SUPPLEMENT TO: TWO-SAMPLE TESTS FOR HIGH-DIMENSION, STRONGLY SPIKED EIGENVALUE MODELS

Makoto Aoshima and Kazuyoshi Yata

University of Tsukuba

Supplementary Material

In this supplement, we give actual data analyses and proofs of the theoretical results in the main work in Aoshima and Yata (2016) together with additional simulations, additional theoretical results and proofs of the additional results. We provide a method to distinguish between the NSSE model defined by (1.4) and the SSE model defined by (1.6). We also give a method to estimate the parameters required in the test procedure (5.5). The equation numbers and the mathematical symbols used in the supplement are the same as those which are made reference to in the main document.

S1 Additional Propositions

In this section, we give two propositions and proofs of the propositions.

S1.1 Proposition S1.1

Proposition S1.1. *Let Θ be the set of positive definite matrices of dimension p. It holds that*

$$\underset{A\in\Theta}{\operatorname{argmax}}\left\{\frac{\Delta(\boldsymbol{A})}{\{K_2(\boldsymbol{A})\}^{1/2}}\right\}=c(\boldsymbol{\Sigma}_1/n_1+\boldsymbol{\Sigma}_2/n_2)^{-1}$$

for any constant $c>0$.

Proof. We assume $\boldsymbol{A}\in\Theta$. Let $\dot{\boldsymbol{\mu}}_A=\boldsymbol{\mu}_A/||\boldsymbol{\mu}_A||$ and $\boldsymbol{\Sigma}_{A\star}=\boldsymbol{\Sigma}_{1,A}/n_1+\boldsymbol{\Sigma}_{2,A}/n_2$. Then, we have that

$$2\Delta(\boldsymbol{A})/\{K_2(\boldsymbol{A})\}^{1/2}=||\boldsymbol{\mu}_A||/(\dot{\boldsymbol{\mu}}_A^T\boldsymbol{\Sigma}_{A\star}\dot{\boldsymbol{\mu}}_A)^{1/2}.$$

The eigen-decomposition of $\boldsymbol{\Sigma}_{A\star}$ is given by $\boldsymbol{\Sigma}_{A\star}=\boldsymbol{H}_A\boldsymbol{\Lambda}_A\boldsymbol{H}_A^T$, where $\boldsymbol{\Lambda}_A=\operatorname{diag}(\lambda_{1,A},...,\lambda_{p,A})$ is a diagonal matrix of eigenvalues, $\lambda_{1,A}\ge\cdots\ge\lambda_{p,A}>0$, and $\boldsymbol{H}_A=[\boldsymbol{h}_{1,A},...,\boldsymbol{h}_{p,A}]$ is an orthogonal matrix of the corresponding eigenvectors. There exist some constants $c_1,...,c_p$ such that $\dot{\boldsymbol{\mu}}_A=\sum_{j=1}^p c_j\boldsymbol{h}_{j,A}$ and $\sum_{j=1}^p c_j^2=1$. From Schwarz's inequality, it holds that $(\dot{\boldsymbol{\mu}}_A^T\boldsymbol{\Sigma}_{A\star}\dot{\boldsymbol{\mu}}_A)(\dot{\boldsymbol{\mu}}_A^T\boldsymbol{\Sigma}_{A\star}^{-1}\dot{\boldsymbol{\mu}}_A)=(\sum_{j=1}^p c_j^2\lambda_{j,A})(\sum_{j=1}^p c_j^2\lambda_{j,A}^{-1})\ge 1$, so that

$$||\boldsymbol{\mu}_A||/(\dot{\boldsymbol{\mu}}_A^T\boldsymbol{\Sigma}_{A\star}\dot{\boldsymbol{\mu}}_A)^{1/2}\le(||\boldsymbol{\mu}_A||^2\dot{\boldsymbol{\mu}}_A^T\boldsymbol{\Sigma}_{A\star}^{-1}\dot{\boldsymbol{\mu}}_A)^{1/2}=\{(\boldsymbol{\mu}_1-\boldsymbol{\mu}_2)^T(\boldsymbol{\Sigma}_1/n_1+\boldsymbol{\Sigma}_2/n_2)^{-1}(\boldsymbol{\mu}_1-\boldsymbol{\mu}_2)\}^{1/2}.$$

