Hypothesis tests for high-dimensional covariance structures

Aki Ishii · Kazuyoshi Yata · Makoto Aoshima

Abstract We consider hypothesis testing for high-dimensional covariance structures in which the covariance matrix is a (i) scaled identity matrix, (ii) diagonal matrix, or (iii) intraclass covariance matrix. Our purpose is to systematically establish a nonparametric approach for testing the high-dimensional covariance structures (i) - (iii). We produce a new common test statistic for each covariance structure and show that the test statistic is an unbiased estimator of its corresponding test parameter. We prove that the test statistic establishes the asymptotic normality. We propose a new test procedure for (i) - (iii) and evaluate its asymptotic size and power theoretically when both the dimension and sample size increase. We investigate the performance of the proposed test procedure in simulations. As an application of testing the covariance structures, we give a test procedure to identify an eigenvector. Finally, we demonstrate the proposed test procedure by using a microarray data set.

Aki Ishii

Department of Information Sciences, Tokyo University of Science, 2641 Yamazaki, Noda-shi, Chiba 278-8510, Japan
E-mail: a.ishii@rs.tus.ac.jp

Kazuyoshi Yata

Institute of Mathematics, University of Tsukuba, 1-1-1 Tennodai, Tsukuba, Ibaraki 305-8571, Japan
E-mail: yata@math.tsukuba.ac.jp

Makoto Aoshima (Corresponding author)

Institute of Mathematics, University of Tsukuba, 1-1-1 Tennodai, Tsukuba, Ibaraki 305-8571, Japan
E-mail: aoshima@math.tsukuba.ac.jp

The online version of this article contains supplementary material.

1 Introduction

Suppose we take samples, $\boldsymbol{x}_j = (x_{1j}, \ldots, x_{pj})^T$, $j = 1, \ldots, n$, of size n (≥ 4), which are independent and identically distributed (i.i.d.) as a p (≥ 2)-variate distribution. We assume that $\boldsymbol{x}_j$ has an unknown mean vector $\boldsymbol{\mu}$ and unknown (positive-semidefinite) covariance matrix $\boldsymbol{\Sigma}$. We introduce the spectral decomposition $\boldsymbol{\Sigma} = \boldsymbol{H}\boldsymbol{\Lambda}\boldsymbol{H}^T$, where $\boldsymbol{\Lambda} = \mathrm{diag}(\lambda_1, \ldots, \lambda_p)$ is a diagonal matrix of eigenvalues of $\boldsymbol{\Sigma}$, $\lambda_1 \geq \cdots \geq \lambda_p \geq 0$, and $\boldsymbol{H} = (\boldsymbol{h}_1, \ldots, \boldsymbol{h}_p)$ is an orthogonal matrix of the corresponding eigenvectors. Let $\boldsymbol{x}_j = \boldsymbol{H}\boldsymbol{\Lambda}^{1/2}\boldsymbol{z}_j + \boldsymbol{\mu}$, where $\boldsymbol{z}_j = (z_{1j}, \ldots, z_{pj})^T$ is a random vector having zero mean and identity covariance matrix. Let $\sigma = \mathrm{tr}(\boldsymbol{\Sigma})/p$. Let σ_{ij} be the (i,j)th element of $\boldsymbol{\Sigma}$ for $i, j = 1, \ldots, p$. We assume that $\sigma_{jj} \in (0, \infty)$ as $p \to \infty$ for all j. For a function, $f(\cdot)$, "$f(p) \in (0, \infty)$ as $p \to \infty$" implies that $\liminf_{p \to \infty} f(p) > 0$ and $\limsup_{p \to \infty} f(p) < \infty$. Then, it holds that $\sigma \in (0, \infty)$ as $p \to \infty$. Let $\rho = \sum_{i \neq j}^p \sigma_{ij}/\{\sigma p(p-1)\}$. Note that

$$\frac{\boldsymbol{1}_p^T \boldsymbol{\Sigma} \boldsymbol{1}_p}{p} = \sigma\{1 + \rho(p-1)\} \tag{1}$$

and $\rho \in [-(p-1)^{-1}, 1]$, where $\boldsymbol{1}_p = (1, \ldots, 1)^T$. We denote the identity matrix of dimension p by $\boldsymbol{I}_p$.

In this paper, we consider testing

$$H_0 : \boldsymbol{\Sigma} = \boldsymbol{\Sigma}_* \quad \text{vs.} \quad H_1 : \boldsymbol{\Sigma} \neq \boldsymbol{\Sigma}_*, \tag{2}$$

where $\boldsymbol{\Sigma}_*$ is a candidate (positive-semidefinite) covariance matrix. For $\boldsymbol{\Sigma}_*$ we mainly consider the following covariance structures: (i) scaled identity matrix, (ii) diagonal matrix, and (iii) intraclass covariance matrix. Let

$$\boldsymbol{\Sigma}_{\mathrm{S}} = \sigma \boldsymbol{I}_p, \ \ \boldsymbol{\Sigma}_{\mathrm{D}} = \mathrm{diag}(\sigma_{11}, \ldots, \sigma_{pp}) \ \ \text{and} \ \ \boldsymbol{\Sigma}_{\mathrm{IC}} = \sigma\{(1-\rho)\boldsymbol{I}_p + \rho \boldsymbol{1}_p \boldsymbol{1}_p^T\}.$$

Ledoit and Wolf (2002) derived test procedures for

$$H_0 : \boldsymbol{\Sigma} = \boldsymbol{I}_p \quad \text{vs.} \quad H_1 : \boldsymbol{\Sigma} \neq \boldsymbol{I}_p \tag{3}$$

and

$$H_0 : \boldsymbol{\Sigma} = \boldsymbol{\Sigma}_{\mathrm{S}} \quad \text{vs.} \quad H_1 : \boldsymbol{\Sigma} \neq \boldsymbol{\Sigma}_{\mathrm{S}} \tag{4}$$

when $p/n \to c > 0$ and $\boldsymbol{x}_j$ is Gaussian. Schott (2005) and Bao et al. (2015) derived test procedures for

$$H_0 : \boldsymbol{\Sigma} = \boldsymbol{\Sigma}_{\mathrm{D}} \quad \text{vs.} \quad H_1 : \boldsymbol{\Sigma} \neq \boldsymbol{\Sigma}_{\mathrm{D}} \tag{5}$$

when $p/n \to c > 0$ and $\boldsymbol{x}_j$ is Gaussian. Srivastava et al. (2011) considered test procedures for (3) - (5) when $n/p \to 0$ under an assumption that is stronger than (A-ii) given in Section 2. On the other hand, Srivastava and Reid (2012) derived a test procedure for

$$H_0 : \boldsymbol{\Sigma} = \boldsymbol{\Sigma}_{\mathrm{IC}} \quad \text{vs.} \quad H_1 : \boldsymbol{\Sigma} \neq \boldsymbol{\Sigma}_{\mathrm{IC}} \tag{6}$$

when $n/p \to 0$ and $\boldsymbol{x}_j$ is Gaussian. Meanwhile, Zhong et al. (2017) considered a high-dimensional regression model and testing (6) for the covariance matrix associated with error vectors when the error vectors are Gaussian. However, it is known that those test statistics do not always give a preferable performance unless $\boldsymbol{x}_j$ is Gaussian. As for a nonparametric approach, Chen et al. (2010) considered test statistics based on a U-statistic for (3) and (4). In the current paper, we take a different nonparametric approach and produce new test statistics for (2). We utilize the extended cross-data-matrix (ECDM) method developed by Yata and Aoshima (2013) which is an extension of the cross-data-matrix methodology created by Yata and Aoshima (2010). The ECDM method is a nonparametric method to produce an unbiased estimator for a function of $\boldsymbol{\Sigma}$ at a low computational cost even for ultra high-dimensional data. In addition, the ECDM method possesses a high versatility in high-dimensional data analysis. See Yata and Aoshima (2016) for the details. In this paper, we consider test statistics, for (2), derived by the ECDM method.

When $\boldsymbol{\Sigma}_* = \boldsymbol{\Sigma}_{\mathrm{S}}$, $\boldsymbol{\Sigma}_* = \boldsymbol{\Sigma}_{\mathrm{D}}$ or $\boldsymbol{\Sigma}_* = \boldsymbol{\Sigma}_{\mathrm{IC}}$, the eigenstructures are identified, however, they involve unknown parameters. Hence, we consider testing the high-dimensional covariance structures by using the following model: Let $\boldsymbol{A}_j$ be a $p \times p$ known idempotent matrix with rank r_j (≥ 1) for $j = 1, \ldots, q$, such that $\sum_{j=1}^{q} r_j = p$ and $\sum_{j=1}^{q} \boldsymbol{A}_j = \boldsymbol{I}_p$, where $r_1 \leq \cdots \leq r_q$ when $q \geq 2$. Note that $\mathrm{tr}(\boldsymbol{A}_j) = r_j$, $\boldsymbol{A}_j^2 = \boldsymbol{A}_j$ and $\boldsymbol{A}_j \boldsymbol{A}_{j'} = \boldsymbol{O}$ for all j $(\neq j')$. Let κ_j (≥ 0) be an unknown scalar such that $\mathrm{tr}(\boldsymbol{\Sigma} \boldsymbol{A}_j) = r_j \kappa_j$ for all j. We assume that $\boldsymbol{\Sigma}_*$ has the following structure:

$$\boldsymbol{\Sigma}_* = \kappa_1 \boldsymbol{A}_1 + \cdots + \kappa_q \boldsymbol{A}_q. \tag{7}$$

One can summarize as follows:

(I) $\boldsymbol{A}_1 = \boldsymbol{I}_p$, $\kappa_1 = \sigma$, $r_1 = p$ and $q = 1$ when $\boldsymbol{\Sigma}_* = \boldsymbol{\Sigma}_{\mathrm{S}}$;

(II) $\boldsymbol{A}_j = \mathrm{diag}(0, \ldots, 0, 1, 0, \ldots, 0)$ whose j-th diagonal element is 1, $\kappa_j = \sigma_{jj}$, $r_j = 1$ for all j and $q = p$ when $\boldsymbol{\Sigma}_* = \boldsymbol{\Sigma}_{\mathrm{D}}$;

(III) $\boldsymbol{A}_1 = \boldsymbol{1}_p \boldsymbol{1}_p^T / p$, $\boldsymbol{A}_2 = \boldsymbol{I}_p - \boldsymbol{1}_p \boldsymbol{1}_p^T / p$, $\kappa_1 = \sigma\{1 + (p-1)\rho\}$, $\kappa_2 = \sigma(1 - \rho)$, $r_1 = 1$, $r_2 = p - 1$ and $q = 2$ when $\boldsymbol{\Sigma}_* = \boldsymbol{\Sigma}_{\mathrm{IC}}$.

In this paper, we consider constructing new test procedures for (2), including (3), (4), (5) and (6). In Section 2, we produce a test statistic when $\boldsymbol{\Sigma}_*$ is known such as $\boldsymbol{\Sigma}_* = \boldsymbol{I}_p$. We show that the test statistic is an unbiased estimator of its test parameter even in a high-dimensional setting. In Section 3, we produce a test statistic for the structure (7). We propose a new test procedure based on the test statistic and evaluate its asymptotic size and power theoretically when both p and n increase. In Section 4, we apply the new test procedure to testing (4) - (6). In Sections 5 and 6, we investigate the performance of the proposed test procedure in simulations and actual data analyses. We also give a test procedure to identify an eigenvector in Appendix B of the online supplementary material.

