Testing Block-Diagonal Covariance Structure for High-Dimensional Data Under Non-normality

Yuki Yamada[†], Masashi Hyodo[††] and Takahiro Nishiyama[†††]

[†] *Department of Mathematical Information Science, Tokyo University of Science, 1-3, Kagurazaka, Shinjuku-ku, Tokyo, 162-8601, Japan, E-Mail:y.y.rfpfg.t.o.d.2006.s.m.e@gmail.com.*

[††] *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*

[†††] *Department of Business Administration, Senshu University, 2-1-1 Higashimita, Tama-ku, Kawasaki-shi, Kanagawa 214-8580, Japan. E-Mail: nishiyama@isc.senshu-u.ac.jp*

Abstract

In this article, we propose a test for making an inference about the block-diagonal covariance structure of a covariance matrix in non-normal high-dimensional data. We prove that the limiting null distribution of the proposed test is normal under mild conditions when its dimension is substantially larger than its sample size. We further study the local power of the proposed test. Finally, we study the finite sample performance of the proposed test via Monte Carlo simulations. We demonstrate the relevance and benefits of the proposed approach for a number of alternative covariance structures.

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

Key words:
Tests of covariance structure, high dimension, statistical hypothesis testing, non-normality.

1. Introduction

In recent years, a number of statistical methods for high-dimensional data such as DNA microarray gene expressions, where the number of feature variables, p, exceeds the sample size, N, have been discussed by many aouthors. For analysing high-dimensional data, the $p \times p$ covariance matrix Σ is an important measure for dependence among the components of a high-dimensional random vector $\boldsymbol{x}$, and thus plays a major role in many statistical procedures. Accordingly, the statistical inference of Σ is an important problem. Incidentally, a number of procedures for testing a high-dimensional covariance matrix, which include testing identity, sphericity and diagonal structure, have been proposed by Schott (2005), Srivastava (2005), Akita, Jin and Wakaki (2010), Srivastava and Reid (2012), and so on. All these tests assume either normality or moderate dimensionality such that $p/n \to c$ for a finite constant c or both assumptions. There has also been substantial research on the inference of hte covariance matrix under non-normal high-dimensional distributions: see, for example, Chen, Zhang and Zhong (2010), Srivastava, Kollo and Rosen (2011), Zhong and Chen (2011), Li and Chen (2012), Qiu and Chen (2012), Yata and Aoshima (2013), Srivastava, Yanagihara and Kubokawa (2014) and Himeno and Yamada (2014).

In this paper, we provide tests for covariance matrices without the normality assumption while allowing the dimension p to be much larger than the sample size N. Let $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_N$ $(N \geq 4)$ be independent and identically distributed $(i.i.d.)$ p-dimensional random vectors with $\mathrm{E}[\boldsymbol{x}_i] = \boldsymbol{\mu}$ and $\mathrm{Var}[\boldsymbol{x}_i] = \Sigma$. Now, we partition $\boldsymbol{x}_i$, $\boldsymbol{\mu}$, and Σ into q components:

$$\boldsymbol{x}_i = \begin{pmatrix} \boldsymbol{x}_i^{(1)} \\ \boldsymbol{x}_i^{(2)} \\ \vdots \\ \boldsymbol{x}_i^{(q)} \end{pmatrix}, \ \boldsymbol{\mu} = \begin{pmatrix} \boldsymbol{\mu}^{(1)} \\ \boldsymbol{\mu}^{(2)} \\ \vdots \\ \boldsymbol{\mu}^{(q)} \end{pmatrix}, \ \Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} & \cdots & \Sigma_{1q} \\ \Sigma_{21} & \Sigma_{22} & \cdots & \Sigma_{2q} \\ \vdots & \vdots & \ddots & \vdots \\ \Sigma_{q1} & \Sigma_{q2} & \cdots & \Sigma_{qq} \end{pmatrix},$$

where $\boldsymbol{x}_i^{(g)}$ and $\boldsymbol{\mu}^{(g)}$ are $p_g \times 1$ vectors and Σ_{gh} is a $p_g \times p_h$ matrix, $g, h = 1, \ldots, q$. Note that $p = \sum_{g=1}^{q} p_g$. Our interest is to test structure for covariance:

$$H_0 : \Sigma = \Sigma_d \ \text{ vs. } \ H_A : \Sigma \neq \Sigma_d, \tag{1.1}$$

where $\Sigma_d = \mathrm{diag}(\Sigma_{11}, \Sigma_{22}, \ldots, \Sigma_{qq})$.

When $p \gg N$, to verify the block-diagonal sturucture is very important; if H_0 holds, then the total number of unknown parameters to be estimated for Σ reduces to $\sum_{i=1}^{q} p_i(p_i+1)/2$ instead of $p(p+1)/2$. Further, it is interesting to note that the block-diagonal covariance structure attracts much attention across the fields of convex optimization and machine learning, with the main forcus being the Gaussian graph structure learning procedures.

For testing hypothesis (1.1), Hyodo et al. (2015) proposed a test statistic based on the normalized Frobenious matrix norm under normal assumption. They used an estimator of $\mathrm{tr}\Sigma^2$ proposed by Bai and Saranadasa (1996) and Srivastava (2005). Under normality assumption, this estimator has unbaisedness, but this estimator is not generally unbiased. Thus, in this paper, we propose a new test statistic based on an unbiased estimator proposed by Himeno and Yamada (2014) and Srivastava, Yanagihara and Kubokawa (2014) that does not need the normality assumption.

Hypothesis (1.1) includes some important hypotheses. For $p_g = 1$, that is, for a test of the covariance being diagonal ($H_0 : \Sigma = \mathrm{diag}(\sigma_1, \cdots, \sigma_p)$), Schott (2005) and Srivastava (2005) proposed an approximate test for a high-dimensional diagonal matrix under the multivariate normal assumption. Besides, Chen, Zhang and Zhong (2010) proposed approximate test without the normality assumption. Further, for $q = 2$, that is, for a test of the independence of two sub-vectors ($H_0 : \Sigma = \mathrm{diag}(\Sigma_{11}, \Sigma_{22})$), Srivastava and Reid (2012) proposed a test procedure under the multivariate normal assumption, while Yata and Aoshima (2015) proposed a test procedure without the normality assumption. Our testing procedure extends the results derived by Chen, Zhang and Zhong (2010) and Yata and Aoshima (2015).

The rest of the paper is organized as follows. Section 2 presents test procedure after establishing the asymptotic normality of the test statistics, and derives asymptotic powers in Section 3. Further, in Section 4, the attained significance levels and powers of the suggested test are empirically analyzed. Finally, Section 5 concludes this paper. Some technical details are relegated to the Appendix.

2. Test procedure

2.1. Assumptions

In this section, we mention the several assumptions to made construct a test for (1.1). Suppose $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_N$ are *i.i.d.* p-dimensional random vectors such that

$$\boldsymbol{x}_i = \Gamma \boldsymbol{z}_i + \boldsymbol{\mu} \quad \text{for } i = 1, \ldots, N, \tag{2.1}$$

where $\Gamma = (\Gamma^{(1)}, \Gamma^{(2)}, \cdots, \Gamma^{(q)})$ is a $p \times m$ constant matrix so that $\Gamma\Gamma' = \Sigma$, and $\boldsymbol{z}_1, \boldsymbol{z}_2, \ldots, \boldsymbol{z}_N$ are *i.i.d.* m-dimensional random vectors such that $\mathrm{E}[\boldsymbol{z}_i] = \mathbf{0}$ and $\mathrm{Var}[\boldsymbol{z}_i] = I_m$, an $m \times m$ identity matrix. Note that $\boldsymbol{x}_i^{(g)} = \Gamma^{(g)} \boldsymbol{z}_i + \boldsymbol{\mu}^{(g)}$.

Further, we use the following assumptions as necessary.

(A1) Let $\boldsymbol{z}_i = (z_{i1}, \ldots, z_{im})'$, with each z_{ij} having a uniformly bounded 4th moment, and there existing finite constants κ_3, κ_4 such that $\mathrm{E}[z_{ij}^3] = \kappa_3$, $\mathrm{E}[z_{ij}^4] = \kappa_4 + 3$ and for any positive integers r and α_l such that $\alpha_l \leq 4$, $\sum_{l=1}^r \alpha_l \leq 8$, $\mathrm{E}[z_{ij_1}^{\alpha_1} z_{ij_2}^{\alpha_2} \cdots z_{ij_r}^{\alpha_r}] = \mathrm{E}[z_{ij_1}^{\alpha_1}]\mathrm{E}[z_{ij_2}^{\alpha_2}] \cdots \mathrm{E}[z_{ij_r}^{\alpha_r}]$, whenever $j_1, j_2, \ldots, j_r$ are distinct indices.

(A1') Let $\boldsymbol{z}_i = (z_{i1}, \ldots, z_{im})'$, with each z_{ij} having a uniformly bounded 8th moment, and there existing finite constants κ_3, κ_4 such that $\mathrm{E}[z_{ij}^3] = \kappa_3$, $\mathrm{E}[z_{ij}^4] = \kappa_4 + 3$ and for any positive integers r and α_l such that $\sum_{l=1}^r \alpha_l \leq 8$, $\mathrm{E}[z_{ij_1}^{\alpha_1} z_{ij_2}^{\alpha_2} \cdots z_{ij_r}^{\alpha_r}] = \mathrm{E}[z_{ij_1}^{\alpha_1}]\mathrm{E}[z_{ij_2}^{\alpha_2}] \cdots \mathrm{E}[z_{ij_r}^{\alpha_r}]$, whenever $j_1, j_2, \ldots, j_r$ are distinct indices.

(A2) We assume one of the following asymptotic frameworks:

(i) $N \to \infty$, q is fixed and at least one of $p_1, \ldots, p_q$ goes to infinity.

(ii) $N \to \infty$, $q \to \infty$ and $p_1, \ldots, p_q$ are fixed.

(A3) Let $\tau = \sum_{\substack{g,h=1 \\ g \neq h}}^{q} \mathrm{tr}\Sigma_{gg}^2 \mathrm{tr}\Sigma_{hh}^2$. Then under (A2),

$$\tau^{-1} \min \left\{ \mathrm{tr}\Sigma^4, \ \left(\sum_{\substack{g,h=1 \\ g \neq h}}^{q} (\mathrm{tr}\Sigma_{gg}^4)^{1/4} (\mathrm{tr}\Sigma_{hh}^4)^{1/4} \right)^2 \right\} = o(1).$$

(A4) Under (A2),

$$\tau^{-1} \sum_{\substack{g,h,g',h'=1 \\ g \neq g' \neq h' \neq h \neq g}}^{q} \mathrm{tr}(\Sigma_{gh}\Sigma_{hg})\mathrm{tr}(\Sigma_{g'h'}\Sigma_{h'g'}) = o(1), \ \ \tau^{-1} N \mathrm{tr}\left(\Sigma^2 - \Sigma_d^2\right)^2 = o(1),$$

$$\tau^{-1} \sum_{\substack{g,g',h'=1 \\ g \neq g' \neq h' \neq g}}^{q} \mathrm{tr}(\Sigma_{gg}^2)\mathrm{tr}(\Sigma_{g'h'}\Sigma_{h'g'}) = o(1).$$

(A5) Under (A2),

$$\frac{1}{N^2\{\mathrm{tr}(\Sigma - \Sigma_d)^2\}^2} \sum_{\substack{g,h,g',h'=1 \\ g \neq g', h \neq h'}}^{q} \mathrm{tr}(\Sigma_{gh}\Sigma_{hg})\mathrm{tr}(\Sigma_{g'h'}\Sigma_{h'g'}) = o(1).$$

We state some remarks on the conditions. In assumption (A1), the requirement of the common third and fourth moments of z_{ij}, is not essential and is purely for the sake of simpler notation. Our theory allows different third and fourth moments as long as they are uniformly bounded, which is assured by z_{ij} having a uniformly bounded fourth moment. We sometimes assume (A1') instead of (A1) to investigate the power properties of the test.

Assumption (A2)-(i) means that number of blocks is small but the number of dimensions is large. Conversely, assumption (A2)-(ii) means that the number of blocks is small, but the matrix is large-dimensional.

We discuss some specific examples of assumptions (A3) and (A4). Assumption (A3) is used to derive the asymptotic null distribution, and assumption (A4) is used to derive the asymptotic non-null distribution. Let us consider the case $q = 2$. If $q = 2$, H_0 means the hypothesis that testing the correlation coefficient matrix between $\boldsymbol{x}_i^{(1)}$ and $\boldsymbol{x}_i^{(2)}$. For this hypothesis, Yata and Aoshima (2015) proposed a test statistic using the extended cross-data-matrix methodology and showed the asymptotic normality of this statistic under assumptions (A1),

(C1) $\min_{g=1,2}\left\{\frac{\mathrm{tr}\Sigma_{gg}^4}{(\mathrm{tr}\Sigma_{gg}^2)^2}\right\} \to 0$ as $p \to \infty$,

(C2) $\limsup_{m'\to\infty}\left\{\frac{N^2(\mathrm{tr}\Sigma_{12}\Sigma_{21})^2}{\mathrm{tr}\Sigma_{11}^2\mathrm{tr}\Sigma_{22}^2}\right\} < \infty$,

where $m' = \min\{N, p\}$. Note that if (C1) holds,

$$\frac{\min\left\{\mathrm{tr}\Sigma^4,\ 2(\mathrm{tr}\Sigma_{11}^4)^{1/2}(\mathrm{tr}\Sigma_{22}^4)^{1/2}\right\}}{\tau} \le \frac{2(\mathrm{tr}\Sigma_{11}^4)^{1/2}(\mathrm{tr}\Sigma_{22}^4)^{1/2}}{\tau} \to 0$$

as $p \to \infty$. We also note that if (C1) and (C2) hold,

$$\frac{N\mathrm{tr}\left(\Sigma^2 - \Sigma_d^2\right)^2}{\tau} \le \frac{4N\lambda_{\max}(\Sigma_{11})\lambda_{\max}(\Sigma_{22})\mathrm{tr}\Sigma_{12}\Sigma_{21}}{\mathrm{tr}\Sigma_{11}^2\mathrm{tr}\Sigma_{22}^2} \to 0$$

as $p \to \infty$. Here, $\lambda_{\max}(A)$ denotes the largest eigenvalue of matrix A. From these results, (A3) and (A4) quite similar to but somewhat weaker than conditions (C1) and (C2) formulated by Yata and Aoshima (2015). Thus, we can compare our statistics with those proposed by Yata and Aoshima (2015) under assumptions (A1), (C1), and (C2).

We now consider specific types of the alternative covariance structure under $p_1 = p_2 = \cdots = p_q$. The following covariance structure is called the "blocked compound symmetric (BCS)" structure:

$$\Sigma = I_q \otimes (\Sigma_* - R) + \mathbf{1}_q \mathbf{1}_q' \otimes R,$$

where Σ_* is a $p_1 \times p_1$ positive definite symmetric matrix, and R is a $p_1 \times p_1$ symmetric matrix. Here, $\mathbf{1}_q$ denotes a q-dimensional vector with all elements being one. If $q = p$, $\Sigma_* = \sigma_*$ and $R = \sigma_* \rho$, where $-1/(p-1) \leq \rho \leq 1$, $\Sigma = \sigma_* \{(1-\rho) I_p + \rho \mathbf{1}_p \mathbf{1}_p'\}$. This covariance structure is called the intraclass correlation structure. Note that

$$\begin{aligned}
\tau &= q(q-1)(\mathrm{tr}\Sigma_*^2)^2, \\
\left(\sum_{\substack{g,h=1 \\ g \neq h}}^{q} (\mathrm{tr}\Sigma_{gg}^4)^{1/4} (\mathrm{tr}\Sigma_{hh}^4)^{1/4} \right)^2 &= q^2 (q-1)^2 \mathrm{tr}\Sigma_*^4, \\
N \mathrm{tr} \left(\Sigma^2 - \Sigma_d^2 \right)^2 &= N\{ q(q-1)\mathrm{tr}(R\Sigma_*)^2 + 2q(q-1)(q-2)\mathrm{tr}(R^3 \Sigma_*) \\
&\quad + q(q-1)(q^2 - 3q + 3)\mathrm{tr}R^4 \}.
\end{aligned}$$

Using the results for the correlation matrix in Yata and Aoshima (2015), we get

$$\begin{aligned}
&\mathrm{tr}(R\Sigma_*)^2 \leq (\lambda_{\max}(\Sigma_*))^2 \mathrm{tr}R^2, \\
&\mathrm{tr}R^3 \Sigma_* \leq \sqrt{\mathrm{tr}(R\Sigma_*)^2} \sqrt{\mathrm{tr}R^4} \leq (\lambda_{\max}(\Sigma_*))^2 \mathrm{tr}R^2, \\
&\mathrm{tr}R^4 \leq (\lambda_{\max}(\Sigma_*))^2 \mathrm{tr}R^2.
\end{aligned}$$

Suppose that

(A3') $\frac{\mathrm{tr}\Sigma_*^4}{(\mathrm{tr}\Sigma_*^2)^2} \to 0$ as $p \to \infty$,

(A4') $\limsup_{m' \to \infty} \left\{ \frac{N^2 (\mathrm{tr}R^2)^2}{(\mathrm{tr}\Sigma_*^2)^2} \right\} < \infty$.

If (A2)-(i), (A3'), and (A4') are assumed, (A3) and (A4) holds. Further, if we assume (A2)-(ii) and

$$r_{ij} \asymp \frac{1}{\sqrt{Nq}},$$

(A3) and (A4) hold. Here, $a \asymp b$ means that $a = O(b)$ and $b = O(a)$, and r_{ij} denotes each element of the correlation matrix R.

2.2. Test statistic

We note that H_0 is valid if and only if $\Sigma_{gh} = O$ for $g \neq h$, and the latter implies that $\mathrm{tr}(\Sigma - \Sigma_d)^2 = 0$. A test for the hypothesis H_0 vs. H_A can be based on the unbiased estimator of $\mathrm{tr}(\Sigma - \Sigma_d)^2 = \mathrm{tr}\Sigma^2 - \mathrm{tr}\Sigma_d^2$ and the same can be used to develop the test statistic. We assume (2.1), and let

$$\overline{\boldsymbol{x}} = \frac{1}{N}\sum_{i=1}^{N}\boldsymbol{x}_i,\ S = \frac{1}{N-1}\sum_{i=1}^{N}(\boldsymbol{x}_i - \overline{\boldsymbol{x}})(\boldsymbol{x}_i - \overline{\boldsymbol{x}})',\ Q = \frac{1}{N-1}\sum_{i=1}^{N}\left((\boldsymbol{x}_i - \overline{\boldsymbol{x}})'(\boldsymbol{x}_i - \overline{\boldsymbol{x}})\right)^2;$$

then, the unbiased estimator of $\mathrm{tr}\Sigma^2$ is

$$\widehat{\mathrm{tr}\Sigma^2} = \frac{N-1}{N(N-2)(N-3)}\{(N-1)(N-2)\mathrm{tr}S^2 + (\mathrm{tr}S)^2 - NQ\}.$$

This estimator has been proposed in Himeno and Yamada (2014) and Srivastava, Yanagihara and Kubokawa (2014). From these results, it is straightforward to propose the unbiased estimator of $\mathrm{tr}\Sigma_d^2$ as

$$\widehat{\mathrm{tr}\Sigma_d^2} = \frac{N-1}{N(N-2)(N-3)}\sum_{g=1}^{q}\{(N-1)(N-2)\mathrm{tr}S_g^2 + (\mathrm{tr}S_g)^2 - NQ_g\},$$

where

$$\overline{\boldsymbol{x}}^{(g)} = \frac{1}{N}\sum_{i=1}^{N}\boldsymbol{x}_i^{(g)},\ S_g = \frac{1}{N-1}\sum_{i=1}^{N}(\boldsymbol{x}_i^{(g)} - \overline{\boldsymbol{x}}^{(g)})(\boldsymbol{x}_i^{(g)} - \overline{\boldsymbol{x}}^{(g)})'$$

and

$$Q_g = \frac{1}{N-1}\sum_{i=1}^{N}\{(\boldsymbol{x}_i^{(g)} - \overline{\boldsymbol{x}}^{(g)})'(\boldsymbol{x}_i^{(g)} - \overline{\boldsymbol{x}}^{(g)})\}^2.$$

From these facts, we construct the test statistic as follows:

$$T = \widehat{\mathrm{tr}\Sigma^2} - \widehat{\mathrm{tr}\Sigma_d^2}$$

which is an unbiased estimator of $\mathrm{tr}(\Sigma - \Sigma_d)^2$ under (2.1). The following theorem shows that statistic T is the rate-consistent estimator of $\mathrm{tr}(\Sigma - \Sigma_d)^2$.