Note that $||\boldsymbol{\mu}_A||/(\dot{\boldsymbol{\mu}}_A^T\boldsymbol{\Sigma}_{A\star}\dot{\boldsymbol{\mu}}_A)^{1/2}=\{(\boldsymbol{\mu}_1-\boldsymbol{\mu}_2)^T(\boldsymbol{\Sigma}_1/n_1+\boldsymbol{\Sigma}_2/n_2)^{-1}(\boldsymbol{\mu}_1-\boldsymbol{\mu}_2)\}^{1/2}$ when $\boldsymbol{A}=c(\boldsymbol{\Sigma}_1/n_1+\boldsymbol{\Sigma}_2/n_2)^{-1}$ for any constant $c>0$. It concludes the result. □

S1.2 Proposition S1.2

Let us write that $\boldsymbol{\mu}_{A_{12}} = \boldsymbol{A}_1^{1/2}\boldsymbol{\mu}_1 - \boldsymbol{A}_2^{1/2}\boldsymbol{\mu}_2$ and $\boldsymbol{\Sigma}_{i,A_i} = \boldsymbol{A}_i^{1/2}\boldsymbol{\Sigma}_i\boldsymbol{A}_i^{1/2}$, $i = 1,2$. Let $\Delta(\boldsymbol{A}_1,\boldsymbol{A}_2) = ||\boldsymbol{\mu}_{A_{12}}||^2$ and $K(\boldsymbol{A}_1,\boldsymbol{A}_2) = K_1(\boldsymbol{A}_1,\boldsymbol{A}_2) + K_2(\boldsymbol{A}_1,\boldsymbol{A}_2)$, where $K_1(\boldsymbol{A}_1,\boldsymbol{A}_2) = 2\sum_{i=1}^2 \mathrm{tr}(\boldsymbol{\Sigma}_{i,A_i}^2)/\{n_i(n_i-1)\}+4\mathrm{tr}(\boldsymbol{\Sigma}_{1,A_i}\boldsymbol{\Sigma}_{2,A_i})/(n_1n_2)$ and $K_2(\boldsymbol{A}_1,\boldsymbol{A}_2) = 4\sum_{i=1}^2 \boldsymbol{\mu}_{A_{12}}^T\boldsymbol{\Sigma}_{i,A}\boldsymbol{\mu}_{A_{12}}/n_i$. Note that $E\{T(\boldsymbol{A}_1,\boldsymbol{A}_2)\} = \Delta(\boldsymbol{A}_1,\boldsymbol{A}_2)$ and $\mathrm{Var}\{T(\boldsymbol{A}_1,\boldsymbol{A}_2)\} = K(\boldsymbol{A}_1,\boldsymbol{A}_2)$. Then, we have the following result.

Proposition S1.2. *Assume (A-i) and the following conditions:*

(S-i) $\dfrac{\{\lambda_{\max}(\boldsymbol{\Sigma}_{i,A_i})\}^2}{tr(\boldsymbol{\Sigma}_{i,A_i}^2)} \to 0$ *as* $p\to\infty$ *for* $i = 1,2$;

(S-ii) $\dfrac{\{\Delta(\boldsymbol{A}_1,\boldsymbol{A}_2)\}^2}{K_1(\boldsymbol{A}_1,\boldsymbol{A}_2)} \to 0$ *as* $m\to\infty$ *under* H_0.

Then, it holds that as $m\to\infty$

$$P\Big(\frac{T(\boldsymbol{A}_1,\boldsymbol{A}_2)}{\{K_1(\boldsymbol{A}_1,\boldsymbol{A}_2)\}^{1/2}} > z_\alpha\Big) = \alpha + o(1) \quad \text{under } H_0.$$

Proof. From Theorem 2 and Lemma 1, the result is obtained straightforwardly. □

Note that (S-i) is naturally met when $\boldsymbol{A}_i = \boldsymbol{\Sigma}_i^{-1}$, $i = 1,2$, because $\boldsymbol{\Sigma}_{i,A_i} = \boldsymbol{I}_p$ when $\boldsymbol{A}_i = \boldsymbol{\Sigma}_i^{-1}$. However, (S-ii) is difficult to meet when $\boldsymbol{\Sigma}_1 \neq \boldsymbol{\Sigma}_2$ and $\boldsymbol{A}_i = \boldsymbol{\Sigma}_i^{-1}$, $i = 1,2$. For example, when $\boldsymbol{\Sigma}_1 = c\boldsymbol{\Sigma}_2 = \boldsymbol{I}_p$ $(c > 1)$ and $\boldsymbol{\mu}_1 = \boldsymbol{\mu}_2 = (1,...,1)^T$, it follows that $\Delta(\boldsymbol{\Sigma}_1^{-1},\boldsymbol{\Sigma}_2^{-1}) = (1-c^{1/2})^2p$. Then, (S-ii) does not hold because $K_1(\boldsymbol{\Sigma}_1^{-1},\boldsymbol{\Sigma}_2^{-1}) = O(p/n_{\min}^2)$. Hence, we do not recommend to choose $\boldsymbol{A}_i = \boldsymbol{\Sigma}_i^{-1}$, $i = 1,2$. In addition, it is difficult to estimate $\boldsymbol{\Sigma}_i^{-1}$s for high-dimension, non-sparse data.

