# Tests for high-dimensional covariance structures under the non-strongly spiked eigenvalue model

Aki Ishii
Department of Information Sciences
Tokyo University of Science

Kazuyoshi Yata
Institute of Mathematics
University of Tsukuba

Makoto Aoshima
Institute of Mathematics
University of Tsukuba

**Abstract**

In this paper, we consider tests of high-dimensional covariance structures under the non-strongly spiked eigenvalue (NSSE) model. we produce an ECDM test statistic and propose a new test procedure based on the test statistic. We evaluate its asymptotic size and power theoretically.

## 1 Introduction

Suppose we take samples, $\boldsymbol{x}_j = (x_{1j}, \dots, x_{pj})^T,\ j = 1, \dots, n$, of size $n\ (\geq 4)$, which are independent and identically distributed (i.i.d.) as a $p\ (\geq 2)$-variate distribution. We assume that $\boldsymbol{x}_j$ has an unknown mean vector $\boldsymbol{\mu}$ and unknown (positive-semidefinite) covariance matrix $\boldsymbol{\Sigma}$. We have that $\boldsymbol{\Sigma} = \boldsymbol{H}\boldsymbol{\Lambda}\boldsymbol{H}^T$, where $\boldsymbol{\Lambda} = \mathrm{diag}(\lambda_1, ..., \lambda_p)$ is a diagonal matrix of eigenvalues of $\boldsymbol{\Sigma}$, $\lambda_1 \geq \cdots \geq \lambda_p \geq 0$, and $\boldsymbol{H} = (\boldsymbol{h}_1, ..., \boldsymbol{h}_p)$ is an orthogonal matrix of the corresponding eigenvectors. Let $\boldsymbol{x}_j = \boldsymbol{H}\boldsymbol{\Lambda}^{1/2}\boldsymbol{z}_j + \boldsymbol{\mu}$, where $\boldsymbol{z}_j = (z_{1j}, \dots, z_{pj})^T$ is considered as a sphered data vector having the zero mean vector and identity covariance matrix. Let $\sigma = \mathrm{tr}(\boldsymbol{\Sigma})/p$. Let $\sigma_{ij}$ be the $(i, j)$ element of $\boldsymbol{\Sigma}$ for $i, j = 1, \dots, p$. 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 $p \to \infty$. We denote the identity matrix of dimension $p$ by $\boldsymbol{I}_p$.

In this paper, we consider testing

$$H_0 : \boldsymbol{\Sigma} = \boldsymbol{\Sigma}_* \quad \text{vs.} \quad H_1 : \boldsymbol{\Sigma} \neq \boldsymbol{\Sigma}_*, \tag{1}$$

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

$$\boldsymbol{\Sigma}_{\mathrm{S}} = \sigma \boldsymbol{I}_p \quad \text{and} \quad \boldsymbol{\Sigma}_{\mathrm{D}} = \mathrm{diag}(\sigma_{11}, \ldots, \sigma_{pp}).$$

Ledoit and Wolf [10] gave test procedures for

$$H_0 : \boldsymbol{\Sigma} = \boldsymbol{I}_p \quad \text{vs.} \quad H_1 : \boldsymbol{\Sigma} \neq \boldsymbol{I}_p \tag{2}$$

and

$$H_0 : \boldsymbol{\Sigma} = \boldsymbol{\Sigma}_{\mathrm{S}} \quad \text{vs.} \quad H_1 : \boldsymbol{\Sigma} \neq \boldsymbol{\Sigma}_{\mathrm{S}} \tag{3}$$

when $p/n \to c > 0$ and $\boldsymbol{x}_j$ is Gaussian. Scott [11] gave test procedures for

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

when $p/n \to c > 0$ and $\boldsymbol{x}_j$ is Gaussian. Srivastava et al. [12] considered test procedures for (2) to (4) when $n/p \to 0$ under an assumption that is stronger than (A-i) given in Section 2. As for a nonparametric approach, Chen et al. [6] considered test statistics based on the U-statistic for (2) and (3). In the current paper, we take a different nonparametric approach and produce a new test statistic for (1). We utilize the extended cross-data-matrix (ECDM) method developed by Yata and Aoshima [14] which is an extension of the cross-data-matrix methodology created by Yata and Aoshima [13]. 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 [15] for the details.