Theorem 2.1. *Under assumptions (A1), (A3), and (A5),*

$$\frac{T}{\mathrm{tr}(\Sigma - \Sigma_d)^2} \xrightarrow{p} 1.$$

Here, " $\xrightarrow{p}$ *" denotes convergence in probability.*

(Proof) The assertion follows immediately from Lemma A.3.

2.3. Asymptotic normality of the test statistic

We first introduce the asymptotic variance of T:

$$\begin{aligned}\sigma^2 &= \frac{4}{N(N-1)} \sum_{\substack{g,h,g',h'=1 \\ g\neq g', h\neq h'}}^{q} \mathrm{tr}(\Sigma_{gh}\Sigma_{hg})\mathrm{tr}(\Sigma_{g'h'}\Sigma_{h'g'}) + \frac{8}{N}\mathrm{tr}\left(\Sigma^2 - \Sigma_d^2\right)^2 \\ &\quad + \frac{4\kappa_4}{N}\mathrm{tr}\left(\{\Gamma'(\Sigma-\Sigma_d)\Gamma\} \odot \{\Gamma'(\Sigma-\Sigma_d)\Gamma\}\right),\end{aligned}$$

where $\odot$ denotes the Hadamard product, that is, $(A \odot B)_{ij} = (A)_{ij} \cdot (B)_{ij}$. (See Lemma A.3 for details.)

Remark 2.1. *Under* $H_0 : \Sigma = \Sigma_d$*,*

$$\sigma_{H_0}^2 = \frac{4}{N(N-1)} \sum_{\substack{g,h=1 \\ g\neq h}}^{q} \mathrm{tr}\Sigma_{gg}^2 \mathrm{tr}\Sigma_{hh}^2.$$

Note that for any symmetric matrix A, $\mathrm{tr}(A \odot A) \leq \mathrm{tr}A^2$. Hence,

$$\mathrm{tr}\left(\{\Gamma'(\Sigma-\Sigma_d)\Gamma\} \odot \{\Gamma'(\Sigma-\Sigma_d)\Gamma\}\right) \leq \mathrm{tr}\left(\Sigma^2 - \Sigma_d^2\right)^2.$$

Thus, under assumptions (A1) - (A4),

$$\frac{\sigma^2}{\sigma_{H_0}^2} \to 1.$$

The following theorems establish the asymptotic normality of T.

Theorem 2.2. *Under assumptions (A1) - (A4),*

$$\frac{T - \mathrm{tr}(\Sigma - \Sigma_d)^2}{\sigma_{H_0}} \xrightarrow{d} \mathcal{N}(0,1),$$

where $\xrightarrow{d}$ *denotes the convergence in distribution.*

(Proof) See Appendix A.2.

To test H_0, it is necessary to estimate $\sigma_{H_0}^2$. Using the unbiased estimator of $\mathrm{tr}\Sigma_{gg}^2$,

$$\widehat{\mathrm{tr}\Sigma_{gg}^2} = \frac{N-1}{N(N-2)(N-3)}\{(N-1)(N-2)\mathrm{tr}S_g^2 + (\mathrm{tr}S_g)^2 - NQ_g\},$$

we estimate $\sigma^2_{H_0}$ by

$$\widehat{\sigma}^2_{H_0} = \frac{4}{N(N-1)} \sum_{\substack{g,h=1 \\ g \neq h}}^{q} \widehat{\mathrm{tr}\Sigma^2_{gg}} \widehat{\mathrm{tr}\Sigma^2_{hh}}.$$

The following theorem shows that $\widehat{\sigma}^2_{H_0}$ is the ratio-consistent estimator of $\sigma^2_{H_0}$.

Theorem 2.3. *Under H_0 and assumptions (A1) and (A2),*

$$\frac{\widehat{\sigma}^2_{H_0}}{\sigma^2_{H_0}} \xrightarrow{p} 1.$$

(Proof) See Appendix A.3.

Applying Theorems 2.2 and 2.3, under H_0,

$$\frac{T}{\widehat{\sigma}_{H_0}} \xrightarrow{d} \mathcal{N}(0,1).$$

Hence, we propose a test based on the test statistic T and reject H_0 if $T > \widehat{\sigma}_{H_0} z_\alpha$, where z_α is the upper critical value of the standard normal distribution at significance level α.

3. Asymptotic power

In this section, we discuss the asymptotic power of our statistic T. We define the power of the test under $H_A : \Sigma \neq \Sigma_d$ as

$$\mathrm{Power}_{YHN} = \Pr\left(T > \widehat{\sigma}_{H_0} z_\alpha | \Sigma \neq \Sigma_d\right).$$

First, we compare the asymptotic power of our statistic and the asymptotic power of Yata and Aoshima's statistic.

From Theorems 2.2 and 2.3, we obtain the following corollary.

Corollary 3.1. *We assume $q = 2$, (C1), and (C2). Then,*

$$\frac{T - \mathrm{tr}(\Sigma - \Sigma_d)^2}{\widehat{\sigma}_{H_0}} \xrightarrow{d} \mathcal{N}(0,1).$$

Remark 3.1. *Let*

$$\delta = \frac{\mathrm{tr}(\Sigma_{12}\Sigma_{21})}{\sqrt{2/\{N(N-1)\}\mathrm{tr}\Sigma_{11}^2\mathrm{tr}\Sigma_{22}^2}}.$$

From Corollary 3.1, under $q=2$, (A1), (A2), (C1), and (C2),

$$\mathrm{Power}_{YHN} \to \begin{cases} 1-\Phi\left(z_\alpha-\delta_0\right) & if \quad \delta \to \delta_0 < \infty, \\ 1 \quad if \quad \delta \to \infty. \end{cases}$$

From the above results and Theorem 4.1 in Yata and Aoshima (2015), the power of statistic T when $q=2$ is equivalent to that of Yata and Aoshima's statistic asymptotically.

Second, we derive the asymptotic power of our test under a specific type of alternative covariance structure, the BSC structure. Let

$$\tilde{\delta} = \frac{\sqrt{q(q-1)N(N-1)}\mathrm{tr}R^2}{2\mathrm{tr}\Sigma_*^2}.$$

Using Theorem 2.2, under (A1), (A2)-(i), (A3'), and (A4'); or (A1), (A2)-(ii), and $r_{ij} \asymp 1/\sqrt{nq}$,

$$\frac{T-q(q-1)\mathrm{tr}R^2}{\widehat{\sigma}_{H_0}} \xrightarrow{d} \mathcal{N}(0,1).$$

From the above results, we obtain the asymptotic power of our statistic as follows:

$$\mathrm{Power}_{YHN} \to \begin{cases} 1-\Phi\left(z_\alpha-\tilde{\delta}_0\right) & if \quad \tilde{\delta} \to \tilde{\delta}_0 < \infty, \\ 1 \quad if \quad \tilde{\delta} \to \infty. \end{cases}$$

Finally, we derive the asymptotic distribution of T when we cannot assume (A4). Moreover, we consider the asymptotic power of our test under $H_A : \Sigma \neq \Sigma_d$. To derive the asymptotic distribution of T, we assume assumption (A1') instead of assumption (A1). In assumption (A1'), the requirement of common third and fourth moments of z_{ij} is not essential and is purely for the sake of simpler notation. Our theory allows different third and fourth moments as long as they are uniformly bounded, which is assured by z_{ij} having a uniformly bounded eighth moment. The following theorem establishes the asymptotic normality of T.

Theorem 3.1. *Under assumptions (A1'), (A2), and (A3),*

$$\frac{T - \mathrm{tr}(\Sigma - \Sigma_d)^2}{\sigma} \xrightarrow{d} \mathcal{N}(0,1).$$

(Proof) See Appendix A.4.

We define

$$\nu = \frac{\mathrm{tr}(\Sigma - \Sigma_d)^2}{\sigma},$$

which may be viewed as a signal-to-noise ratio for the testing problem. This is because $\mathrm{tr}(\Sigma - \Sigma_d)^2$ is the square of the Frobenius norm of the difference between Σ and Σ_d. The following corollary expresses the asymptotic power of the test more clearly.

Corollary 3.2. *Under assumptions (A1'),(A2), and (A3),*

$$\mathrm{Power}_{YHN} \to \begin{cases} 1 - \Phi\left(\frac{\sigma_{H_0}}{\sigma} z_\alpha - \nu_0\right) & if \quad \nu \to \nu_0 < \infty, \\ 1 \quad if \quad \nu \to \infty. \end{cases}$$

Thus, if the difference between Σ and Σ_d is not too small in that $\mathrm{tr}(\Sigma - \Sigma_d)^2$ is of the same order as σ or of a larger order, the test will be powerful. Conversely, if the difference between Σ and Σ_d is so small that $\mathrm{tr}(\Sigma - \Sigma_d)^2$ is of a smaller order than σ, the test will not be powerful and cannot distinguish H_0 from H_A.

4. Simulation studies

We now investigate the effectiveness of the proposed test statistic by numerical studies.

We first provide the numerical evaluation by simulating the attained significant level (ASL), or size of the test. ASL is defined by

$$\mathrm{ASL}(\alpha) = \frac{\sharp\left(T_0 > \widehat{\sigma}_{H_0} z_\alpha\right)}{r},$$

where T_0 is the value of the test statistic T calculated from N independent observations under the null hypothesis H_0, r is the number of replications, and z_α is the upper 100α percentiles of the standard normal distribution. In our simulation studies, we set $N = 10, 50, 100$ and $r = 10000$. We evaluate ASL under the diagonal structure:

$$\Sigma = \text{diag}\left(\left\{0.5 + \frac{1}{p+1}\right\}^{\frac{1}{2}}, \cdots, \left\{0.5 + \frac{p}{p+1}\right\}^{\frac{1}{2}}\right),$$

for $p = 100, 200, 400, 1000$, and under the block diagonal structure:

$$\Sigma = I_q \otimes \Sigma^*(\equiv \Sigma_q)\ (q = 2, 5).$$

Here,

$$\Sigma^* = B(0.3^{|i-j|^{1/3}})B = B \begin{pmatrix} 1 & 0.3^{|1-2|^{1/3}} & \cdots & 0.3^{|1-d|^{1/3}} \\ 0.3^{|2-1|^{1/3}} & 1 & \cdots & 0.3^{|2-d|^{1/3}} \\ & \cdots & & \\ 0.3^{|d-1|^{1/3}} & 0.3^{|d-2|^{1/3}} & \cdots & 1 \end{pmatrix} B,$$

where $B = \text{diag}(\{0.5 + 1/(d+1)\}^{\frac{1}{2}}, \cdots, \{0.5 + d/(d+1)\}^{\frac{1}{2}})$. Through simulations, for $j = 1, \ldots, N$, $\boldsymbol{z}_j = (z_{ij})$ in assumption (A1) emerges for the following cases:

$$\begin{array}{ll} \text{(Case1)} & \boldsymbol{z}_j \sim \mathcal{N}_p(0, I_p), \\ \text{(Case2)} & z_{ij} = (u_{ij} - 1)/\sqrt{2} \text{ for } u_{ij} \sim \chi_1^2, \\ \text{(Case3)} & z_{ij} \sim SN(1), \\ \text{(Case4)} & z_{ij} = u_{ij}/\sqrt{3} \text{ for } u_{ij} \sim t_3, \\ \text{(Case5)} & z_{ij} = u_{ij}/\sqrt{5} \text{ for } u_{ij} \sim t_5. \end{array}$$

Further, we set $\alpha = 0.05$ and

$$(q, d) = (2, 50),\ (2, 100),\ (2, 200),\ (2, 500),\ (5, 20),\ (5, 40),\ (5, 80),\ (5, 200).$$

Tables 1-3 give the results for ASL. Overall, it may be noted that even if power is low for $N = 10$, when the values of p and N become large, approximation accuracy becomes good. The procedures proposed by Srivasta (2005) and Srivasta and Reid (2012) need the assumption of normality. Thus, these procedures are not helpful for Cases 2-5. However, we note that the accuracy of approximation for our procedure is almost the same for Cases 1-5. Further, because Yata and Aoshima (2015) provide a procedure under

HDLSS framework, for this simulation setting, we observe that our procedure is better than that of Yata and Aoshima (2015).

Further, we compute empirical power (EP) under certain alternative covariance structure. EP is defined as

$$\mathrm{EP}(\alpha) = \frac{\sharp\left(T_1 > \widehat{\sigma}_{H_0} z_\alpha | \Sigma \neq \Sigma_d\right)}{r},$$

where T_1 is the value of test statistic T calculated from N independent observations under the alternative hypothesis H_1. When $H_0 : \Sigma$ is a diagonal covariance, we set the alternative hypothesis as follows:

$$\text{(i) } H_1 : \Sigma = (\sigma_{ij}),\ \sigma_{ij} = \delta_{ij} \left\{0.5 + \frac{i}{p+1}\right\}^{\frac{1}{2}} + 0.03 \times I(|i-j| \leq 20),$$

for $i, j = 1, 2, \ldots, p$, and

$$\text{(ii) } H_1 : \Sigma = \mathrm{diag}\left(\left\{0.5 + \frac{1}{p+1}\right\}^{\frac{1}{2}}, \cdots, \left\{0.5 + \frac{p}{p+1}\right\}^{\frac{1}{2}}\right) + 0.03 \times (\mathbf{1}_p \mathbf{1}_p' - I_p).$$

When $H_0 : \Sigma = \Sigma_q$ $(q = 2, 5)$, we set the following alternative hypothesis:

$$\text{(i) } H_1 \ : \ \Sigma = I_q \otimes (\Sigma^* - 0.03 \times \mathbf{1}_d \mathbf{1}_d') + (\mathbf{1}_q \mathbf{1}_q') \otimes (0.03 \times \mathbf{1}_d \mathbf{1}_d'),$$

and

$$\text{(ii) } H_1 \ : \ \Sigma = (I(|i-j| \leq 1)) \otimes (0.03 \times \mathbf{1}_d \mathbf{1}_d') + I_q \otimes (\Sigma^* - (0.03 \times \mathbf{1}_d \mathbf{1}_d')).$$

Tables 4-9 give the results for EP. As a whole, it may be noted when the values of p and N become large, power becomes high. However, for case (ii) and $q = 2$, power becomes low when $p = 1000$. Besides, for $q = 2$, we observe that the power of our procedure is almost the same as that of Yata and Aoshima (2015) and Srivasta and Reid (2012) for Case 1 and 3. Further, we note that the power for $q = 2$ is higher than that for $q = 5$ in the case (i), and the power for $q = 5$ is higher than that for $q = 2$ in case (ii). When $H_0 : \Sigma$ is a diagonal covariance, we observe that the power of our procedure is almost the same as that of Srivasta (2005) for Case 1 and 3. The proposed test are slightly more powerful than Srivasta (2005) in Case $2, 4$ and 5.

5. Concluding remarks

In this paper, we discuss the test for a block-diagonal covariance structure of the covariance matrix in non-normal and high-dimensional settings. This hypothesis includes some important hypotheses, for example, when each block size is one, that is, in a test for the covariance being diagonal, and when there are two blocks, that is, in a test for the independence of two subvectors. Further, the validity of commonly used high-dimensional statistical procedures including diagonal (or block diagonal) linear classifiers requires the assumption of diagonal (or block diagonal) covariance matrices.

A test for the block-diagonal covariance structure of the covariance matrix usually used the likelihood ratio test. Akita, Jin and Wakaki (2010) derived the Edgeworth expansion of a likelihood ratio test statistic when both sample size and dimension are large, and simulate the goodness of fit of its approximation. However, the likelihood ratio cannot use $p > n$. Motivated by an unbiased estimator of the Frobenius norm of $\Sigma - \Sigma_d$, Hyodo et al. (2015) proposed an approximate test under a multivariate normal population. The difference between our study and theirs lies in the population distribution. Under the non-normal population, we proposed an approximate test based on the asymptotic normality of the unbiased estimator of $\Sigma - \Sigma_d$ that does not need the normality assumption, and also studied its asymptotic power. The type-I error and power of our test have been investigated under some non-normal settings by simulation. Through the simulation results, we have confirmed that the approximate test is not bad in most cases except for in the case of $N = 10$.

However, some multivariate non-normal distributions will not satisfy assumption (A1) (or (A1')). For instance, elliptical distributions except for the multivariate normal model seldom satisfy assumption (A1) (or (A1')). For these reasons, it is interesting to investigate the robustness of our test under elliptical distributions. This is an important topic for future research that is of both theoretical and practical interest.

Acknowledgments.
The research of the third 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] T. Akita, J. Jin, H. Wakaki, High-dimensional Edgeworth expansion of a test statistic on independence and its error bound. *J. Multivariate Anal.*, **101** (2010) 1806–1813.

[2] Z. Bai, H. Saranadasa, Effect of high dimension: by an example of a two sample problem. *Statist. Sinica*, **6** (1996) 311–329.

[3] S. X. Chen, L. X. Zhang, P. S. Zhong, Tests for high-dimensional covariance matrices. *J. Amer. Statist. Assoc.*, **105** (2010) 810–819.

[4] 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.

[5] M. Hyodo, N. Shutoh, T. Nishiyama, T. Pavlenko, Testing block-diagonal covariance structure for high-dimensional data. *Stat. Neerl.*, **69** (2015) 460–482.

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

[7] Y. Qiu, S. X. Chen, Test for bandedness of high-dimensional covariance matrices and bandwidth estimation. *Ann. Statist.*, **40** (2012) 1285–1314.

[8] J. R. Schott, Testing for complete independence in high dimensions. *Biometrika*, **92** (2005) 951–956.

[9] M. S. Srivastava, Some tests concerning the covariance matrix in high-dimensional data. *J. Japan Statist. Soc.*, **35** (2005) 251–272.

[10] M. S. Srivastava, T. Kollo, D. von Rosen, Some tests for the covariance matrix with fewer observations than the dimension under non-normality. *J. Multivariate Anal.*, **102** (2011) 1090–1103.

[11] M. S. Srivastava, N. Reid, Testing the structure of the covariance matrix with fewer observations than the dimension. *J. Multivariate Anal.*, **112** (2012) 156–171.

[12] 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.

[13] K. Yata, M. Aoshima, High-dimensional inference on covariance structures via the extended cross-data-matrix methodology. *Submitted (arXiv:1503.06492)*, (2015).

[14] K. Yata, M. Aoshima, Correlation tests for high-dimensional data using extended cross-data-matrix methodology. *J. Multivariate Anal.*, **117** (2013) 313–331.

[15] P. S. Zhong, S. X. Chen, Tests for high-dimensional regression coefficients with factorial designs. *J. Amer. Statist. Assoc.*, **106** (2011) 260–274.