Remark 1 We consider the following model: Let σ_1 and σ_2 be unknown non-negative scalars, and $\boldsymbol{\Theta} = (\theta_{ij})$ be a known symmetric matrix. We assume

$$\boldsymbol{\Sigma}_* = \sigma_1 \boldsymbol{I}_p + \sigma_2 \boldsymbol{\Theta}. \tag{8}$$

Then, (8) implies (7) because the eigenstructure of $\boldsymbol{\Theta}$ can be identified. Note that (8) includes the case that $\boldsymbol{\Sigma}_* = \boldsymbol{\Sigma}_{\mathrm{IC}}$. In other examples, (8) includes the case of a moving-average model of order 1 (MA(1)) since $\theta_{ij} = 1$ when $|i - j| = 1$, otherwise $\theta_{ij} = 0$ for MA(1). Also, (8) includes the case of an autoregressive model of order 1 (AR(1)) when the autoregressive parameter ϕ is known since $\theta_{ii} = 0$ and $\theta_{ij} = \phi^{|i-j|}$ when $i \neq j$ for AR(1). Thus one can apply hypothesis testing under (7) to test some covariance structures for time series models.

2 A test procedure for (2) when $\boldsymbol{\Sigma}_*$ is known

In this section, we propose a test procedure for (2) when $\boldsymbol{\Sigma}_*$ is known and evaluate its asymptotic size and power theoretically. Let

$$\Delta = \|\boldsymbol{\Sigma} - \boldsymbol{\Sigma}_*\|_F^2 = \mathrm{tr}\{(\boldsymbol{\Sigma} - \boldsymbol{\Sigma}_*)^2\},$$

where $\|\cdot\|_F$ is the Frobenius norm. Note that $\Delta = 0$ under H_0 and $\Delta > 0$ under H_1. We regard Δ as a test parameter and construct a test procedure for (2) by using an estimator of Δ.

2.1 Unbiased estimator of Δ

We first give an unbiased estimator of Δ by using the ECDM method. Let $n_{(1)} = \lceil n/2 \rceil$ and $n_{(2)} = n - n_{(1)}$, where $\lceil x \rceil$ denotes the smallest integer $\geq x$. Let

$$\boldsymbol{V}_{n(1)(k)} = \begin{cases} \{\lfloor k/2 \rfloor - n_{(1)} + 1, \ldots, \lfloor k/2 \rfloor\} & \text{if } \lfloor k/2 \rfloor \geq n_{(1)}, \\ \{1, \ldots, \lfloor k/2 \rfloor\} \cup \{\lfloor k/2 \rfloor + n_{(2)} + 1, \ldots, n\} & \text{otherwise;} \end{cases}$$

$$\boldsymbol{V}_{n(2)(k)} = \begin{cases} \{\lfloor k/2 \rfloor + 1, \ldots, \lfloor k/2 \rfloor + n_{(2)}\} & \text{if } \lfloor k/2 \rfloor \leq n_{(1)}, \\ \{1, \ldots, \lfloor k/2 \rfloor - n_{(1)}\} \cup \{\lfloor k/2 \rfloor + 1, \ldots, n\} & \text{otherwise} \end{cases}$$

for $k = 3, \ldots, 2n - 1$, where $\lfloor x \rfloor$ denotes the largest integer $\leq x$. Let $\#\boldsymbol{S}$ denote the number of elements in a set $\boldsymbol{S}$. Note that $\#\boldsymbol{V}_{n(l)(k)} = n_{(l)}$, $l = 1, 2$, $\boldsymbol{V}_{n(1)(k)} \cap \boldsymbol{V}_{n(2)(k)} = \emptyset$ and $\boldsymbol{V}_{n(1)(k)} \cup \boldsymbol{V}_{n(2)(k)} = \{1, \ldots, n\}$ for $k = 3, \ldots, 2n - 1$. Also, note that $i \in \boldsymbol{V}_{n(1)(i+j)}$ and $j \in \boldsymbol{V}_{n(2)(i+j)}$ for $i < j \, (\leq n)$. Let

$$\overline{\boldsymbol{x}}_{(1)(k)} = n_{(1)}^{-1} \sum_{j \in \boldsymbol{V}_{n(1)(k)}} \boldsymbol{x}_j \quad \text{and} \quad \overline{\boldsymbol{x}}_{(2)(k)} = n_{(2)}^{-1} \sum_{j \in \boldsymbol{V}_{n(2)(k)}} \boldsymbol{x}_j$$

for $k = 3, \ldots, 2n - 1$. Let

$$\boldsymbol{y}_{ij(1)} = \frac{n_{(1)}^{1/2}(\boldsymbol{x}_i - \overline{\boldsymbol{x}}_{(1)(i+j)})}{(n_{(1)} - 1)^{1/2}} \quad \text{and} \quad \boldsymbol{y}_{ij(2)} = \frac{n_{(2)}^{1/2}(\boldsymbol{x}_j - \overline{\boldsymbol{x}}_{(2)(i+j)})}{(n_{(2)} - 1)^{1/2}} \quad (9)$$

for all $i < j$. We note that $E(\boldsymbol{y}_{ij(l)}) = \boldsymbol{0}$, $E(\boldsymbol{y}_{ij(l)}\boldsymbol{y}_{ij(l)}^T) = \boldsymbol{\Sigma}$ for $l = 1, 2$, and

$$\boldsymbol{y}_{ij(1)} \text{ and } \boldsymbol{y}_{ij(2)} \text{ are independent for all } 1 \le i < j \le n.$$

For example, Yata and Aoshima (2013) gave an estimator of $\mathrm{tr}(\boldsymbol{\Sigma}^2)$ as

$$W_n = \frac{2}{n(n-1)} \sum_{i<j}^{n} (\boldsymbol{y}_{ij(1)}^T \boldsymbol{y}_{ij(2)})^2 \quad (10)$$

by the ECDM method. Then, it holds that $E(W_n) = \mathrm{tr}(\boldsymbol{\Sigma}^2)$.

Remark 2 One can save the computational cost of W_n by using previously calculated $\overline{\boldsymbol{x}}_{(i)(k)}$, $k = 3, \ldots, 2n - 1$; $i = 1, 2$. The computational cost of W_n is of the order, $O(n^2 p)$.

We provide in Appendix C of the online supplementary material a program in R-code to calculate the $\boldsymbol{y}_{ij(l)}$.

Now, we can give an unbiased estimator of Δ as

$$\widehat{\Delta}_n = 2 \sum_{i<j}^{n} \frac{\mathrm{tr}\{(\boldsymbol{y}_{ij(1)}\boldsymbol{y}_{ij(1)}^T - \boldsymbol{\Sigma}_*)(\boldsymbol{y}_{ij(2)}\boldsymbol{y}_{ij(2)}^T - \boldsymbol{\Sigma}_*)\}}{n(n-1)} \quad (11)$$

by the ECDM method. Note that $E(\widehat{\Delta}_n) = \Delta$. Here, we write that

$$\widehat{\Delta}_n = W_n + \mathrm{tr}(\boldsymbol{\Sigma}_*^2) - 2 \sum_{i<j}^{n} \frac{\left(\boldsymbol{y}_{ij(1)}^T \boldsymbol{\Sigma}_* \boldsymbol{y}_{ij(1)} + \boldsymbol{y}_{ij(2)}^T \boldsymbol{\Sigma}_* \boldsymbol{y}_{ij(2)}\right)}{n(n-1)}. \quad (12)$$

Since one can avoid calculating $p \times p$ matrices in (12), the computational cost of $\widehat{\Delta}_n$ by (12) is much lower than that by (11) when $n = o(p)$.

2.2 Asymptotic properties of $\widehat{\Delta}_n$

We assume the following model:

$$\boldsymbol{x}_j = \boldsymbol{\Gamma}\boldsymbol{w}_j + \boldsymbol{\mu},$$

where $\boldsymbol{\Gamma} = (\boldsymbol{\gamma}_1, \ldots, \boldsymbol{\gamma}_d)$ is a $p \times d$ matrix for some $d > 0$ such that $\boldsymbol{\Gamma}\boldsymbol{\Gamma}^T = \boldsymbol{\Sigma}$, and $\boldsymbol{w}_j = (w_{1j}, \ldots, w_{dj})^T$, $j = 1, \ldots, n$, are i.i.d. random vectors having $E(\boldsymbol{w}_j) = \boldsymbol{0}$ and $\mathrm{Var}(\boldsymbol{w}_j) = \boldsymbol{I}_d$. Let $\mathrm{Var}(w_{sj}^2) = M_s$ for $s = 1, \ldots, d$. We assume that $\limsup_{p\to\infty} M_s < \infty$ for all s. Similar to Bai and Saranadasa (1996), Chen and Qin (2010) and Aoshima and Yata (2015), we assume that

(A-i) $E(w_{sj}^2 w_{tj}^2) = E(w_{sj}^2)E(w_{tj}^2) = 1$ and $E(w_{sj}w_{tj}w_{uj}w_{vj}) = 0$ for all $s \neq t, u, v$.

We make the following assumption instead of (A-i) as necessary:

(A-ii) $E(w_{s_1 j}^{\alpha_1} w_{s_2 j}^{\alpha_2} \cdots w_{s_v j}^{\alpha_v}) = E(w_{s_1 j}^{\alpha_1})E(w_{s_2 j}^{\alpha_2}) \cdots E(w_{s_v j}^{\alpha_v})$ for all $s_1 \neq s_2 \neq \cdots \neq s_v \in [1, d]$ and $\alpha_i \in [1, 4]$, $i = 1, \ldots, v$, where $v \leq 8$ and $\sum_{i=1}^v \alpha_i \leq 8$.

Note that (A-ii) implies (A-i). When $\boldsymbol{x}_j$ is Gaussian and $\boldsymbol{\Gamma} = \boldsymbol{H}\boldsymbol{\Lambda}^{1/2}$, it holds that $\boldsymbol{w}_j = \boldsymbol{z}_j$ and $\boldsymbol{z}_j$ is distributed as $N_p(\boldsymbol{0}, \boldsymbol{I}_p)$, so that (A-ii) is naturally satisfied.

For $\boldsymbol{\Sigma}$ we assume the following condition as necessary:

(C-i) $\dfrac{\mathrm{tr}(\boldsymbol{\Sigma}^4)}{\mathrm{tr}(\boldsymbol{\Sigma}^2)^2} \to 0$ as $p \to \infty$.

Note that (C-i) is equivalent to "$\lambda_1/\mathrm{tr}(\boldsymbol{\Sigma}^2)^{1/2} \to 0$ as $p \to \infty$". Aoshima and Yata (2018) called (C-i) the "non-strongly spiked eigenvalue (NSSE) model". When $\boldsymbol{\Sigma} = \boldsymbol{\Sigma}_{\mathrm{S}}$ or $\boldsymbol{\Sigma}_{\mathrm{D}}$, (C-i) holds. On the other hand, when $\boldsymbol{\Sigma} = \boldsymbol{\Sigma}_{\mathrm{IC}}$ with $\liminf_{p \to \infty} \rho > 0$, (C-i) does not hold since it follows that

$$\liminf_{p \to \infty} \left\{ \frac{\lambda_1}{\mathrm{tr}(\boldsymbol{\Sigma}^2)^{1/2}} \right\} > 0 \tag{13}$$

from the facts that $\lambda_1 = \sigma\{1 + (p-1)\rho\}$ and $\mathrm{tr}(\boldsymbol{\Sigma}^2) = O(p^2)$. Aoshima and Yata (2018) called (13) the "strongly spiked eigenvalue (SSE) model". For instance, let us consider a spiked model as

$$\lambda_j = a_j p^{\alpha_j} \ (j = 1, \ldots, g) \quad \text{and} \quad \lambda_j = c_j \ (j = g+1, \ldots, p), \tag{14}$$

where a_js, c_js and α_js are positive (fixed) constants, and g is a positive (fixed) integer. For (14), it is a NSSE model when $\alpha_1 < 1/2$ and a SSE model when $\alpha_1 \geq 1/2$. In Section 3.3, we consider a test procedure for the SSE model (13).