In this paper, we consider constructing new test procedures for (1), including (2), (3) and (4), by using the ECDM method. In Section 2, we produce an ECDM test statistic when $\boldsymbol{\Sigma}_*$ is known. We show that the ECDM test statistic is an unbiased estimator of its test parameter even in a high-dimensional setting. In Section 3, we produce an ECDM test statistic when $\boldsymbol{\Sigma}_*$ involves unknown parameters. We propose a new test procedure based on the test statistic and evaluate its asymptotic size and power theoretically. In Section 4, we apply the new test procedure to testing (3) and (4).

## 2 Test procedure for (1) when $\boldsymbol{\Sigma}_*$ is known

In this section, we propose a test procedure for (1) when $\boldsymbol{\Sigma}_*$ is known and evaluate its asymptotic size and power theoretically. Let

$$\boldsymbol{\Sigma}_0 = \boldsymbol{\Sigma} - \boldsymbol{\Sigma}_* \quad \text{and} \quad \Delta = \|\boldsymbol{\Sigma}_0\|_F^2 = \mathrm{tr}(\boldsymbol{\Sigma}_0^2),$$

where $\|\cdot\|_F$ is the Frobenius norm. Note that $\Delta = 0$ under $H_0$ and $\Delta > 0$ under $H_1$. We regard $\Delta$ as a test parameter and construct a test procedure for (1) by using an estimator of $\Delta$.

## 2.1 Unbiased estimator of $\Delta$

We first give an unbiased estimator of $\Delta$ 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

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

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

for $k = 3, \dots, 2n-1$, 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 $\boldsymbol{V}_{n(1)(k)} \cup \boldsymbol{V}_{n(2)(k)} = \{1, \dots, n\}$ for $k = 3, \dots, 2n-1$. 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 $k = 3, \dots, 2n-1$. Let $u_{n(l)} = n_{(l)}/(n_{(l)} - 1)$ for $l = 1, 2$,

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

for all $i < j$. We note that $u_{n(l)} E(\boldsymbol{y}_{ij(l)} \boldsymbol{y}_{ij(l)}^T) = \boldsymbol{\Sigma}$ for $l = 1, 2$, and $\boldsymbol{y}_{ij(1)}$ and $\boldsymbol{y}_{ij(2)}$ are independent for all $i < j$. For example, Yata and Aoshima [14] gave an estimator of tr$(\boldsymbol{\Sigma}^2)$ as

$$W_n = \frac{2 u_{n(1)} u_{n(2)}}{n(n-1)} \sum_{i<j}^{n} (\boldsymbol{y}_{ij(1)}^T \boldsymbol{y}_{ij(2)})^2 \tag{5}$$

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

**Remark 2.1.** *One can save the computational cost of $W_n$ by using previously calculated $\overline{\boldsymbol{x}}_{(i)(k)},\ k = 3, ..., 2n-1;\ i = 1, 2$. Then, the computational cost of $W_n$ is of the order, $O(n^2 p)$.*

Now, we can give an unbiased estimator of $\Delta$ as

$$\widehat{\Delta}_n = 2 \sum_{i<j}^{n} \frac{\text{tr}\big\{(u_{n(1)} \boldsymbol{y}_{ij(1)} \boldsymbol{y}_{ij(1)}^T - \boldsymbol{\Sigma}_*)(u_{n(2)} \boldsymbol{y}_{ij(2)} \boldsymbol{y}_{ij(2)}^T - \boldsymbol{\Sigma}_*)\big\}}{n(n-1)} \tag{6}$$

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

$$\widehat{\Delta}_n = W_n + \text{tr}(\boldsymbol{\Sigma}_*^2) - 2 \sum_{i<j}^{n} \frac{\big(u_{n(1)} \boldsymbol{y}_{ij(1)}^T \boldsymbol{\Sigma}_* \boldsymbol{y}_{ij(1)} + u_{n(2)} \boldsymbol{y}_{ij(2)}^T \boldsymbol{\Sigma}_* \boldsymbol{y}_{ij(2)}\big)}{n(n-1)}. \tag{7}$$

The computational cost of $\widehat{\Delta}_n$ by (7) is much lower than that by (6) 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 $\boldsymbol{\Gamma} = (\boldsymbol{\gamma}_1, \dots, \boldsymbol{\gamma}_d)$ is a $p \times d$ matrix for some $d > 0$ such that $\boldsymbol{\Gamma}\boldsymbol{\Gamma}^T = \boldsymbol{\Sigma}$, and $\boldsymbol{w}_j = (w_{1j}, \dots, w_{dr})^T$, $j = 1, \dots, n$, are i.i.d. random vectors having $E(\boldsymbol{w}_j) = \boldsymbol{0}$ and Var$(\boldsymbol{w}_j) = \boldsymbol{I}_d$. Let Var$(w_{sj}^2) = M_s$ for $s = 1, \dots, d$. We assume that $\limsup_{p\to\infty} M_s < \infty$ for all $s$. Similar to Bai and Saranadasa [4], Chen and Qin [5] and Aoshima and Yata [1], we assume that

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

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-i) is naturally satisfied.