A. Appendix

A.1. Preliminary results

Lemma A.1. *Let A, B, C and D are $m \times m$ symmetric matrices and E, F are $m \times m$ matrices. Under assumption (A1), we have:*

$$
\begin{aligned}
&\text{(i)} \quad \mathrm{E}[\boldsymbol{z}_1' E \boldsymbol{z}_1 \boldsymbol{z}_1' F \boldsymbol{z}_1] = \kappa_4 \mathrm{tr}(E \odot F) + \mathrm{tr}E\mathrm{tr}F + \mathrm{tr}(EF) + \mathrm{tr}(EF'),\\
&\text{(ii)} \quad \mathrm{E}[\boldsymbol{z}_1' A \boldsymbol{z}_2 \boldsymbol{z}_1' B \boldsymbol{z}_2 \boldsymbol{z}_1' E \boldsymbol{z}_2] = \kappa_3^2 \mathrm{tr}\left((A \odot B)E'\right),\\
&\text{(iii)} \quad \mathrm{E}[\boldsymbol{z}_1' A \boldsymbol{z}_2 \boldsymbol{z}_1' B \boldsymbol{z}_2 \boldsymbol{z}_1' C \boldsymbol{z}_2 \boldsymbol{z}_1' D \boldsymbol{z}_2] = \kappa_4^2 \mathrm{tr}\left((A \odot B)(C \odot D)\right)\\
&\qquad + 2\kappa_4 \mathrm{tr}\left((AB) \odot (CD)\right)\\
&\qquad + 2\kappa_4 \mathrm{tr}\left((AC) \odot (BD)\right)\\
&\qquad + 2\kappa_4 \mathrm{tr}\left((AD) \odot (BC)\right)\\
&\qquad + \mathrm{tr}(AB)\mathrm{tr}(CD) + \mathrm{tr}(AC)\mathrm{tr}(BD)\\
&\qquad + \mathrm{tr}(AD)\mathrm{tr}(BC) + \mathrm{tr}(ABCD)\\
&\qquad + \mathrm{tr}(ABDC) + \mathrm{tr}(ACBD)\\
&\qquad + \mathrm{tr}(ACDB) + \mathrm{tr}(ADBC)\\
&\qquad + \mathrm{tr}(ADCB),\\
&\text{(iv)} \quad \mathrm{E}\left[\left(\sum_{i \neq j}^{m} a_{ij} z_{1i} z_{1j}\right)^4\right] \leq \mathit{const.} \times \left(\sum_{i,j=1}^{m} a_{ij}^2\right)^2.
\end{aligned}
$$

(Proof) See supplemental material.

Lemma A.2. *Let $\Gamma^{(g)\prime}\Gamma^{(g)} = A^{(g)} = (a_{ij}^{(g)})$ and $\sum_{g=1}^{q} A^{(g)} = \Gamma'\Gamma = A = (a_{ij})$. It holds that*

$$
\begin{aligned}
&\text{(i)} \quad \mathrm{tr}\left(\sum_{g \neq h}^{q} A^{(g)} \odot A^{(h)}\right)^2\\
&\qquad \leq \min\left\{\left(\sum_{g \neq h}^{q} (\mathrm{tr}\Sigma_{gg}^4)^{1/4} (\mathrm{tr}\Sigma_{hh}^4)^{1/4}\right)^2,\ 2\mathrm{tr}\Sigma^4\right\},
\end{aligned}
$$

(ii) $$\sum_{\substack{g,h,g',h'=1\\ g\neq h, g'\neq h'}}^{q} \mathrm{tr}(A^{(g)}A^{(g')}A^{(h)}A^{(h')})$$
$$\leq \min\left\{\left(\sum_{g\neq h}^{q}(\mathrm{tr}\Sigma_{gg}^4)^{1/4}(\mathrm{tr}\Sigma_{hh}^4)^{1/4}\right)^2,\ 2\mathrm{tr}\Sigma^4\right\},$$

(iii) $$\sum_{\substack{g,h,g',h'=1\\ g\neq g', h\neq h'}}^{q} \mathrm{tr}\left(\left(A^{(g)}A^{(h)}\right)\odot\left(A^{(g')}A^{(h')}\right)\right)$$
$$\leq \min\left\{\left(\sum_{g\neq h}^{q}(\mathrm{tr}\Sigma_{gg}^4)^{1/4}(\mathrm{tr}\Sigma_{hh}^4)^{1/4}\right)^2,\ 2\mathrm{tr}\Sigma^4\right\},$$

(iv) $$\left|\sum_{\substack{g,h,g',h'=1\\ g\neq h, g'\neq h'}}^{q} \mathrm{tr}\left((A^{(g)}\odot A^{(g')})A^{(h')}A^{(h)}\right)\right|$$
$$\leq \min\left\{\left(\sum_{g\neq h}^{q}(\mathrm{tr}\Sigma_{gg}^4)^{1/4}(\mathrm{tr}\Sigma_{hh}^4)^{1/4}\right)^2,\ 4\mathrm{tr}\Sigma^4\right\},$$

(v) $$\sum_{\substack{g_1\neq h_1, g_2\neq h_2\\ g_3\neq h_3, g_4\neq h_4}}^{q} \mathrm{tr}\left((A^{(g_1)}A^{(g_3)})\odot(A^{(h_1)}A^{(h_3)})\right)$$
$$\left((A^{(g_2)}A^{(g_4)})\odot(A^{(h_2)}A^{(h_4)})\right)$$
$$\leq \min\left\{\left(\sum_{g\neq h}^{q}(\mathrm{tr}\Sigma_{gg}^4)^{1/4}(\mathrm{tr}\Sigma_{hh}^4)^{1/4}\right)^2,\ 18\mathrm{tr}\Sigma^4\right\},$$

(vi) $$\sum_{\substack{g_1\neq h_1, g_2\neq h_2\\ g_3\neq h_3, g_4\neq h_4}}^{q} \mathrm{tr}(A^{(g_1)}A^{(g_2)}A^{(g_3)}A^{(g_4)})$$
$$\times\, \mathrm{tr}(A^{(h_1)}A^{(h_2)}A^{(h_3)}A^{(h_4)})$$
$$\leq \min\left\{\left(\sum_{g\neq h}^{q}(\mathrm{tr}\Sigma_{gg}^4)^{1/4}(\mathrm{tr}\Sigma_{hh}^4)^{1/4}\right)^2,\ 18\mathrm{tr}\Sigma^4\right\},$$

$$\text{(vii)} \quad \mathrm{tr}\left(\sum_{g\neq h}^{q} A^{(g)}\Gamma'(\Sigma-\Sigma_d)\Gamma A^{(h)}\right)^2$$
$$\leq \min\left\{4\mathrm{tr}\Sigma^4, \left(\sum_{g\neq h}^{q}(\mathrm{tr}\Sigma_{gg}^4)^{1/4}(\mathrm{tr}\Sigma_{hh}^4)^{1/4}\right)^2\right\}\mathrm{tr}(\Sigma(\Sigma-\Sigma_d))^2,$$
$$\text{(viii)} \quad \sum_{\substack{g,h,g',h'=1\\ g\neq g',h\neq h'}}^{q} [\mathrm{diag}(\Gamma(\Sigma-\Sigma_d)\Gamma)]'\left\{\left(A^{(g)}A^{(h)}\right)\odot\left(A^{(g')}A^{(h')}\right)\right\}$$
$$[\mathrm{diag}(\Gamma(\Sigma-\Sigma_d)\Gamma)]$$
$$\leq \min\left\{2\mathrm{tr}\Sigma^4, \left(\sum_{g\neq h}^{q}(\mathrm{tr}\Sigma_{gg}^4)^{1/4}(\mathrm{tr}\Sigma_{hh}^4)^{1/4}\right)^2\right\}\mathrm{tr}(\Sigma(\Sigma-\Sigma_d))^2.$$

Here, $[\mathrm{diag}(A)] = (a_{11}, a_{22}, \cdots, a_{mm})$ *for* $m\times m$ *matrix* $A = (a_{ij})$*.*

(Proof) See supplemental material.

Lemma A.3. *Under assumptions (A1)-(A3),*

$$\mathrm{E}[T] = \mathrm{tr}(\Sigma-\Sigma_d)^2, \quad \mathrm{Var}[T] = \sigma^2\left(1+o(1)\right).$$

(Proof) Let

$$U_1 = \frac{1}{N(N-1)}\sum_{\substack{g,h=1\\ g\neq h}}^{q}\sum_{\substack{i,j=1\\ i\neq j}}^{N} \boldsymbol{z}_i'A^{(g)}\boldsymbol{z}_j\boldsymbol{z}_i'A^{(h)}\boldsymbol{z}_j,$$
$$U_2 = \frac{1}{N(N-1)(N-2)}\sum_{\substack{g,h=1\\ g\neq h}}^{q}\sum_{\substack{i,j,k=1\\ i\neq j\neq k\neq i}}^{N} \boldsymbol{z}_i'A^{(g)}\boldsymbol{z}_j\boldsymbol{z}_i'A^{(h)}\boldsymbol{z}_k,$$
$$U_3 = \frac{1}{N(N-1)(N-2)(N-3)}\sum_{\substack{g,h=1\\ g\neq h}}^{q}\sum_{\substack{i,j,k,l=1\\ i\neq j\neq k\neq l\\ j\neq l\neq i\neq k}}^{N} \boldsymbol{z}_i'A^{(g)}\boldsymbol{z}_j\boldsymbol{z}_k'A^{(h)}\boldsymbol{z}_l.$$

Then T can be expressed as $T = U_1 - 2U_2 + U_3$. It is straightforward to show that $\mathrm{E}[T] = \mathrm{E}[U_1] = \mathrm{tr}(\Sigma-\Sigma_d)^2$. By using Lemma A.1, $\mathrm{Var}[U_1]$ is calculate as

$$\mathrm{Var}[U_1] = \sum_{i=1}^{6} V_i^{(1)},$$

where

$$
\begin{aligned}
V_1^{(1)} &= \frac{4}{N(N-1)} \sum_{\substack{g,h,g',h'=1 \\ g\neq g', h\neq h'}}^{q} \mathrm{tr}(A^{(g)}A^{(h)})\mathrm{tr}(A^{(g')}A^{(h')}) \\
&= \frac{4}{N(N-1)} \sum_{\substack{g,h,g',h'=1 \\ g\neq g', h\neq h'}}^{q} \mathrm{tr}(\Sigma_{gh}\Sigma_{hg})\mathrm{tr}(\Sigma_{g'h'}\Sigma_{h'g'}), \\
V_2^{(1)} &= \frac{8}{N}\mathrm{tr}\left(\sum_{g\neq h}^{q} A^{(g)}A^{(h)}\right)^2 \\
&= \frac{8}{N}\mathrm{tr}\left(\Sigma^2 - \Sigma_d^2\right)^2, \\
V_3^{(1)} &= \frac{4\kappa_4}{N}\mathrm{tr}\left(\left(\sum_{g\neq h}^{q} A^{(g)}A^{(h)}\right) \odot \left(\sum_{g\neq h}^{q} A^{(g)}A^{(h)}\right)\right) \\
&= \frac{4\kappa_4}{N}\mathrm{tr}\left(\Gamma(\Sigma-\Sigma_d)\Gamma \odot \Gamma(\Sigma-\Sigma_d)\Gamma\right), \\
V_4^{(1)} &= \frac{2\kappa_4^2}{N(N-1)}\mathrm{tr}\left(\sum_{g\neq h}^{q} A^{(g)} \odot A^{(h)}\right)^2, \\
V_5^{(1)} &= \frac{8\kappa_4}{N(N-1)} \sum_{\substack{g,h,g',h'=1 \\ g\neq g', h\neq h'}}^{q} \mathrm{tr}\left(\left(A^{(g)}A^{(h)}\right) \odot \left(A^{(g')}A^{(h')}\right)\right), \\
V_6^{(1)} &= \frac{4}{N(N-1)} \sum_{\substack{g,h,g',h'=1 \\ g\neq h, g'\neq h'}}^{q} \mathrm{tr}(A^{(g)}A^{(g')}A^{(h)}A^{(h')}).
\end{aligned}
$$

From Lemma A.2, $V_4^{(1)}$, $V_5^{(1)}$ and $V_6^{(1)}$ are negligible under (A1)-(A3). Hence,

$$
\frac{\mathrm{Var}[U_1]}{\sigma^2} = \frac{\sum_{t=1}^{3} V_t^{(1)}}{\sigma^2} + o(1) \tag{A.1}
$$

under (A1)-(A3). From (A.1) and $\sum_{t=1}^{3} V_t^{(1)} = \sigma^2$, we have

$$
\frac{\mathrm{Var}[U_1]}{\sigma^2} \to 1. \tag{A.2}
$$

By using Lemma A.1, $\mathrm{Var}[U_2]$ is calculate as

$$\mathrm{Var}[U_2] = \sum_{i=1}^{5} V_i^{(2)},$$

where

$$\begin{aligned}
V_1^{(2)} &= \frac{2}{N(N-1)(N-2)} \sum_{\substack{g,h,g',h'=1 \\ g\neq g', h\neq h'}}^{q} \mathrm{tr}(A^{(g)}A^{(h)})\mathrm{tr}(A^{(g')}A^{(h')}) \\
&= \frac{2}{N(N-1)(N-2)} \sum_{\substack{g,h,g',h'=1 \\ g\neq g', h\neq h'}}^{q} \mathrm{tr}(\Sigma_{gh}\Sigma_{hg})\mathrm{tr}(\Sigma_{g'h'}\Sigma_{h'g'}), \\
V_2^{(2)} &= \frac{2}{N(N-1)} \mathrm{tr}\left(\sum_{g\neq h}^{q} A^{(g)}A^{(h)}\right)^2 \\
&= \frac{2}{N(N-1)} \mathrm{tr}\left(\Sigma^2 - \Sigma_d^2\right)^2, \\
V_3^{(2)} &= \frac{2\kappa_4}{N(N-1)(N-2)} \sum_{\substack{g,h,g',h'=1 \\ g\neq g', h\neq h'}}^{q} \mathrm{tr}\left(\left(A^{(g)}A^{(h)}\right) \odot \left(A^{(g')}A^{(h')}\right)\right), \\
V_4^{(2)} &= \frac{4\kappa_3^2}{N(N-1)(N-2)} \sum_{\substack{g,h,g',h'=1 \\ g\neq h, g'\neq h'}}^{q} \mathrm{tr}\left((A^{(g)} \odot A^{(g')})A^{(h')}A^{(h)}\right), \\
V_5^{(2)} &= \frac{2}{N(N-1)(N-2)} \sum_{\substack{g,h,g',h'=1 \\ g\neq h, g'\neq h'}}^{q} \mathrm{tr}(A^{(g)}A^{(g')}A^{(h)}A^{(h')}).
\end{aligned}$$

From Lemma A.2, under (A1)-(A3),

$$\frac{V_3^{(2)}}{\sigma^2} \to 0, \; \frac{V_4^{(2)}}{\sigma^2} \to 0, \; \frac{V_5^{(2)}}{\sigma^2} \to 0, \tag{A.3}$$

and

$$\frac{V_1^{(2)}}{V_1^{(1)}} \to 0, \; \frac{V_2^{(2)}}{V_2^{(1)}} \to 0 \tag{A.4}$$

as $N \to \infty$. From (A.3) and (A.4), under (A1)-(A3),

$$\frac{\mathrm{Var}[U_2]}{\sigma^2} = \frac{\sum_{t=1}^{5} V_t^{(2)}}{\sigma^2} \to 0. \tag{A.5}$$

By using Lemma A.1, $\mathrm{Var}[U_3]$ is calculate as

$$\begin{aligned}
\mathrm{Var}[U_3] &= \frac{16}{N^2(N-1)^2(N-2)^2(N-3)^2} \\
&\quad \times \sum_{\substack{g_1,h_1,g_2,h_2=1 \\ g_1\neq h_1, g_2\neq h_2}}^{q} \sum_{\substack{i,j,k,l=1 \\ i\neq j\neq k\neq l \\ j\neq l\neq i\neq k}}^{N} \mathrm{E}[\boldsymbol{z}_i' A^{(g_1)} \boldsymbol{z}_j \boldsymbol{z}_k' A^{(h_1)} \boldsymbol{z}_l \boldsymbol{z}_i' A^{(g_2)} \boldsymbol{z}_k \boldsymbol{z}_j' A^{(h_2)} \boldsymbol{z}_l] \\
&\quad + \frac{8}{N^2(N-1)^2(N-2)^2(N-3)^2} \\
&\quad \times \sum_{\substack{g_1,h_1,g_2,h_2=1 \\ g_1\neq h_1, g_2\neq h_2}}^{q} \sum_{\substack{i,j,k,l=1 \\ i\neq j\neq k\neq l \\ j\neq l\neq i\neq k}}^{N} \mathrm{E}[\boldsymbol{z}_i' A^{(g_1)} \boldsymbol{z}_j \boldsymbol{z}_k' A^{(h_1)} \boldsymbol{z}_l \boldsymbol{z}_i' A^{(g_2)} \boldsymbol{z}_j \boldsymbol{z}_k' A^{(h_2)} \boldsymbol{z}_l] \\
&= \frac{8}{N(N-1)(N-2)(N-3)} \sum_{\substack{g,h,g',h'=1 \\ g\neq h, g'\neq h'}}^{q} \Big(2\mathrm{tr}(A^{(g)} A^{(g')} A^{(h)} A^{(h')}) \\
&\quad + \mathrm{tr}(A^{(g)} A^{(g')}) \mathrm{tr}(A^{(h)} A^{(h')}) \Big).
\end{aligned}$$

From above result, under assumptions (A1)-(A3),

$$\frac{\mathrm{Var}[U_3]}{\sigma^2} \to 0. \tag{A.6}$$

Combining (A.2), (A.5) and (A.6), $\mathrm{Var}[T] = \sigma^2\{1 + o(1)\}$. □

Lemma A.4. *Under assumptions (A1)-(A4), it holds that*

$$T - \mathrm{tr}(\Sigma - \Sigma_d)^2 = \frac{1}{N(N-1)} \sum_{i\neq j}^{N} \sum_{k_1\neq k_2, \ell_1\neq \ell_2}^{m} \beta_{k_1\ell_1k_2\ell_2} z_{ik_1} z_{ik_2} z_{j\ell_1} z_{j\ell_2} + o_p(\sigma_{H_0}),$$

where

$$\beta_{k_1\ell_1k_2\ell_2} = \sum_{g\neq h}^{q} a_{k_1\ell_1}^{(g)} a_{k_2\ell_2}^{(h)}.$$

And under assumptions (A1)-(A3), it holds that

$$T - \mathrm{tr}(\Sigma - \Sigma_d)^2 = \frac{1}{N(N-1)} \sum_{i \neq j}^{N} \sum_{k_1 \neq k_2, \ell_1 \neq \ell_2}^{m} \beta_{k_1 \ell_1 k_2 \ell_2} z_{ik_1} z_{ik_2} z_{j\ell_1} z_{j\ell_2} + \frac{2}{N} \sum_{j=1}^{N} \left\{ \boldsymbol{z}_j' \Gamma' (\Sigma - \Sigma_d) \Gamma \boldsymbol{z}_j - \mathrm{tr}\left(\Sigma^2 - \Sigma_d^2\right) \right\} + o_p(\sigma).$$

(Proof) From (A.5) and (A.6), U_2 and U_3 are negligible under (A1)-(A3). So, we only consider U_1 in $T - \mathrm{tr}(\Sigma - \Sigma_d)^2$. Let

$$U_{11} = \frac{1}{N(N-1)} \sum_{i \neq j}^{N} \sum_{\substack{k_1, k_2, \ell_1, \ell_2 = 1 \\ k_1 \neq k_2, \ell_1 \neq \ell_2}}^{m} \beta_{k_1 \ell_1 k_2 \ell_2} z_{ik_1} z_{ik_2} z_{j\ell_1} z_{j\ell_2},$$

$$U_{12} = \frac{2}{N} \sum_{j=1}^{N} \sum_{k,\ell=1}^{m} \beta_{k\ell k\ell} (z_{j\ell}^2 - 1),$$

$$U_{13} = \frac{2}{N} \sum_{j=1}^{N} \sum_{\substack{k, \ell_1, \ell_2 = 1 \\ \ell_1 \neq \ell_2}}^{m} \beta_{k\ell_1 k\ell_2} z_{j\ell_1} z_{j\ell_2},$$

$$U_{14} = \frac{1}{N(N-1)} \sum_{i \neq j}^{N} \sum_{k,\ell=1}^{m} \beta_{k\ell k\ell} (z_{ik}^2 - 1)(z_{j\ell}^2 - 1),$$

$$U_{15} = \frac{2}{N(N-1)} \sum_{i \neq j}^{N} \sum_{\substack{k_1, k_2, \ell = 1 \\ k_1 \neq k_2}}^{m} \beta_{k_1 \ell k_2 \ell} z_{ik_1} z_{ik_2} (z_{j\ell}^2 - 1).$$