Let

$$m = \min\{p, n\}.$$

We consider the divergence condition as

$$p \to \infty \quad \text{and} \quad n \to \infty,$$

which is equivalent to $m \to \infty$. Note that "$m \to \infty$" includes the cases when $n/p = o(1)$ and $p/n \to c \in [0, \infty)$. In this paper, we mainly consider the case when $n/p = o(1)$ such as $p = O(n^c)$ with $c > 1$ and $n = O(\log p)$. However, we emphasize that our proposed test procedures can handle cases not only when $n/p = o(1)$ but also when $p/n \to c \in [0, \infty)$. Let

$$K = 4\mathrm{tr}(\boldsymbol{\Sigma}^2)^2/n^2 \quad \text{and} \quad K_* = 4\mathrm{tr}(\boldsymbol{\Sigma}_*^2)^2/n^2.$$

We assume one of the following assumptions as necessary:

(C-ii) $\dfrac{K^{1/2}}{\Delta} \to 0$ as $m \to \infty$;   (C-iii) $\limsup_{m \to \infty} \dfrac{\Delta}{K^{1/2}} < \infty$.

Note that (C-iii) holds under H_0 in (2). For $\widehat{\Delta}_n$ in (11), we have the following results.

Lemma 1 *Assume (A-i). Then, it holds that as $m \to \infty$*

$$Var(\widehat{\Delta}_n) = K\{1 + o(1)\} + O\Big(\frac{tr(\boldsymbol{\Sigma}^4)^{1/2}\Delta}{n} + \frac{tr(\boldsymbol{\Sigma}^4)}{n^2}\Big).$$

Furthermore, under (C-i) and (C-iii), it holds that as $m \to \infty$

$$Var(\widehat{\Delta}_n) = K\{1 + o(1)\}.$$

From Lemma 1 we obtain the following result under (C-ii).

Theorem 1 *Assume (A-i) and (C-ii). Then, it holds that as $m \to \infty$*

$$\widehat{\Delta}_n/\Delta = 1 + o_P(1).$$

On the other hand, we obtain the following result under (C-iii).

Theorem 2 *Assume (A-ii), (C-i) and (C-iii). Then, it holds that as $m \to \infty$*

$$\frac{\widehat{\Delta}_n - \Delta}{K^{1/2}} \Rightarrow N(0,1),$$

where "$\Rightarrow$" denotes the convergence in distribution and $N(0,1)$ denotes a random variable distributed as the standard normal distribution.

Remark 3 We note that $K^{1/2}$ is the main term in the standard deviation of $\widehat{\Delta}_n$. The condition (C-ii) means that the intrinsic information about $\boldsymbol{\Sigma} - \boldsymbol{\Sigma}_*$ is larger than the noise. Thus, if the difference between $\boldsymbol{\Sigma}$ and $\boldsymbol{\Sigma}_*$ is quite large enough to claim (C-ii), $\widehat{\Delta}_n$ has the consistency in Theorem 1. On the other hand, the condition (C-iii) means that the noise is as large as the intrinsic information or more. Thus, if the difference is not large enough to claim (C-ii), $\widehat{\Delta}_n$ has the asymptotic normality in Theorem 2.

2.3 A test procedure based on $\widehat{\Delta}_n$

Note that $tr(\boldsymbol{\Sigma}^2) = tr(\boldsymbol{\Sigma}_*^2)$ under H_0 in (2). Let

$$T_n = \frac{n\widehat{\Delta}_n}{2tr(\boldsymbol{\Sigma}_*^2)}. \tag{15}$$

From Theorem 2 we propose a test procedure for (2) by

$$\text{rejecting } H_0 \Longleftrightarrow T_n > z_\alpha, \tag{16}$$

where z_α is a constant such that $P\{N(0,1) > z_\alpha\} = \alpha$ with $\alpha \in (0, 1/2)$. Then, we have the following result.

Theorem 3 *Assume (A-ii) and (C-i). For the test procedure (16), we have that as $m \to \infty$*

$$Size = \alpha + o(1) \quad and \quad Power = \Phi\left(\frac{\Delta}{K^{1/2}} - z_\alpha \frac{K_*^{1/2}}{K^{1/2}}\right) + o(1) \qquad (17)$$

where $\Phi(\cdot)$ denotes the c.d.f. of $N(0,1)$.

We note that Theorem 3 holds not only when $n/p = o(1)$ but also when $p/n \to c \in [0, \infty)$. When (C-ii) is met, we have the following result.

Corollary 1 *Assume (A-i). Assume also (C-ii) under H_1 in (2). For the test procedure (16), we have that as $m \to \infty$*

$$Power \to 1. \qquad (18)$$

From Corollary 1, if the difference between $\boldsymbol{\Sigma}$ and $\boldsymbol{\Sigma}_*$ is large enough to claim (C-ii), the power tends to 1. In the next section, we apply the test procedure (16) to testing the identity structure in (3).

2.4 A test of the identity structure in (3)

We consider the case when $\boldsymbol{\Sigma}_* = \boldsymbol{I}_p$. Note that $\Delta = \mathrm{tr}(\boldsymbol{\Sigma}^2) + p - 2\mathrm{tr}(\boldsymbol{\Sigma})$ when $\boldsymbol{\Sigma}_* = \boldsymbol{I}_p$. From (15) we write that when $\boldsymbol{\Sigma}_* = \boldsymbol{I}_p$,

$$T_n = \frac{n\widehat{\Delta}_n(\boldsymbol{I}_p)}{2p} \quad (= T_n(\boldsymbol{I}_p), \text{ say}),$$

where $\widehat{\Delta}_n(\boldsymbol{I}_p) = W_n + p - 2\sum_{i<j}^n (\|\boldsymbol{y}_{ij(1)}\|^2 + \|\boldsymbol{y}_{ij(2)}\|^2)/\{n(n-1)\}$. Here, $\|\cdot\|$ denotes the Euclidean norm.

Proposition 1 *The condition (C-iii) with $\boldsymbol{\Sigma}_* = \boldsymbol{I}_p$ implies (C-i).*

Thus, from Theorem 3 and Corollary 1, we have the following result.

Corollary 2 *For the test procedure (16) with $T_n = T_n(\boldsymbol{I}_p)$ for (3)*

(i)   (17) holds as $m \to \infty$ under (A-ii),
(ii) (18) holds as $m \to \infty$ under (A-i) and (C-ii) with $\boldsymbol{\Sigma}_* = \boldsymbol{I}_p$.

Remark 4 For (3) Chen et al. (2010) gave a test procedure based on the following U-statistic:

$$\widehat{\Delta}_{\mathrm{CZZ}} = A_n - 2\left(\sum_{j=1}^n \frac{\boldsymbol{x}_j^T \boldsymbol{x}_j}{n} - \sum_{j \neq j'}^n \frac{\boldsymbol{x}_j^T \boldsymbol{x}_{j'}}{n(n-1)}\right) + p,$$

where

$$A_n = \sum_{j \neq j'}^n \frac{(\boldsymbol{x}_j^T \boldsymbol{x}_{j'})^2}{n(n-1)} - 2\sum_{j \neq j' \neq j''}^n \frac{\boldsymbol{x}_{j'}^T \boldsymbol{x}_j \boldsymbol{x}_j^T \boldsymbol{x}_{j''}}{n(n-1)(n-2)}$$

$$+ \sum_{j \neq j' \neq l \neq l'}^n \frac{\boldsymbol{x}_j^T \boldsymbol{x}_{j'} \boldsymbol{x}_l^T \boldsymbol{x}_{l'}}{n(n-1)(n-2)(n-3)}.$$

Note that $E(A_n) = \text{tr}(\boldsymbol{\Sigma}^2)$ and $E(\widehat{\Delta}_{\text{CZZ}}) = \text{tr}(\boldsymbol{\Sigma}^2) + p - 2\text{tr}(\boldsymbol{\Sigma})$. The test procedure by $\widehat{\Delta}_{\text{CZZ}}$ is asymptotically equivalent to (16) with $\boldsymbol{\Sigma}_* = \boldsymbol{I}_p$. It should be noted that the test of the identity structure is obtained as an example of the test procedure (16).

3 Test procedure for (2) under the structure (7)

In this section, we construct an unbiased estimator of Δ under the structure (7) and propose a test procedure by using the unbiased estimator.

3.1 Test procedure

Hereafter, we assume the structure (7) for $\boldsymbol{\Sigma}_*$. Then, we note that

$$\text{tr}(\boldsymbol{\Sigma}_*^2) = \sum_{j=1}^{q} r_j \kappa_j^2 \quad \text{and} \quad \Delta = \text{tr}(\boldsymbol{\Sigma}^2) - \text{tr}(\boldsymbol{\Sigma}_*^2),$$

so that $\text{tr}(\boldsymbol{\Sigma}^2) \geq \text{tr}(\boldsymbol{\Sigma}_*^2)$. Also, note that $\{\text{tr}(\boldsymbol{\Sigma}\boldsymbol{A}_j)\}^2/r_j = r_j \kappa_j^2$ for all j. Then, we give an estimator of $\text{tr}(\boldsymbol{\Sigma}_*^2)$ as

$$U_n = 2 \sum_{s=1}^{q} \sum_{i<j}^{n} \frac{\boldsymbol{y}_{ij(1)}^T \boldsymbol{A}_s \boldsymbol{y}_{ij(1)} \boldsymbol{y}_{ij(2)}^T \boldsymbol{A}_s \boldsymbol{y}_{ij(2)}}{r_s n(n-1)} \tag{19}$$

by the ECDM method. Note that $E(U_n) = \text{tr}(\boldsymbol{\Sigma}_*^2)$. Let

$$\widetilde{\Delta}_n = W_n - U_n,$$

where W_n is given by (10). Then, it holds that $E(\widetilde{\Delta}_n) = \Delta$. Here, we write that

$$\begin{aligned} U_n = &B_n(1) - \text{tr}(\boldsymbol{\Sigma}_*^2) \\ &+ 2 \sum_{i<j}^{n} \frac{\boldsymbol{y}_{ij(1)}^T \boldsymbol{\Sigma}_* \boldsymbol{y}_{ij(1)} + \boldsymbol{y}_{ij(2)}^T \boldsymbol{\Sigma}_* \boldsymbol{y}_{ij(2)}}{n(n-1)}, \end{aligned} \tag{20}$$

