widehat{(P_p\Sigma_g)}^2$ *is the same as that of Himeno and Yamada [5].*

Using (3.2) and (3.3) yields

$$
\begin{aligned}
\widehat{\sigma}^2_{\mathcal{H}_{01}} &= \sum_{g=1}^{a}\left(1-\frac{1}{a}\right)^2 \frac{2\mathrm{tr}\{\widehat{(P_p\Sigma_g)^2}\}}{n_g(n_g-1)} + \sum_{g=2}^{a}\sum_{h=1}^{g-1}\frac{4\mathrm{tr}(\widehat{P_p\Sigma_g P_p}\Sigma_h)}{a^2 n_g n_h}, \\
\widehat{\sigma}^2_{\mathcal{H}_{02}} &= \sum_{g=1}^{a}\frac{2\mathrm{tr}\{\widehat{(P_p\Sigma_g)^2}\}}{a^2 n_g(n_g-1)} + \sum_{g=2}^{a}\sum_{h=1}^{g-1}\frac{4\mathrm{tr}(\widehat{P_p\Sigma_g P_p}\Sigma_h)}{a^2 n_g n_h},
\end{aligned}
$$

as the unbiased estimators of $\sigma^2_{\mathcal{H}_{01}}$ and $\sigma^2_{\mathcal{H}_{02}}$, respectively.

Since $\sqrt{\widehat{\sigma}^2_{\mathcal{H}_{0x}}}/\sigma_{\mathcal{H}_{0x}} = 1 + o_p(1)$ as $p \to \infty$ under (A1) and (A2), Lemma 2.1 and Slutsky's theorem complete the following null asymptotic normality. For $x \in \{1, 2\}$, under (A1), (A2), (A3), and $\mathcal{H}_{0x}$,

$$\frac{T_{0x}}{\sqrt{\widehat{\sigma}^2_{\mathcal{H}_{0x}}}} \rightsquigarrow \mathcal{N}(0, 1) \tag{3.4}$$

as $p \to \infty$.

Based on the asymptotic normality (3.4), we propose the following approximation tests:

$$\text{rejecting } \mathcal{H}_{01} \iff T_{01} \geq \sqrt{\widehat{\sigma}^2_{\mathcal{H}_{01}}} z_\alpha, \tag{3.5}$$

$$\text{rejecting } \mathcal{H}_{02} \iff T_{02} \geq \sqrt{\widehat{\sigma}^2_{\mathcal{H}_{02}}} z_\alpha, \tag{3.6}$$

where z_α denotes upper 100α percentile of standard normal distribution.

3.2. Asymptotic size and power

In this subsection, we investigate sizes and powers of test (3.5) and (3.6).

First, we investigate the asymptotic sizes of tests (3.5) and (3.6). Using (3.4), for $x \in \{1, 2\}$, under (A1), (A2), and (A3) yields

$$\Pr\left(T_{0x} \geq \sqrt{\widehat{\sigma}^2_{\mathcal{H}_{0x}}} z_\alpha | \mathcal{H}_{0x}\right) = \alpha + o(1)$$

as $p \to \infty$.

Next, we investigate the asymptotic powers of tests (3.5) and (3.6). Using if Lemma 2.1 and Lemma 3.1, we have the following theorem summarizing the asymptotic powers of tests (3.5) and (3.6).

Theorem 3.1. *For $x \in \{1, 2\}$, under (A1), (A2), and (A3),*

$$\Pr\left(T_{0x} \geq \sqrt{\widehat{\sigma}^2_{\mathcal{H}_{0x}}} z_\alpha | \mathcal{A}_{0x}\right) = \Phi\left(\frac{\boldsymbol{\mu}^\top K_{0x} \boldsymbol{\mu}}{\sigma_{\mathcal{A}_{0x}}} - \frac{\sigma_{\mathcal{H}_{0x}}}{\sigma_{\mathcal{A}_{0x}}} z_\alpha\right) + o(1)$$

as $p \to \infty$, where $\Phi(\cdot)$ denotes the cumulative distribution function (CDF) of the standard normal distribution.

Proof. See, Appendix A.4. □

Thus, if the difference between $\mathcal{H}_{0x}$ and $\mathcal{A}_{0x}$ is not too small in that $\boldsymbol{\mu}^\top K_{0x} \boldsymbol{\mu}$ is of the same order as $\sigma_{\mathcal{A}_{0x}}$ or of a higher order, the test will be powerful. Conversely, if the difference between $\mathcal{H}_{0x}$ and $\mathcal{A}_{0x}$ is so small that $\boldsymbol{\mu}^\top K_{0x} \boldsymbol{\mu}$ is of a lower order than $\sigma_{\mathcal{A}_{0x}}$, the test will not be powerful and cannot distinguish $\mathcal{H}_{0x}$ from $\mathcal{A}_{0x}$.

4. Simulation and real example

4.1. Simulation

In this section, we perform Monte Carlo simulation for some selected parameters in order to verify the superiority of our test as compared to Harrar and Kong's tests for (1.1) and (1.2) when the kurtosis parameter is not 0.

In our simulation, we compare the empirical size and power of the proposed tests and Harrar and Kong's tests. We generated data from the following model:

$$\boldsymbol{X}_{gi} = \Sigma_g^{1/2} \boldsymbol{Z}_{gi} + \boldsymbol{\mu}_g \quad \text{for } i \in \{1, \ldots, n_g\}, \tag{4.1}$$

where

$$\boldsymbol{\mu}_g = (g - 1)\mathbf{1}_p, \quad \Sigma_g = (1 - 0.1g) I_p + (0.1g) \mathbf{1}_p \mathbf{1}_p^\top.$$

for $g \in \{1, \ldots, a\}$. We deal with $a \in \{2, 3, 4, 5\}$. Under this model, the null hypotheses $\mathcal{H}_{01}$ and $\mathcal{H}_{02}$ hold. For the distribution of $\boldsymbol{Z}_{gj}$ in (4.1), we set one of the following distributions:

(a) Multivariate normal distribution,

(b) Contaminated normal distribution with $(\epsilon, \eta) = (0.1, 5)$,

(c) Multivariate t distribution with five degrees of freedom.

The kurtosis parameters of (a), (b), and (c) are 0, approximately 4.48, and 2, respectively. The sizes calculated with 10,000 replications are listed in Tables 1 and 2. Here, the nominal significance level is $\alpha = 0.05$. The settings of a, p and $\boldsymbol{n}$ are also checked in these tables.

Please insert Tables 1 and 2 approximately here.

The empirical sizes of our proposed test and Harrar and Kong's test are presented in Tables 1 and 2, respectively. As can be seen in Table 1, our approximate test shows only approximately a 0.01 difference from the nominal significance level $\alpha = 0.05$ regardless of the population distribution setting when the dimension p is 200 or 400. From Table 2, we see that Harrar and Kong's test has the same tendency only when the population distribution is a multivariate normal distribution. However, the empirical size of Harrar and Kong's test is significantly less than the nominal significance level when the distribution of $\boldsymbol{Z}_{gj}$ is (b) or (c).

For the alternative hypothesis, we choose $\boldsymbol{\mu}_g$ in (4.1) as follows:

$$\boldsymbol{\mu}_g = \begin{cases} (g-1)\mathbf{1}_p, & g \in \{1, \ldots, a-1\}, \\ (a-1)(\mathbf{1}_{\lfloor 0.99p \rfloor}^\top,\ 0.7 \times \mathbf{1}_{p-\lfloor 0.99p \rfloor}^\top)^\top, & g = a, \end{cases}$$

where $\lfloor \cdot \rfloor$ denotes the floor function. The settings of a, p, $\boldsymbol{n}$, the covariance matrix and the distribution of $\boldsymbol{Z}_{gj}$ are the same as the settings for the null hypothesis. Under these models, both $\mathcal{H}_{01}$ and $\mathcal{H}_{02}$ do not hold. The power calculated using 10,000 replications is listed in Tables 3 and 4. Here, the nominal significance level is $\alpha = 0.05$. The settings of a, p, and $\boldsymbol{n}$ are also checked in these tables. The empirical power of our proposed test and Harrar and Kong's test are presented in Tables 3 and 4, respectively. In Table 3, the asymptotic approximation of the power of our proposed test,

$$approx = \Phi\left(\frac{\boldsymbol{\mu}^\top K_{0x}\boldsymbol{\mu}}{\sigma_{\mathcal{A}_{0x}}} - \frac{\sigma_{\mathcal{H}_{0x}}}{\sigma_{\mathcal{A}_{0x}}} z_\alpha\right)$$

is also calculated in each setting. This approximation is based on the result of Theorem 3.1.

Please insert Tables 3 and 4 approximately here.