S2 How to Check SSE Models and Estimate Parameters

In this section, we provide a method to distinguish between the NSSE model defined by (1.4) and the SSE model defined by (1.6). We also give a method to estimate the parameters required in the test procedure (5.5).

S2.1 Checking Whether (1.4) Holds or Not

As discussed in Section 3, we recommend to use the test by (3.1) with $\boldsymbol{A} = \boldsymbol{I}_p$ when (A-ii) is met, otherwise the test by (5.5). It is crucial to check whether (1.4) holds or not (that is, whether (1.6) holds).

Let $\hat{\eta}_i = \tilde{\lambda}_{i1}^2/W_{in_i}$ for $i = 1,2$, where W_{in_i}s are defined in Section 2.2 and $\tilde{\lambda}_{ij}$s are defined by (5.2). Then, we have the following result.

Proposition S2.1. *Assume (A-i). It holds that as* $m \to \infty$

$$\hat{\eta}_i = o_P(1) \quad \text{for } i = 1, 2, \text{ under (1.4)};$$
$$P(\hat{\eta}_i > c) \to 1 \quad \text{with some fixed constant } c \in (0, 1) \text{ for some } i \text{ under (1.6)}.$$

By using Proposition S2.1, one can distinguish between (1.4) and (1.6). One may claim (1.4) if both $\hat{\eta}_1$ and $\hat{\eta}_2$ are sufficiently small, otherwise (1.6). In addition, we have the following result for $\hat{\eta}_i$.

Proposition S2.2. *Assume (A-viii). Assume also* $\lambda_{i1}^2/tr(\boldsymbol{\Sigma}_i^2) = O(n_i^{-c})$ *as* $m \to \infty$ *with some fixed constant* $c > 1/2$ *for* $i = 1, 2$. *It holds as* $m \to \infty$

$$P\Big(\hat{\eta}_i < \kappa(n_i)\Big) \to 1 \quad \text{for } i = 1, 2,$$

where $\kappa(n_i)$ *is a function such that* $\kappa(n_i) \to 0$ *and* $n_i^{1/2}\kappa(n_i) \to \infty$ *as* $n_i \to \infty$.

From Proposition S2.2 one may claim (1.4) if $\hat{\eta}_i < \kappa(n_i)$ both for $i = 1, 2$, otherwise (1.6). One can choose $\kappa(n_i)$ such as $(n_i^{-1}\log n_i)^{1/2}$ or n_i^{-c} with $c \in (0, 1/2)$. In Section S3, we use $\kappa(n_i) = (n_i^{-1}\log n_i)^{1/2}$ in actual data analyses.

S2.2 Estimation of $\Psi_{i(j)}$ and k_i

Let $n_{i(1)} = \lceil n_i/2 \rceil$ and $n_{i(2)} = n_i - n_{i(1)}$. Let $\boldsymbol{X}_{i1} = [\boldsymbol{x}_{i1}, ..., \boldsymbol{x}_{in_{i(1)}}]$ and $\boldsymbol{X}_{i2} = [\boldsymbol{x}_{in_{i(1)}+1}, ..., \boldsymbol{x}_{in_i}]$ for $i = 1, 2$. We define

$$\boldsymbol{S}_{iD(1)} = \{(n_{i(1)} - 1)(n_{i(2)} - 1)\}^{-1/2}(\boldsymbol{X}_{i1} - \overline{\boldsymbol{X}}_{i1})^T(\boldsymbol{X}_{i2} - \overline{\boldsymbol{X}}_{i2})$$