where

$$B_n(t) = 2 \sum_{i<j}^{n} \sum_{s=t}^{q} \frac{(\boldsymbol{y}_{ij(1)}^T \boldsymbol{A}_s \boldsymbol{y}_{ij(1)} - \kappa_s r_s)(\boldsymbol{y}_{ij(2)}^T \boldsymbol{A}_s \boldsymbol{y}_{ij(2)} - \kappa_s r_s)}{r_s n(n-1)}$$

for $t = 1, \ldots, q$. By combining (12) and (20), we have that

$$\widetilde{\Delta}_n = \widehat{\Delta}_n - B_n(1). \tag{21}$$

Note that $E\{B_n(t)\} = 0$ for all t. Let us consider an asymptotic variance of $\widetilde{\Delta}_n$. Let $q_\star$ be the maximum integer such that

$$r_1 = \cdots = r_{q_\star} = 1 < r_{q_\star+1} \leq \cdots \leq r_q.$$

If $r_1 = \cdots = r_q = 1$, we set $q_\star = q$. We set $B_n(q_\star + 1) = 0$ when $q_\star = q$. If $r_j \geq 2$ for all j, we set $q_\star = 0$. Let

$$\boldsymbol{Y}_{ij(l),s} = \boldsymbol{y}_{ij(l)}\boldsymbol{y}_{ij(l)}^T \boldsymbol{A}_s - \kappa_s \boldsymbol{A}_s$$

for all i, j, l, s. Then, from (11), it follows that

$$\widetilde{\Delta}_n = \begin{cases} 2\sum_{i<j}^n \left(\sum_{s \neq s'}^q \dfrac{\mathrm{tr}(\boldsymbol{Y}_{ij(1),s}\boldsymbol{Y}_{ij(2),s'})}{n(n-1)} + \sum_{s=q_\star+1}^q \dfrac{\mathrm{tr}(\boldsymbol{Y}_{ij(1),s}\boldsymbol{Y}_{ij(2),s})}{n(n-1)} \right) \\ \quad -B_n(q_\star + 1) \qquad\qquad\qquad\qquad (q_\star < q), \\ 2\sum_{i<j}^n \sum_{s \neq s'}^q \dfrac{\mathrm{tr}(\boldsymbol{Y}_{ij(1),s}\boldsymbol{Y}_{ij(2),s'})}{n(n-1)} \qquad (q_\star = q). \end{cases}$$

We have the following result.

Proposition 2 *Assume (A-i). Under H_0 in (2), it holds that as $m \to \infty$*

$$Var\{\widetilde{\Delta}_n + B_n(q_\star + 1)\} = \frac{4\Psi}{n^2}\{1 + o(1)\} + O\left(\frac{tr(\boldsymbol{\Sigma}_\ast^4)}{n^2}\right),$$

where $\Psi = tr(\boldsymbol{\Sigma}_\ast^2)^2 - \sum_{s=1}^{q_\star} \kappa_s^4$ when $q_\star \geq 1$ and $\Psi = tr(\boldsymbol{\Sigma}_\ast^2)^2$ when $q_\star = 0$.

Since $B_n(q_\star + 1)$ is a redundant term, we can regard the term "$4\Psi/n^2$" as an asymptotic variance of $\widetilde{\Delta}_n$ under H_0 in (2). We note that

$$E\left(2\sum_{i<j}^n \frac{\boldsymbol{y}_{ij(1)}^T \boldsymbol{A}_s \boldsymbol{y}_{ij(1)} \boldsymbol{y}_{ij(2)}^T \boldsymbol{A}_s \boldsymbol{y}_{ij(2)}}{r_s n(n-1)} \right) = r_s \kappa_s^2$$

in view of (19). We give an estimator of Ψ by

$$\widetilde{\Psi}_n = \begin{cases} U_n^2 - \sum_{s=1}^{q_\star} \left(2\sum_{i<j}^n \dfrac{\boldsymbol{y}_{ij(1)}^T \boldsymbol{A}_s \boldsymbol{y}_{ij(1)} \boldsymbol{y}_{ij(2)}^T \boldsymbol{A}_s \boldsymbol{y}_{ij(2)}}{n(n-1)} \right)^2 & (q_\star \geq 1), \\ U_n^2 & (q_\star = 0) \end{cases}$$

in view of $r_1 = \cdots = r_{q_\star} = 1$ when $q_\star \geq 1$. Note that $P(\widetilde{\Psi}_n \geq 0) = 1$. Let

$$\widetilde{T}_n = \frac{n\widetilde{\Delta}_n}{2\widetilde{\Psi}_n^{1/2}}. \tag{22}$$

Then, for (2) with (7), we propose a test procedure by

$$\text{rejecting } H_0 \Longleftrightarrow \widetilde{T}_n > z_\alpha. \tag{23}$$

In Sections 3.2 and 3.3, we investigate the test procedure (23) under the NSSE model (C-i) and the SSE model (13), respectively.

3.2 The test procedure (23) under the NSSE model

We consider the test procedure (23) under the NSSE model (C-i). For $B_n(t)$, we have the following result.

Lemma 2 *Assume (A-i). It holds that as $m \to \infty$*

$$Var\{B_n(t)\} = O\left(\sum_{j,j'=t}^{q} \frac{tr(\boldsymbol{\Sigma A}_j \boldsymbol{\Sigma A}_{j'})^2 + (\sum_{s=1}^{d} \boldsymbol{\gamma}_s^T \boldsymbol{A}_j \boldsymbol{\gamma}_s \boldsymbol{\gamma}_s^T \boldsymbol{A}_{j'} \boldsymbol{\gamma}_s)^2}{r_j r_{j'} n^2} \right)$$

for $t \leq q$.

Then, under (A-i) and

$$\text{(C-iv)} \quad \sum_{j,j'=1}^{q} \frac{tr(\boldsymbol{\Sigma A}_j \boldsymbol{\Sigma A}_{j'})^2 + (\sum_{s=1}^{d} \boldsymbol{\gamma}_s^T \boldsymbol{A}_j \boldsymbol{\gamma}_s \boldsymbol{\gamma}_s^T \boldsymbol{A}_{j'} \boldsymbol{\gamma}_s)^2}{r_j r_{j'} tr(\boldsymbol{\Sigma}_*^2)^2} \to 0 \quad \text{as } p \to \infty,$$

it follows from Chebyshev's inequality that as $m \to \infty$

$$B_n(1) = o_P(K_*^{1/2}).$$

Note that $tr(\boldsymbol{\Sigma}^2) \geq tr(\boldsymbol{\Sigma}_*^2)$. From (21) and Theorem 2, under (A-ii), (C-i), (C-iii) and (C-iv), it holds that as $m \to \infty$

$$\frac{\widetilde{\Delta}_n - \Delta}{K^{1/2}} = \frac{\widehat{\Delta}_n - \Delta}{K^{1/2}} + o_P(1) \Rightarrow N(0,1). \tag{24}$$

We have the following result.

Lemma 3 *Assume (A-i) and (C-iv). It holds that $\widetilde{\Psi}_n / tr(\boldsymbol{\Sigma}_*^2)^2 = 1 + o_P(1)$ as $m \to \infty$.*

From (24) and Lemma 3, we have the following results.

Theorem 4 *Assume (A-ii), (C-i) and (C-iv). For the test procedure (23), we have (17) as $m \to \infty$.*

Corollary 3 *Assume (A-i). Assume also (C-ii) under H_1 in (2). For the test by (23), we have (18) as $m \to \infty$.*

3.3 The test procedure (23) under the SSE model

For the SSE model (13), we focus on the following model:

$$\frac{\lambda_1}{tr(\boldsymbol{\Sigma}^2)^{1/2}} \to 1 \quad \text{as } p \to \infty. \tag{25}$$

Note that (25) is one of the SSE models. When $\boldsymbol{\Sigma} = \boldsymbol{\Sigma}_{\text{IC}}$ and $\liminf_{p \to \infty} \rho > 0$, (25) is met. When $\alpha_1 > \max\{\alpha_2, 1/2\}$ in the spiked model (14), (25) is met. We call (25) the "uni-SSE (USSE) model". See Ishii et al. (2016, 2019) for several statistical inferences under the USSE model.

We consider the test procedure (23) under the USSE model (25). One may suppose $r_1 = 1$. We assume the following condition:

(C-v) $\quad \dfrac{\kappa_1}{\mathrm{tr}(\boldsymbol{\Sigma}^2)^{1/2}} \to 1 \quad$ as $p \to \infty$.

Note that (25) holds under (C-v) from the fact that $\lambda_1 \geq \kappa_1 = \mathrm{tr}(\boldsymbol{\Sigma}\boldsymbol{A}_1)$. Let $\boldsymbol{A}_{(1)} = \boldsymbol{I}_p - \boldsymbol{A}_1$, $\boldsymbol{\Omega}_1 = \boldsymbol{A}_1\boldsymbol{\Sigma}\boldsymbol{A}_{(1)}$ and $\boldsymbol{\Omega}_2 = \boldsymbol{A}_{(1)}\boldsymbol{\Sigma}\boldsymbol{A}_{(1)}$. Note that (C-v) is equivalent to

$$\mathrm{tr}(\boldsymbol{\Omega}_2^2)^{1/2}/\kappa_1 \to 0 \ \text{ as } p \to \infty$$

from the facts that $\mathrm{tr}(\boldsymbol{\Sigma}^2) = \kappa_1^2 + \mathrm{tr}(\boldsymbol{\Omega}_2^2) + 2\|\boldsymbol{\Omega}_1\|_F^2$ and $\|\boldsymbol{\Omega}_1\|_F^2 \leq \kappa_1\mathrm{tr}(\boldsymbol{\Omega}_2^2)^{1/2}$. As for $\boldsymbol{\Omega}_2$, we assume the following model:

(C-i') $\quad \dfrac{\mathrm{tr}(\boldsymbol{\Omega}_2^4)}{\mathrm{tr}(\boldsymbol{\Omega}_2^2)^2} \to 0 \ \text{ as } p \to \infty$.

Note that (C-i') holds when $\boldsymbol{\Sigma} = \boldsymbol{\Sigma}_{\mathrm{IC}}$ and $\limsup_{p\to\infty} \rho < 1$ because $\boldsymbol{\Omega}_2 = \sigma(1-\rho)(\boldsymbol{I}_p - \boldsymbol{1}_p\boldsymbol{1}_p^T/p)$ and $\mathrm{tr}(\boldsymbol{\Omega}_2^4)/\mathrm{tr}(\boldsymbol{\Omega}_2^2)^2 = 1/\mathrm{tr}(\boldsymbol{I}_p - \boldsymbol{1}_p\boldsymbol{1}_p^T/p) = 1/(p-1)$ when $\boldsymbol{\Sigma} = \boldsymbol{\Sigma}_{\mathrm{IC}}$. Here, we write that

$$\widetilde{\Delta}_n = 2\sum_{i<j}^{n} \left(\sum_{s\neq s'}^{q} \frac{\mathrm{tr}(\boldsymbol{Y}_{ij(1),s}\boldsymbol{Y}_{ij(2),s'})}{n(n-1)} + \sum_{s=2}^{q} \frac{\mathrm{tr}(\boldsymbol{Y}_{ij(1),s}\boldsymbol{Y}_{ij(2),s})}{n(n-1)} \right) - B_n(2).$$

Let

$$\Upsilon = 2\kappa_1^2\mathrm{tr}(\boldsymbol{\Omega}_2^2) + \mathrm{tr}(\boldsymbol{\Omega}_2^2)^2 + 2\|\boldsymbol{\Omega}_1\|_F^4 + 4\|\boldsymbol{\Omega}_1\|_F^2\mathrm{tr}(\boldsymbol{\Omega}_2^2) \ \text{ and}$$
$$L = 4\Upsilon/n^2.$$

Note that

$$\Upsilon = 2\kappa_1^2\sum_{s=2}^{q} r_s\kappa_s^2 + \left(\sum_{s=2}^{q} r_s\kappa_s^2 \right)^2 = \mathrm{tr}(\boldsymbol{\Sigma}_*^2)^2 - \kappa_1^4 \ \ (= \Psi_1, \text{ say})$$

when $\boldsymbol{\Sigma} = \boldsymbol{\Sigma}_*$. Then, we have the following results.