It can be confirmed that these asymptotic approximations are accurate. Therefore, the power of the proposed tests can be roughly estimated as the value obtained by dividing the distance between the null hypothesis and the alternative hypothesis by the variance of T_{0x} under the alternative hypothesis. From Tables 3 and 4, we see that the powers of both tests are almost same when $\boldsymbol{Z}_{gj}$ follows a multivariate normal distribution. On the other hand, it can also be seen that the proposed test is more powerful than Harrar and Kong's test when the distribution of $\boldsymbol{Z}_{gj}$ is (b) or (c).

From these simulation results, we can see that our tests are more robust against the effects of non-normality as compared to Harrar and Kong's test. The difference between Harrar and Kong's test and our test appears in the estimator of $\mathrm{tr}\{(P_g\Sigma_g)^2\}$. Thus, we compare the bias of Harrar and Kong's estimator $\mathrm{tr}\{\widetilde{(P_p\Sigma_g)^2}\}$ divided by the true parameter $\mathrm{tr}\{(P_g\Sigma_g)^2\}$ and one of our estimators $\mathrm{tr}\{\widehat{(P_p\Sigma_g)}^2\}$ divided by the true parameter $\mathrm{tr}\{(P_g\Sigma_g)^2\}$ under the model (4.1) with $\boldsymbol{\mu}_g = \mathbf{0}$ and $\Sigma_g = (1-0.5)I_p + 0.5\mathbf{1}_p\mathbf{1}_p^\top$. The biases of these estimators are calculated using 10,000 replications in each setting. The biases of $\mathrm{tr}\{\widetilde{(P_p\Sigma_g)}^2\}/\mathrm{tr}\{(P_g\Sigma_g)^2\}$ and $\mathrm{tr}\{\widehat{(P_p\Sigma_g)}^2\}/\mathrm{tr}\{(P_g\Sigma_g)^2\}$ are presented in Table 5. The settings of p and n_g can be checked in Table 5.

Please insert Table 5 approximately here.

From Table 5, we can see that $\mathrm{tr}\{\widetilde{(P_p\Sigma_g)^2}\}$ remarkably overestimates $\mathrm{tr}\{(P_p\Sigma_g)^2\}$ when the distribution of $\boldsymbol{Z}_{gj}$ is (b) or (c). This overestimation contributes to raising the critical value of the test. Since raising the critical value is equivalent to difficulty in rejecting the null hypothesis, we expect that the size for Harrar and Kong's test will be lower than the nominal significance level. This observation is consistent with the results of Table 2. On the other hand, we can understand that $\mathrm{tr}\{\widehat{(P_p\Sigma_g)}^2\}$ is almost unaffected by change in distribution. Therefore, we can understand that changing the estimator of $\mathrm{tr}\{(P_p\Sigma_g)^2\}$ is never a minor correction.

4.2. Real example

We apply our test to the following dataset analyzed by Takahashi and Shutoh [13]. The data consist of $a = 2$ groups, with a body weight of $n_g = 10$ rats for each group. The weights of the total 20 rats were observed every week for $p = 22$ weeks.

We applied our tests (3.5) and (3.6) to the parallelism hypothesis $\mathcal{H}_{01}$ and flatness hypothesis $\mathcal{H}_{02}$ with the significance level α=0.05. Since $T_{01}/\sqrt{\widehat{\sigma}^2_{\mathcal{H}_{01}}}$ and z_α are calculated as

$$T_{01}/\sqrt{\widehat{\sigma}^2_{\mathcal{H}_{01}}} \approx -0.657 < z_\alpha \approx 1.6449,$$

we can see that the parallelism hypothesis $\mathcal{H}_{01}$ is retained. Since $T_{02}/\sqrt{\widehat{\sigma}^2_{\mathcal{H}_{02}}}$ and z_α are calculated as

$$T_{02}/\sqrt{\widehat{\sigma}^2_{\mathcal{H}_{02}}} \approx 250.264 > z_\alpha \approx 1.6449,$$

we can see that the flatness hypothesis $\mathcal{H}_{02}$ is rejected.

5. Discussion and Conclusion

In this paper, we proposed new approximation tests for the parallelism and flatness hypotheses in profile analysis for high-dimensional elliptical populations with unequal covariance matrices, and we derived the asymptotic sizes and power of these proposed tests. We showed that the asymptotic sizes of the proposed tests are at a nominal significance level. However, Harrar and Kong's approximation tests do not necessarily have similar properties for elliptical populations, because the unbiasedness of their estimator of asymptotic variance depends on kurtosis parameters. Furthermore, we found that the asymptotic power depends on the value obtained by dividing the distance between the null and alternative hypotheses by the variance under the alternative hypothesis.

Furthermore, we compared the proposed tests and Harrar and Kong's tests numerically in simulation studies. We found that our tests and Harrar and Kong's tests had approximately the same accuracy when the population distribution is a multivariate normal distribution, and we confirmed our expectation that our tests are superior to Harrar and Kong's tests for elliptical distributions other than the multivariate normal distribution. We also confirmed that this superiority is attributable to the estimator used in asymptotic variance.

In addition, we applied the tests to real data. However, we have not determined whether the assumption of an elliptical population is appropriate. The solution to this problem is to find the validity of the elliptical distribution from the data or to guarantee accuracy with a wider range of distribution families than the elliptical distribution. To achieve the former, it is necessary to extend the method proposed by Batsidis and Zografos [1] to high-dimensional settings. To achieve the latter, it will first be necessary to investigate situations in which the symmetry of the distribution is not assumed, such as when a skew elliptical distribution is used. This change is expected to complicate the estimation of the asymptotic variance. These two tasks are left to future work.

Acknowledgments

We would like to thank Mr. Ogawa for the Monte Carlo simulation used to obtain the numerical results. We also would like to thank the Editor-in-Chief, an Associate Editor, and two reviewers for many valuable comments and helpful suggestions which led to an improved version of this paper. The research of the author was supported in part by a Grant-in-Aid for Young Scientists (B) (17K14238) from the Japan Society for the Promotion of Science.

References

[1] A. Batsidis, K. Zografos A necessary test of fit of specific elliptical distributions based on an estimator of Song's measure, *J.Multivariate Anal.*, **113** (2013) 91–105.

[2] 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.

[3] P. Hall, C.C. Heyde, *Martingale Limit Theory and its Application*, Academic Press, New York, 1980.

[4] S. W. Harrar, X. Kong, High-dimensional multivariate repeated measures analysis with unequal covariance matrices, *J.Multivariate Anal.*, **145** (2016) 1–21.

[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] Y. Kano An asymptotic expansion of the distribution of Hotelling's T^2-statistic under general distributions. *Amer. J. Math. and Manage. Sci.*, **15** (1995) 317-341.

[7] Y. Maruyama, Asymptotic expansions of the null distributions of some test statistics for profile analysis in general conditions, *J.Statist. Plann.Inference*, **137** (2007) 506–526.

[8] A.M. Mathai, S.B. Provost, T. Hayakawa, *Bilinear forms and zonal polynomials, Lecture Notes in Statistics*, 102, Springer-Verlag, New York, 1995.

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

[10] N. Okamoto, N. Miura, T. Seo, On the distributions of some test statistics for profile analysis in elliptical populations, *Amer. J. Math. Manag. Sci.*, **26** (2006) 1–31.

[11] A.N. Shiryaev, *Probability, second ed.*, Springer-Verlag, 1984.

[12] M. S. Srivastava, Profile analysis of several groups, *Commun. Stat.-Theory Meth.*, **16**(3) (1987) 909–926.

[13] S. Takahashi, N. Shutoh, Tests for parallelism and flatness hypotheses of two mean vectors in high-dimensional settings, *J.Stat. Comput.Simul.*, **86** (2016) 1150–1165.