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

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

Note that (C-i) is equivalent to "$\lambda_1/\text{tr}(\boldsymbol{\Sigma}^2)^{1/2} \to 0$ as $p \to \infty$". Aoshima and Yata [2, 3] called (C-i) the "non-strongly spiked eigenvalue (NSSE) model". When $\boldsymbol{\Sigma} = \boldsymbol{\Sigma}_{\text{S}}$ or $\boldsymbol{\Sigma}_{\text{D}}$, (C-i) holds. On the other hand, Aoshima and Yata [2, 3] proposed the following model.

$$\liminf_{p\to\infty} \left\{ \frac{\lambda_1}{\text{tr}(\boldsymbol{\Sigma}^2)^{1/2}} \right\} > 0 \tag{8}$$

from the facts that $\lambda_1 = \sigma\{1 + (p-1)\rho\}$ and $\text{tr}(\boldsymbol{\Sigma}^2) = O(p^2)$. Aoshima and Yata [2, 3] called (8) the "strongly spiked eigenvalue (SSE) model". For instance, let us consider a spiked model as

$$\lambda_j = a_j p^{\alpha_j} \ (j = 1, \dots, g) \quad \text{and} \quad \lambda_j = c_j \ (j = g+1, \dots, p), \tag{9}$$

where $a_j$s, $c_j$s and $\alpha_j$s are positive (fixed) constants, and $g$ is a positive (fixed) integer. For (9), it is a NSSE model when $\alpha_1 < 1/2$ and a SSE model when $\alpha_1 \ge 1/2$. Ishii et al. [7, 8] gave test procedures for equality tests of high-dimensional covariance matrices under the SSE model (8). Also, Ishii et al. [9] gave a test procedure for (1) under the SSE model (8).

Let

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

We consider the divergence condition as

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

which is equivalent to $m \to \infty$. Let

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

We assume one of the following assumptions as necessary:

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

Note that (C-ii) holds under $H_0$ in (1). Ishii et al. [9] gave the following result.

**Theorem 2.1** ([9])**.** *Assume (A-i), (C-i) and (C-ii). 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.*

## 2.3 Test procedure based on $\widehat{\Delta}_n$

Note that $\mathrm{tr}(\boldsymbol{\Sigma}^2) = \mathrm{tr}(\boldsymbol{\Sigma}_*^2)$ under $H_0$ in (1). Let

$$T_n = \frac{n\widehat{\Delta}_n}{2\mathrm{tr}(\boldsymbol{\Sigma}_*^2)}. \tag{10}$$

From Theorem 2.1 we propose a test procedure for (1) by

$$\text{rejecting } H_0 \Longleftrightarrow T_n > z_\alpha, \tag{11}$$

where $z_\alpha$ is a constant such that $P\{N(0,1) > z_\alpha\} = \alpha$ with $\alpha \in (0, 1/2)$. Then, Ishii et al. [9] gave the following result.

**Theorem 2.2** ([9])**.** *Assume (A-i) and (C-i). For the test procedure (11), we have that*

$$\textit{Size} = \alpha + o(1) \quad \textit{and} \quad \textit{Power} = \Phi\left(\frac{\Delta}{K^{1/2}} - z_\alpha \frac{K_*^{1/2}}{K^{1/2}}\right) + o(1) \ \ \textit{as } m \to \infty, \tag{12}$$

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

In the next section, we apply the test procedure (11) to testing the identity structure in (2).

## 2.4 Test of the identity structure in (2)

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 (10) 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. Ishii et al. [9] gave the following result.