Then these random variables are uncorrelated, namely, $\mathrm{Cov}(U_{1t_1}, U_{1t_2}) = 0$ for $t_1 \neq t_2$, and U_1 can be expressed as

$$U_1 = \mathrm{tr}(\Sigma - \Sigma_d)^2 + \sum_{t=1}^{5} U_{1t}.$$

At first we show first statement in Lemma A.4. For the proof of first statement, we need to show that

$$\frac{\mathrm{Var}[\sum_{t=2}^{5} U_{1t}]}{\sigma_{H_0}^2} \to 0$$

under (A1)-(A4). The variance of U_{11} is caluclated as

$$\begin{aligned}\mathrm{Var}[U_{11}] &= \frac{2}{N(N-1)}\mathrm{E}\left[\left(\sum_{k_1\neq k_2,\ell_1\neq\ell_2}^{m}\beta_{k_1\ell_1k_2\ell_2}z_{ik_1}z_{ik_2}z_{j\ell_1}z_{j\ell_2}\right)^2\right] \\ &= \frac{4}{N(N-1)}\sum_{k_1\neq k_2,\ell_1\neq\ell_2}^{m}(\beta^2_{k_1\ell_1k_2\ell_2}+\beta_{k_1\ell_1k_2\ell_2}\beta_{k_2\ell_1k_1\ell_2})\end{aligned}$$

$$\begin{aligned}&= \frac{4}{N(N-1)}\left[\sum_{\substack{g,h,g',h'=1\\ g\neq g',h\neq h'}}^{q}\mathrm{tr}(A^{(g)}A^{(h)})\mathrm{tr}(A^{(g')}A^{(h')})\right. \\ &+\sum_{\substack{g,h,g',h'=1\\ g\neq g',h\neq h'}}^{q}\mathrm{tr}(A^{(g)}A^{(h)}A^{(g')}A^{(h')})+2\mathrm{tr}\left(\sum_{g\neq h}^{q}A^{(g)}\odot A^{(h)}\right)^2 \\ &\left.-4\sum_{\substack{g,h,g',h'=1\\ g\neq g',h\neq h'}}^{q}\mathrm{tr}\left(\left(A^{(g)}A^{(h)}\right)\odot\left(A^{(g')}A^{(h')}\right)\right)\right].\end{aligned}$$

Thus under (A1)-(A4),

$$\frac{\mathrm{Var}[U_{11}]}{V_1^{(1)}}\to 1. \tag{A.7}$$

We also note that under (A1)-(A4),

$$\frac{\mathrm{Var}[U_1]}{V_1^{(1)}}\to 1,\ \frac{V_1^{(1)}}{\sigma^2_{H_0}}\to 1. \tag{A.8}$$

From (A.7) and (A.8), under (A1)-(A4),

$$\frac{\mathrm{Var}[\sum_{t=2}^{5}U_{1t}]}{\sigma^2_{H_0}}=\frac{\mathrm{Var}[U_1]-\mathrm{Var}[U_{11}]}{V_1^{(1)}}\times\frac{V_1^{(1)}}{\sigma^2_{H_0}}\to 0.$$

This proves the first statement of the lemma.

Next we show first statement in Lemma A.4. For the proof of first statement, we need to show that

$$\frac{\mathrm{Var}[\sum_{t=4}^{5} U_{1t}]}{\sigma^2} \to 0$$

under (A1)-(A3). Note that

$$U_{12} + U_{13} = \frac{2}{N} \sum_{j=1}^{N} \left\{ \boldsymbol{z}_j' \Gamma' (\Sigma - \Sigma_d) \Gamma \boldsymbol{z}_j - \mathrm{tr}\left(\Sigma^2 - \Sigma_d^2\right) \right\}.$$

By using Lemma A.1, the variance of $U_{12} + U_{13}$ is caluclated as

$$\begin{aligned}
\mathrm{Var}[U_{12} + U_{13}] &= \frac{4}{N^2} \mathrm{Var}\left[\sum_{j=1}^{N} [\boldsymbol{z}_j' \{\Gamma'(\Sigma - \Sigma_d)\Gamma\} \boldsymbol{z}_j - \mathrm{tr}\left(\Sigma^2 - \Sigma_d^2\right)] \right] \\
&= \frac{4\kappa_4}{N} \mathrm{tr}[\{\Gamma'(\Sigma - \Sigma_d)\Gamma\} \odot \{\Gamma'(\Sigma - \Sigma_d)\Gamma\}] + \frac{8}{N} \mathrm{tr}\left(\Sigma^2 - \Sigma_d^2\right)^2.
\end{aligned}$$

Thus under (A1)-(A3),

$$\frac{\mathrm{Var}[\sum_{t=1}^{3} U_{1t}]}{\sigma^2} = \frac{\mathrm{Var}[U_{11}] + \mathrm{Var}[U_{12} + U_{13}]}{\sigma^2} \to 1. \tag{A.9}$$

We also note that under (A1)-(A3),

$$\frac{\mathrm{Var}[U_1]}{\sigma^2} \to 1. \tag{A.10}$$

From (A.9) and (A.10), under (A1)-(A3),

$$\frac{\mathrm{Var}[\sum_{t=4}^{5} U_{1t}]}{\sigma^2} = \frac{\mathrm{Var}[U_1] - \mathrm{Var}[\sum_{t=1}^{3} U_{1t}]}{\sigma^2} \to 0.$$

This proves the second statement of the lemma. □

A.2. Proof of Theorem 2.2.

From Lemma A.4, Under assumptions (A1)-(A4), the leading order term in $\mathrm{Var}[T]$ is contributed by U_{11}. Hence, we only need to show the asymptotic normality of U_{11}. Let

$$\varepsilon_j = \frac{2}{N(N-1)} \sum_{i=1}^{j-1} \sum_{k_1 \neq k_2, \ell_1 \neq \ell_2}^{m} \beta_{k_1 \ell_1 k_2 \ell_2} z_{ik_1} z_{ik_2} z_{j\ell_1} z_{j\ell_2}.$$

Then $U_{11} = \sum_{j=1}^{N} \varepsilon_j$. Define $\mathcal{F}_0 = \{\emptyset, \Omega\}$, $\mathcal{F}_j = \sigma\{\boldsymbol{z}_1, \boldsymbol{z}_2, \ldots, \boldsymbol{z}_j\}$ for $j = 1, \ldots, N$ be the σ-algebra generated by the sequence of random vectors $\boldsymbol{z}_1, \boldsymbol{z}_2, \ldots, \boldsymbol{z}_j$. Then it is straightforward to show that $\mathrm{E}[\varepsilon_j] = 0$ and $\mathrm{E}[\varepsilon_j|\mathcal{F}_{j-1}] = 0$. So ε_j is a martingale difference sequence. To apply martingale central limit theorem, we need to show that

$$\frac{\sum_{j=1}^{N} \mathrm{E}[\varepsilon_j^2|\mathcal{F}_{j-1}]}{\sigma_{H_0}^2} \xrightarrow{p} 1 \quad \text{and} \quad \frac{\sum_{j=1}^{N} \mathrm{E}[\varepsilon_j^4]}{\sigma_{H_0}^4} \to 0. \tag{A.11}$$

We first show the first part of (A.11). Note that

$$\mathrm{E}[\varepsilon_j^2|\mathcal{F}_{j-1}] = \frac{4}{N^2(N-1)^2} \sum_{i_1,i_2=1}^{j-1} \sum_{k_1 \neq k_2, k_3 \neq k_4}^{m} \gamma_{k_1k_2k_3k_4} z_{i_1k_1} z_{i_1k_2} z_{i_2k_3} z_{i_2k_4},$$

where

$$\gamma_{k_1k_2k_3k_4} = \sum_{l_1 \neq l_2}^{m} \beta_{k_1l_1k_2l_2}\beta_{k_3l_1k_4l_2} + \beta_{k_1l_1k_2l_2}\beta_{k_3l_2k_4l_1}.$$

To show the first part of (A.11) we decompose $\sum_{j=1}^{N} \mathrm{E}[\varepsilon_j^2|\mathcal{F}_{j-1}]$ into sum of ten parts

$$\sum_{j=1}^{N} \sigma_j^2 = \mathrm{Var}[U_{11}] + \sum_{t=1}^{9} R_t,$$

where

$$\begin{aligned}
R_1 &= \frac{8}{N^2(N-1)^2} \sum_{j=2}^{N} \sum_{i=1}^{j-1} \sum_{k_1 \neq k_2}^{m} \gamma_{k_1k_2k_1k_2} (z_{ik_1}^2 - 1)(z_{ik_2}^2 - 1), \\
R_2 &= \frac{8}{N^2(N-1)^2} \sum_{j=2}^{N} \sum_{i=1}^{j-1} \sum_{k_1 \neq k_2}^{m} \gamma_{k_1k_2k_1k_2} (z_{ik_1}^2 - 1), \\
R_3 &= \frac{8}{N^2(N-1)^2} \sum_{j=2}^{N} \sum_{i=1}^{j-1} \sum_{k_1 \neq k_2}^{m} \gamma_{k_1k_2k_1k_2} (z_{ik_2}^2 - 1), \\
R_4 &= \frac{8}{N^2(N-1)^2} \sum_{j=2}^{N} \sum_{i=1}^{j-1} \sum_{k_1 \neq k_2 \neq k_3 \neq k_1}^{m} \gamma_{k_1k_2k_1k_3} (z_{ik_1}^2 - 1) z_{ik_2} z_{ik_3},
\end{aligned}$$

$$R_5 = \frac{8}{N^2(N-1)^2}\sum_{j=2}^{N}\sum_{i=1}^{j-1}\sum_{k_1\neq k_2\neq k_3\neq k_1}^{m}\gamma_{k_1k_2k_1k_3}z_{ik_2}z_{ik_3},$$

$$R_6 = \frac{8}{N^2(N-1)^2}\sum_{j=2}^{N}\sum_{i=1}^{j-1}\sum_{k_1\neq k_2\neq k_3\neq k_1}^{m}\gamma_{k_1k_2k_2k_3}(z_{ik_2}^2-1)z_{ik_1}z_{ik_3},$$

$$R_7 = \frac{8}{N^2(N-1)^2}\sum_{j=2}^{N}\sum_{i=1}^{j-1}\sum_{k_1\neq k_2\neq k_3\neq k_1}^{m}\gamma_{k_1k_2k_2k_3}z_{ik_1}z_{ik_3},$$

$$R_8 = \frac{4}{N^2(N-1)^2}\sum_{j=2}^{N}\sum_{i=1}^{j-1}\sum_{\substack{k_1\neq k_2\neq k_3\neq k_4\\ k_2\neq k_4\neq k_1\neq k_3}}^{m}\gamma_{k_1k_2k_2k_3}z_{ik_1}z_{ik_2}z_{ik_3}z_{ik_4},$$

$$R_9 = \frac{8}{N^2(N-1)^2}\sum_{j=2}^{N}\sum_{i_1<i_2}^{j-1}\sum_{k_1\neq k_2,k_3\neq k_4}^{m}\gamma_{k_1k_2k_2k_3}z_{i_1k_1}z_{i_1k_2}z_{i_2k_3}z_{i_2k_4}.$$

Then we need to show $\mathrm{E}[R_t^2]=o(\sigma_{H_0}^4)$ for $t=1,\ldots,9$ since $\mathrm{Var}[U_{11}]/\sigma_{H_0}^2\to 1$. By applying Cauchy-Schwarz inequality, we obtain

$$\begin{aligned}
\mathrm{E}[R_1^2] &= \frac{32(\kappa_4+2)^2}{N^2(N-1)^3}\sum_{k_1\neq k_2}^{m}(\gamma_{k_1k_2k_1k_2}^2+\gamma_{k_1k_2k_1k_2}\gamma_{k_2k_1k_2k_1})\\
&\le \frac{64(\kappa_4+2)^2}{N^2(N-1)^3}\sum_{k_1\neq k_2}^{m}\gamma_{k_1k_2k_1k_2}^2\\
&= \frac{64(\kappa_4+2)^2}{N^2(N-1)^3}\sum_{k_1\neq k_2}^{m}\left(\sum_{\ell_1\neq\ell_2}^{m}\beta_{k_1\ell_1k_2\ell_2}^2+\sum_{\ell_1\neq\ell_2}^{m}\beta_{k_1\ell_1k_2\ell_2}\beta_{k_1\ell_2k_1\ell_1}\right)^2\\
&\le \frac{256(\kappa_4+2)^2}{N^2(N-1)^3}\sum_{k_1\neq k_2}^{m}\left(\sum_{\ell_1\neq\ell_2}^{m}\beta_{k_1\ell_1k_2\ell_2}^2\right)^2
\end{aligned}$$

$$\begin{aligned}
&\le \frac{256(\kappa_4+2)^2}{N^2(N-1)^3}\left(\sum_{k_1,k_2,\ell_1,\ell_2=1}^{m}\beta_{k_1\ell_1k_2\ell_2}^2\right)^2\\
&= \frac{256(\kappa_4+2)^2}{N^3(N-1)^2}\left(\sum_{\substack{g,h,g',h'=1\\ g\neq g',h\neq h'}}^{q}\mathrm{tr}(\Sigma_{gh}\Sigma_{hg})\mathrm{tr}(\Sigma_{g'h'}\Sigma_{h'g'})\right)^2.
\end{aligned}$$

Simiraly, we obtain

$$\begin{aligned}
\mathrm{E}[R_2^2] &\leq \frac{512(\kappa_4+2)}{N^3(N-1)^2}\left(\sum_{\substack{g,h,g',h'=1\\ g\neq g',h\neq h'}}^{q}\mathrm{tr}(\Sigma_{gh}\Sigma_{hg})\mathrm{tr}(\Sigma_{g'h'}\Sigma_{h'g'})\right)^2,\\
\mathrm{E}[R_3^2] &\leq \frac{512(\kappa_4+2)}{N^3(N-1)^2}\left(\sum_{\substack{g,h,g',h'=1\\ g\neq g',h\neq h'}}^{q}\mathrm{tr}(\Sigma_{gh}\Sigma_{hg})\mathrm{tr}(\Sigma_{g'h'}\Sigma_{h'g'})\right)^2,\\
\mathrm{E}[R_4^2] &\leq \frac{256(\kappa_4+2\kappa_3+2)}{N^3(N-1)^2}\left(\sum_{\substack{g,h,g',h'=1\\ g\neq g',h\neq h'}}^{q}\mathrm{tr}(\Sigma_{gh}\Sigma_{hg})\mathrm{tr}(\Sigma_{g'h'}\Sigma_{h'g'})\right)^2,\\
\mathrm{E}[R_5^2] &\leq \frac{2560}{N^3(N-1)^2}\left(\sum_{\substack{g,h,g',h'=1\\ g\neq g',h\neq h'}}^{q}\mathrm{tr}(\Sigma_{gh}\Sigma_{hg})\mathrm{tr}(\Sigma_{g'h'}\Sigma_{h'g'})\right)^2,\\
\mathrm{E}[R_6^2] &\leq \frac{256(\kappa_4+2\kappa_3+2)}{N^3(N-1)^2}\left(\sum_{\substack{g,h,g',h'=1\\ g\neq g',h\neq h'}}^{q}\mathrm{tr}(\Sigma_{gh}\Sigma_{hg})\mathrm{tr}(\Sigma_{g'h'}\Sigma_{h'g'})\right)^2,\\
\mathrm{E}[R_7^2] &\leq \frac{2560}{N^3(N-1)^2}\left(\sum_{\substack{g,h,g',h'=1\\ g\neq g',h\neq h'}}^{q}\mathrm{tr}(\Sigma_{gh}\Sigma_{hg})\mathrm{tr}(\Sigma_{g'h'}\Sigma_{h'g'})\right)^2,\\
\mathrm{E}[R_8^2] &\leq \frac{784}{N^3(N-1)^2}\left(\sum_{\substack{g,h,g',h'=1\\ g\neq g',h\neq h'}}^{q}\mathrm{tr}(\Sigma_{gh}\Sigma_{hg})\mathrm{tr}(\Sigma_{g'h'}\Sigma_{h'g'})\right)^2.
\end{aligned}$$

From these results and $V_1^{(1)^2} \leq \sigma^4$, under (A1)-(A3),

$$\frac{\mathrm{E}[R_t^2]}{\sigma^4} \leq \frac{\mathrm{E}[R_t^2]}{V_1^{(1)^2}} = O(N^{-1}) \quad (t=1,\dots,8). \tag{A.12}$$

Also, we evaluate $\mathrm{E}[R_9^2]$. The expectation of R_9^2 is evaluated as following:

$$
\begin{aligned}
\mathrm{E}[R_9^2] &= \frac{32(N-2)}{3N^2(N-1)^3} \sum_{k_1 \neq k_2, k_3 \neq k_4}^{m} (\gamma_{k_1k_2k_3k_4}^2 + \gamma_{k_1k_2k_3k_4}\gamma_{k_1k_2k_4k_3} \\
&\quad + \gamma_{k_1k_2k_3k_4}\gamma_{k_2k_1k_3k_4} + \gamma_{k_1k_2k_3k_4}\gamma_{k_2k_1k_4k_3}) \\
&\leq \frac{128(N-2)}{3N^2(N-1)^3} \sum_{k_1,k_2,k_3,k_4=1}^{m} \gamma_{k_1k_2k_3k_4}^2 \\
&= \frac{128(N-2)}{3N^2(N-1)^3} \sum_{k_1,k_2,k_3,k_4=1}^{m} \left(\sum_{\ell_1 \neq \ell_2}^{m} \beta_{k_1\ell_1k_2\ell_2}\beta_{k_3\ell_1k_4\ell_2} + \beta_{k_1\ell_1k_2\ell_2}\beta_{k_3\ell_2k_4\ell_1} \right)^2 \\
&\leq \frac{256(N-2)}{3N^2(N-1)^3} \sum_{k_1,k_2,k_3,k_4=1}^{m} \left(\sum_{\ell_1 \neq \ell_2}^{m} \beta_{k_1\ell_1k_2\ell_2}\beta_{k_3\ell_1k_4\ell_2} \right)^2 \\
&\quad + \frac{256(N-2)}{3N^2(N-1)^3} \sum_{k_1,k_2,k_3,k_4=1}^{m} \left(\sum_{\ell_1 \neq \ell_2}^{m} \beta_{k_1\ell_1k_2\ell_2}\beta_{k_3\ell_2k_4\ell_1} \right)^2 \\
&\leq \frac{512(N-2)}{3N^2(N-1)^3} \sum_{k_1,k_2,k_3,k_4=1}^{m} \left(\sum_{\ell_1,\ell_2=1}^{m} \beta_{k_1\ell_1k_2\ell_2}\beta_{k_3\ell_1k_4\ell_2} \right)^2 \\
&\quad + \frac{512(N-2)}{3N^2(N-1)^3} \sum_{k_1,k_2,k_3,k_4=1}^{m} \left(\sum_{\ell_1,\ell_2=1}^{m} \beta_{k_1\ell_1k_2\ell_2}\beta_{k_3\ell_2k_4\ell_1} \right)^2 \\
&\quad + \frac{1024(N-2)}{3N^2(N-1)^3} \sum_{k_1,k_2,k_3,k_4=1}^{m} \left(\sum_{l=1}^{m} \beta_{k_1\ell k_2\ell}\beta_{k_3\ell k_4\ell} \right)^2 \\
&= \frac{1024(N-2)}{3N^2(N-1)^3} \sum_{\substack{g_1 \neq h_1, g_2 \neq h_2 \\ g_3 \neq h_3, g_4 \neq h_4}}^{q} \mathrm{tr}(A^{(g_1)}A^{(g_2)}A^{(g_3)}A^{(g_4)})\mathrm{tr}(A^{(h_1)}A^{(h_2)}A^{(h_3)}A^{(h_4)}) \\
&\quad + \frac{1024(N-2)}{3N^2(N-1)^3} \\
&\quad \times \sum_{\substack{g_1 \neq h_1, g_2 \neq h_2 \\ g_3 \neq h_3, g_4 \neq h_4}}^{q} \mathrm{tr}\left((A^{(g_1)}A^{(g_3)}) \odot (A^{(h_1)}A^{(h_3)}) \right) \left((A^{(g_2)}A^{(g_4)}) \odot (A^{(h_2)}A^{(h_4)}) \right).
\end{aligned}
$$

From Lemma A.2, under (A1)-(A3),

$$\frac{\mathrm{E}[R_9^2]}{\sigma^4} = o(1). \tag{A.13}$$

From (A.12), (A.13) and $\sigma^2/\sigma_{H_0}^2 \to 1$ under (A1)-(A4), we obtain $\mathrm{E}[R_t^2] = o(\sigma_{H_0}^4)$. This proves first part of (A.11).