A. Appendix

A.1. Some moments

Lemma A.1. *Let be $\boldsymbol{X} \sim C_p(\xi, \boldsymbol{0}, \Lambda)$, and let A and B be $p \times p$ symmetric real matrices.*

$$\begin{aligned}
&\text{(i)} \quad \mathrm{E}(\boldsymbol{X}^\top A\boldsymbol{X}) = \mathrm{tr}(A\Sigma),\\
&\text{(ii)} \quad \mathrm{E}(\boldsymbol{X}^\top A\boldsymbol{X}\boldsymbol{X}^\top B\boldsymbol{X}) = (\kappa+1)\{\mathrm{tr}(A\Sigma)\mathrm{tr}(B\Sigma) + 2\mathrm{tr}(A\Sigma B\Sigma)\},\\
&\text{(iii)} \quad \mathrm{var}(\boldsymbol{X}^\top A\boldsymbol{X}) = \kappa\{\mathrm{tr}(A\Sigma)\}^2 + 2(\kappa+1)\mathrm{tr}(A\Sigma)^2,\\
&\text{(iv)} \quad \mathrm{cov}(\boldsymbol{X}^\top A\boldsymbol{X}, \boldsymbol{X}^\top B\boldsymbol{X}) = \kappa\mathrm{tr}(A\Sigma)\mathrm{tr}(B\Sigma) + 2(\kappa+1)\mathrm{tr}(A\Sigma B\Sigma),
\end{aligned}$$

where

$$\Sigma := -2\xi'(0)\Lambda, \ \kappa := \frac{\xi''(0)}{(\xi'(0))^2} - 1 \left(= \frac{\mathrm{E}\{(\boldsymbol{X}'\Sigma^{-1}\boldsymbol{X})^2\}}{p(p+2)} - 1 \right).$$

Proof. See Mathai et al. [8]. □

A.2. Proof of Lemma 2.1

We define $\boldsymbol{Y}_{gi} = \boldsymbol{X}_{gi} - \boldsymbol{\mu}_g$ for $g \in \{1, \ldots, a\}$, $i \in \{1, \ldots, n_g\}$. Note that $\boldsymbol{Y}_{gi}$ is distributed according to an elliptical distribution with $\mathrm{E}(\boldsymbol{Y}_{gi}) = \boldsymbol{0}$ and $\mathrm{var}(\boldsymbol{Y}_{gi}) = \Sigma_g$.

Then the statistic T can be rewritten as

$$T = \sum_{g=1}^{a} \frac{d_g}{n_g(n_g-1)} \sum_{\substack{i=1 \\ i \neq j}}^{n_g} \boldsymbol{Y}_{gi}^\top P_p \boldsymbol{Y}_{gj} + \sum_{g \neq h}^{a} \psi\delta_g\delta_h \overline{\boldsymbol{Y}}_g^\top P_p \overline{\boldsymbol{Y}}_h + 2\boldsymbol{\mu}^\top(R_a \otimes P_p)\overline{\boldsymbol{Y}} + \boldsymbol{\mu}^\top(R_a \otimes P_p)\boldsymbol{\mu},$$

where $\overline{\boldsymbol{Y}} = (\overline{\boldsymbol{Y}}_1^\top, \ldots, \overline{\boldsymbol{Y}}_a^\top)^\top$. We note that $\boldsymbol{\mu}^\top(R_a \otimes P_p)\overline{\boldsymbol{Y}} = 0$ and $\boldsymbol{\mu}^\top(R_a \otimes P_p)\boldsymbol{\mu} = 0$ if $(R_a \otimes P_p)\boldsymbol{\mu} = \boldsymbol{0}$. That is, the distribution of T does not depend on $\boldsymbol{\mu}$ as long as $(R_a \otimes P_p)\boldsymbol{\mu} = \boldsymbol{0}$.

Let $n_{(0)} = 0$, $n_{(g)} = \sum_{\ell=1}^{g} n_\ell$ for $g \in \{1, \ldots, a\}$, and $i' = i - n_{(g-1)}$. We define

$$\varepsilon_i = \frac{2}{\sigma n_g(n_g-1)} \boldsymbol{Y}_{gi'}^\top P_p \boldsymbol{a}_{gi'}$$

for $g \in \{1, \ldots, a\}$, $i \in \{n_{(g-1)} + 1, \ldots, n_{(g)}\}$. Here,

$$\boldsymbol{a}_{gi'} = I(i' \geq 2) d_g \sum_{j=1}^{i'-1} \boldsymbol{Y}_{gj} + I(g \geq 2)(n_g - 1) \sum_{h=1}^{g-1} \psi\delta_g\delta_h \overline{\boldsymbol{Y}}_h + (n_g - 1) \sum_{h=1}^{a} \psi\delta_g\delta_h \boldsymbol{\mu}_h,$$

where $I(\cdot)$ denotes an indicator function. Then,

$$\frac{T - \boldsymbol{\mu}^\top(R_a \otimes P_p)\boldsymbol{\mu}}{\sigma} = \sum_{i=1}^{n_{(a)}} \varepsilon_i.$$

Define $\mathcal{F}_0 = \{\emptyset, \Omega\}$, and let $\mathcal{F}_i$ for $i \in \mathbb{N}$ be the σ-algebra generated by the random variables $(\varepsilon_1, \ldots, \varepsilon_i)$. Then we find that

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

and $\mathrm{E}(\varepsilon_i | \mathcal{F}_{i-1}) = 0$. Thus, $\{\varepsilon_i\}$ is a martingale difference sequence.

We show the asymptotic normality of $\sum_{i=1}^{n_{(a)}} \varepsilon_i$ by adapting the martingale difference central limit theorem (see Shiryaev [11] or Hall and Heyde [3]). It is necessary to check the following two conditions to apply this theorem:

$$\text{(I)} \quad \sum_{i=1}^{n_{(a)}} \mathrm{E}(\varepsilon_i^2 | \mathcal{F}_{i-1}) = 1 + o_p(1) \text{ as } p \to \infty,$$

$$\text{(II)} \quad \sum_{i=1}^{n_{(a)}} \mathrm{E}(\varepsilon_i^4) = o(1) \text{ as } p \to \infty$$

under (A1), (A2), and (A3).

First, we check condition (I). We rewrite

$$\sum_{i=1}^{n_{(a)}} \mathrm{E}(\varepsilon_i^2|\mathcal{F}_{i-1}) = 1 + \sum_{j=1}^{7} A_j,$$

where

$$\begin{aligned}
A_1 &= \sum_{g=1}^{a} \frac{4d_g^2}{\sigma^2 n_g^2(n_g-1)^2} \sum_{i=1}^{n_g} (n_g - i)[\boldsymbol{Y}_{gi}^\top P_p \Sigma_g P_p \boldsymbol{Y}_{gi} - \mathrm{tr}\{(\Sigma_g P_p)^2\}], \\
A_2 &= \sum_{g=1}^{a} \frac{8d_g^2}{\sigma^2 n_g^2(n_g-1)^2} \sum_{i=2}^{n_g} \sum_{j=1}^{i-1} (n_g - i)\boldsymbol{Y}_{gi}^\top P_p \Sigma_g P_p \boldsymbol{Y}_{gj}, \\
A_3 &= \sum_{g=2}^{a} \sum_{h=1}^{g-1} \frac{8d_g\psi\delta_g\delta_h}{\sigma^2 n_g^2(n_g-1)} \sum_{i=1}^{n_g} (n_g - i)\boldsymbol{Y}_{gi}^\top P_p \Sigma_g P_p \overline{\boldsymbol{Y}}_h, \\
A_4 &= \sum_{g=1}^{a} \sum_{h=1}^{a} \frac{8d_g\psi\delta_g\delta_h}{\sigma^2 n_g^2(n_g-1)} \sum_{i=1}^{n_g} (n_g - i)\boldsymbol{Y}_{gi}^\top P_p \Sigma_g P_p \boldsymbol{\mu}_h, \\
A_5 &= \sum_{g=2}^{a} \sum_{h=1}^{g-1} \frac{4\psi^2\delta_g^2\delta_h^2}{\sigma^2 n_g} \left\{ \overline{\boldsymbol{Y}}_h^\top P_p \Sigma_g P_p \overline{\boldsymbol{Y}}_h - \frac{\mathrm{tr}(P_p\Sigma_g P_p \Sigma_h)}{n_h} \right\}, \\
A_6 &= \sum_{g=3}^{a} \sum_{h=2}^{g-1} \sum_{\ell=1}^{h-1} \frac{8\psi^2\delta_g^2\delta_h\delta_\ell}{n_g} \overline{\boldsymbol{Y}}_h^\top P_p \Sigma_g P_p \overline{\boldsymbol{Y}}_\ell, \\
A_7 &= \sum_{g=2}^{a} \sum_{h=1}^{g-1} \sum_{\ell=1}^{a} \frac{8\psi^2\delta_g^2\delta_h\delta_\ell}{\sigma^2 n_g} \overline{\boldsymbol{Y}}_h^\top P_p \Sigma_g P_p \boldsymbol{\mu}_\ell.
\end{aligned}$$

It is straightforward to show that $\mathrm{E}(\sum_{j=1}^{7} A_j) = 0$. Holder's inequality yields

$$\mathrm{var}\left\{\sum_{i=1}^{n_{(a)}} \mathrm{E}(\varepsilon_i^2|\mathcal{F}_{i-1})\right\} = \mathrm{E}\left\{\left(\sum_{i=1}^{7} A_i\right)^2\right\} \le 7\sum_{i=1}^{7} \mathrm{E}(A_i^2). \tag{A.1}$$