**Corollary 2.1** ([9]). *Assume (A-i). For the test procedure (11) with $T_n = T_n(\boldsymbol{I}_p)$ for (2), we have (12).*

**Remark 2.2.** *For (2) Chen et al. [6] gave a test procedure based on the following U statistic:*

$$T_{\text{czz}} = A_n - \Big(\sum_{j=1}^{n} \frac{\boldsymbol{x}_j^T \boldsymbol{x}_j}{n} - \sum_{j\neq j'}^{n} \frac{\boldsymbol{x}_j^T \boldsymbol{x}_{j'}}{n(n-1)}\Big) + p,$$

*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 $E(A_n) = \text{tr}(\boldsymbol{\Sigma}^2)$. The test procedure by $T_{\text{czz}}$ is asymptotically equivalent to (11) 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 (11).*

# 3 Test procedure for (1) when $\boldsymbol{\Sigma}_*$ involves unknown parameters

When $\boldsymbol{\Sigma}_* = \boldsymbol{\Sigma}_{\text{S}}$ or $\boldsymbol{\Sigma}_* = \boldsymbol{\Sigma}_{\text{D}}$, the eigenstructures are identified. Otherwise they involve unknown parameters. In this section, we construct an unbiased estimator of $\Delta$ through the eigenstructures and propose a test procedure by using the unbiased estimator.

Let $\boldsymbol{A}_j$ be a $p \times p$ known idempotent matrix with rank $r_j$ $(\geq 1)$ for $j = 1, \dots, q$, 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 $\text{tr}(\boldsymbol{A}_j) = r_j$, $\boldsymbol{A}_j^2 = \boldsymbol{A}_j$ and $\boldsymbol{A}_j \boldsymbol{A}_{j'} = \boldsymbol{O}$ for all $j$ $(\neq j')$. Let $\kappa_j$ $(\geq 0)$ be an unknown scalar such that $\text{tr}(\boldsymbol{\Sigma}\boldsymbol{A}_j) = r_j\kappa_j$ for all $j$. Hereafter, we assume that $\boldsymbol{\Sigma}_*$ has the following structure:

$$\boldsymbol{\Sigma}_* = \kappa_1 \boldsymbol{A}_1 + \cdots + \kappa_q \boldsymbol{A}_q. \tag{13}$$

Note that $\text{tr}(\boldsymbol{\Sigma}_*^2) = \sum_{j=1}^q r_j \kappa_j^2$ and $\Delta = \text{tr}(\boldsymbol{\Sigma}^2) - \text{tr}(\boldsymbol{\Sigma}_*^2)$, so that $\text{tr}(\boldsymbol{\Sigma}^2) \geq \text{tr}(\boldsymbol{\Sigma}_*^2)$. One can summarize as follows:

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

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

We 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

$$U_n = 2\sum_{s=1}^{q}\sum_{i<j}^{n} \frac{u_{n(1)}u_{n(2)}\boldsymbol{y}_{ij(1)}^T \boldsymbol{A}_s \boldsymbol{y}_{ij(1)} \boldsymbol{y}_{ij(2)}^T \boldsymbol{A}_s \boldsymbol{y}_{ij(2)}}{r_s n(n-1)} \tag{14}$$

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 (5). Then, it holds that $E(\widetilde{\Delta}_n) = \Delta$. Let $q_\star$ be the maximum integer such that

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

If $r_1 = \cdots = r_q = 1$, we set $q_\star = q$.

Let $\Psi = \text{tr}(\boldsymbol{\Sigma}_*^2)^2 - \sum_{s=1}^{q_\star} \kappa_s^4$ when $q_\star \ge 1$ and $\Psi = \text{tr}(\boldsymbol{\Sigma}_*^2)^2$ when $q_\star = 0$. Then, we give an estimator of $\Psi$ by

$$\widetilde{\Psi}_n = \begin{cases} U_n^2 - \displaystyle\sum_{s=1}^{q_\star} \Big( 2 \sum_{i<j}^{n} \frac{u_{n(1)} u_{n(2)} \boldsymbol{y}_{ij(1)}^T \boldsymbol{A}_s \boldsymbol{y}_{ij(1)} \boldsymbol{y}_{ij(2)}^T \boldsymbol{A}_s \boldsymbol{y}_{ij(2)}}{n(n-1)} \Big)^2 & (q_\star \ge 1), \\ U_n^2 & (q_\star = 0). \end{cases}$$

Note that $P(\widetilde{\Psi}_n \ge 0) = 1$. Let

$$\widetilde{T}_n = \frac{n\widetilde{\Delta}_n}{2\widetilde{\Psi}_n^{1/2}}. \tag{15}$$

Then, for (1) with (13), we propose a test procedure by

$$\text{rejecting } H_0 \Longleftrightarrow \widetilde{T}_n > z_\alpha. \tag{16}$$

We consider the test procedure (16) under the NSSE model (C-i). We asssume the following condition.