Next we show the second part of (A.11). There exists constants c_1 and c_2 such that

$$\begin{aligned}
\mathrm{E}[\varepsilon_j^4] &\leq \frac{c_1(j-1)(j-2)}{N^4(N-1)^4}\mathrm{E}\left[\left\{\sum_{k_1,k_2=1}^{m}\left(\sum_{\ell_1\neq\ell_2}^{m}\beta_{k_1\ell_1k_2\ell_2}z_{j\ell_1}z_{j\ell_2}\right)^2\right\}^2\right] \\
&\leq \frac{c_2(j-1)(j-2)}{N^4(N-1)^4}\left(\sum_{\substack{g,h,g',h'=1\\ g\neq g',h\neq h'}}^{q}\mathrm{tr}(\Sigma_{gh}\Sigma_{hg})\mathrm{tr}(\Sigma_{g'h'}\Sigma_{h'g'})\right)^2.
\end{aligned}$$

From the above results,

$$\frac{\sum_{j=1}^{N}\mathrm{E}[\varepsilon_j^4]}{\sigma^4} \leq \frac{\sum_{j=1}^{N}\mathrm{E}[\varepsilon_j^4]}{V_1^{(1)2}} = O(N^{-1}). \tag{A.14}$$

From (A.14) and $\sigma^2/\sigma_{H_0}^2 \to 1$ under (A1)-(A4), we obtain

$$\sum_{j=1}^{N}\mathrm{E}[\varepsilon_j^4] = o(\sigma_{H_0}^4).$$

This proves second part of (A.11) and completes the proof of Theorem 2.1. □

A.3. Proof of Theorem 2.3.

Under assumptions (A1) and (A2), it holds that

$$\widehat{\mathrm{tr}\Sigma_{gg}^2} = \mathrm{tr}\Sigma_{gg}^2(1+O_p(N^{-1/2})).$$

Hence

$$\widehat{\sigma}_{H_0}^2 = \sigma_{H_0}^2(1+O_p(N^{-1/2}))$$

and it hold

$$\frac{\widehat{\sigma}_{H_0}^2}{\sigma_{H_0}^2} = 1 + O_p(N^{-1/2}).$$

From the above results, it completes the proof of Theorem 2.2. □

A.4. Proof of Theorem 3.1.

From Lemma A.4, under assumptions (A1')-(A3), the leading order term in $\mathrm{Var}[T]$ is contributed by $U_{11} + U_{12} + U_{13}$. Hence, we only need to show the asymptotic normality of $U_{11} + U_{12} + U_{13}$. Let

$$\psi_j \;=\; \varepsilon_j + \frac{2}{N}\left\{\boldsymbol{z}_j'\Gamma'(\Sigma - \Sigma_d)\Gamma\boldsymbol{z}_j - \mathrm{tr}\left(\Sigma^2 - \Sigma_d^2\right)\right\}.$$

Then $\sum_{j=1}^N \psi_j = U_{11} + U_{12} + U_{13}$, and it is straightforward to show that $\mathrm{E}[\psi_j] = 0$ and $\mathrm{E}[\psi_j|\mathcal{F}_{j-1}] = 0$. So ψ_j is a martingale difference sequence. To apply martingale central limit theorem, we need to show that

$$\frac{\sum_{j=1}^N \mathrm{E}[\psi_j^2|\mathcal{F}_{j-1}]}{\sigma^2} \xrightarrow{p} 1 \qquad \text{and} \qquad \frac{\sum_{j=1}^N \mathrm{E}[\psi_j^4]}{\sigma^4} \to 0. \tag{A.15}$$

We first show the first part of (A.15). To show the first part of (A.15) we decompose $\sum_{j=1}^N \mathrm{E}[\psi_j^2|\mathcal{F}_{j-1}]$ into sum of ten parts

$$\mathrm{E}[\psi_j^2|\mathcal{F}_{j-1}] \;=\; \mathrm{Var}[U_{11}] + V_2^{(1)} + V_3^{(1)} + \sum_{t=1}^{10} R_t,$$

where R_t $(t = 1, 2, \ldots, 9)$ is defined in (A.1), and

$$R_{10} \;=\; \frac{8}{N^2(N-1)} \sum_{i=1}^{j-1} \sum_{\substack{k_1,k_2,k_3=1\\ k_1\neq k_2}}^{m} \gamma_{k_1k_2k_3k_3} z_{ik_1} z_{ik_2}.$$

Then we need to show $\mathrm{E}[R_t^2] = o(\sigma^4)$ for $t = 1, \ldots, 10$ since under (A1')-(A3),

$$\frac{\mathrm{Var}[U_{11}] + V_2^{(1)} + V_3^{(1)}}{\sigma^2} \to 1.$$

From (A.12) and (A.13), under (A1')-(A3),

$$\frac{\mathrm{E}[R_t^2]}{\sigma^4} = o(1) \;\; (t = 1, 2, \ldots, 9). \tag{A.16}$$

Also, we evaluate $\mathrm{E}[R_{10}^2]$. The expectation of R_{10}^2 is evaluated as following:

$$\begin{aligned}
\mathrm{E}[R_{10}^2] &= \frac{32}{N^2(N-1)}\sum_{k_1\neq k_2}^{m}\left\{\left(\sum_{k_3=1}^{m}\gamma_{k_1k_2k_3k_3}\right)^2+\left(\sum_{k_3=1}^{m}\gamma_{k_1k_2k_3k_3}\right)\left(\sum_{k_3=1}^{m}\gamma_{k_2k_1k_3k_3}\right)^2\right\}\\
&\leq \frac{64}{N^2(N-1)}\sum_{k_1\neq k_2}^{m}\left(\sum_{k_3=1}^{m}\gamma_{k_1k_2k_3k_3}\right)^2\\
&= \frac{256}{N^2(N-1)}\sum_{k_1\neq k_2}^{m}\left(\sum_{\ell_1,\ell_2,k_3=1}^{m}\beta_{k_1\ell_1k_2\ell_2}\beta_{k_3\ell_1k_3\ell_2}-\sum_{\ell,k_3=1}^{m}\beta_{k_1\ell k_2\ell}\beta_{k_3\ell k_3\ell}\right)^2\\
&\leq \frac{256}{N^2(N-1)}\sum_{k_1,k_2=1}^{m}\left(\sum_{\ell_1,\ell_2,k_3=1}^{m}\beta_{k_1\ell_1k_2\ell_2}\beta_{k_3\ell_1k_3\ell_2}\right)^2\\
&\quad+\frac{256}{N^2(N-1)}\sum_{k_1,k_2=1}^{m}\left(\sum_{\ell,k_3=1}^{m}\beta_{k_1\ell k_2\ell}\beta_{k_3\ell k_3\ell}\right)^2
\end{aligned}$$

$$\begin{aligned}
&= \frac{256}{N^2(N-1)}\mathrm{tr}\left(\sum_{g\neq h}^{q}A^{(g)}\Gamma'(\Sigma-\Sigma_d)\Gamma A^{(h)}\right)^2\\
&\quad+\frac{256}{N^2(N-1)}\\
&\quad\times\sum_{\substack{g,h,g',h'=1\\ g\neq g',h\neq h'}}^{q}[\mathrm{diag}(\Gamma(\Sigma-\Sigma_d)\Gamma)]'\left(A^{(g)}A^{(h)}\right)\odot\left(A^{(i)}A^{(j)}\right)\mathrm{diag}(\Gamma(\Sigma-\Sigma_d)\Gamma).
\end{aligned}$$

From above result and Lemma A.2,

$$\mathrm{E}[R_{10}^2] \leq \frac{512}{N^2(N-1)}\min\left\{4\mathrm{tr}\Sigma^4,\left(\sum_{g\neq h}^{q}(\mathrm{tr}\Sigma_{gg}^4)^{1/4}(\mathrm{tr}\Sigma_{hh}^4)^{1/4}\right)^2\right\}\mathrm{tr}(\Sigma^2-\Sigma_d^2)^2.$$

From above result, under (A1')-(A3),

$$\frac{\mathrm{E}[R_{10}^2]}{\sigma^4}=o(1). \tag{A.17}$$

Combining (A.16) and (A.17), we prove the first part of (A.15).

Next we show the second part of (A.15). By applying Hölder's inequality,

$$\mathrm{E}[\psi_j^4] \leq 8\mathrm{E}[\varepsilon_j^4] + \frac{128}{N^4}\mathrm{E}\left[\left(\boldsymbol{z}_j'\Gamma'(\Sigma - \Sigma_d)\Gamma\boldsymbol{z}_j - \mathrm{tr}\left(\Sigma^2 - \Sigma_d^2\right)\right)^4\right]. \quad \text{(A.18)}$$

Note that for any symmetric matrix A,

$$\mathrm{E}[(\boldsymbol{z}_1'A\boldsymbol{z}_1 - \mathrm{tr}A)^4] \leq const. \times \{\mathrm{tr}A^2\}^2.$$

Hence, there exists a constant c_3 such that

$$\mathrm{E}[\left\{\boldsymbol{z}_j'\Gamma'(\Sigma - \Sigma_d)\Gamma\boldsymbol{z}_j - \mathrm{tr}(\Sigma^2 - \Sigma_d^2)\right\}^4] \leq c_3\left(\mathrm{tr}\left(\Sigma^2 - \Sigma_d^2\right)^2\right)^2. \quad \text{(A.19)}$$

From (A.18) and (A.19), $\sum_{j=1}^N \mathrm{E}[\psi_j^4] = o(\sigma^4)$. This proves second part of (A.15) and completes the proof of Theorem 2.3. □

Table 1: Attained significant level in the case of diagonal matrix.

p	N	Case 1	Case 2	Case 3	Case 4	Case 5
		our procedure				
100	10	0.081	0.078	0.087	0.088	0.088
	50	0.063	0.061	0.056	0.058	0.051
	100	0.062	0.052	0.053	0.053	0.056
200	10	0.080	0.081	0.069	0.096	0.077
	50	0.055	0.057	0.044	0.058	0.052
	100	0.047	0.054	0.057	0.061	0.054
400	10	0.075	0.071	0.083	0.082	0.081
	50	0.053	0.058	0.048	0.056	0.062
	100	0.049	0.056	0.047	0.057	0.047
1000	10	0.076	0.076	0.079	0.087	0.079
	50	0.056	0.051	0.053	0.053	0.058
	100	0.054	0.054	0.055	0.055	0.056
		Srivastava (2005)				
100	10	0.064	0.094	0.053	0.086	0.076
	50	0.054	0.109	0.047	0.061	0.061
	100	0.042	0.098	0.053	0.152	0.065
200	10	0.057	0.107	0.046	0.124	0.096
	50	0.054	0.111	0.038	0.153	0.062
	100	0.044	0.103	0.056	0.183	0.060
400	10	0.067	0.101	0.059	0.115	0.086
	50	0.044	0.095	0.044	0.169	0.064
	100	0.048	0.096	0.043	0.206	0.061
1000	10	0.073	0.102	0.057	0.145	0.091
	50	0.051	0.112	0.048	0.178	0.065
	100	0.049	0.103	0.046	0.216	0.063

Table 2: Attained significant level in the case of $q = 2$.

p	N	Case 1	Case 2	Case 3	Case 4	Case 5
		our procedure				
	10	0.078	0.078	0.074	0.079	0.078
100	50	0.056	0.061	0.058	0.062	0.061
	100	0.057	0.059	0.057	0.056	0.058
	10	0.079	0.077	0.073	0.078	0.073
200	50	0.059	0.056	0.059	0.055	0.059
	100	0.061	0.057	0.059	0.057	0.060
	10	0.082	0.076	0.081	0.077	0.072
400	50	0.058	0.056	0.055	0.058	0.055
	100	0.055	0.054	0.052	0.054	0.052
	10	0.078	0.076	0.073	0.077	0.076
1000	50	0.052	0.051	0.054	0.057	0.057
	100	0.055	0.054	0.056	0.054	0.053
		Yata and Aoshima (2015)				
	10	0.113	0.113	0.105	0.110	0.111
100	50	0.060	0.064	0.066	0.067	0.066
	100	0.058	0.060	0.058	0,059	0.061
	10	0.110	0.108	0.107	0.109	0.103
200	50	0.066	0.062	0.065	0.060	0.063
	100	0.063	0.060	0.063	0.059	0.063
	10	0.114	0.109	0.116	0.111	0.107
400	50	0.064	0.061	0.060	0.063	0.061
	100	0.058	0.057	0.056	0.058	0.055
	10	0.111	0.110	0.105	0.107	0.111
1000	50	0.058	0.058	0.062	0.062	0.060
	100	0.058	0.057	0.059	0.057	0.054
		Srivastava and Reid (2012)				
	10	0.047	0.081	0.047	0.094	0.059
100	50	0.052	0.084	0.053	0.109	0.062
	100	0.055	0.074	0.055	0.103	0.061
	10	0.052	0.084	0.044	0.098	0.053
200	50	0.053	0.081	0.056	0.122	0.060
	100	0.058	0.075	0.056	0.118	0.062
	10	0.046	0.085	0.047	0.118	0.054
400	50	0.052	0.080	0.052	0.140	0.060
	100	0.054	0.072	0.051	0.143	0.056
	10	0.048	0.089	0.043	0.126	0.053
1000	50	0.048	0.077	0.050	0.163	0.058
	100	0.053	0.073	0.054	0.172	0.053

Table 3: Attained significant level in the case of $q = 5$.

p	N	Case 1	Case 2	Case 3	Case 4	Case 5
	10	0.084	0.078	0.084	0.083	0.082
100	50	0.064	0.061	0.057	0.057	0.056
	100	0.056	0.058	0.061	0.053	0.057
	10	0.077	0.084	0.088	0.078	0.081
200	50	0.058	0.059	0.057	0.055	0.058
	100	0.054	0.056	0.058	0.054	0.057
	10	0.081	0.085	0.083	0.081	0.083
400	50	0.056	0.060	0.054	0.060	0.057
	100	0.058	0.052	0.054	0.051	0.053
	10	0.079	0.078	0.082	0.084	0.086
1000	50	0.054	0.057	0.057	0.057	0.054
	100	0.054	0.054	0.051	0.055	0.052

Table 4: Empirical power in the case of diagonal matrix for (i).

			our procedure			
p	N	Case 1	Case 2	Case 3	Case 4	Case 5
	10	0.165	0.165	0.167	0.173	0.158
100	50	0.615	0.611	0.606	0.625	0.607
	100	0.942	0.941	0.943	0.940	0.941
	10	0.261	0.257	0.259	0.263	0.268
200	50	0.905	0.904	0.903	0.907	0.904
	100	0.999	0.999	0.999	0.999	0.999
	10	0.419	0.416	0.422	0.432	0.417
400	50	0.994	0.994	0.994	0.994	0.995
	100	1.000	1.000	1.000	1.000	1.000
	10	0.487	0.476	0.491	0.495	0.489
1000	50	0.999	0.998	0.999	0.999	0.998
	100	1.000	1.000	1.000	1.000	1.000
			Srivastava (2005)			
	10	0.134	0.154	0.132	0.151	0.124
100	50	0.610	0.499	0.601	0.535	0.573
	100	0.940	0.859	0.942	0.793	0.932
	10	0.227	0.201	0.225	0.194	0.203
200	50	0.906	0.772	0.903	0.753	0.875
	100	0.999	0.991	0.999	0.958	0.998
	10	0.398	0.275	0.397	0.258	0.333
400	50	0.994	0.956	0.994	0.871	0.991
	100	1.000	0.999	1.000	0.993	1.000
	10	0.476	0.379	0.475	0.356	0.413
1000	50	0.999	0.977	0.999	0.913	0.996
	100	1.000	1.000	1.000	1.000	1.000

Table 5: Empirical power in the case of diagonal matrix for (ii).

		our procedure				
p	N	Case 1	Case 2	Case 3	Case 4	Case 5
100	10	0.104	0.105	0.102	0.111	0.106
	50	0.212	0.213	0.209	0.216	0.205
	100	0.455	0.461	0.449	0.461	0.457
200	10	0.107	0.112	0.107	0.113	0.111
	50	0.220	0.222	0.207	0.222	0.212
	100	0.478	0.483	0.483	0.491	0.482
400	10	0.100	0.109	0.109	0.107	0.106
	50	0.215	0.217	0.215	0.226	0.215
	100	0.507	0.508	0.518	0.508	0.507
1000	10	0.105	0.113	0.112	0.109	0.110
	50	0.218	0.221	0.219	0.232	0.221
	100	0.531	0.528	0.536	0.530	0.527
		Srivastava (2005)				
100	10	0.073	0.112	0.074	0.126	0.089
	50	0.198	0.235	0.202	0.222	0.203
	100	0.454	0.423	0.448	0.373	0.440
200	10	0.075	0.122	0.075	0.135	0.085
	50	0.208	0.249	0.205	0.253	0.209
	100	0.477	0.447	0.482	0.378	0.462
400	10	0.073	0.121	0.074	0.145	0.091
	50	0.204	0.243	0.212	0.274	0.212
	100	0.501	0.469	0.511	0.387	0.479
1000	10	0.072	0.128	0.076	0.151	0.095
	50	0.207	0.245	0.214	0.287	0.221
	100	0.525	0.496	0.532	0.432	0.491

Table 6: Empirical power in the case of $q = 2$ for (i).

p	N	Case 1	Case 2	Case 3	Case 4	Case 5
		our procedure				
100	10	0.090	0.098	0.096	0.098	0.094
	50	0.154	0.151	0.149	0.155	0.152
	100	0.274	0.277	0.284	0.281	0.284
200	10	0.108	0.108	0.104	0.106	0.108
	50	0.264	0.263	0.265	0.271	0.262
	100	0.551	0.549	0.550	0.569	0.547
400	10	0.131	0.138	0.135	0.140	0.137
	50	0.543	0.554	0.556	0.564	0.551
	100	0.926	0.929	0.925	0.927	0.928
1000	10	0.162	0.160	0.164	0.164	0.153
	50	0.732	0.743	0.740	0.736	0.739
	100	0.989	0.991	0.989	0.988	0.991
		Yata and Aoshima (2015)				
100	10	0.130	0.130	0.128	0.133	0.132
	50	0.161	0.159	0.156	0.163	0.162
	100	0.278	0.280	0.283	0.285	0.287
200	10	0.143	0.141	0.138	0.143	0.143
	50	0.274	0.272	0.271	0.277	0.272
	100	0.554	0.552	0.554	0.570	0.549
400	10	0.173	0.175	0.176	0.180	0.175
	50	0.547	0.558	0.559	0.568	0.557
	100	0.926	0.928	0.924	0.926	0.929
1000	10	0.200	0.201	0.211	0.207	0.195
	50	0.733	0.743	0.737	0.736	0.740
	100	0.988	0.990	0.989	0.989	0.990
		Srivastava and Reid (2012)				
100	10	0.061	0.111	0.063	0.099	0.071
	50	0.147	0.179	0.143	0.168	0.150
	100	0.270	0.284	0.278	0.259	0.276
200	10	0.074	0.111	0.069	0.116	0.078
	50	0.260	0.270	0.254	0.244	0.243
	100	0.548	0.511	0.551	0.434	0.532
400	10	0.090	0.129	0.090	0.145	0.102
	50	0.538	0.487	0.548	0.391	0.517
	100	0.925	0.874	0.924	0.721	0.914
1000	10	0.116	0.150	0.120	0.187	0.116
	50	0.730	0.670	0.739	0.552	0.717
	100	0.989	0.976	0.990	0.884	0.990