The expectations $\mathrm{E}(A_1^2)$ through $\mathrm{E}(A_7^2)$ are evaluated as follows:

$$\mathrm{E}(A_1^2) \le \sum_{g=1}^{a} \frac{2(2n_g-1)(3\kappa_g+2)}{3n_g(n_g-1)} = o(1), \tag{A.2}$$

$$\mathrm{E}(A_2^2) \le \sum_{g=1}^{a} \frac{4(n_g-2)}{3n_g} \frac{\mathrm{tr}\{(P_p\Sigma_g)^4\}}{[\mathrm{tr}\{(P_p\Sigma_g)^2\}]^2} = o(1), \tag{A.3}$$

$$\mathrm{E}(A_3^2) \le \sum_{g=2}^{a} \sum_{h=1}^{g-1} \frac{2(2n_g-1)}{3n_g} \sqrt{\frac{\mathrm{tr}\{(P_p\Sigma_g)^4\}}{[\mathrm{tr}\{(P_p\Sigma_g)^2\}]^2}} \sqrt{\frac{\mathrm{tr}\{(P_p\Sigma_g P_p\Sigma_h)^2\}}{\{\mathrm{tr}(P_p\Sigma_g P_p\Sigma_h)\}^2}} = o(1), \tag{A.4}$$

$$\mathrm{E}(A_4^2) \le \sum_{g=1}^{a} \frac{2(2n_g-1)}{3n_g} \sqrt{\frac{\mathrm{tr}\{(P_p\Sigma_g)^4\}}{[\mathrm{tr}\{(P_p\Sigma_g)^2\}]^2}} = o(1), \tag{A.5}$$

$$\mathrm{E}(A_5^2) \le a(a-1)\sum_{g=2}^{a} \sum_{h=1}^{a-1} \left[\frac{\kappa_h}{2n_h} + \frac{(\kappa_h+n_h)\mathrm{tr}\{(P_p\Sigma_g P_p\Sigma_h)^2\}}{n_h\{\mathrm{tr}(P_p\Sigma_g P_p\Sigma_h)\}^2}\right] = o(1), \tag{A.6}$$

$$\mathrm{E}(A_6^2) \leq a(a-1)(a-2)\sum_{g=3}^{a}\sum_{h=2}^{g-1}\sum_{\ell=1}^{h-1}\frac{2\sqrt{\mathrm{tr}\{(P_p\Sigma_g P_p\Sigma_h)^2\}\mathrm{tr}\{(P_p\Sigma_g P_p\Sigma_\ell)^2\}}}{3\mathrm{tr}(P_p\Sigma_g P_p\Sigma_h)\mathrm{tr}(P_p\Sigma_g P_p\Sigma_\ell)} = o(1), \tag{A.7}$$

$$\mathrm{E}(A_7^2) \leq a(a-1)\sum_{g=2}^{a}\sum_{h=1}^{g-1}\frac{\sqrt{\mathrm{tr}\{(P_p\Sigma_g P_p\Sigma_h)^2\}}}{\mathrm{tr}(P_p\Sigma_g P_p\Sigma_h)} = o(1). \tag{A.8}$$

The details of these inequalities are described in Supplemental Materials 1. Substituting (A.2)-(A.8) into (A.1) yields

$$\mathrm{var}\left\{\sum_{i=1}^{n_{(a)}}\mathrm{E}(\varepsilon_i^2|\mathcal{F}_{i-1})\right\} = o(1).$$

Condition (I) follows.

Next, we check condition (II). Define

$$\varepsilon_i^{(1)} = \frac{2I(i'\geq 2)d_g}{\sigma n_g(n_g-1)}\boldsymbol{Y}_{gi'}^\top P_p\sum_{j=1}^{i'-1}\boldsymbol{Y}_{gj},\ \varepsilon_i^{(2)} = \frac{2I(g\geq 2)}{\sigma n_g}\boldsymbol{Y}_{gi'}^\top P_p\sum_{h=1}^{\ell-1}\psi\delta_g\delta_h\overline{\boldsymbol{Y}}_h,\ \varepsilon_i^{(3)} = \frac{2}{\sigma n_g}\boldsymbol{Y}_{gi'}^\top P_p\sum_{h=1}^{a}\psi\delta_g\delta_h\boldsymbol{\mu}_h.$$

Then, Holder's inequality yields

$$\sum_{i=1}^{n_{(a)}}\mathrm{E}(\varepsilon_i^4) = \sum_{g=1}^{a}\sum_{i=n_{(g-1)}+1}^{n_{(g)}}\mathrm{E}\left\{\left(\sum_{j=1}^{3}\varepsilon_i^{(j)}\right)^4\right\} \leq \sum_{g=1}^{a}\sum_{i=n_{(g-1)}+1}^{n_{(g)}}\mathrm{E}\left(3^3\sum_{j=1}^{3}\varepsilon_i^{(j)^4}\right) = 3^3\sum_{g=1}^{a}\sum_{i=n_{(g-1)}+1}^{n_{(g)}}\sum_{j=1}^{3}\mathrm{E}\left(\varepsilon_i^{(j)^4}\right).$$

Thus, it is sufficient to show that $\sum_{i=n_{(g-1)}+1}^{n_{(g)}}\mathrm{E}(\varepsilon_i^{(j)^4}) = o(1)$ for $j\in\{1,2,3\}$. We have following expectations:

$$\sum_{i=n_{(g-1)}+1}^{n_{(g)}}\mathrm{E}\left(\varepsilon_i^{(1)^4}\right) \leq \frac{18(\kappa_g+1)^2}{n_g(n_g-1)} + \frac{4(\kappa_g+1)(n_g-2)}{n_g(n_g-1)} + \frac{8(\kappa_g+1)(n_g-2)}{n_g(n_g-1)} = O(n_g^{-1}),$$

$$\sum_{i=n_{(g-1)}+1}^{n_{(g)}}\mathrm{E}\left(\varepsilon_i^{(2)^4}\right) \leq \sum_{h=1}^{g-1}\frac{9(\kappa_g+1)(\kappa_h+1)}{n_g n_h} + \sum_{h=1}^{g-1}\frac{3(\kappa_g+1)(n_h-1)}{n_g n_h} + \sum_{h\neq h'}^{g-1}\frac{3(\kappa_g+1)}{n_g} + \sum_{h\neq h'}^{g-1}\frac{6(\kappa_g+1)}{n_g} = O(n_g^{-1}),$$

$$\sum_{i=n_{(g-1)}+1}^{n_{(g)}}\mathrm{E}\left(\varepsilon_i^{(3)^4}\right) \leq \frac{3}{n_g} = O(n_g^{-1}).$$

The details of these inequalities are described in Supplemental Materials 2. The above results complete the proof of (II). □

A.3. Proof of Lemma 3.1

First, we show the unbiasedness and consistency of $\widehat{\mathrm{tr}(P_p\Sigma_g)^2}$. From Lemma A.1, it follows that

$$\mathrm{E}[\mathrm{tr}\{(P_pS_g)^2\}] = \frac{\kappa_g+1}{n_g}\left[\{\mathrm{tr}(P_p\Sigma_g)\}^2 + 2\mathrm{tr}\{(P_p\Sigma_g)^2\}\right] + \frac{n_g^2-2n_g+2}{n_g(n_g-1)}\mathrm{tr}\{(P_p\Sigma_g)^2\} + \frac{1}{n_g(n_g-1)}\{\mathrm{tr}(P_p\Sigma_g)\}^2,$$

$$\mathrm{E}[\{\mathrm{tr}(P_pS_g)\}^2] = \frac{\kappa_g+1}{n_g}\left[\{\mathrm{tr}(P_p\Sigma_g)\}^2 + 2\mathrm{tr}\{(P_p\Sigma_g)^2\}\right] + \frac{2}{n_g(n_g-1)}\mathrm{tr}\{(P_p\Sigma_g)^2\} + \frac{n_g-1}{n_g}\{\mathrm{tr}(P_p\Sigma_g)\}^2,$$

$$\mathrm{E}(M_g) = \frac{(\kappa_g+1)(n_g^2-3n_g+3)}{n_g^2}\left[\{\mathrm{tr}(P_p\Sigma_g)\}^2 + 2\mathrm{tr}\{(P_p\Sigma_g)^2\}\right] + \frac{4n_g-6}{n_g^2}\mathrm{tr}\{(P_p\Sigma_g)^2\} + \frac{2n_g-3}{n_g^2}\{\mathrm{tr}(P_p\Sigma_g)\}^2.$$

Then we solve simultaneously for $\mathrm{tr}\{(P_p\Sigma_g)^2\}$, $\{\mathrm{tr}(P_p\Sigma_g)\}^2$ and κ_g. The solutions of the simultaneous equations can be obtained easily , with the result that

$$\mathrm{tr}\{(P_p\Sigma_g)^2\} = \frac{n_g-1}{n_g(n_g-2)(n_g-3)}\left\{(n_g-1)(n_g-2)\mathrm{E}[\mathrm{tr}\{(P_pS_g)^2\}] + \mathrm{E}[\{\mathrm{tr}(P_pS_g)\}^2] - n_g\mathrm{E}(M_g)\right\}.$$