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

Then, Ishii et al. [9] gave the following result.

**Theorem 3.1** ([9])**.** *Assume (A-i), (C-i) and (C-iii). For the test procedure (16), we have (12).*

# 4 Applications of the test procedure (16) to testing (3) and (4)

In this section, we apply the test procedure (16) to testing (3) and (4).

## 4.1 Scaled identity structure (3)

We consider the case when $\boldsymbol{\Sigma}_* = \boldsymbol{\Sigma}_{\text{S}}$. 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 (15) we write that

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

where

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

Note that $E(U_{n(\mathrm{S})})=p\sigma^2$. From Theorem 3.1, Ishii et al. [9] gave the following result.

**Corollary 4.1** ([9]). *Assume (A-i). For the test procedure (16) with $\widetilde{T}_n=\widetilde{T}_{n(S)}$ for (3), we have (12).*

**Remark 4.1.** *For (3) Chen et al. [6] gave a test procedure based on a U-statistic. Although the test procedure by Chen et al. [6] is asymptotically equivalent to (16) with $\widetilde{T}_n=\widetilde{T}_{n(S)}$, the latter is more applicable to the sequential analysis ensuring prespecified accuracy. See Yata et al. [16] for the details.*

## 4.2 Diagonal structure (4)

We consider the case when $\boldsymbol{\Sigma}_*=\boldsymbol{\Sigma}_{\mathrm{D}}$. Note that $\Delta=\mathrm{tr}(\boldsymbol{\Sigma}^2)-\sum_{j=1}^{p}\sigma_{jj}^2$ (= $\Delta_{\mathrm{D}}$, say) and $q_\star=p$. Also, note that $\Psi=\mathrm{tr}(\boldsymbol{\Sigma}_{\mathrm{D}}^2)^2-\sum_{j=1}^{p}\sigma_{jj}^4$. Let $\boldsymbol{y}_{ij(l)}=(y_{1ij(l)},...,y_{pij(l)})^T$ for all $i,j,l$. From (15) we write that

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

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

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

with

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

From Theorem 3.1, Ishii et al. [9] gave the following result.

**Corollary 4.2** ([9]). *Assume (A-i). For the test procedure (16) with $\widetilde{T}_n=\widetilde{T}_{n(D)}$ for (4), we have (12).*

**Remark 4.2.** *For (4) Srivastava et al. [12] gave the following test statistic:*

$$T_s=\frac{(n-1)[c_n\{\mathrm{tr}(\boldsymbol{S}_n^2)-\mathrm{tr}(\boldsymbol{S}_n)^2/(n-1)\}-c_n\sum_{j=1}^{p}s_j^2]}{2\sqrt{\Big(c_n\sum_{j=1}^{p}s_j^2\Big)^2-\sum_{j=1}^{p}s_j^4}},\tag{17}$$

*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 (4), they showed that as $m\to\infty$*

$$T_s\Rightarrow N(0,1)$$

*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* $c_n\{\mathrm{tr}(\boldsymbol{S}_n^2) - \mathrm{tr}(\boldsymbol{S}_n)^2/(n-1)\} - c_n\sum_{j=1}^p s_j^2$ $(= \widehat{\Delta}_s,$ *say) is an estimator of* $\Delta_D$*. It should be noted that* $\widehat{\Delta}_s$ *is heavily biased unless* $\boldsymbol{x}_j$ *is Gaussian. In addition, one cannot claim* $\mathrm{Var}(\widehat{\Delta}_s/\Delta_D) < \infty$ *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.*

## Acknowledgements

Research of the first author was partially supported by Grant-in-Aid for Young Scientists, Japan Society for the Promotion of Science (JSPS), under Contract Number 18K18015. Research of the second author was partially supported by Grant-in-Aid for Scientific Research (C), JSPS, under Contract Number 18K03409. Research of the third author was partially supported by Grants-in-Aid for Scientific Research (A) and Challenging Research (Exploratory), JSPS, under Contract Numbers 20H00576 and 19K22837.

### References

[1] 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.

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

[3] 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.

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

[5] 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.

[6] 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.

[7] 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.

[8] 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.

[9] Ishii, A., Yata, K., Aoshima, M. (2019). Tests for high-dimensional covariance structures, submitted.

[10] 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.

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

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

[13] 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.

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

[15] 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.

[16] 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.

Department of Information Sciences
Tokyo University of Science
Chiba 278-8510
Japan
E-mail address: `a.ishii@rs.tus.ac.jp`