Lemma 4 *Assume (A-i). It holds that as $m \to \infty$*

$$Var\{\widetilde{\Delta}_n + B_n(2)\} = L\{1 + o(1)\} + O\Big(\frac{\Delta \mathrm{tr}(\boldsymbol{\Omega}^4)^{1/4}\{\mathrm{tr}(\boldsymbol{\Omega}^4)^{1/4} + \kappa_1\}}{n} \Big)$$
$$+ O\Big(\frac{\mathrm{tr}(\boldsymbol{\Omega}_2^4)^{1/2}\{\kappa_1^2 + \mathrm{tr}(\boldsymbol{\Omega}_2^4)^{1/2}\}}{n^2} \Big).$$

Furthermore, under (C-i') and

(C-iii') $\quad \limsup_{m\to\infty} \dfrac{\Delta}{L^{1/2}} < \infty$,

it holds that as $m \to \infty$

$$Var\{\widetilde{\Delta}_n + B_n(2)\} = L\{1 + o(1)\}.$$

Lemma 5 *Assume (A-ii), (C-i') and (C-v). Assume also (C-iii'). It holds that as $m \to \infty$*

$$\frac{\widetilde{\Delta}_n + B_n(2) - \Delta}{L^{1/2}} \Rightarrow N(0,1).$$

See Appendix A of the online supplementary material for the details of the asymptotic normality in Lemma 5.

We note that $\sum_{j,j'=2}^{q}(\sum_{s=1}^{d}\gamma_s^T \boldsymbol{A}_j\gamma_s\gamma_s^T \boldsymbol{A}_{j'}\gamma_s)^2 \leq \{\sum_{s=1}^{d}(\gamma_s^T \boldsymbol{A}_{(1)}\gamma_s)^2\}^2 \leq \text{tr}\{(\boldsymbol{\Sigma}\boldsymbol{A}_{(1)})^2\}^2 = \text{tr}(\boldsymbol{\Omega}_2^2)^2$ and $\sum_{j,j'=2}^{q}\text{tr}(\boldsymbol{\Sigma}\boldsymbol{A}_j\boldsymbol{\Sigma}\boldsymbol{A}_{j'})^2 \leq \text{tr}(\boldsymbol{\Omega}_2^2)^2$. Then, from Lemma 2, under (A-i) and (C-v), it holds that as $m \to \infty$

$$\text{Var}\{B_n(2)\} = o(L). \tag{26}$$

Thus, from Lemma 5, we have the following result.

Theorem 5 *Assume (A-ii), (C-i'), (C-iii') and (C-v). It holds that as $m \to \infty$*

$$\frac{\widetilde{\Delta}_n - \Delta}{L^{1/2}} \Rightarrow N(0,1).$$

For $\widetilde{\Psi}_n$, we have the following result.

Lemma 6 *Assume (A-i) and (C-v). It holds that $\widetilde{\Psi}_n/\Psi_1 = 1 + o_P(1)$ as $m \to \infty$.*

From Theorem 5 and Lemma 6, we have the following results.

Theorem 6 *Assume (A-ii), (C-i') and (C-v). For the test procedure (23), we have that as $m \to \infty$*

$$Size = \alpha + o(1) \quad and \quad Power = \Phi\left(\frac{\Delta}{L^{1/2}} - z_\alpha \frac{L_*^{1/2}}{L^{1/2}}\right) + o(1) \tag{27}$$

where $L_* = 4\Psi_1/n^2$.

Corollary 4 *Assume (A-i) and (C-v). Assume also*

(C-ii') $\dfrac{L^{1/2}}{\Delta} \to 0$ *as $m \to \infty$*

under H_1 in (2). For the test procedure (23), we have (18) as $m \to \infty$.

4 Applications of the test procedure (23) to testing (4) - (6)

In this section, we apply the test procedure (23) to testing (4) - (6).

4.1 The scaled identity structure (4)

We consider the case when $\boldsymbol{\Sigma}_* = \boldsymbol{\Sigma}_S$. Note that $\Delta = \text{tr}(\boldsymbol{\Sigma}^2) - p\sigma^2$, $q_\star = 0$ and $q = 1$, so that $\Psi^{1/2} = \text{tr}(\boldsymbol{\Sigma}_*^2) = p\sigma^2$ and $\widetilde{\Psi}_n^{1/2} = U_n$. From (22) we write that

$$\widetilde{T}_n = \frac{nW_n}{2U_{n(S)}} - n/2 \quad (= \widetilde{T}_{n(S)}, \text{ say}),$$

where

$$U_{n(S)} = 2\sum_{i<j}^{n} \frac{\|\boldsymbol{y}_{ij(1)}\|^2\|\boldsymbol{y}_{ij(2)}\|^2}{pn(n-1)}.$$

Note that $E(U_{n(S)}) = p\sigma^2$.

Proposition 3 *The condition (C-iii) implies (C-i) when $\boldsymbol{\Sigma}_* = \boldsymbol{\Sigma}_S$.*

We note that (C-iv) holds under (C-iii) when $\boldsymbol{\Sigma}_* = \boldsymbol{\Sigma}_S$ because $\mathrm{tr}(\boldsymbol{\Sigma}^2) = O\{\mathrm{tr}(\boldsymbol{\Sigma}_S^2)\}$ under (C-iii) and

$$\sum_{j,j'=1}^{q} \frac{\mathrm{tr}(\boldsymbol{\Sigma A}_j \boldsymbol{\Sigma A}_{j'})^2 + (\sum_{s=1}^{d} \boldsymbol{\gamma}_s^T \boldsymbol{A}_j \boldsymbol{\gamma}_s \boldsymbol{\gamma}_s^T \boldsymbol{A}_{j'} \boldsymbol{\gamma}_s)^2}{r_j r_{j'}} = O\{\mathrm{tr}(\boldsymbol{\Sigma}^2)^2/p^2\} \quad (28)$$

when $\boldsymbol{\Sigma}_* = \boldsymbol{\Sigma}_S$. From Theorem 4 and Corollary 3, we have the following result.

Corollary 5 *For the test procedure (23) with $\widetilde{T}_n = \widetilde{T}_{n(S)}$ for (4)*

(i) (17) holds as $m \to \infty$ under (A-ii),
(ii) (18) holds as $m \to \infty$ under (A-i) and (C-ii) with $\boldsymbol{\Sigma}_* = \boldsymbol{\Sigma}_S$.

Remark 5 For (4) Chen et al. (2010) gave the following test statistic:

$$T_{\mathrm{CZZ}} = \frac{npA_n}{2\Big(n^{-1}\sum_{i=1}^{n} \boldsymbol{x}_i^T \boldsymbol{x}_i - \{n(n-1)\}^{-1}\sum_{i \neq j}^{n} \boldsymbol{x}_i^T \boldsymbol{x}_j\Big)^2} - \frac{n}{2}, \qquad (29)$$

where A_n is defined in Remark 4. Under H_0 in (4), they showed that as $m \to \infty$

$$T_{\mathrm{CZZ}} \Rightarrow N(0,1)$$

under a similar condition of (A-ii). Although the test procedure by Chen et al. (2010) is asymptotically equivalent to (23) with $\widetilde{T}_n = \widetilde{T}_{n(S)}$, the latter is more applicable to the sequential analysis ensuring prespecified accuracy. See Yata et al. (2018) for the details.

4.2 The diagonal structure (5)

We consider the case when $\boldsymbol{\Sigma}_* = \boldsymbol{\Sigma}_D$. Note that $\Delta = \mathrm{tr}(\boldsymbol{\Sigma}^2) - \sum_{j=1}^{p} \sigma_{jj}^2$ ($= \Delta_D$, say) and $q_\star = p$. Also, note that $\Psi = \mathrm{tr}(\boldsymbol{\Sigma}_D^2)^2 - \sum_{j=1}^{p} \sigma_{jj}^4$. Let $\boldsymbol{y}_{ij(l)} = (y_{1ij(l)}, ..., y_{pij(l)})^T$ for all i, j, l. From (22) we write that

$$\widetilde{T}_n = \frac{n\widetilde{\Delta}_{n(D)}}{2\widetilde{\Psi}_{n(D)}^{1/2}} \quad (= \widetilde{T}_{n(D)}, \text{ say}),$$

where $\widetilde{\Delta}_{n(D)} = W_n - U_{n(D)}$ and

$$\widetilde{\Psi}_{n(D)} = U_{n(D)}^2 - \sum_{s=1}^{p} \Big(2\sum_{i<j}^{n} \frac{y_{sij(1)}^2 y_{sij(2)}^2}{n(n-1)}\Big)^2$$

with

$$U_{n(D)} = 2\sum_{i<j}^{n} \sum_{s=1}^{p} \frac{y_{sij(1)}^2 y_{sij(2)}^2}{n(n-1)}.$$

Proposition 4 *The condition (C-iii) implies (C-i) and (C-iv) when $\boldsymbol{\Sigma}_* = \boldsymbol{\Sigma}_D$.*

From Theorem 4 and Corollary 3, we have the following result.

Corollary 6 *For the test procedure (23) with $\widetilde{T}_n = \widetilde{T}_{n(D)}$ for (5)*

(i) *(17) holds as $m \to \infty$ under (A-ii),*

(ii) *(18) holds as $m \to \infty$ under (A-i) and (C-ii) with $\boldsymbol{\Sigma}_* = \boldsymbol{\Sigma}_D$.*

Remark 6 For (5) Srivastava et al. (2011) gave the following test statistic:

$$T_{\mathrm{S}} = \frac{(n-1)[c_n\{\mathrm{tr}(\boldsymbol{S}_n^2) - \mathrm{tr}(\boldsymbol{S}_n)^2/(n-1)\} - c_n \sum_{j=1}^p s_j^2]}{2\sqrt{\left(c_n \sum_{j=1}^p s_j^2\right)^2 - \sum_{j=1}^p s_j^4}}, \qquad (30)$$

where $\boldsymbol{S}_n$ is the sample covariance matrix, s_j is the j-th diagonal element of $\boldsymbol{S}_n$ and $c_n = (n-1)^2/\{(n-2)(n+1)\}$. Under H_0 in (5), they showed that as $m \to \infty$

$$T_{\mathrm{S}} \Rightarrow N(0,1)$$

under the assumptions that z_{ij}s are i.i.d., $E(z_{ij}^8)$s are uniformly bounded and some regularity conditions. Note that $c_n\{\mathrm{tr}(\boldsymbol{S}_n^2) - \mathrm{tr}(\boldsymbol{S}_n)^2/(n-1)\} - c_n \sum_{j=1}^p s_j^2$ $(= \widehat{\Delta}_{\mathrm{S}}$, say) is an estimator of Δ_D. It should be noted that $\widehat{\Delta}_{\mathrm{S}}$ is heavily biased unless $\boldsymbol{x}_j$ is Gaussian. In addition, one cannot claim $\mathrm{Var}(\widehat{\Delta}_{\mathrm{S}}/\Delta_D) < \infty$ unless $E(z_{ij}^8)$s are uniformly bounded. Contrary to that, the proposed estimator, $\widetilde{\Delta}_{n(D)}$, is robust against the Gaussian assumption and one can claim that $E(\widetilde{\Delta}_{n(D)}) = \Delta_D$ without any assumptions. See Section 5 for numerical comparisons.