Table 7: Empirical power in the case of $q = 2$ for (ii).

p	N	Case 1	Case 2	Case 3	Case 4	Case 5
		our procedure				
100	10	0.106	0.109	0.104	0.107	0.118
	50	0.271	0.272	0.266	0.284	0.280
	100	0.539	0.529	0.531	0.545	0.525
200	10	0.135	0.137	0.136	0.136	0.139
	50	0.529	0.536	0.526	0.541	0.528
	100	0.889	0.895	0.891	0.892	0.876
400	10	0.196	0.202	0.196	0.204	0.174
	50	0.854	0.845	0.852	0.847	0.846
	100	0.998	0.998	0.997	0.997	0.999
1000	10	0.129	0.127	0.132	0.132	0.126
	50	0.474	0.462	0.465	0.477	0.433
	100	0.827	0.821	0.825	0.812	0.818
		Yata and Aoshima (2015)				
100	10	0.141	0.139	0.140	0.142	0.136
	50	0.277	0.280	0.276	0.287	0.292
	100	0.545	0.532	0.534	0.548	0.531
200	10	0.173	0.175	0.171	0.177	0.172
	50	0.533	0.543	0.528	0.543	0.531
	100	0.890	0.894	0.891	0.892	0.873
400	10	0.236	0.242	0.234	0.246	0.224
	50	0.852	0.844	0.852	0.846	0.850
	100	0.997	0.998	0.997	0.997	1.000
1000	10	0.171	0.163	0.165	0.172	0.183
	50	0.478	0.468	0.472	0.477	0.437
	100	0.827	0.821	0.825	0.811	0.819
		Srivastava and Reid (2012)				
100	10	0.075	0.122	0.075	0.109	0.090
	50	0.263	0.277	0.263	0.251	0.272
	100	0.537	0.500	0.527	0.451	0.515
200	10	0.096	0.137	0.098	0.130	0.104
	50	0.525	0.475	0.523	0.403	0.481
	100	0.889	0.827	0.892	0.746	0.866
400	10	0.154	0.167	0.154	0.189	0.141
	50	0.853	0.757	0.854	0.672	0.828
	100	0.998	0.991	0.997	0.958	0.999
1000	10	0.091	0.128	0.090	0.185	0.110
	50	0.466	0.462	0.458	0.499	0.444
	100	0.826	0.796	0.824	0.800	0.823

Table 8: Empirical power in the case of $q = 5$ for (i).

p	N	Case 1	Case 2	Case 3	Case 4	Case 5
	10	0.088	0.090	0.089	0.091	0.088
100	50	0.102	0.104	0.095	0.096	0.097
	100	0.155	0.155	0.153	0.156	0.161
	10	0.096	0.095	0.095	0.098	0.092
200	50	0.144	0.144	0.145	0.154	0.142
	100	0.279	0.283	0.277	0.283	0.289
	10	0.109	0.112	0.111	0.117	0.112
400	50	0.268	0.271	0.271	0.271	0.253
	100	0.599	0.603	0.588	0.600	0.594
	10	0.160	0.159	0.162	0.169	0.159
1000	50	0.708	0.711	0.699	0.713	0.705
	100	0.992	0.992	0.994	0.990	0.991

Table 9: Empirical power in the case of $q = 5$ for (ii).

p	N	Case 1	Case 2	Case 3	Case 4	Case 5
	10	0.127	0.124	0.118	0.124	0.132
100	50	0.357	0.364	0.361	0.362	0.371
	100	0.687	0.687	0.683	0.699	0.678
	10	0.168	0.169	0.169	0.176	0.166
200	50	0.643	0.637	0.643	0.641	0.637
	100	0.947	0.944	0.940	0.947	0.942
	10	0.267	0.259	0.261	0.276	0.260
400	50	0.914	0.916	0.916	0.919	0.922
	100	1.000	0.999	0.999	0.999	0.999
	10	0.495	0.499	0.501	0.508	0.501
1000	50	0.999	0.999	0.998	0.999	0.999
	100	1.000	1.000	1.000	1.000	1.000

Supplemental Material:Testing Block-Diagonal Covariance Structure for High-Dimensional Under Non-normality

Yuki Yamada[*], Masashi Hyodo[†] and Takahiro Nishiyama[‡]

March 30, 2016

Abstract

The supplement provides full proof of Lemma A.1 and 2, which are omitted in the original paper.

1 Proof of Lemma A.1

First, we show (i).

$$\begin{aligned}
\mathrm{E}[\boldsymbol{z}_1' E\boldsymbol{z}_1\boldsymbol{z}_1' F\boldsymbol{z}_1] &= \mathrm{E}\left[\left(\sum_{i=1}^{m} e_{ii}z_{1i}^2 + \sum_{\substack{i,j=1\\ i\neq j}}^{m} e_{ij}z_{1i}z_{1j}\right)\left(\sum_{i=1}^{m} f_{ii}z_{1i}^2 + \sum_{\substack{i,j=1\\ i\neq j}}^{m} f_{ij}z_{1i}z_{1j}\right)\right] \\
&= \sum_{i=1}^{m} e_{ii}f_{ii}\mathrm{E}[z_{1i}^4] + \sum_{\substack{i,j=1\\ i\neq j}}^{m} e_{ii}f_{jj}\mathrm{E}[z_{1i}^2z_{1j}^2] + \sum_{\substack{i,j=1\\ i\neq j}}^{m} (e_{ij}f_{ij}+e_{ij}f_{ji})\mathrm{E}[z_{1i}^2z_{1j}^2] \\
&= \kappa_4 \mathrm{tr}(E\odot F) + \mathrm{tr}E\mathrm{tr}F + \mathrm{tr}(EF) + \mathrm{tr}(EF').
\end{aligned}$$

Second, we show (ii).

$$\begin{aligned}
\mathrm{E}[\boldsymbol{z}_1' A\boldsymbol{z}_2\boldsymbol{z}_1' B\boldsymbol{z}_2\boldsymbol{z}_1' E\boldsymbol{z}_2] &= \mathrm{E}\left[\left(\sum_{i,j=1}^{m} a_{ij}z_{1i}z_{2j}\right)\left(\sum_{i,j=1}^{m} b_{ij}z_{1i}z_{2j}\right)\left(\sum_{i,j=1}^{m} e_{ij}z_{1i}z_{2j}\right)\right] \\
&= \sum_{i,j=1}^{m} a_{ij}b_{ij}e_{ij}\mathrm{E}[z_{1i}^3z_{2j}^3] \\
&= \kappa_3^2\mathrm{tr}\left((A\odot B)E'\right)
\end{aligned}$$

[*]Department of Mathematical Information Science, Tokyo University of Science, 1-3, Kagurazaka, Shinjuku-ku, Tokyo, 162-8601, Japan, E-Mail:y.y.rfpfg.t.o.d.2006.s.m.e@gmail.com.

[†]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.

Third, we show (iii).

$$
\begin{aligned}
\mathrm{E} \quad & [\boldsymbol{z}_1' A \boldsymbol{z}_2 \boldsymbol{z}_1' B \boldsymbol{z}_2 \boldsymbol{z}_1' C \boldsymbol{z}_2 \boldsymbol{z}_1' D \boldsymbol{z}_2] \\
= \quad & \mathrm{E}\left[\left(\sum_{i,j=1}^{m} a_{ij} z_{1i} z_{2j}\right)\left(\sum_{i,j=1}^{m} b_{ij} z_{1i} z_{2j}\right)\left(\sum_{i,j=1}^{m} c_{ij} z_{1i} z_{2j}\right)\left(\sum_{i,j=1}^{m} d_{ij} z_{1i} z_{2j}\right)\right] \\
= \quad & \mathrm{E}\left[\left(\sum_{i,j=1}^{m} a_{ij} b_{ij} z_{1i}^2 z_{2j}^2 + \sum_{\substack{i,j,k=1\\ j\neq k}}^{m} a_{ij} b_{ik} z_{1i}^2 z_{2i} z_{2k} + \sum_{\substack{i,j,k=1\\ i\neq j}}^{m} a_{ik} b_{jk} z_{1i} z_{1j} z_{2k}^2\right.\right. \\
& \left. + \sum_{\substack{i,j,k,l=1\\ i\neq j,k\neq l}}^{m} a_{ij} b_{kl} z_{1i} z_{1j} z_{2k} z_{2l}\right) \times \left(\sum_{i,j=1}^{m} c_{ij} d_{ij} z_{1i}^2 z_{2j}^2 + \sum_{\substack{i,j,k=1\\ j\neq k}}^{m} c_{ij} d_{ik} z_{1i}^2 z_{2i} z_{2k}\right. \\
& \left.\left. + \sum_{\substack{i,j,k=1\\ i\neq j}}^{m} c_{ik} d_{jk} z_{1i} z_{1j} z_{2k}^2 + \sum_{\substack{i,j,k,l=1\\ i\neq j,k\neq l}}^{m} c_{ij} d_{kl} z_{1i} z_{1j} z_{2k} z_{2l}\right)\right] \\
= \quad & \sum_{i,j=1}^{m} a_{ij} b_{ij} c_{ij} d_{ij} \mathrm{E}[z_{1i}^4 z_{2j}^4] + \sum_{\substack{i,j,k=1\\ i\neq k}}^{m} a_{ij} b_{ij} c_{kj} d_{kj} \mathrm{E}[z_{1i}^2 z_{1k}^2 z_{2j}^4] \\
& + \sum_{\substack{i,j,k=1\\ j\neq k}}^{m} a_{ij} b_{ij} c_{ik} d_{ik} \mathrm{E}[z_{1i}^4 z_{2j}^2 z_{2k}^2] + \sum_{\substack{i,j,k,l=1\\ i\neq k,j\neq l}}^{m} a_{ij} b_{ij} c_{kl} d_{kl} \mathrm{E}[z_{1i}^2 z_{1k}^2 z_{2j}^2 z_{2l}^2] \\
& + \sum_{\substack{i,j,k=1\\ j\neq k}}^{m} a_{ij} b_{ik} c_{ij} d_{ik} \mathrm{E}[z_{1i}^4 z_{2j}^2 z_{2k}^2] + \sum_{\substack{i,j,k=1\\ j\neq k}}^{m} a_{ij} b_{ik} c_{ik} d_{ij} \mathrm{E}[z_{1i}^4 z_{2j}^2 z_{2k}^2] \\
& + \sum_{\substack{i,j,k,l=1\\ i\neq j,k\neq l}}^{m} a_{ik} b_{il} c_{jk} d_{jl} \mathrm{E}[z_{1i}^2 z_{1k}^2 z_{2j}^2 z_{2l}^2] + \sum_{\substack{i,j,k,l=1\\ i\neq j,k\neq l}}^{m} a_{ik} b_{il} c_{jl} d_{jk} \mathrm{E}[z_{1i}^2 z_{1k}^2 z_{2j}^2 z_{2l}^2] \\
& + \sum_{\substack{i,j,k=1\\ i\neq j}}^{m} a_{ik} b_{jk} c_{ik} d_{jk} \mathrm{E}[z_{1i}^2 z_{1j}^2 z_{2k}^4] + \sum_{\substack{i,j,k=1\\ i\neq j}}^{m} a_{ik} b_{jk} c_{jk} d_{ik} \mathrm{E}[z_{1i}^2 z_{1j}^2 z_{2k}^4] \\
& + \sum_{\substack{i,j,k,l=1\\ i\neq j,k\neq l}}^{m} a_{ik} b_{jk} c_{il} d_{jl} \mathrm{E}[z_{1i}^2 z_{1j}^2 z_{2k}^2 z_{2l}^2] + \sum_{\substack{i,j,k,l=1\\ i\neq j,k\neq l}}^{m} a_{ik} b_{jk} c_{jl} d_{il} \mathrm{E}[z_{1i}^2 z_{1j}^2 z_{2k}^2 z_{2l}^2] \\
& + \sum_{\substack{i,j,k,l=1\\ i\neq j,k\neq l}}^{m} a_{ik} b_{jl} c_{ik} d_{jl} \mathrm{E}[z_{1i}^2 z_{1k}^2 z_{2j}^2 z_{2l}^2] + \sum_{\substack{i,j,k,l=1\\ i\neq j,k\neq l}}^{m} a_{ik} b_{jl} c_{il} d_{ik} \mathrm{E}[z_{1i}^2 z_{1k}^2 z_{2j}^2 z_{2l}^2] \\
& + \sum_{\substack{i,j,k,l=1\\ i\neq j,k\neq l}}^{m} a_{ik} b_{jl} c_{jk} d_{il} \mathrm{E}[z_{1i}^2 z_{1k}^2 z_{2j}^2 z_{2l}^2] + \sum_{\substack{i,j,k,l=1\\ i\neq j,k\neq l}}^{m} a_{ik} b_{jl} c_{jl} d_{ik} \mathrm{E}[z_{1i}^2 z_{1k}^2 z_{2j}^2 z_{2l}^2]
\end{aligned}
$$

$$
\begin{aligned}
&= \kappa_4^2 \sum_{i,j=1}^m a_{ij}b_{ij}c_{ij}d_{ij} + 2\kappa_4 \sum_{i,j,k=1}^m a_{ij}b_{ij}c_{kj}d_{kj} + 2\kappa_4 \sum_{i,j,k=1}^m a_{ij}b_{ik}c_{ij}d_{ik} \\
&\quad +2\kappa_4 \sum_{i,j,k=1}^m a_{ij}b_{ik}c_{ik}d_{ij} + \sum_{i,j,k,l=1}^m a_{ij}b_{ij}c_{kl}d_{kl} + \sum_{i,j,k,l=1}^m a_{ik}b_{il}c_{jk}d_{jl} + \sum_{i,j,k,l=1}^m a_{ik}b_{il}c_{jl}d_{jk} \\
&\quad + \sum_{i,j,k,l=1}^m a_{ik}b_{jk}c_{il}d_{jl} + \sum_{i,j,k,l=1}^m a_{ik}b_{jk}c_{jl}d_{il} + \sum_{i,j,k,l=1}^m a_{ik}b_{jl}c_{ik}d_{jl} \\
&\quad + \sum_{i,j,k,l=1}^m a_{ik}b_{jl}c_{il}d_{ik} + \sum_{i,j,k,l=1}^m a_{ik}b_{jl}c_{jk}d_{il} + \sum_{i,j,k,l=1}^m a_{ik}b_{jl}c_{jl}d_{ik} \\
&= \kappa_4^2 \mathrm{tr}\left((A \odot B)(C \odot D)\right) 2\kappa_4 \mathrm{tr}\left((AB) \odot (CD)\right) + 2\kappa_4 \mathrm{tr}\left((AC) \odot (BD)\right) \\
&\quad +2\kappa_4 \mathrm{tr}\left((AD) \odot (BC)\right) + \mathrm{tr}(AB)\mathrm{tr}(CD) + \mathrm{tr}(AC)\mathrm{tr}(BD) + \mathrm{tr}(AD)\mathrm{tr}(BC) \\
&\quad +\mathrm{tr}(ABCD) + \mathrm{tr}(ABDC) + \mathrm{tr}(ACBD) + \mathrm{tr}(ACDB) + \mathrm{tr}(ADBC) + \mathrm{tr}(ADCB).
\end{aligned}
$$

Finally, we show (iv).

$$
\begin{aligned}
\mathrm{E}\left[\left(\sum_{i\neq j}^m a_{ij}z_{1i}z_{1j}\right)^4\right] &= \mathrm{E}\left[\left(2\sum_{i\neq j}^m a_{ij}^2 z_{1i}^2 z_{1j}^2 + 4\sum_{i\neq j\neq k\neq i}^m a_{ij}a_{ik}z_{1i}^2 z_{1j}z_{1k}\right.\right. \\
&\quad \left.\left. + \sum_{\substack{i\neq j\neq k\neq l \\ j\neq l\neq i\neq k}}^m a_{ij}a_{kl}z_{1i}z_{1j}z_{1k}z_{1l}\right)^2\right] \\
&\leq 12\mathrm{E}\left[\left(\sum_{i\neq j}^m a_{ij^2} z_{1i}^2 z_{1j}^2\right)^2\right] + 48\mathrm{E}\left[\left(\sum_{i\neq j\neq k\neq i}^m a_{ij}a_{ik}z_{1i}^2 z_{1j}z_{1k}\right)^2\right] \\
&\quad +3\mathrm{E}\left[\left(\sum_{\substack{i\neq j\neq k\neq l \\ j\neq l\neq i\neq k}}^m a_{ij}a_{kl}z_{1i}z_{1j}z_{1k}z_{1l}\right)^2\right]. \qquad (1.1)
\end{aligned}
$$

The each term on right hand side in (1.1) is evaluated as

$$
\begin{aligned}
\mathrm{E}\left[\left(\sum_{\substack{i\neq j\neq k\neq l \\ j\neq l\neq i\neq k}}^m a_{ij}a_{kl}z_{1i}z_{1j}z_{1k}z_{1l}\right)^2\right] &= \sum_{\substack{i\neq j\neq k\neq l \\ j\neq l\neq i\neq k}}^m (8a_{ij}^2 a_{kl}^2 + 16a_{ij}a_{kl}a_{ik}a_{jl})\mathrm{E}[z_{1i}^2 z_{1j}^2 z_{1k}^2 z_{1l}^2] \\
&\leq 24\sum_{\substack{i\neq j\neq k\neq l \\ j\neq l\neq i\neq k}}^m a_{ij}^2 a_{kl}^2 \\
&\leq 24\left(\sum_{i,j=1}^m a_{ij}^2\right)^2, \qquad (1.2) \\
& \qquad (1.3)
\end{aligned}
$$

$$
\begin{aligned}
\mathrm{E}\left[\left(\sum_{i\neq j\neq k\neq i}^{m} a_{ij}a_{ik}z_{1i}^2 z_{1j}z_{1k}\right)^2\right] &= 2\sum_{i\neq j\neq k\neq i}^{m} a_{ij}^2 a_{ik}^2 \mathrm{E}[z_{1i}^4 z_{1j}^2 z_{1k}^2] \\
&\quad +4\sum_{i\neq j\neq k\neq i}^{m} a_{ij}a_{ik}a_{ji}a_{jk}\mathrm{E}[z_{1i}^3 z_{1j}^3 z_{1k}^2] \\
&\quad +2\sum_{\substack{i\neq j\neq k\neq l\\ j\neq l\neq i\neq k}}^{m} a_{ij}a_{ik}a_{lj}a_{lk}\mathrm{E}[z_{1i}^2 z_{1j}^2 z_{1k}^2 z_{1l}^2] \\
&\le 2(\kappa_4+3)\sum_{i\neq j\neq k\neq i}^{m} a_{ij}^2 a_{ik}^2 + 4\kappa_3^2 \sum_{i\neq j\neq k\neq i}^{m} a_{ij}a_{ik}a_{ji}a_{jk} \\
&\quad +2\sum_{\substack{i\neq j\neq k\neq l\\ j\neq l\neq i\neq k}}^{m} a_{ij}a_{ik}a_{lj}a_{lk} \\
&\le (2\kappa_4+4\kappa_3^2+8)\left(\sum_{i,j=1}^{m} a_{ij}^2\right)^2, \qquad (1.4)
\end{aligned}
$$

and

$$
\begin{aligned}
\mathrm{E}\left[\left(\sum_{i\neq j}^{m} a_{ij}^2 z_{1i}^2 z_{1j}^2\right)^2\right] &= 2\sum_{i\neq j}^{m} a_{ij}^4 \mathrm{E}[z_{1i}^4 z_{1j}^4] + 4\sum_{i\neq j\neq k\neq i}^{m} a_{ij}^2 a_{ik}^2 \mathrm{E}[z_{1i}^4 z_{1j}^2 z_{1k}^2] \\
&\quad + \sum_{\substack{i\neq j\neq k\neq l\\ j\neq l\neq i\neq k}}^{m} a_{ij}^2 a_{kl}^2 \mathrm{E}[z_{1i}^2 z_{1j}^2 z_{1k}^2 z_{1l}^2] \\
&\le \left\{2(\kappa_4+3)^2 + 4(\kappa_4+3)+1\right\}\left(\sum_{i,j=1}^{m} a_{ij}^2\right)^2. \qquad (1.5)
\end{aligned}
$$

Combining (1.1)-(1.4), we obtain (iv).