Thus, the unbiased estimator of $\mathrm{tr}\{(P_p\Sigma_g)^2\}$ is $\mathrm{tr}\{\widehat{(P_p\Sigma_g)}^2\}$. Also, variance of $\mathrm{tr}\{\widehat{(P_p\Sigma_g)}^2\}$ is

$$\begin{aligned}\mathrm{var}[\mathrm{tr}\{\widehat{(P_p\Sigma_g)}^2\}] &= \left\{\frac{4(n_g^2-6n_g+11)}{n_g(n_g-1)(n_g-2)(n_g-3)}+\frac{2\kappa_g}{n_g(n_g-1)}\right\}\left[\mathrm{tr}\{(P_p\Sigma_g)^2\}\right]^2 \\ &+\left\{\frac{4(2n_g^3-12n_g^2+21n_g-5)}{n_g(n_g-1)(n_g-2)(n_g-3)}+\frac{\kappa_1^2}{n_g(n_g-1)}+\frac{4n_g^2-11n_g+2}{n_g(n_g-1)(n_g-2)}\kappa_g\right\}\mathrm{tr}\{(P_p\Sigma_g)^4\}.\end{aligned}$$

Thus under (A1) and (A2),

$$\frac{\mathrm{tr}\widehat{(P_p\Sigma_g)}^2}{\mathrm{tr}(P_p\Sigma_g)^2}=1+o_p(1)\ \ as\ \ p\to\infty.$$

Next, we show the unbiasedness and consistency of $\mathrm{tr}(\widehat{P_p\Sigma_gP_p}\Sigma_h)$. We can rewrite the estimator $\mathrm{tr}(\widehat{P_p\Sigma_gP_p}\Sigma_h)$ as

$$\begin{aligned}\mathrm{tr}(\widehat{P_p\Sigma_gP_p}\Sigma_h) &= \frac{1}{n_gn_h}\sum_{i=1}^{n_g}\sum_{j=1}^{n_h}(\boldsymbol{Y}_{gi}^\top P_p\boldsymbol{Y}_{hj})^2-\frac{1}{n_gn_h(n_h-1)}\sum_{i=1}^{n_g}\sum_{\substack{j,k=1\\ j\neq k}}^{n_h}\boldsymbol{Y}_{gi}^\top P_p\boldsymbol{Y}_{hj}\boldsymbol{Y}_{gi}^\top P_p\boldsymbol{Y}_{hk} \\ &-\frac{1}{n_g(n_h-1)n_h}\sum_{\substack{i,j=1\\ i\neq j}}^{n_g}\sum_{k=1}^{n_h}\boldsymbol{Y}_{gi}^\top P_p\boldsymbol{Y}_{hk}\boldsymbol{Y}_{gj}^\top P_p\boldsymbol{Y}_{hk}+\frac{1}{n_gn_h(n_g-1)(n_h-1)}\sum_{\substack{i,j=1\\ i\neq j}}^{n_g}\sum_{\substack{k,\ell=1\\ k\neq \ell}}^{n_h}\boldsymbol{Y}_{gi}^\top P_p\boldsymbol{Y}_{hk}\boldsymbol{Y}_{gj}^\top P_p\boldsymbol{Y}_{h\ell}.\end{aligned}$$

Using Lemma A.1, we get $\mathrm{E}\{\mathrm{tr}(\widehat{P_p\Sigma_gP_p}\Sigma_h)\}=\mathrm{tr}(P_p\Sigma_gP_p\Sigma_h)$. Also, second moment of $\mathrm{tr}(\widehat{P_p\Sigma_gP_p}\Sigma_h)$ is

$$\begin{aligned}\mathrm{E}[\{\mathrm{tr}(\widehat{P_p\Sigma_gP_p}\Sigma_h)\}^2] &= \left\{1+\frac{3\kappa_g\kappa_h}{n_gn_h}+\frac{\kappa_g(n_h+1)}{n_g(n_h-1)}+\frac{\kappa_h(n_g+1)}{n_h(n_g-1)}+\frac{2}{(n_g-1)(n_h-1)}\right\}\{\mathrm{tr}(P_p\Sigma_gP_p\Sigma_h)\}^2 \\ &+2\left\{\frac{3\kappa_g\kappa_h}{n_gn_h}+\frac{\kappa_g(n_h+1)}{n_g(n_h-1)}+\frac{\kappa_h(n_g+1)}{(n_g-1)n_h}+\frac{1}{(n_g-1)(n_h-1)}+\frac{1}{n_g-1}+\frac{1}{n_h-1}\right\}\mathrm{tr}\{(P_p\Sigma_gP_p\Sigma_h)^2\}.\end{aligned}$$

Thus under (A1) and (A2),

$$\frac{\mathrm{tr}(\widehat{P_p\Sigma_gP_p}\Sigma_h)}{\mathrm{tr}(P_p\Sigma_gP_p\Sigma_h)}=1+o_p(1)\ \ as\ \ p\to\infty.\ \square$$

A.4. Proof of Theorem 3.1

We assume $\mathcal{A}_{0x}$. From Lemma 3.1 and $\sigma^2_{\mathcal{H}_{0x}}/\sigma^2_{\mathcal{A}_{0x}}<1$, under (A1) and (A2), $\widehat{\sigma}^2_{\mathcal{H}_{0x}}/\sigma^2_{\mathcal{A}_{0x}}=\sigma^2_{\mathcal{H}_{0x}}/\sigma^2_{\mathcal{A}_{0x}}+o_p(1)$ as $p\to\infty$. Thus, under (A1) and (A2),

$$\Pr\left(T_{0x}\geq\sqrt{\widehat{\sigma}^2_{\mathcal{H}_{0x}}}z_\alpha\right)=\Pr\left(\frac{T_{0x}-\boldsymbol{\mu}^\top K_{0x}\boldsymbol{\mu}}{\sigma_{\mathcal{A}_{0x}}}\geq\frac{\sigma_{\mathcal{H}_{0x}}}{\sigma_{\mathcal{A}_{0x}}}z_\alpha-\frac{\boldsymbol{\mu}^\top K_{0x}\boldsymbol{\mu}}{\sigma_{\mathcal{A}_{0x}}}\right)+o(1)$$

as $p\to\infty$. Furthermore, from Lemma 2.1, under (A1), (A2), and (A3),

$$\Pr\left(\frac{T_{0x}-\boldsymbol{\mu}^\top K_{0x}\boldsymbol{\mu}}{\sigma_{\mathcal{A}_{0x}}}\geq\frac{\sigma_{\mathcal{H}_{0x}}}{\sigma_{\mathcal{A}_{0x}}}z_\alpha-\frac{\boldsymbol{\mu}^\top K_{0x}\boldsymbol{\mu}}{\sigma_{\mathcal{A}_{0x}}}\right)=\Phi\left(\frac{\boldsymbol{\mu}^\top K_{0x}\boldsymbol{\mu}}{\sigma_{\mathcal{A}_{0x}}}-\frac{\sigma_{\mathcal{H}_{0x}}}{\sigma_{\mathcal{A}_{0x}}}z_\alpha\right)+o(1)$$

as $p\to\infty$. Combining these two equations yields the theorem.$\square$

Table 1: The empirical size of proposed test

p	$\boldsymbol{n}^\top$	Parallelism			Flatness		
		(a)	(b)	(c)	(a)	(b)	(c)
	(25,25)	0.056	0.059	0.053	0.054	0.062	0.056
100	(10,20,20)	0.075	0.062	0.068	0.060	0.062	0.062
	(10,10,10,20)	0.076	0.072	0.069	0.062	0.060	0.054
	(10,10,10,10,10)	0.064	0.065	0.064	0.053	0.055	0.074
	(50,50)	0.057	0.058	0.057	0.053	0.058	0.058
200	(20,40,40)	0.057	0.055	0.059	0.057	0.060	0.056
	(20,20,20,40)	0.057	0.059	0.056	0.058	0.059	0.055
	(40,40,40,40,40)	0.058	0.053	0.059	0.056	0.057	0.054
	(100,100)	0.053	0.051	0.057	0.052	0.052	0.057
400	(40,80,80)	0.056	0.053	0.052	0.055	0.056	0.054
	(40,40,40,80)	0.054	0.054	0.057	0.054	0.053	0.054
	(40,40,40,40,40)	0.052	0.056	0.053	0.056	0.057	0.055