4.3 The intraclass covariance structure (6)

We consider the case when $\boldsymbol{\Sigma}_* = \boldsymbol{\Sigma}_{\mathrm{IC}}$. Note that $q_\star = 1$ and $q = 2$. Let $\boldsymbol{A}_{1(\mathrm{IC})} = \boldsymbol{1}_p \boldsymbol{1}_p^T/p$, $\boldsymbol{A}_{2(\mathrm{IC})} = \boldsymbol{I}_p - \boldsymbol{A}_{1(\mathrm{IC})}$, $\kappa_{1(\mathrm{IC})} = \sigma\{1 + (p-1)\rho\}$, $\kappa_{2(\mathrm{IC})} = \sigma(1-\rho)$,

$$\boldsymbol{\Omega}_{1(\mathrm{IC})} = \boldsymbol{A}_{1(\mathrm{IC})}\boldsymbol{\Sigma}\boldsymbol{A}_{2(\mathrm{IC})} \quad \text{and} \quad \boldsymbol{\Omega}_{2(\mathrm{IC})} = \boldsymbol{A}_{2(\mathrm{IC})}\boldsymbol{\Sigma}\boldsymbol{A}_{2(\mathrm{IC})}.$$

Note that $\mathrm{tr}(\boldsymbol{\Sigma}^2) = \mathrm{tr}[\{\boldsymbol{\Sigma}(\boldsymbol{A}_{1(\mathrm{IC})} + \boldsymbol{A}_{2(\mathrm{IC})})\}^2]$ and $\mathrm{tr}\{(\boldsymbol{\Sigma}\boldsymbol{A}_{1(\mathrm{IC})})^2\} = \kappa_{1(\mathrm{IC})}^2$ from (1). Then, we write that

$$\Delta = 2\|\boldsymbol{\Omega}_{1(\mathrm{IC})}\|_F^2 + \mathrm{tr}(\boldsymbol{\Omega}_{2(\mathrm{IC})}^2) - (p-1)\kappa_{2(\mathrm{IC})}^2 \quad (= \Delta_{\mathrm{IC}}, \text{ say}) \quad \text{and}$$

$$\Upsilon = 2\kappa_{1(\mathrm{IC})}^2\mathrm{tr}(\boldsymbol{\Omega}_{2(\mathrm{IC})}^2) + \mathrm{tr}(\boldsymbol{\Omega}_{2(\mathrm{IC})}^2)^2 + 2\|\boldsymbol{\Omega}_{1(\mathrm{IC})}\|_F^4 + 4\|\boldsymbol{\Omega}_{1(\mathrm{IC})}\|_F^2\mathrm{tr}(\boldsymbol{\Omega}_{2(\mathrm{IC})}^2).$$

Note that (C-i') holds when $\boldsymbol{\Sigma} = \boldsymbol{\Sigma}_{\mathrm{IC}}$ because

$$\frac{\mathrm{tr}(\boldsymbol{\Omega}_{2(\mathrm{IC})}^4)}{\mathrm{tr}(\boldsymbol{\Omega}_{2(\mathrm{IC})}^2)^2} = \frac{\mathrm{tr}(\boldsymbol{A}_{2(\mathrm{IC})}^4)}{\mathrm{tr}(\boldsymbol{A}_{2(\mathrm{IC})}^2)^2} = \frac{1}{\mathrm{tr}(\boldsymbol{A}_{2(\mathrm{IC})})} = \frac{1}{p-1}$$

when $\boldsymbol{\Sigma} = \boldsymbol{\Sigma}_{\mathrm{IC}}$. Here, we consider the following condition:

(C-v') $\rho \dfrac{p}{\mathrm{tr}(\boldsymbol{\Omega}_{2(\mathrm{IC})}^2)^{1/2}} \to \infty$ as $p \to \infty$ and $\limsup\limits_{p \to \infty} \rho < 1$.

Note that (C-v') implies (C-v) with $\boldsymbol{\Sigma}_* = \boldsymbol{\Sigma}_{\mathrm{IC}}$ from the fact that $\kappa_{1(\mathrm{IC})} \geq \sigma p \rho$ when $\boldsymbol{\Sigma}_* = \boldsymbol{\Sigma}_{\mathrm{IC}}$. Also, note that the first condition of (C-v') is met when $\boldsymbol{\Sigma} = \boldsymbol{\Sigma}_{\mathrm{IC}}$ and $\rho p^{1/2} \to \infty$ as $p \to \infty$. From (22) we write that

$$\widetilde{T}_n = \frac{n \widetilde{\Delta}_{n(\mathrm{IC})}}{2 \widetilde{\Psi}_{n(\mathrm{IC})}^{1/2}} \quad (= \widetilde{T}_{n(\mathrm{IC})}, \text{ say}),$$

where $\widetilde{\Delta}_{n(\mathrm{IC})} = W_n - U_{n(\mathrm{IC})}$ and

$$\widetilde{\Psi}_{n(\mathrm{IC})} = U_{n(\mathrm{IC})}^2 - \left(2 \sum_{i<j}^n \frac{\boldsymbol{y}_{ij(1)}^T \boldsymbol{A}_{1(\mathrm{IC})} \boldsymbol{y}_{ij(1)} \boldsymbol{y}_{ij(2)}^T \boldsymbol{A}_{1(\mathrm{IC})} \boldsymbol{y}_{ij(2)}}{n(n-1)} \right)^2$$

$$= U_{n(\mathrm{IC})}^2 - \left(2 \sum_{i<j}^n \frac{(\boldsymbol{y}_{ij(1)}^T \mathbf{1}_p)^2 (\boldsymbol{y}_{ij(2)}^T \mathbf{1}_p)^2}{p^2 n(n-1)} \right)^2$$

with

$$U_{n(\mathrm{IC})} = 2 \sum_{i<j}^n \frac{\boldsymbol{y}_{ij(1)}^T \boldsymbol{A}_{1(\mathrm{IC})} \boldsymbol{y}_{ij(1)} \boldsymbol{y}_{ij(2)}^T \boldsymbol{A}_{1(\mathrm{IC})} \boldsymbol{y}_{ij(2)}}{n(n-1)}$$

$$\quad + 2 \sum_{i<j}^n \frac{\boldsymbol{y}_{ij(1)}^T \boldsymbol{A}_{2(\mathrm{IC})} \boldsymbol{y}_{ij(1)} \boldsymbol{y}_{ij(2)}^T \boldsymbol{A}_{2(\mathrm{IC})} \boldsymbol{y}_{ij(2)}}{(p-1)n(n-1)}$$

$$= 2 \sum_{i<j}^n \frac{(\boldsymbol{y}_{ij(1)}^T \mathbf{1}_p)^2 (\boldsymbol{y}_{ij(2)}^T \mathbf{1}_p)^2}{p^2 n(n-1)}$$

$$\quad + 2 \sum_{i<j}^n \frac{\{\|\boldsymbol{y}_{ij(1)}\|^2 - (\boldsymbol{y}_{ij(1)}^T \mathbf{1}_p)^2/p\}\{\|\boldsymbol{y}_{ij(2)}\|^2 - (\boldsymbol{y}_{ij(2)}^T \mathbf{1}_p)^2/p\}}{(p-1)n(n-1)}.$$

Note that $E(\widetilde{\Delta}_{n(\mathrm{IC})}) = \Delta_{\mathrm{IC}}$. From Theorem 6 and Corollary 4, we have the following result.

Corollary 7 *For the test procedure (23) with $\widetilde{T}_n = \widetilde{T}_{n(IC)}$ for (6)*

(i)   (27) holds as $m \to \infty$ under (A-ii), (C-i') with $\boldsymbol{\Sigma}_* = \boldsymbol{\Sigma}_{IC}$ and (C-v'),
(ii)   (18) holds as $m \to \infty$ under (A-i), (C-ii') with $\boldsymbol{\Sigma}_* = \boldsymbol{\Sigma}_{IC}$ and (C-v').

Remark 7 For (6) Srivastava and Reid (2012) gave the following test statistic:

$$T_{\mathrm{SR}} = \frac{n-1}{\sqrt{2}} \left(\frac{\hat{\delta}_{1(\mathrm{SR})}}{\sqrt{2\hat{\kappa}_{1(\mathrm{SR})} \hat{\delta}_{2(\mathrm{SR})}}} + \frac{\hat{\delta}_{2(\mathrm{SR})} - \hat{\kappa}_{2(\mathrm{SR})}^2}{2\hat{\delta}_{2(\mathrm{SR})}} \right), \tag{31}$$

where $\hat{\delta}_{1(\mathrm{SR})} = \mathrm{tr}(\boldsymbol{A}_{1(\mathrm{IC})} \boldsymbol{S}_n \boldsymbol{A}_{2(\mathrm{IC})} \boldsymbol{S}_n) - \hat{\kappa}_{1(\mathrm{SR})} \hat{\kappa}_{2(\mathrm{SR})}/(n-1)$ and $\hat{\delta}_{2(\mathrm{SR})} = \mathrm{tr}\{(\boldsymbol{A}_{2(\mathrm{IC})} \boldsymbol{S}_n)^2\} - \hat{\kappa}_{2(\mathrm{SR})}^2/(n-1)$ with $\hat{\kappa}_{1(\mathrm{SR})} = \mathrm{tr}(\boldsymbol{A}_{1(\mathrm{IC})} \boldsymbol{S}_n)$ and $\hat{\kappa}_{2(\mathrm{SR})} = \mathrm{tr}(\boldsymbol{A}_{2(\mathrm{IC})} \boldsymbol{S}_n)$. Under H_0 in (6), they showed that as $m \to \infty$

$$T_{\mathrm{SR}} \Rightarrow N(0,1)$$

under the assumption that $\boldsymbol{x}_j$ is Gaussian and some regularity conditions. It should be noted that the test procedure by T_{SR} cannot ensure accuracy unless $\boldsymbol{x}_j$ is Gaussian. However, the test procedure (23) with $\widetilde{T}_{n(\mathrm{IC})}$ can ensure the accuracy even in non-Gaussian situations. See Section 5 for numerical comparisons.

5 Simulation studies

In this section, we check the performance of the proposed test procedure in simulations.

Throughout this section, we set $\alpha = 0.05$, $n = 2\lceil p^{1/2}\rceil$ and $p = 2^s$ for $s = 7,\ldots,12$. We handled the following three cases:

(i) $\boldsymbol{x}_j$ is $N_p(\boldsymbol{0}, \boldsymbol{\Sigma})$;

(ii) $z_{sj} = (v_{sj} - 10)/20^{1/2}$ $(s = 1,\ldots,p)$ in which v_{sj}s are i.i.d. as the chi-squared distribution with 10 degrees of freedom; and

(iii) $(z_{1j},\ldots,z_{pj})^T$s are i.i.d. as p-variate t-distribution, $t_p(\boldsymbol{I}_p, \nu)$, with mean zero, covariance matrix $\boldsymbol{I}_p$ and degrees of freedom $\nu = 20$.