2 Proof of Lemma A.2

First, we show (i). By applying Cauchy-Schwarz inequality

$$\begin{aligned}
\operatorname{tr}\left(\sum_{g\neq h}^{q} A^{(g)}\odot A^{(h)}\right)^2 &= \sum_{i,j=1}^{m}\left(\sum_{g\neq h}^{q} a_{ij}^{(g)}a_{ij}^{(h)}\right)^2 \\
&\le \sum_{\substack{g,h,g',h'=1\\ g\neq h,g'\neq h'}}^{q}\left(\sum_{i,j=1}^{m} a_{ij}^{(g)^2}a_{ij}^{(h)^2}\right)^{1/2}\left(\sum_{i,j=1}^{m} a_{ij}^{(g')^2}a_{ij}^{(h')^2}\right)^{1/2} \\
&\le \sum_{\substack{g,h,g',h'=1\\ g\neq h,g'\neq h'}}^{q}\left(\sum_{i,j=1}^{m} a_{ij}^{(g)^4}\right)^{1/4}\left(\sum_{i,j=1}^{m} a_{ij}^{(h)^4}\right)^{1/4} \\
&\qquad \left(\sum_{i,j=1}^{m} a_{ij}^{(g')^4}\right)^{1/4}\left(\sum_{i,j=1}^{m} a_{ij}^{(h')^4}\right)^{1/4} \\
&\le \sum_{\substack{g,h,g',h'=1\\ g\neq h,g'\neq h'}}^{q}\left(\operatorname{tr}A^{(g)^4}\right)^{1/4}\left(\operatorname{tr}A^{(h)^4}\right)^{1/4}\left(\operatorname{tr}A^{(g')^4}\right)^{1/4}\left(\operatorname{tr}A^{(h')^4}\right)^{1/4} \\
&= \left(\sum_{g\neq h}^{q}(\operatorname{tr}\Sigma_{gg}^4)^{1/4}(\operatorname{tr}\Sigma_{hh}^4)^{1/4}\right)^2. \qquad (2.1)
\end{aligned}$$

On the other hand, we note that

$$\begin{aligned}
\operatorname{tr}\left(\sum_{g\neq h}^{q} A^{(g)}\odot A^{(h)}\right)^2 &= \sum_{i,j=1}^{m}\sum_{\substack{g,h,g',h'=1\\ g\neq h,g'\neq h'}}^{q} a_{ij}^{(g)}a_{ij}^{(h)}a_{ij}^{(g')}a_{ij}^{(h')} \\
&= \sum_{i,j=1}^{m} a_{ij}^4 - 2\sum_{i,j=1}^{m}\sum_{g=1}^{m} a_{ij}^2 a_{ij}^{(g)^2} + \sum_{i,j=1}^{m}\left(\sum_{g=1}^{q} a_{ij}^{(g)^2}\right)^2 \\
&\le 2\operatorname{tr}\Sigma^4, \qquad (2.2)
\end{aligned}$$

since

$$\sum_{i,j=1}^{m} a_{ij}^4 = \operatorname{tr}(A\odot A)^2 \le \operatorname{tr}(A^2\odot A^2) \le \operatorname{tr}\Sigma^4,$$

$$\begin{aligned}
\sum_{i,j=1}^{m}\left(\sum_{g=1}^{q} a_{ij}^{(g)^2}\right)^2 &= \sum_{g,h=1}^{q}\operatorname{tr}\left((A^{(g)}\odot A^{(g)})(A^{(h)}\odot A^{(h)})\right) \\
&\le \sum_{g,h=1}^{q}\operatorname{tr}(A^{(g)^2}\odot A^{(h)^2}) = \operatorname{tr}(\Gamma'\Sigma_d\Gamma\odot\Gamma'\Sigma_d\Gamma) \\
&\le \operatorname{tr}(\Sigma\Sigma_d)^2 \le \operatorname{tr}\Sigma^2\Sigma_d^2 \le (\operatorname{tr}\Sigma^4)^{1/2}(\operatorname{tr}\Sigma_d^4)^{1/2} = \operatorname{tr}\Sigma^4.
\end{aligned}$$

From (2.1) and (2.2), we obtain

$$\mathrm{tr}\left(\sum_{g\neq h}^{q} A^{(g)} \odot A^{(h)}\right)^2 \le \min\left\{\left(\sum_{g\neq h}^{q} (\mathrm{tr}\Sigma_{gg}^4)^{1/4}(\mathrm{tr}\Sigma_{hh}^4)^{1/4}\right)^2,\ 2\mathrm{tr}\Sigma^4\right\}.$$

Second, we show (ii). By applying Cauchy-Schwarz inequality

$$\begin{aligned}\sum_{\substack{g,h,g',h'=1\\ g\neq h, g'\neq h'}}^{q} \mathrm{tr}A^{(g)}A^{(g')}A^{(h)}A^{(h')} &\le \sum_{\substack{g,h,g',h'=1\\ g\neq h, g'\neq h'}}^{q} (\mathrm{tr}A^{(g)^2}A^{(g')^2})^{1/2}(\mathrm{tr}A^{(h)^2}A^{(h')^2})^{1/2}\\ &\le \left(\sum_{g\neq h}^{q} (\mathrm{tr}\Sigma_{gg}^4)^{1/4}(\mathrm{tr}\Sigma_{hh}^4)^{1/4}\right)^2. \qquad (2.3)\end{aligned}$$

On the other hand, we note that

$$\begin{aligned}\sum_{\substack{g,h,g',h'=1\\ g\neq h, g'\neq h'}}^{q} \mathrm{tr}A^{(g)}A^{(g')}A^{(h)}A^{(h')} &= \mathrm{tr}A^4 - 2\sum_{g=1}^{q}\mathrm{tr}(A^{(g)}A)^2 + \sum_{g,h=1}^{q}\mathrm{tr}(A^{(g)}A^{(h)})^2\\ &\le \mathrm{tr}A^4 + \sum_{g,h=1}^{q}\mathrm{tr}(A^{(g)}A^{(h)})^2\\ &\le 2\mathrm{tr}\Sigma^4, \qquad (2.4)\end{aligned}$$

since

$$\sum_{g,h=1}^{q}\mathrm{tr}(A^{(g)}A^{(h)})^2 \le \sum_{g,h=1}^{q}\mathrm{tr}A^{(g)^2}A^{(h)^2} = \mathrm{tr}(\Sigma\Sigma_d)^2 \le \mathrm{tr}\Sigma^4.$$

From (2.3) and (2.4), we obtain

$$\sum_{\substack{g,h,g',h'=1\\ g\neq h, g'\neq h'}}^{q} \mathrm{tr}A^{(g)}A^{(g')}A^{(h)}A^{(h')} \le \min\left\{\left(\sum_{g\neq h}^{q} (\mathrm{tr}\Sigma_{gg}^4)^{1/4}(\mathrm{tr}\Sigma_{hh}^4)^{1/4}\right)^2,\ 2\mathrm{tr}\Sigma^4\right\}.$$

Third, we show (iii). By applying Cauchy-Schwarz inequality

$$\begin{aligned}
\sum_{\substack{g,h,g',h'=1 \\ g\neq g', h\neq h'}}^{q} \operatorname{tr}\left(\left(A^{(g)}A^{(h)}\right)\odot\left(A^{(g')}A^{(h')}\right)\right) &= \sum_{\substack{g_1,h_1,g_2,h_2=1 \\ g_1\neq h_1, g_2\neq h_2}}^{q} \sum_{i=1}^{m}\left(\sum_{j=1}^{m} a_{ij}^{(g_1)}a_{ij}^{(g_2)}\right)\left(\sum_{j=1}^{m} a_{ij}^{(h_1)}a_{ij}^{(h_2)}\right) \\
&\le \sum_{\substack{g_1,h_1,g_2,h_2=1 \\ g_1\neq h_1, g_2\neq h_2}}^{q} \left(\sum_{i=1}^{m}\left(\sum_{j=1}^{m} a_{ij}^{(g_1)}a_{ij}^{(g_2)}\right)^2\right)^{1/2} \\
&\quad \times \left(\sum_{i=1}^{m}\left(\sum_{j=1}^{m} a_{ij}^{(h_1)}a_{ij}^{(h_2)}\right)^2\right)^{1/2} \\
&\le \sum_{\substack{g_1,h_1,g_2,h_2=1 \\ g_1\neq h_1, g_2\neq h_2}}^{q} \left(\sum_{i,j=1}^{m} a_{ij}^{(g_1)^2}a_{ij}^{(g_2)^2}\right)^{1/2} \\
&\quad \times \left(\sum_{i,j=1}^{m} a_{ij}^{(h_1)^2}a_{ij}^{(h_2)^2}\right)^{1/2} \\
&\le \left(\sum_{g\neq h}^{q} (\operatorname{tr}\Sigma_{gg}^4)^{1/4}(\operatorname{tr}\Sigma_{hh}^4)^{1/4}\right)^2. \qquad (2.5)
\end{aligned}$$

On the other hand, we note that

$$\begin{aligned}
\sum_{\substack{g,h,g',h'=1 \\ g\neq g', h\neq h'}}^{q} \operatorname{tr}\left(\left(A^{(g)}A^{(h)}\right)\odot\left(A^{(g')}A^{(h')}\right)\right) &= \operatorname{tr}(A^2\odot A^2) - 2\sum_{g=1}^{q}\operatorname{tr}\left((A^{(g)}A)\odot(A^{(g)}A)\right) \\
&\quad + \sum_{g,h=1}^{q}\operatorname{tr}\left((A^{(g)}A^{(h)})\odot(A^{(g)}A^{(h)})\right) \\
&\le \operatorname{tr}(A^2\odot A^2) + \sum_{g,h=1}^{q}\operatorname{tr}\left((A^{(g)}A^{(h)})\odot(A^{(g)}A^{(h)})\right) \\
&\le 2\operatorname{tr}\Sigma^4, \qquad (2.6)
\end{aligned}$$

since

$$\operatorname{tr}(A^2\odot A^2) \le \operatorname{tr}\Sigma^4,$$
$$\sum_{g,h=1}^{q}\operatorname{tr}\left((A^{(g)}A^{(h)})\odot(A^{(g)}A^{(h)})\right) \le \sum_{g,h=1}^{q}\operatorname{tr}A^{(g)^2}A^{(h)^2} = \operatorname{tr}(\Sigma\Sigma_d)^2 \le \operatorname{tr}\Sigma^4.$$

From (2.5) and (2.6), we obtain

$$\sum_{\substack{g,h,g',h'=1 \\ g\neq g', h\neq h'}}^{q} \operatorname{tr}\left(\left(A^{(g)}A^{(h)}\right)\odot\left(A^{(g')}A^{(h')}\right)\right) \le \min\left\{\left(\sum_{g\neq h}^{q}(\operatorname{tr}\Sigma_{gg}^4)^{1/4}(\operatorname{tr}\Sigma_{hh}^4)^{1/4}\right)^2,\ 2\operatorname{tr}\Sigma^4\right\}.$$

Fourth, we show (iv). By applying Cauchy-Schwarz inequality

$$\begin{aligned}
\left|\sum_{\substack{g,h,g',h'=1\\ g\neq h,g'\neq h'}}^{q} \mathrm{tr}\left((A^{(g)}\odot A^{(g')})A^{(h')}A^{(h)}\right)\right| &\leq \sum_{\substack{g,h,g',h'=1\\ g\neq h,g'\neq h'}}^{q} \left(\mathrm{tr}(A^{(g)}\odot A^{(g')})^2\right)^{1/2}\left(\mathrm{tr}A^{(h')^2}A^{(h)^2}\right)^{1/2} \\
&\leq \sum_{\substack{g,h,g',h'=1\\ g\neq h,g'\neq h'}}^{q} \left(\mathrm{tr}(A^{(g)^2}\odot A^{(g')^2})\right)^{1/2} \\
&\quad \times\left((\mathrm{tr}A^{(h')^4})^{1/2}(\mathrm{tr}A^{(h)^4})^{1/2}\right)^{1/2} \\
&\leq \sum_{\substack{g,h,g',h'=1\\ g\neq h,g'\neq h'}}^{q} \left((\mathrm{tr}A^{(g)^4})^{1/2}(\mathrm{tr}A^{(g')^4})^{1/2}\right)^{1/2} \\
&\quad \times\left((\mathrm{tr}A^{(h')^4})^{1/2}(\mathrm{tr}A^{(h)^4})^{1/2}\right)^{1/2} \\
&= \left(\sum_{g\neq h}^{q}(\mathrm{tr}\Sigma_{gg}^4)^{1/4}(\mathrm{tr}\Sigma_{hh}^4)^{1/4}\right)^2. \qquad (2.7)
\end{aligned}$$

On the other hand, we note that

$$\begin{aligned}
\left|\sum_{\substack{g,h,g',h'=1\\ g\neq h,g'\neq h'}}^{q} \mathrm{tr}\left((A^{(g)}\odot A^{(g')})A^{(h')}A^{(h)}\right)\right| &= \Bigg|\mathrm{tr}\left((A\odot A)A^2\right)-2\sum_{g=1}^{q}\mathrm{tr}\left((A^{(g)}\odot A)AA^{(g)}\right) \\
&\quad \sum_{g,h=1}^{q}\mathrm{tr}\left((A^{(g)}\odot A^{(h)})A^{(h)}A^{(g)}\right)\Bigg| \\
&\leq 4\mathrm{tr}\Sigma^4, \qquad (2.8)
\end{aligned}$$

since

$$\begin{aligned}
\left|\mathrm{tr}\left((A\odot A)A^2\right)\right| &\le \left(\mathrm{tr}(A\odot A)^2\right)^{1/2}(\mathrm{tr}A^4)^{1/2}\le \mathrm{tr}\Sigma^4,\\
\left|\sum_{g=1}^q \mathrm{tr}\left((A^{(g)}\odot A)AA^{(g)}\right)\right| &= \left|\sum_{g=1}^q\sum_{i,j,k=1}^m a_{ij}^{(g)}a_{ij}a_{jk}a_{ki}^{(g)}\right|\\
&\le \left(\sum_{g=1}^q\sum_{i,j=1}^m a_{ij}^{(g)^2}a_{ij}^2\right)^{1/2}\left(\sum_{g=1}^q\sum_{i,j=1}^m(\sum_{k=1}^m a_{jk}a_{ki}^{(g)})^2\right)^{1/2}\\
&\le \left(\sum_{i,j=1}^m a_{ij}^4\right)^{1/4}\left(\sum_{i,j=1}^m(\sum_{g=1}^q a_{ij}^{(g)^2})^2\right)^{1/4}\left(\sum_{g=1}^q \mathrm{tr}A^{(g)^2}A^2\right)^{1/2}\\
&\le \left(\mathrm{tr}\Sigma^4\right)^{1/4}(\mathrm{tr}(\Sigma\Sigma_d)^2)^{1/4}(\mathrm{tr}\Sigma^3\Sigma_d)^{1/2}\\
&\le \mathrm{tr}\Sigma^4,\\
\left|\sum_{g,h=1}^q \mathrm{tr}\left((A^{(g)}\odot A^{(h)})A^{(h)}A^{(g)}\right)\right| &= \left|\sum_{g,h=1}^q\sum_{i,j,k=1}^m a_{ij}^{(g)}a_{ij}^{(h)}a_{jk}^{(h)}a_{ki}^{(g)}\right|\\
&= \left|\sum_{g,h=1}^q\sum_{i,j=1}^m a_{ij}^{(g)}a_{ij}^{(h)}\sum_{k=1}^m a_{jk}^{(h)}a_{ki}^{(g)}\right|\\
&\le \left(\sum_{g,h=1}^q\sum_{i,j=1}^m a_{ij}^{(g)^2}a_{ij}^{(h)^2}\right)^{1/2}\left(\sum_{g,h=1}^q\sum_{i,j=1}^m(\sum_{k=1}^m a_{jk}^{(h)}a_{ki}^{(g)})^2\right)^{1/2}\\
&\le \mathrm{tr}(\Sigma\Sigma_d)^2\le \mathrm{tr}\Sigma^4.
\end{aligned}$$

From (2.7) and (2.8), we obtain

$$\left|\sum_{\substack{g,h,g',h'=1\\ g\neq h,g'\neq h'}}^q \mathrm{tr}\left((A^{(g)}\odot A^{(g')})A^{(h')}A^{(h)}\right)\right|\le \min\left\{\left(\sum_{g\neq h}^q(\mathrm{tr}\Sigma_{gg}^4)^{1/4}(\mathrm{tr}\Sigma_{hh}^4)^{1/4}\right)^2,\ 4\mathrm{tr}\Sigma^4\right\}.$$

Fifth, we show (v). Note that

$$\begin{aligned}
&\mathrm{tr}\left(\{(A^{(g_1)}A^{(g_3)})\odot(A^{(h_1)}A^{(h_3)})\}\{(A^{(g_2)}A^{(g_4)})\odot(A^{(h_2)}A^{(h_4)})\}\right)\\
&\le \left(\mathrm{tr}\left(\{(A^{(g_1)}A^{(g_3)})\odot(A^{(h_1)}A^{(h_3)})\}\{(A^{(g_3)}A^{(g_1)})\odot(A^{(h_3)}A^{(h_1)})\}\right)\right)^{1/2}\\
&\times\left(\mathrm{tr}\left(\{(A^{(g_2)}A^{(g_4)})\odot(A^{(h_2)}A^{(h_4)})\}\{(A^{(g_4)}A^{(g_2)})\odot(A^{(h_4)}A^{(h_2)})\}\right)\right)^{1/2}\\
&\le \left(\mathrm{tr}(A^{(g_1)^2}A^{(g_3)^2})\mathrm{tr}(A^{(h_1)^2}A^{(h_3)^2})\right)^{1/2}\left(\mathrm{tr}(A^{(g_2)^2}A^{(g_4)^2})\mathrm{tr}(A^{(h_2)^2}A^{(h_4)^2})\right)^{1/2}\\
&\le \prod_{t=1}^4\left(\mathrm{tr}A^{(g_t)^4}\right)^{1/4}\left(\mathrm{tr}A^{(h_t)^4}\right)^{1/4}.
\end{aligned}$$

Hence,

$$\sum_{\substack{g_1\neq h_1,g_2\neq h_2\\ g_3\neq h_3,g_4\neq h_4}}^{q} \operatorname{tr}\left(\{(A^{(g_1)}A^{(g_3)})\odot(A^{(h_1)}A^{(h_3)})\}\{(A^{(g_2)}A^{(g_4)})\odot(A^{(h_2)}A^{(h_4)})\}\right)$$
$$\leq \left(\sum_{g\neq h}^{q}(\operatorname{tr}\Sigma_{gg}^4)^{1/4}(\operatorname{tr}\Sigma_{hh}^4)^{1/4}\right)^4. \tag{2.9}$$