Table 2: The empirical size of Harrar and Kong's test

p	$\boldsymbol{n}^\top$	Parallelism			Flatness		
		(a)	(b)	(c)	(a)	(b)	(c)
	(25,25)	0.057	0.000	0.006	0.054	0.000	0.017
100	(10,20,20)	0.071	0.000	0.005	0.062	0.000	0.010
	(10,10,10,20)	0.075	0.000	0.003	0.056	0.000	0.007
	(10,10,10,10,10)	0.064	0.000	0.001	0.055	0.000	0.011
	(50,50)	0.057	0.000	0.005	0.053	0.000	0.008
200	(20,40,40)	0.057	0.000	0.001	0.057	0.000	0.005
	(20,20,20,40)	0.057	0.000	0.000	0.058	0.000	0.005
	(40,40,40,40,40)	0.058	0.000	0.000	0.056	0.000	0.006
	(100,100)	0.053	0.000	0.004	0.052	0.000	0.005
400	(40,80,80)	0.056	0.000	0.007	0.055	0.000	0.002
	(40,40,40,80)	0.054	0.000	0.000	0.054	0.000	0.002
	(40,40,40,40,40)	0.052	0.000	0.000	0.057	0.000	0.005

Table 3: The empirical power of proposed test

p	$\boldsymbol{n}^\top$	Parallelism				Flatness			
		(a)	(b)	(c)	*approx*	(a)	(b)	(c)	*approx*
100	(25,25)	0.069	0.076	0.073	0.063	0.071	0.076	0.072	0.063
	(10,20,20)	0.093	0.104	0.094	0.078	0.079	0.087	0.084	0.072
	(10,10,10,20)	0.117	0.140	0.126	0.102	0.093	0.103	0.097	0.082
	(10,10,10,10,10)	0.166	0.201	0.167	0.149	0.107	0.111	0.107	0.096
200	(50,50)	0.095	0.104	0.094	0.089	0.098	0.097	0.088	0.089
	(20,40,40)	0.153	0.191	0.161	0.152	0.131	0.145	0.138	0.123
	(20,20,20,40)	0.266	0.320	0.275	0.265	0.162	0.187	0.170	0.161
	(40,40,40,40,40)	0.482	0.532	0.500	0.486	0.212	0.244	0.222	0.221
400	(100,100)	0.200	0.204	0.192	0.195	0.197	0.208	0.198	0.195
	(40,80,80)	0.491	0.535	0.507	0.501	0.348	0.379	0.345	0.349
	(40,40,40,80)	0.858	0.861	0.863	0.862	0.505	0.523	0.513	0.513
	(40,40,40,40,40)	0.995	0.991	0.994	0.995	0.713	0.727	0.716	0.717

Table 4: The empirical power of Harrar and Kong's method

p	$\boldsymbol{n}^\top$	Parallelism			Flatness		
		(a)	(b)	(c)	(a)	(b)	(c)
100	(25,25)	0.069	0.001	0.014	0.071	0.001	0.014
	(10,20,20)	0.093	0.001	0.007	0.079	0.001	0.016
	(10,10,10,20)	0.117	0.001	0.006	0.093	0.002	0.018
	(10,10,10,10,10)	0.167	0.003	0.005	0.107	0.001	0.029
200	(50,50)	0.095	0.000	0.014	0.098	0.000	0.011
	(20,40,40)	0.152	0.000	0.007	0.131	0.000	0.018
	(20,20,20,40)	0.266	0.000	0.007	0.162	0.001	0.028
	(40,40,40,40,40)	0.482	0.000	0.016	0.212	0.000	0.049
400	(100,100)	0.200	0.000	0.027	0.197	0.000	0.028
	(40,80,80)	0.491	0.000	0.034	0.344	0.000	0.054
	(40,40,40,80)	0.858	0.000	0.098	0.505	0.001	0.124
	(40,40,40,40,40)	0.995	0.000	0.394	0.713	0.011	0.307

Table 5: Bias of $\mathrm{tr}\{\widetilde{(P_p\Sigma_1})^2\}/\mathrm{tr}\{(P_p\Sigma_1)^2\}$ and of $\mathrm{tr}\{\widehat{(P_p\Sigma_1})^2\}/\mathrm{tr}\{(P_p\Sigma_1)^2\}$

		$\mathrm{tr}\{\widetilde{(P_p\Sigma_1})^2\}/\mathrm{tr}\{(P_p\Sigma_1)^2\}$			$\mathrm{tr}\{\widehat{(P_p\Sigma_1})^2\}/\mathrm{tr}\{(P_p\Sigma_1)^2\}$		
p	n_1	(a)	(b)	(c)	(a)	(b)	(c)
100	10	0.00	37.80	14.92	0.00	0.05	0.00
	25	0.00	16.91	7.67	0.00	0.00	0.00
200	20	0.00	40.57	14.76	0.00	-0.01	-0.01
	50	0.00	17.33	8.12	0.00	0.00	0.00
400	40	0.00	43.03	16.31	0.00	0.02	-0.01
	100	0.00	17.54	6.75	0.00	-0.01	0.00

· **The evaluation of** $\mathrm{E}(A_1^2)$

$$\begin{aligned}
\mathrm{E}(A_1^2) &= \sum_{g=1}^{a}\frac{8d_g^4(2n_g-1)}{3\sigma_g^4 n_g^3(n_g-1)^3}\mathrm{var}(\boldsymbol{Y}_{g1}^\top P_p\Sigma_g P_p \boldsymbol{Y}_{g1}) \\
&= \sum_{g=1}^{a}\frac{8d_g^4(2n_g-1)}{3\sigma_g^4 n_g^3(n_g-1)^3}\left[\kappa_g[\mathrm{tr}\{(P_p\Sigma_g)^2\}]^2+2(\kappa_g+1)\mathrm{tr}\{(P_p\Sigma_g)^4\}\right] \\
&\leq \sum_{g=1}^{a}\frac{8d_g^4(2n_g-1)(3\kappa_g+2)}{3\sigma_g^4 n_g^3(n_g-1)^3}[\mathrm{tr}\{(P_p\Sigma_g)^2\}]^2 \\
&\leq \sum_{g=1}^{a}\frac{2(2n_g-1)(3\kappa_g+2)}{3n_g(n_g-1)} \\
&= o(1).
\end{aligned}$$

· **The evaluation of** $\mathrm{E}(A_2^2)$

$$\begin{aligned}
\mathrm{E}(A_2^2) &= \sum_{g=1}^{a}\frac{16d_g^4(n_g-2)}{3\sigma^4 n_g^3(n_g-1)^2}\mathrm{E}\{(\boldsymbol{Y}_{g1}^\top P_p\Sigma_g P_p \boldsymbol{Y}_{g2})^2\} \\
&= \sum_{g=1}^{a}\frac{16d_g^4(n_g-2)}{3\sigma^4 n_g^3(n_g-1)^2}\mathrm{tr}\{(P_p\Sigma_g)^4\} \\
&\leq \sum_{g=1}^{a}\frac{4(n_g-2)}{3n_g}\frac{\mathrm{tr}\{(P_p\Sigma_g)^4\}}{[\mathrm{tr}\{(P_p\Sigma_g)^2\}]^2} \\
&= o(1).
\end{aligned}$$

· **The evaluation of** $\mathrm{E}(A_3^2)$

$$\begin{aligned}
\mathrm{E}(A_3^2) &= \sum_{g=2}^{a}\sum_{h=1}^{g-1}\frac{32(2n_g-1)d_g^2\psi^2\delta_g^2\delta_h^2}{3\sigma^4 n_g^3(n_g-1)n_h}\mathrm{E}\{(\boldsymbol{Y}_{g1}^\top P_p\Sigma_g P_p \boldsymbol{Y}_{h1})^2\} \\
&= \sum_{g=2}^{a}\sum_{h=1}^{g-1}\frac{32(2n_g-1)d_g^2\psi^2\delta_g^2\delta_h^2}{3\sigma^4 n_g^3(n_g-1)n_h}\mathrm{tr}\{(P_p\Sigma_g)^3P_p\Sigma_h\} \\
&\leq \sum_{g=2}^{a}\sum_{h=1}^{g-1}\frac{32(2n_g-1)d_g^2\psi^2\delta_g^2\delta_h^2}{3\sigma^4 n_g^3(n_g-1)n_h}\sqrt{\mathrm{tr}\{(P_p\Sigma_g)^4\}}\sqrt{\mathrm{tr}\{(P_p\Sigma_g P_p\Sigma_h)^2\}} \\
&\leq \sum_{g=2}^{a}\sum_{h=1}^{g-1}\frac{2(2n_g-1)}{3n_g}\sqrt{\frac{\mathrm{tr}\{(P_p\Sigma_g)^4\}}{[\mathrm{tr}\{(P_p\Sigma_g)^2\}]^2}}\sqrt{\frac{\mathrm{tr}\{(P_p\Sigma_g P_p\Sigma_h)^2\}}{\{\mathrm{tr}(P_p\Sigma_g P_p\Sigma_h)\}^2}} \\
&= o(1).
\end{aligned}$$