Note that (A-ii) holds for (i) and (ii). However, neither (A-i) nor (A-ii) hold for (iii).

5.1 The scaled identity structure (4)

For (4), we compared the performance of the test procedure (23) between T_{CZZ} in (29) and $\widetilde{T}_{n(\mathrm{S})}$. We considered $\boldsymbol{\Sigma} = \boldsymbol{I}_p$ for H_0. As for H_1, we considered $\boldsymbol{\Sigma} = \mathrm{diag}(1, ..., 1, 2, ..., 2)$ of which the last $\lceil p^{2/3}\rceil$ elements were 2 and the remaining elements were 1, so that $\liminf_{p\to\infty} \Delta/p^{2/3} > 0$. Thus, from Corollary 5, (18) holds because $\mathrm{tr}(\boldsymbol{\Sigma}^2) = O(p)$ under H_1. For each case in (i) to (iii), we checked the performance by 2000 replications. We defined $P_r = 1$ (or 0) when H_0 was falsely rejected (or not) for $r = 1,\ldots,2000$, and calculated $\bar{\alpha} = \sum_{r=1}^{2000} P_r/2000$ to estimate the size. We also defined $P_r = 1$ (or 0) when H_1 was falsely rejected (or not) for $r = 1,\ldots,2000$, and calculated $1 - \bar{\beta} = 1 - \sum_{r=1}^{2000} P_r/2000$ to estimate the power. Note that their standard deviations are less than 0.011.

In Fig. 1, we plotted $\bar{\alpha}$ in the left panel and $1 - \bar{\beta}$ in the right panel for (i) - (iii). We observed that $\widetilde{T}_{n(\mathrm{S})}$ and T_{CZZ} give similar and preferable performances both for the size and power in (i) - (iii). This is because the asymptotic alternative distribution of $\widetilde{T}_{n(\mathrm{S})}$ is equivalent to that of T_{CZZ}. See Section 3 in Chen et al. (2010) for the asymptotic distribution of T_{CZZ}.

[**Fig. 1 should be inserted here**]

5.2 The diagonal structure (5)

For (5), we compared the performance of the test procedure (23) between T_{S} in (30) and $\widetilde{T}_{n(\mathrm{D})}$. We considered $\boldsymbol{\Sigma} = \boldsymbol{I}_p$ for H_0. As for H_1, we considered $\boldsymbol{\Sigma} = (0.2^{|i-j|^{1/3}})$, so that $\liminf_{p \to \infty} \Delta/p > 0$. Thus, from Corollary 6, (18) holds because $\mathrm{tr}(\boldsymbol{\Sigma}^2) = O(p)$ under H_1. Similar to Section 5.1, we checked the performance by 2000 replications and estimated the size and power.

In Fig. 2, we plotted $\bar{\alpha}$ in the left panel and $1 - \bar{\beta}$ in the right panel for (i) - (iii). We observed that $\widetilde{T}_{n(\mathrm{D})}$ gives preferable performances both for the size and power in (i) - (iii). On the other hand, T_{S} gave poor performances because of huge bias in high dimension.

[**Fig. 2 should be inserted here**]

5.3 The intraclass covariance structure (6)

For (6), we compared the performance of the test procedure (23) between T_{SR} in (31) and $\widetilde{T}_{n(\mathrm{IC})}$. We considered $\boldsymbol{\Sigma} = 0.5\boldsymbol{I}_p + 0.5\mathbf{1}_p\mathbf{1}_p^T$ for H_0. As for H_1, we considered

$$\boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\Sigma}_1 & \boldsymbol{O}_{(p-2,2)} \\ \boldsymbol{O}_{(2,p-2)} & \boldsymbol{I}_2 \end{pmatrix},$$

where $\boldsymbol{\Sigma}_1 = 0.5\boldsymbol{I}_{p-2} + 0.5\mathbf{1}_{p-2}\mathbf{1}_{p-2}^T$ and $\boldsymbol{O}_{(p-2,2)}$ denotes the $(p-2) \times 2$ zero matrix. Similar to Section 5.1, we checked the performance by 2000 replications and estimated the size and power.

In Fig. 3, we plotted $\overline{\alpha}$ in the left panel and $1 - \overline{\beta}$ in the right panel for (i) - (iii). We observed that $\widetilde{T}_{n(\mathrm{IC})}$ gives preferable performances both for the size and power in (i) - (iii). On the other side, T_{SR} gave poor performances for (ii) and (iii) because T_{SR} assumes that $\boldsymbol{x}_j$ is Gaussian.

[**Fig. 3 should be inserted here**]

6 Data analysis

In this section, we demonstrate the test procedure (23) by using gene expression data. Since the SSE model often appears when we analyze a microarray data set, we considered testing (6). First, we compared the performance between T_{SR} in (31) and $\widetilde{T}_{n(\mathrm{IC})}$. It should be noted that high correlation among components of $\boldsymbol{\Sigma}$ is one of the reasons why a microarray data set has the SSE model. Actually, we can find several gene clusters in a high-dimensional space and genes in a cluster are usually highly correlated to each other. See Aoshima and Yata (2018, 2019) for the details.

We used microarray data sets of colon cancer with $2000(= p)$ genes. The data sets consist of two classes: π_1: tumor (40 samples) and π_2: normal colon (22 samples). See Alon et al. (1999) for the details. The data sets are

available at Jeffery's web page (URL: http://www.bioinf.ucd.ie/people/ian/).
In order to determine whether the data sets belong to the NSSE model or the
SSE model, we calculated $\lambda_1/\mathrm{tr}(\boldsymbol{\Sigma}^2)^{1/2}$. We estimated $\mathrm{tr}(\boldsymbol{\Sigma}^2)$ by W_n in (10)
and λ_1 by using the noise-reduction methodology given by Yata and Aoshima
(2012). Then, we obtained that the estimates of $\lambda_1/\mathrm{tr}(\boldsymbol{\Sigma}^2)^{1/2}$ were 0.732 for
π_1 and 0.839 for π_2. We confirmed that each class fits the SSE model in (13).
Since the data sets fit the SSE model, we considered $\boldsymbol{\Sigma}_* = \boldsymbol{\Sigma}_{\mathrm{IC}}$ as a candidate
covariance structure.

We tested (6) at a significance level 0.05. Then, H_0 was rejected by $\widetilde{T}_{n(\mathrm{IC})}$
for π_1 and π_2. The results were summarized in Table 1.

Table 1. Test of the intraclass covariance structure (6) by T_{SR} and $\widetilde{T}_{n(\mathrm{IC})}$. We used two
data sets of colon cancer with $2000 (= p)$ genes in Alon et al. (1999). We set $\alpha = 0.05$, so
that $z_{0.05} = 1.64$.

	T_{SR}	$\widetilde{T}_{n(\mathrm{IC})}$
π_1 : tumor $(n = 40)$	13.75	1858
π_2 : normal colon $(n = 22)$	8.11	827.9

Next, we considered testing the following structure:

$$H_0 : \boldsymbol{\Sigma}\boldsymbol{A}_{1(\mathrm{IC})} = \kappa_{1(\mathrm{IC})}\boldsymbol{A}_{1(\mathrm{IC})} \quad \text{vs.} \quad H_1 : \boldsymbol{\Sigma}\boldsymbol{A}_{1(\mathrm{IC})} \neq \kappa_{1(\mathrm{IC})}\boldsymbol{A}_{1(\mathrm{IC})}. \qquad (32)$$

We used the same data sets and applied the test procedure (B.2) in Appendix
B of the online supplementary material to the data sets. We tested (32) at a
significance level 0.05. The value of the test statistic, $\widetilde{T}_{n(V)}$, for (32) is -0.154
for π_1 and 0.98 for π_2. Thus, H_0 was accepted both for π_1 and π_2. From
this data analysis we can conclude both for π_1 and π_2 that $\mathbf{1}_p/p^{1/2}$ is an
eigenvector of $\boldsymbol{\Sigma}$. Remember that H_0 in (6) was rejected both for π_1 and π_2.
We can conclude that $\boldsymbol{\Sigma}\boldsymbol{A}_{2(\mathrm{IC})} \neq \kappa_{2(\mathrm{IC})}\boldsymbol{A}_{2(\mathrm{IC})}$.

7 Conclusion

In this paper, we considered testing covariance structures systematically. By
using the ECDM method, we constructed a common test procedure for a
(i) scaled identity matrix, (ii) diagonal matrix, or (iii) intraclass covariance
matrix in a non-parametric approach. We emphasize that its test statistic can
be calculated at a low computational cost. Any eigenstructure of a covariance
matrix is classified into the SSE model or the NSSE model given by Aoshima
and Yata (2018). We showed that the proposed test procedure can establish
its asymptotic normality under both of the above models. We evaluated the
asymptotic size and power of the test procedure theoretically and numerically.
In conclusion, we recommend using the test procedure (23).

Supplementary Material

We give details of the asymptotic normality in Lemma 5, an additional test procedure, an R-code to calculate $\boldsymbol{y}_{ij(l)}$s and proofs of the theoretical results in the online supplementary material.

References

Alon, U., Barkai, N., Notterman, D.A., Gish, K., Ybarra, S., Mack, D., Levine, A.J. (1999). Broad patterns of gene expression revealed by clustering analysis of tumor and normal colon tissues probed by oligonucleotide arrays. *Proceedings of the National Academy of Sciences of the United States of America, 96*, 6745–6750.

Aoshima, M., Yata, K. (2015). Asymptotic normality for inference on multisample, high-dimensional mean vectors under mild conditions. *Methodology and Computing in Applied Probability, 17*, 419–439.

Aoshima, M., Yata, K.(2018). Two-sample tests for high-dimension, strongly spiked eigenvalue models. *Statistica Sinica, 28*, 43–62.

Aoshima, M., Yata, K. (2019). Distance-based classifier by data transformation for high-dimension, strongly spiked eigenvalue models. *Annals of the Institute of Statistical Mathematics, 71*, 473–503.

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

Bao, Z., Lin, L.-C., Pan, G., Zhou, W. (2015). Spectral statistics of large dimensional Spearman's rank correlation matrix and its application. *Annals of Statistics, 43*, 2588–2623.

Chen, S.X., Qin, Y.-L. (2010). A two-sample test for high-dimensional data with applications to gene-set testing. *The Annals of Statistics, 38*, 808–835.

Chen, S.X., Zhang, L.-X., Zhong, P.-S. (2010). Tests for high-dimensional covariance matrices. *Journal of the American Statistical Association, 105*, 810–819.

Ishii, A., Yata, K., Aoshima, M. (2016). Asymptotic properties of the first principal component and equality tests of covariance matrices in high-dimension, low-sample-size context. *Journal of Statistical Planning and Inference, 170*, 186–199.

Ishii, A., Yata, K., Aoshima, M. (2019). Equality tests of high-dimensional covariance matrices under the strongly spiked eigenvalue model. *Journal of Statistical Planning and Inference, 202*, 99–111.

Ledoit, O., Wolf, M. (2002). Some hypothesis tests for the covariance matrix when the dimension is large compared to the sample size. *Annals of Statistics, 30*, 1081–1102.