By applying Holder's inequality,

$$\sum_{\substack{g_1\neq h_1,g_2\neq h_2\\ g_3\neq h_3,g_4\neq h_4}}^{q} \operatorname{tr}\left(\{(A^{(g_1)}A^{(g_3)})\odot(A^{(h_1)}A^{(h_3)})\}\{(A^{(g_2)}A^{(g_4)})\odot(A^{(h_2)}A^{(h_4)})\}\right)$$
$$= \sum_{k_1,k_2,k_3,k_4=1}^{m}\left(\sum_{\ell=1}^{m}a_{k_1\ell}a_{k_2\ell}a_{k_3\ell}a_{k_4\ell} - 2\sum_{\ell=1}^{m}\sum_{g=1}^{q}a_{k_1\ell}^{(g)}a_{k_2\ell}^{(g)}a_{k_3\ell}a_{k_4\ell}\right.$$
$$\left. + \sum_{\ell=1}^{m}\sum_{g,h=1}^{g}a_{k_1\ell}^{(g)}a_{k_2\ell}^{(g)}a_{k_3\ell}^{(h)}a_{k_4\ell}^{(h)}\right)^2$$
$$\leq 3\sum_{k_1,k_2,k_3,k_4=1}^{m}\left(\sum_{\ell=1}^{m}a_{k_1\ell}a_{k_2\ell}a_{k_3\ell}a_{k_4\ell}\right)^2 + 12\sum_{k_1,k_2,k_3,k_4=1}^{m}\left(\sum_{\ell=1}^{m}\sum_{g=1}^{q}a_{k_1\ell}^{(g)}a_{k_2\ell}^{(g)}a_{k_3\ell}a_{k_4\ell}\right)^2$$
$$+ 3\sum_{k_1,k_2,k_3,k_4=1}^{m}\left(\sum_{\ell=1}^{m}\sum_{g,h=1}^{g}a_{k_1\ell}^{(g)}a_{k_2\ell}^{(g)}a_{k_3\ell}^{(h)}a_{k_4\ell}^{(h)}\right)^2 \tag{2.10}$$

We also note that

$$\sum_{k_1,k_2,k_3,k_4=1}^{m}\left(\sum_{\ell=1}^{m}a_{k_1\ell}a_{k_2\ell}a_{k_3\ell}a_{k_4\ell}\right)^2 = \operatorname{tr}(A^2\odot A^2)^2 \leq \operatorname{tr}\Sigma^8 \leq \left(\operatorname{tr}\Sigma^4\right)^2, \tag{2.11}$$

$$\begin{aligned}\sum_{k_1,k_2,k_3,k_4=1}^{m}\left(\sum_{\ell=1}^{m}\sum_{g=1}^{q}a_{k_1\ell}^{(g)}a_{k_2\ell}^{(g)}a_{k_3\ell}a_{k_4\ell}\right)^2 &= \operatorname{tr}\left((A^2\odot A^2)\left(\sum_{g,h=1}^{q}(A^{(g)}A^{(h)})\odot(A^{(g)}A^{(h)})\right)\right)\\ &\leq \left(\operatorname{tr}(A^2\odot A^2)^2\right)^{1/2}\\ &\quad\times\left(\operatorname{tr}\left(\sum_{g,h=1}^{q}(A^{(g)}A^{(h)})\odot(A^{(g)}A^{(h)})\right)^2\right)^{1/2}\\ &\leq \operatorname{tr}\Sigma^4\operatorname{tr}(\Sigma\Sigma_d)^2 \leq \left(\operatorname{tr}\Sigma^4\right)^2\end{aligned} \tag{2.12}$$

$$\begin{aligned}\sum_{k_1,k_2,k_3,k_4=1}^{m}\left(\sum_{\ell=1}^{m}\sum_{g,h=1}^{g}a_{k_1\ell}^{(g)}a_{k_2\ell}^{(g)}a_{k_3\ell}^{(h)}a_{k_4\ell}^{(h)}\right)^2 &= \operatorname{tr}\left(\sum_{g,h=1}^{q}(A^{(g)}A^{(h)})\odot(A^{(g)}A^{(h)})\right)^2\\ &\leq \sum_{g_1,h_1,g_2,h_2=1}^{q}\operatorname{tr}(A^{(g_1)^2}A^{(g_2)^2})\odot(A^{(h_1)^2}A^{(h_2)^2})\\ &\leq \operatorname{tr}(\Sigma\Sigma_d)^4 \leq (\operatorname{tr}(\Sigma\Sigma_d)^2)^2 \leq \left(\operatorname{tr}\Sigma^4\right)^2.\end{aligned} \tag{2.13}$$

Combining (2.10)-(2.13), we obtain

$$\sum_{\substack{g_1\neq h_1,g_2\neq h_2\\ g_3\neq h_3,g_4\neq h_4}}^{q} \mathrm{tr}\left(\{(A^{(g_1)}A^{(g_3)})\odot(A^{(h_1)}A^{(h_3)})\}\{(A^{(g_2)}A^{(g_4)})\odot(A^{(h_2)}A^{(h_4)})\}\right) \leq 18\left(\mathrm{tr}\Sigma^4\right)^2. \quad (2.14)$$

From (2.9) and (2.14), we obtain

$$\begin{aligned}\sum_{\substack{g_1\neq h_1,g_2\neq h_2\\ g_3\neq h_3,g_4\neq h_4}}^{q} & \mathrm{tr}\left(\{(A^{(g_1)}A^{(g_3)})\odot(A^{(h_1)}A^{(h_3)})\}\{(A^{(g_2)}A^{(g_4)})\odot(A^{(h_2)}A^{(h_4)})\}\right)\\ &\leq \min\left\{\left(\sum_{g\neq h}^{q}(\mathrm{tr}\Sigma_{gg}^4)^{1/4}(\mathrm{tr}\Sigma_{hh}^4)^{1/4}\right)^4, 18\left(\mathrm{tr}\Sigma^4\right)^2\right\}.\end{aligned}$$

Sixth, we show (vi). Note that

$$\begin{aligned}&\mathrm{tr}(A^{(g_1)}A^{(g_2)}A^{(g_3)}A^{(g_4)})\mathrm{tr}(A^{(h_1)}A^{(h_2)}A^{(h_3)}A^{(h_4)})\\ &\leq\left(\mathrm{tr}(A^{(g_1)^2}A^{(g_2)^2})\mathrm{tr}(A^{(g_3)^2}A^{(g_4)^2})\right)^{1/2}\left(\mathrm{tr}(A^{(h_1)^2}A^{(h_2)^2})\mathrm{tr}(A^{(h_3)^2}A^{(h_4)^2})\right)^{1/2}\\ &\leq\prod_{t=1}^{4}\left(\mathrm{tr}A^{(g_t)^4}\right)^{1/4}\left(\mathrm{tr}A^{(h_t)^4}\right)^{1/4}.\end{aligned}$$

Hence,

$$\sum_{\substack{g_1\neq h_1,g_2\neq h_2\\ g_3\neq h_3,g_4\neq h_4}}^{q}\mathrm{tr}(A^{(g_1)}A^{(g_2)}A^{(g_3)}A^{(g_4)})\mathrm{tr}(A^{(h_1)}A^{(h_2)}A^{(h_3)}A^{(h_4)}) \leq\left(\sum_{g\neq h}^{q}(\mathrm{tr}\Sigma_{gg}^4)^{1/4}(\mathrm{tr}\Sigma_{hh}^4)^{1/4}\right)^4. \quad (2.15)$$

By applying Holder's inequality,

$$\begin{aligned}&\sum_{\substack{g_1\neq h_1,g_2\neq h_2\\ g_3\neq h_3,g_4\neq h_4}}^{q}\mathrm{tr}(A^{(g_1)}A^{(g_2)}A^{(g_3)}A^{(g_4)})\mathrm{tr}(A^{(h_1)}A^{(h_2)}A^{(h_3)}A^{(h_4)})\\ &=\sum_{k_1,k_2,k_3,k_4=1}^{m}\left(\sum_{\ell_1,\ell_2=1}^{m}a_{k_1\ell_1}a_{k_2\ell_2}a_{k_3\ell_1}a_{k_4\ell_2}-2\sum_{\ell_1,\ell_2=1}^{m}\sum_{g=1}^{q}a_{k_1\ell_1}^{(g)}a_{k_2\ell_2}^{(g)}a_{k_3\ell_1}a_{k_4\ell_2}\right.\\ &\qquad\left.+\sum_{\ell_1,\ell_2=1}^{m}\sum_{g,h=1}^{g}a_{k_1\ell_1}^{(g)}a_{k_2\ell_2}^{(g)}a_{k_3\ell_1}^{(h)}a_{k_4\ell_2}^{(h)}\right)^2\\ &\leq 3\sum_{k_1,k_2,k_3,k_4=1}^{m}\left(\sum_{\ell_1,\ell_2=1}^{m}a_{k_1\ell_1}a_{k_2\ell_2}a_{k_3\ell_1}a_{k_4\ell_2}\right)^2+12\sum_{k_1,k_2,k_3,k_4=1}^{m}\left(\sum_{\ell_1,\ell_2=1}^{m}\sum_{g=1}^{q}a_{k_1\ell_1}^{(g)}a_{k_2\ell_2}^{(g)}a_{k_3\ell_1}a_{k_4\ell_2}\right)^2\\ &\quad+3\sum_{k_1,k_2,k_3,k_4=1}^{m}\left(\sum_{\ell_1,\ell_2=1}^{m}\sum_{g,h=1}^{g}a_{k_1\ell_1}^{(g)}a_{k_2\ell_2}^{(g)}a_{k_3\ell_1}^{(h)}a_{k_4\ell_2}^{(h)}\right)^2. \quad (2.16)\end{aligned}$$

We also note that

$$\sum_{k_1,k_2,k_3,k_4=1}^{m}\left(\sum_{\ell_1,\ell_2=1}^{m} a_{k_1\ell_1}a_{k_2\ell_2}a_{k_3\ell_1}a_{k_4\ell_2}\right)^2 = (\mathrm{tr}\Sigma^4)^2, \tag{2.17}$$

$$\begin{aligned}\sum_{k_1,k_2,k_3,k_4=1}^{m}\left(\sum_{\ell_1,\ell_2=1}^{m}\sum_{g=1}^{q} a^{(g)}_{k_1\ell_1}a^{(g)}_{k_2\ell_2}a_{k_3\ell_1}a_{k_4\ell_2}\right)^2 &= \sum_{g,h=1}^{q}\left(\mathrm{tr}(A^2A^{(g)}A^{(h)})\right)^2\\ &\le \sum_{g,h=1}^{q}\mathrm{tr}(A^4\mathrm{tr}A^{(g)^2}A^{(h)^2})\\ &= \mathrm{tr}\Sigma^4\mathrm{tr}(\Sigma\Sigma_d)^2 \le \left(\mathrm{tr}\Sigma^4\right)^2 \end{aligned} \tag{2.18}$$

$$\begin{aligned}\sum_{k_1,k_2,k_3,k_4=1}^{m}\left(\sum_{\ell_1,\ell_2=1}^{m}\sum_{g,h=1}^{g} a^{(g)}_{k_1\ell_1}a^{(g)}_{k_2\ell_2}a^{(h)}_{k_3\ell_1}a^{(h)}_{k_4\ell_2}\right)^2 &= \sum_{g_1,g_2,h_1,h_2=1}^{q}\left(\mathrm{tr}(A^{(g_1)}A^{(h_1)}A^{(h_2)}A^{(g_2)})\right)^2\\ &\le \left(\sum_{g_1,g_2=1}^{q}\mathrm{tr}(A^{(g_1)^2}A^{(g_2)^2})\right)^2\\ &= \left(\mathrm{tr}(\Sigma\Sigma_d)^2\right)^2 \le \left(\mathrm{tr}\Sigma^4\right)^2. \end{aligned} \tag{2.19}$$

Combining (2.16)-(2.19), we obtain

$$\sum_{\substack{g_1\neq h_1,g_2\neq h_2\\ g_3\neq h_3,g_4\neq h_4}}^{q}\mathrm{tr}(A^{(g_1)}A^{(g_2)}A^{(g_3)}A^{(g_4)})\mathrm{tr}(A^{(h_1)}A^{(h_2)}A^{(h_3)}A^{(h_4)}) \le 18\left(\mathrm{tr}\Sigma^4\right)^2. \tag{2.20}$$

From (2.15) and (2.20), we obtain

$$\begin{aligned}&\sum_{\substack{g_1\neq h_1,g_2\neq h_2\\ g_3\neq h_3,g_4\neq h_4}}^{q}\mathrm{tr}(A^{(g_1)}A^{(g_2)}A^{(g_3)}A^{(g_4)})\mathrm{tr}(A^{(h_1)}A^{(h_2)}A^{(h_3)}A^{(h_4)})\\ &\le \min\left\{\left(\sum_{g\neq h}^{q}(\mathrm{tr}\Sigma_{gg}^4)^{1/4}(\mathrm{tr}\Sigma_{hh}^4)^{1/4}\right)^4, 18\left(\mathrm{tr}\Sigma^4\right)^2\right\}.\end{aligned}$$

Seventh, we show (vii). By applying Cauchy-Schwarz inequality

$$
\begin{aligned}
&\mathrm{tr}\left(\left\{\sum_{g\neq h}^{q} A^{(g)}\Gamma'(\Sigma-\Sigma_d)\Gamma A^{(h)}\right\}^2\right)\\
&= \sum_{g_1\neq h_1, g_2\neq h_2}^{q} \mathrm{tr}\left(A^{(h_2)}A^{(g_1)}\Gamma'(\Sigma-\Sigma_d)\Gamma A^{(h_1)}A^{(g_2)}\Gamma'(\Sigma-\Sigma_d)\Gamma\right)\\
&\leq \sum_{g_1\neq h_1, g_2\neq h_2}^{q} \left(\mathrm{tr}\left(A^{(g_1)}A^{(h_2)^2}A^{(g_1)}(\Gamma'(\Sigma-\Sigma_d)\Gamma)^2\right)\right)^{1/2}\left(\mathrm{tr}\left(A^{(g_2)}A^{(h_1)^2}A^{(g_2)}(\Gamma'(\Sigma-\Sigma_d)\Gamma)^2\right)\right)^{1/2}\\
&\leq \sum_{g_1\neq h_1, g_2\neq h_2}^{q} \left(\mathrm{tr}(A^{(g_1)^2}A^{(h_2)^2})^2\mathrm{tr}(\Gamma'(\Sigma-\Sigma_d)\Gamma)^4\right)^{1/4}\left(\mathrm{tr}(A^{(g_2)^2}A^{(h_1)^2})^2\mathrm{tr}(\Gamma'(\Sigma-\Sigma_d)\Gamma)^4\right)^{1/4}\\
&\leq \mathrm{tr}(\Gamma'(\Sigma-\Sigma_d)\Gamma)^2 \sum_{g_1\neq h_1, g_2\neq h_2}^{q} \mathrm{tr}(A^{(g_1)^4}A^{(h_2)^4})^{1/4}\mathrm{tr}(A^{(g_2)^4}A^{(h_1)^4})^{1/4}\\
&\leq \mathrm{tr}(\Gamma'(\Sigma-\Sigma_d)\Gamma)^2\left(\sum_{g\neq h}^{q}(\mathrm{tr}\Sigma_{gg}^4)^{1/4}(\mathrm{tr}\Sigma_{hh}^4)^{1/4}\right)^2 \qquad (2.21)
\end{aligned}
$$

On the other hand, we note that

$$
\begin{aligned}
\mathrm{tr}\left(\left\{\sum_{g\neq h}^{q} A^{(g)}\Gamma'(\Sigma-\Sigma_d)\Gamma A^{(h)}\right\}^2\right) &= \mathrm{tr}(A\Gamma'(\Sigma-\Sigma_d)\Gamma A)^2\\
&\quad -2\mathrm{tr}(A\Gamma'(\Sigma-\Sigma_d)\Gamma A)\left(\sum_{g=1}^{q} A^{(g)}\Gamma'(\Sigma-\Sigma_d)\Gamma A^{(g)}\right)\\
&\quad +\mathrm{tr}\left(\sum_{g=1}^{q} A^{(g)}\Gamma'(\Sigma-\Sigma_d)\Gamma A^{(g)}\right)^2\\
&\leq 4\mathrm{tr}\Sigma^4\mathrm{tr}(\Sigma^2-\Sigma_d^2)^2, \qquad (2.22)
\end{aligned}
$$

since

$$
\begin{aligned}
\mathrm{tr}(A\Gamma'(\Sigma-\Sigma_d)\Gamma A)^2 &= \mathrm{tr}(\Sigma^3(\Sigma-\Sigma_d))^2 \leq \mathrm{tr}\Sigma^4\mathrm{tr}(\Sigma^2-\Sigma_d^2)^2,\\
\mathrm{tr}\left(\sum_{g=1}^{q} A^{(g)}\Gamma'(\Sigma-\Sigma_d)\Gamma A^{(g)}\right)^2 &= \sum_{g,h=1}^{q} \mathrm{tr}\left(A^{(g)}A^{(h)}\Gamma'(\Sigma-\Sigma_d)\Gamma\right)^2\\
&\leq \sum_{g,h=1}^{q} \mathrm{tr}A^{(g)^2}A^{(h)^2}\mathrm{tr}\left(\Sigma^2-\Sigma_d^2\right)^2\\
&= \mathrm{tr}(\Sigma\Sigma_d)^2\mathrm{tr}\left(\Sigma^2-\Sigma_d^2\right)^2 \leq \mathrm{tr}\Sigma^4\mathrm{tr}\left(\Sigma^2-\Sigma_d^2\right)^2.
\end{aligned}
$$

Combining (2.21)-(2.22), we obtain

$$\begin{aligned}
&\mathrm{tr}\left(\left\{\sum_{g\neq h}^{q} A^{(g)}\Gamma'(\Sigma-\Sigma_d)\Gamma A^{(h)}\right\}^2\right)\\
&\leq \min\left\{4\mathrm{tr}\Sigma^4, \left(\sum_{g\neq h}^{q}(\mathrm{tr}\Sigma_{gg}^4)^{1/4}(\mathrm{tr}\Sigma_{hh}^4)^{1/4}\right)^2\right\}\mathrm{tr}(\Sigma^2-\Sigma_d^2)^2
\end{aligned}$$

Finally, we show (viii). Note that

$$\begin{aligned}
&\sum_{\substack{g,h,g',h'=1\\ g\neq g',h\neq h'}}^{q}[\mathrm{diag}\{\Gamma(\Sigma-\Sigma_d)\Gamma\}]'\left\{\left(A^{(g)}A^{(h)}\right)\odot\left(A^{(g')}A^{(h')}\right)\right\}[\mathrm{diag}\{\Gamma(\Sigma-\Sigma_d)\Gamma\}]\\
&\leq \mathrm{tr}(\Sigma^2-\Sigma_d^2)^2\sum_{g\neq h,g'\neq h'}^{q}\mathrm{tr}(A^{(g)}A^{(h)})\odot(A^{(g')}A^{(h')}).
\end{aligned}$$

From above results and (i), we obtain

$$\begin{aligned}
&\sum_{\substack{g,h,g',h'=1\\ g\neq g',h\neq h'}}^{q}[\mathrm{diag}\{\Gamma(\Sigma-\Sigma_d)\Gamma\}]'\left\{\left(A^{(g)}A^{(h)}\right)\odot\left(A^{(g')}A^{(h')}\right)\right\}[\mathrm{diag}\{\Gamma(\Sigma-\Sigma_d)\Gamma\}]\\
&\leq \min\left\{2\mathrm{tr}\Sigma^4, \left(\sum_{g\neq h}^{q}(\mathrm{tr}\Sigma_{gg}^4)^{1/4}(\mathrm{tr}\Sigma_{hh}^4)^{1/4}\right)^2\right\}\mathrm{tr}(\Sigma^2-\Sigma_d^2)^2.
\end{aligned}$$