· **The evaluation of** $\mathrm{E}(A_4^2)$

$$\begin{aligned}
\mathrm{E}(A_4^2) &= \sum_{g=1}^{a} \frac{32d_g^2(2n_g-1)}{3\sigma^4 n_g^3(n_g-1)} \mathrm{E}\left[\left\{\boldsymbol{Y}_{g1}^\top P_p \Sigma_g P_p \left(\sum_{h=1}^{a} \psi\delta_g\delta_h \boldsymbol{\mu}_h\right)\right\}^2\right] \\
&= \sum_{g=1}^{a} \frac{32d_g^2(2n_g-1)}{3\sigma^4 n_g^3(n_g-1)} \\
&\quad \times \left(\sum_{h=1}^{a} \psi\delta_g\delta_h \boldsymbol{\mu}_h\right)^\top P_p \Sigma_g P_p \Sigma_g P_p \Sigma_g P_p \left(\sum_{h=1}^{a} \psi\delta_g\delta_h \boldsymbol{\mu}_h\right) \\
&\leq \sum_{g=1}^{a} \frac{32d_g^2(2n_g-1)}{3\sigma^4 n_g^3(n_g-1)} \\
&\quad \times \left(\sum_{h=1}^{a} \psi\delta_g\delta_h \boldsymbol{\mu}_h\right)^\top P_p \Sigma_g P_p \left(\sum_{h=1}^{a} \psi\delta_g\delta_h \boldsymbol{\mu}_h\right) \sqrt{\mathrm{tr}\{(P_p\Sigma_g)^4\}} \\
&\leq \sum_{g=1}^{a} \frac{32d_g^2(2n_g-1)}{3\sigma^4 n_g^2(n_g-1)} \sqrt{\mathrm{tr}\{(P_p\Sigma_g)^4\}} \\
&\quad \times \boldsymbol{\mu}^\top (R_a \otimes P_p) \left\{\sum_{g=1}^{a} n_g^{-1} (\boldsymbol{e}_g \boldsymbol{e}_g^\top) \otimes \Sigma_g\right\} (R_a \otimes P_p)\boldsymbol{\mu} \\
&\leq \sum_{g=1}^{a} \frac{2(2n_g-1)}{3n_g} \sqrt{\frac{\mathrm{tr}\{(P_p\Sigma_g)^4\}}{[\mathrm{tr}\{(P_p\Sigma_g)^2\}]^2}} \\
&= o(1).
\end{aligned}$$

· **The evaluation of** $\mathrm{E}(A_5^2)$

$$
\begin{aligned}
\mathrm{E}(A_5^2) &\le a(a-1)\sum_{g=2}^{a}\sum_{h=1}^{g-1}\frac{8\psi^4\delta_g^4\delta_h^4}{\sigma^4 n_g^2}\left\{\frac{\mathrm{var}(\boldsymbol{Y}_{h1}^\top P_p\Sigma_g P_p\boldsymbol{Y}_{h1})}{n_h^3}\right.\\
&\quad\left.+\frac{2(n_h-1)\mathrm{tr}\{(P_p\Sigma_g P_p\Sigma_h)^2\}}{n_h^3}\right\}\\
&= a(a-1)\sum_{g=2}^{a}\sum_{h=1}^{g-1}\frac{8\kappa_h\psi^4\delta_g^4\delta_h^4}{\sigma^4 n_g^2 n_h^3}\{\mathrm{tr}(P_p\Sigma_g P_p\Sigma_h)\}^2\\
&\quad+a(a-1)\sum_{g=2}^{a}\sum_{h=1}^{g-1}\frac{16(n_h+\kappa_h)\psi^4\delta_g^4\delta_h^4}{\sigma^4 n_g^2 n_h^3}\{\mathrm{tr}(P_p\Sigma_g P_p\Sigma_h)\}^2\\
&\le a(a-1)\sum_{g=2}^{a}\sum_{h=1}^{a-1}\frac{\kappa_h}{2n_h}\\
&\quad+a(a-1)\sum_{g=2}^{a}\sum_{h=1}^{a-1}\frac{\kappa_h+n_h}{n_h}\frac{\mathrm{tr}\{(P_p\Sigma_g P_p\Sigma_h)^2\}}{\{\mathrm{tr}(P_p\Sigma_g P_p\Sigma_h)\}^2}\\
&= o(1).
\end{aligned}
$$

· **The evaluation of** $\mathrm{E}(A_6^2)$

$$
\begin{aligned}
\mathrm{E}(A_6^2) &\le a(a-1)(a-2)\sum_{g=3}^{a}\sum_{h=2}^{g-1}\sum_{\ell=1}^{h-1}\frac{32\psi^4\delta_g^4\delta_h^2\delta_\ell^2}{3\sigma^4 n_g^2 n_h n_\ell}\mathrm{E}\{(\boldsymbol{Y}_{h1}^\top P_p\Sigma_g P_p\boldsymbol{Y}_{\ell 1})^2\}\\
&= a(a-1)(a-2)\sum_{g=3}^{a}\sum_{h=2}^{g-1}\sum_{\ell=1}^{h-1}\frac{32\psi^4\delta_g^4\delta_h^2\delta_\ell^2}{3\sigma^4 n_g^2 n_h n_\ell}\mathrm{tr}(P_p\Sigma_g P_p\Sigma_h P_p\Sigma_g P_p\Sigma_\ell)\\
&\le a(a-1)(a-2)\sum_{g=3}^{a}\sum_{h=2}^{g-1}\sum_{\ell=1}^{h-1}\frac{32\psi^4\delta_g^4\delta_h^2\delta_\ell^2}{3\sigma^4 n_g^2 n_h n_\ell}\sqrt{\mathrm{tr}\{(P_p\Sigma_g P_p\Sigma_h)^2\}\mathrm{tr}\{(P_p\Sigma_g P_p\Sigma_\ell)^2\}}\\
&\le a(a-1)(a-2)\sum_{g=3}^{a}\sum_{h=2}^{g-1}\sum_{\ell=1}^{h-1}\frac{2\sqrt{\mathrm{tr}\{(P_p\Sigma_g P_p\Sigma_h)^2\}\mathrm{tr}\{(P_p\Sigma_g P_p\Sigma_\ell)^2\}}}{3\mathrm{tr}(P_p\Sigma_g P_p\Sigma_h)\mathrm{tr}(P_p\Sigma_g P_p\Sigma_\ell)}\\
&= o(1).
\end{aligned}
$$

· **The evaluation of** $\mathrm{E}(A_7^2)$

$$
\begin{aligned}
\mathrm{E}(A_7^2) &\leq a(a-1)\sum_{g=2}^{a}\sum_{h=1}^{g-1}\frac{32\psi^2\delta_g^2\delta_h^2}{\sigma^4 n_g^2 n_h}\mathrm{E}\left[\left\{\boldsymbol{Y}_{h1}^\top P_p\Sigma_g P_p\left(\sum_{\ell=1}^{a}\psi\delta_g\delta_\ell\boldsymbol{\mu}_\ell\right)\right\}^2\right] \\
&= a(a-1)\sum_{g=2}^{a}\sum_{h=1}^{g-1}\frac{32\psi^2\delta_g^2\delta_h^2}{\sigma^4 n_g^2 n_h}\left(\sum_{\ell=1}^{a}\psi\delta_g\delta_\ell\boldsymbol{\mu}_\ell\right)^\top P_p\Sigma_g P_p\Sigma_h P_p\Sigma_g P_p\left(\sum_{\ell=1}^{a}\psi\delta_g\delta_\ell\boldsymbol{\mu}_\ell\right) \\
&\leq a(a-1)\sum_{g=2}^{a}\sum_{h=1}^{g-1}\frac{32\psi^2\delta_g^2\delta_h^2}{\sigma^4 n_g^2 n_h}\sqrt{\mathrm{tr}\{(P_p\Sigma_g P_p\Sigma_h)^2\}} \\
&\qquad\times\left(\sum_{\ell=1}^{a}\psi\delta_g\delta_\ell\boldsymbol{\mu}_\ell\right)^\top P_p\Sigma_g P_p\left(\sum_{\ell=1}^{a}\psi\delta_g\delta_\ell\boldsymbol{\mu}_\ell\right) \\
&\leq a(a-1)\sum_{g=2}^{a}\sum_{h=1}^{g-1}\frac{32\psi^2\delta_g^2\delta_h^2}{\sigma^4 n_g n_h}\sqrt{\mathrm{tr}\{(P_p\Sigma_g P_p\Sigma_h)^2\}} \\
&\qquad\times\boldsymbol{\mu}^\top(R_a\otimes P_p)\left\{\sum_{g=1}^{a}n_g^{-1}(\boldsymbol{e}_g\boldsymbol{e}_g^\top)\otimes\Sigma_g\right\}(R_a\otimes P_p) \\
&\leq a(a-1)\sum_{g=2}^{a}\sum_{h=1}^{g-1}\frac{\sqrt{\mathrm{tr}\{(P_p\Sigma_g P_p\Sigma_h)^2\}}}{\mathrm{tr}(P_p\Sigma_g P_p\Sigma_h)} \\
&= o(1).
\end{aligned}
$$