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

Srivastava, M. S., Kollo, T., Rosen, V.D. (2011). Some tests for the covariance matrix with fewer observations than the dimension under non-normality. *Journal of Multivariate Analysis, 102*, 1090–1103.

Srivastava, M. S., Reid, N. (2012). Testing the structure of the covariance matrix with fewer observations than the dimension. *Journal of Multivariate Analysis, 112*, 156–171.

Yata, K., Aoshima, M. (2010). Effective PCA for high-dimension, low-sample-size data with singular value decomposition of cross data matrix. *Journal of Multivariate Analysis, 101*, 2060–2077.

Yata, K., Aoshima, M. (2012). Effective PCA for high-dimension, low-sample-size data with noise reduction via geometric representations. *Journal of Multivariate Analysis, 105*, 193–215.

Yata, K., Aoshima, M. (2013). Correlation tests for high-dimensional data using extended cross-data-matrix methodology. *Journal of Multivariate Analysis, 117*, 313–331.

Yata, K., Aoshima, M. (2016). High-dimensional inference on covariance structures via the extended cross-data-matrix methodology. *Journal of Multivariate Analysis, 151*, 151–166.

Yata, K., Aoshima, M., Nakayama, Y. (2018). A test of sphericity for high-dimensional data and its application for detection of divergently spiked noise. *Sequential Analysis, 37*, 397–411.

Zhong, P.-S., Lan, W., Song, X.K.P., Tsai, C.H. (2017). Tests for covariance structures with high-dimensional repeated mesurements. *Annals of Statistics, 45*, 1185–1213.
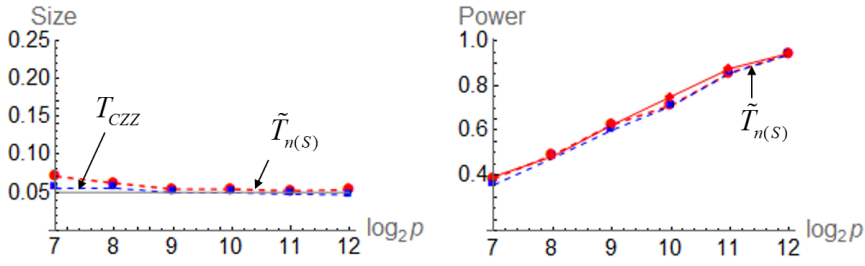
(i) When $\boldsymbol{x}_j$ is $N_p(\boldsymbol{0}, \boldsymbol{\Sigma})$.
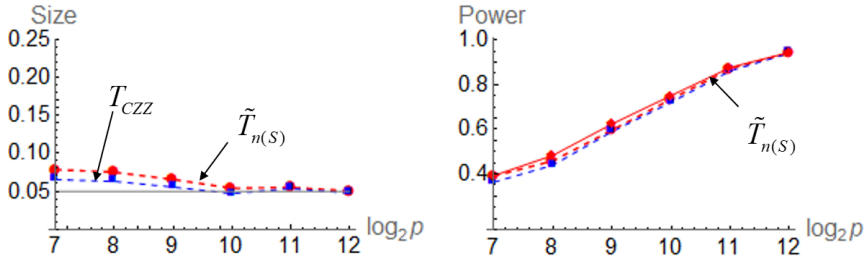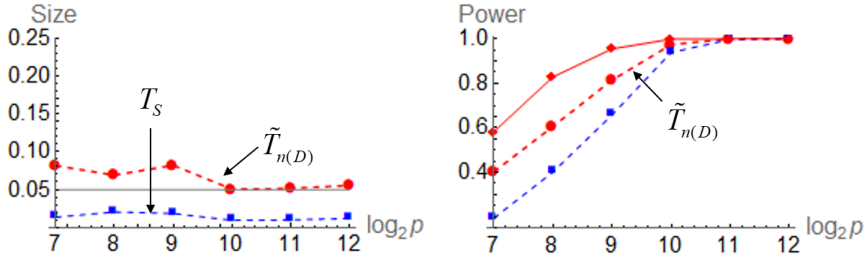

(ii) When $z_{sj} = (v_{sj} - 10)/20^{1/2}$ ($s = 1, \ldots, p$) in which v_{sj}s are i.i.d. as the chi-squared distribution with 10 degrees of freedom.
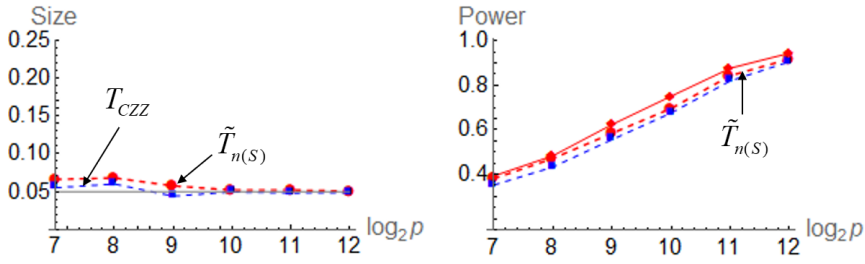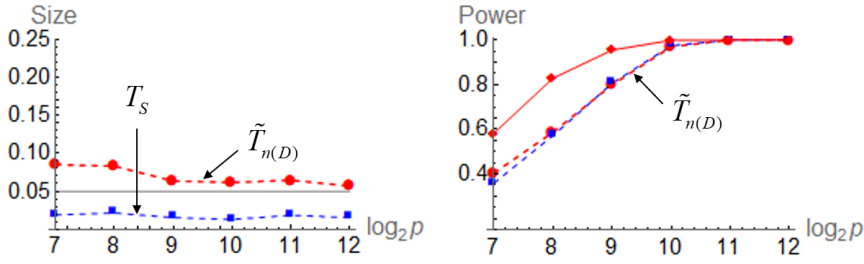

(iii) When $(z_{1j}, \ldots, z_{pj})^T$s are i.i.d. as p-variate t-distribution, $t_p(\boldsymbol{I}_p, \nu)$, with mean zero, covariance matrix $\boldsymbol{I}_p$ and degrees of freedom $\nu = 20$.
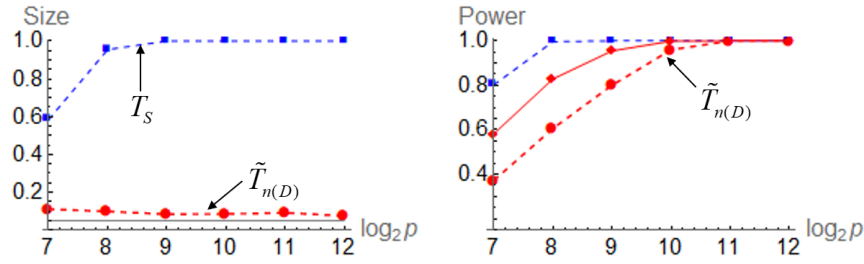
Fig. 1 The performance of the test procedures given by T_{CZZ} and $\widetilde{T}_{n(\mathrm{S})}$ in (i) - (iii). The value of $\overline{\alpha}$ is denoted by the dashed line in the left panels. The value of $1 - \overline{\beta}$ is denoted by the dashed line in the right panels together with the asymptotic power (17) which is in the solid line.

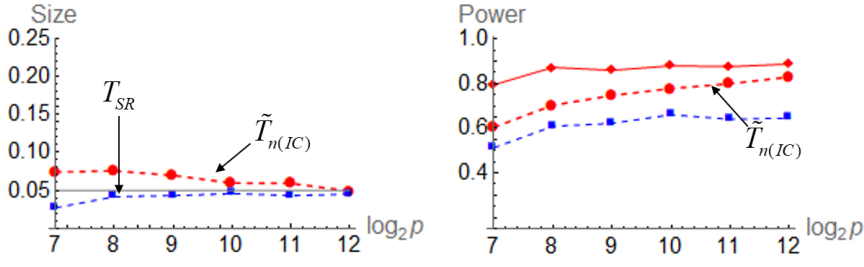(i) When $\boldsymbol{x}_j$ is $N_p(\boldsymbol{0}, \boldsymbol{\Sigma})$.

(ii) When $z_{sj} = (v_{sj} - 10)/20^{1/2}$ $(s = 1, \ldots, p)$ in which v_{sj}s are i.i.d. as the chi-squared distribution with 10 degrees of freedom.
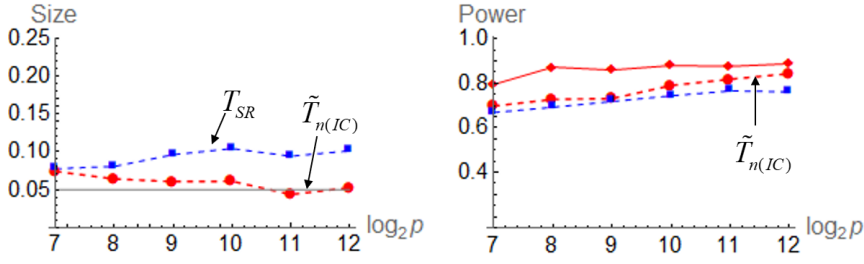

(iii) When $(z_{1j}, \ldots, z_{pj})^T$s are i.i.d. as p-variate t-distribution, $t_p(\boldsymbol{I}_p, \nu)$, with mean zero, covariance matrix $\boldsymbol{I}_p$ and degrees of freedom $\nu = 20$.
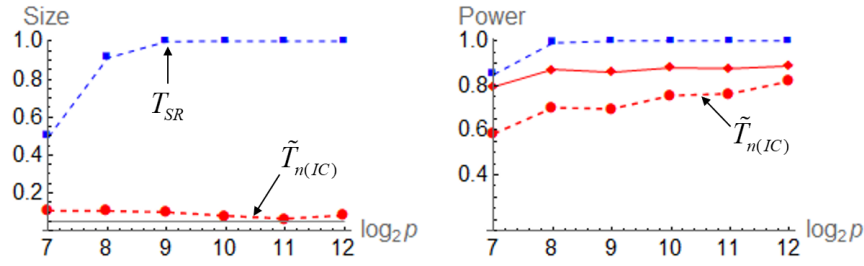
Fig. 2 The performance of the test procedures given by T_{S} and $\widetilde{T}_{n(\mathrm{D})}$ in (i) - (iii). The value of $\overline{\alpha}$ is denoted by the dashed line in the left panels. The value of $1 - \overline{\beta}$ is denoted by the dashed line in the right panels together with the asymptotic power (17) which is in the solid line.

(i) When $\boldsymbol{x}_j$ is Gaussian, $N_p(\boldsymbol{0}, \boldsymbol{\Sigma})$.

(ii) When $z_{sj} = (v_{sj} - 10)/20^{1/2}$ $(s = 1, \ldots, p)$ in which v_{sj}s are i.i.d. as the chi-squared distribution with 10 degrees of freedom.

(iii) When $(z_{1j}, \ldots, z_{pj})^T$s are i.i.d. as p-variate t-distribution, $t_p(\boldsymbol{I}_p, \nu)$, with mean zero, covariance matrix $\boldsymbol{I}_p$ and degrees of freedom $\nu = 20$.

Fig. 3 The performance of the test procedures by T_{SR} and $\widetilde{T}_{n(\mathrm{IC})}$ in (i) - (iii). The value of $\overline{\alpha}$ is denoted by the dashed line in the left panels. The value of $1 - \overline{\beta}$ is denoted by the dashed line in the right panels together with the asymptotic power (27) which is in the solid line.