· **The evaluation of** $\mathrm{E}\left(\varepsilon_i^{(1)^4}\right)$

$$
\begin{aligned}
\sum_{i=n_{(g-1)}+1}^{n_{(g)}} \mathrm{E}\left(\varepsilon_i^{(1)^4}\right) &= \frac{16d_g^4}{\sigma^4 n_g^4(n_g-1)^4}\sum_{i=2}^{n_g}\mathrm{E}\left\{\left(\boldsymbol{Y}_{gi}^\top P_p \sum_{j=1}^{i-1}\sum_{k=1}^{i-1}\boldsymbol{Y}_{gj}\boldsymbol{Y}_{gk}^\top P_p \boldsymbol{Y}_{gi}\right)^2\right\} \\
&\leq \frac{48(\kappa_g+1)d_g^4}{\sigma^4 n_g^4(n_g-1)^4}\sum_{i=2}^{n_g}\mathrm{E}\left\{\left(\sum_{j=1}^{i-1}\sum_{k=1}^{i-1}\boldsymbol{Y}_{gk}^\top P_p\Sigma_g P_p\boldsymbol{Y}_{gj}\right)^2\right\} \\
&\leq \frac{48(\kappa_g+1)d_g^4}{\sigma^4 n_g^4(n_g-1)^4}\sum_{i=1}^{n_g}(i-1)\mathrm{E}\{(\boldsymbol{Y}_{g1}^\top P_p\Sigma_g P_p\boldsymbol{Y}_{g1})^2\} \\
&\quad +\frac{48(\kappa_g+1)d_g^4}{\sigma^4 n_g^4(n_g-1)^4}\sum_{i=1}^{n_g}(i-1)(i-2)[\mathrm{tr}\{(P_p\Sigma_g)^2\}]^2 \\
&\quad +\frac{96(\kappa_g+1)d_g^4}{\sigma^4 n_g^4(n_g-1)^4}\sum_{i=1}^{n_g}(i-1)(i-2)\mathrm{tr}\{(P_p\Sigma_g)^4\} \\
&\leq \frac{72(\kappa_g+1)^2 d_g^4}{\sigma^4 n_g^3(n_g-1)^3}[\mathrm{tr}\{(P_p\Sigma_g)^2\}]^2 \\
&\quad +\frac{16(\kappa_g+1)(n_g-2)d_g^4}{\sigma^4 n_g^3(n_g-1)^3}[\mathrm{tr}\{(P_p\Sigma_g)^2\}]^2 \\
&\quad +\frac{32(\kappa_g+1)(n_g-2)d_g^4}{\sigma^4 n_g^3(n_g-1)^3}\mathrm{tr}\{(P_p\Sigma_g)^4\} \\
&= \frac{18(\kappa_g+1)^2}{n_g(n_g-1)}+\frac{4(\kappa_g+1)(n_g-2)}{n_g(n_g-1)} \\
&\quad +\frac{8(\kappa_g+1)(n_g-2)}{n_g(n_g-1)} \\
&= O(n_g^{-1}).
\end{aligned}
$$

· **The evaluation of** $\mathrm{E}\left(\varepsilon_i^{(2)^4}\right)$

$$
\begin{aligned}
\sum_{i=n_{(g-1)}+1}^{n_{(g)}} \mathrm{E}\left(\varepsilon_i^{(2)^4}\right) &= \frac{16}{\sigma^4 n_g^3} \sum_{i=1}^{n_{(g)}} \mathrm{E}\left\{\left(\boldsymbol{Y}_{gi}^\top P_p \sum_{h=1}^{g-1} v_{gh} \overline{\boldsymbol{Y}}_h\right)^4\right\} \\
&\le \frac{48(\kappa_g+1)}{\sigma^4 n_g^3} \left\{\sum_{h=1}^{g-1} (\psi\delta_g\delta_h)^4 \mathrm{E}\{(\overline{\boldsymbol{Y}}'_h P_p \Sigma_g P_p \overline{\boldsymbol{Y}}_h)^2\}\right. \\
&\quad + \sum_{h\neq h'}^{g-1} (\psi\delta_g\delta_h)^2 (\psi\delta_g\delta_{h'})^2 \mathrm{E}(\overline{\boldsymbol{Y}}_h^\top P_p \Sigma_g P_p \overline{\boldsymbol{Y}}_h) \mathrm{E}(\overline{\boldsymbol{Y}}_{h'}^\top P_p \Sigma_g P_p \overline{\boldsymbol{Y}}_{h'}) \\
&\quad \left. +2 \sum_{h\neq h'}^{g-1} (\psi\delta_g\delta_h)^2 (\psi\delta_g\delta_{h'})^2 \mathrm{E}\{(\overline{\boldsymbol{Y}}_h^\top P_p \Sigma_g P_p \overline{\boldsymbol{Y}}_{h'})^2\}\right\} \\
&\le \sum_{h=1}^{g-1} \frac{144(\kappa_g+1)(\kappa_h+1)(\psi\delta_g\delta_h)^4}{\sigma^4 n_g^3 n_h^3} \{\mathrm{tr}(P_p\Sigma_g P_p \Sigma_h)\}^2 \\
&\quad + \sum_{h=1}^{g-1} \frac{48(\kappa_g+1)(n_h-1)(\psi\delta_g\delta_h)^4}{\sigma^4 n_g^3 n_h^3} \{\mathrm{tr}(P_p\Sigma_g P_p \Sigma_h)\}^2 \\
&\quad + \sum_{h\neq h'}^{g-1} \frac{48(\kappa_g+1)(\psi\delta_g\delta_h)^2(\psi\delta_g\delta_{h'})^2}{\sigma^4 n_g^3 n_h n_{h'}} \mathrm{tr}(P_p\Sigma_g P_p \Sigma_h) \mathrm{tr}(P_p\Sigma_g P_p \Sigma_{h'}) \\
&\quad + \sum_{h\neq h'}^{g-1} \frac{96(\kappa_g+1)(\psi\delta_g\delta_h)^2(\psi\delta_g\delta_{h'})^2}{\sigma^4 n_g^3 n_h n_{h'}} \mathrm{tr}(P_p\Sigma_g P_p \Sigma_h P_p \Sigma_g P_p \Sigma_{h'}) \\
&\le \sum_{h=1}^{g-1} \frac{9(\kappa_g+1)(\kappa_h+1)}{n_g n_h} + \sum_{h=1}^{g-1} \frac{3(\kappa_g+1)(n_h-1)}{n_g n_h} \\
&\quad + \sum_{h\neq h'}^{g-1} \frac{3(\kappa_g+1)}{n_g} + \sum_{h\neq h'}^{g-1} \frac{6(\kappa_g+1)}{n_g} \\
&= O(n_g^{-1}).
\end{aligned}
$$

· **The evaluation of** $\mathrm{E}\left(\varepsilon_i^{(3)^4}\right)$

$$
\begin{aligned}
\sum_{i=n_{(g-1)}+1}^{n_{(g)}} \mathrm{E}\left[\varepsilon_i^{(3)^4}\right] &= \frac{16}{\sigma^4 n_g^3}\mathrm{E}\left[\left\{\boldsymbol{Y}_{g1}^\top P_p \left(\sum_{h=1}^{a}\psi\delta_g\delta_h\boldsymbol{\mu}_h\right)\left(\sum_{h=1}^{a}\psi\delta_g\delta_h\boldsymbol{\mu}_h\right)^\top P_p \boldsymbol{Y}_{g1}\right\}^2\right] \\
&\leq \frac{48(\kappa_g+1)}{\sigma^4 n_g^3}\left\{\left(\sum_{h=1}^{a}\psi\delta_g\delta_h\boldsymbol{\mu}_h\right)^\top P_p\Sigma_g P_p\left(\sum_{h=1}^{a}\psi\delta_g\delta_h\boldsymbol{\mu}_h\right)\right\}^2 \\
&\leq \frac{3}{n_g} = O(n_g^{-1}).
\end{aligned}
$$