for $i = 1, 2$, where $\overline{\boldsymbol{X}}_{ij} = [\overline{\boldsymbol{x}}_{in_i(j)}, ..., \overline{\boldsymbol{x}}_{in_i(j)}]$ with $\overline{\boldsymbol{x}}_{in_i(1)} = \sum_{l=1}^{n_{i(1)}} \boldsymbol{x}_{il}/n_{i(1)}$ and $\overline{\boldsymbol{x}}_{in_i(2)} = \sum_{l=n_{i(1)}+1}^{n_i} \boldsymbol{x}_{il}/n_{i(2)}$. By using the cross-data-matrix (CDM) methodology by Yata and Aoshima (2010), we estimate λ_{ij} by the j-th singular value, $\acute{\lambda}_{ij}$, of $\boldsymbol{S}_{iD(1)}$, where $\acute{\lambda}_{i1} \geq \cdots \geq \acute{\lambda}_{in_{i(2)}-1} \geq 0$. Yata and Aoshima (2010, 2013) showed that $\acute{\lambda}_{ij}$ has several consistency properties for high-dimensional non-Gaussian data. Aoshima and Yata (2011) applied the CDM methodology to obtaining an unbiased estimator of $\mathrm{tr}(\boldsymbol{\Sigma}_i^2)$ by $\mathrm{tr}(\boldsymbol{S}_{iD(1)}\boldsymbol{S}_{iD(1)}^T)$, $i = 1, 2$. Note that $E\{\mathrm{tr}(\boldsymbol{S}_{iD(1)}\boldsymbol{S}_{iD(1)}^T)\} = \mathrm{tr}(\boldsymbol{\Sigma}_i^2)$. Based on the CDM methodology, we consider estimating $\Psi_{i(j)}$ as follows: Let $\widehat{\Psi}_{i(1)} = \mathrm{tr}(\boldsymbol{S}_{iD(1)}\boldsymbol{S}_{iD(1)}^T)$ and

$$\widehat{\Psi}_{i(j)} = \mathrm{tr}(\boldsymbol{S}_{iD(1)}\boldsymbol{S}_{iD(1)}^T) - \sum_{l=1}^{j-1} \acute{\lambda}_{il}^2 \quad \text{for } i = 1, 2;\ j = 2, ..., n_{i(2)}. \tag{S2.1}$$

Note that $\widehat{\Psi}_{i(j)} \geq 0$ w.p.1 for $j = 1, ..., n_{i(2)}$. Then, we have the following result.

Lemma S2.1. *Assume (A-i) and (A-vi). Then, it holds that* $\widehat{\Psi}_{i(j)}/\Psi_{i(j)} = 1 + o_P(1)$ *as* $m \to \infty$ *for* $i = 1, 2;\ j = 1, ..., k_i + 1$.

Let $\hat{\tau}_{i(j)} = \widehat{\Psi}_{i(j+1)}/\widehat{\Psi}_{i(j)}$ $(= 1 - \acute{\lambda}_{ij}^2/\widehat{\Psi}_{i(j)})$ for $i = 1, 2$. Note that $\hat{\tau}_{i(j)} \in [0, 1)$ for $\acute{\lambda}_{ij} > 0$. Then, we have the following result.

Proposition S2.3. *Assume (A-i) and (A-vi). It holds for $i = 1, 2$ that as $m \to \infty$*

$$P(\hat{\tau}_{i(j)} < 1 - c_j) \to 1 \quad \text{with some fixed constant } c_j \in (0,1) \text{ for } j = 1, ..., k_i;$$
$$\hat{\tau}_{i(k_i+1)} = 1 + o_P(1).$$

From Proposition S2.3, one may choose k_i as the first integer j such that $1 - \hat{\tau}_{i(j+1)}$ is sufficiently small. In addition, we have the following result for $\hat{\tau}_{i(k_i+1)}$.

Proposition S2.4. *Assume (A-vi), (A-viii) and (A-ix). Assume also $\lambda_{ik_i+1}^2/\Psi_{i(k_i+1)} = O(n_i^{-c})$ as $m \to \infty$ with some fixed constant $c > 1/2$ for $i = 1, 2$. It holds for $i = 1, 2$ that as $m \to \infty$*

$$P\Big(\hat{\tau}_{i(k_i+1)} > \{1 + (k_i + 1)\kappa(n_i)\}^{-1}\Big) \to 1,$$

where $\kappa(n_i)$ is defined in Proposition S2.2.

From Propositions S2.3 and S2.4, if one can assume the conditions in Proposition S2.4, one may consider k_i as the first integer j $(= \hat{k}_{oi},$ say) such that

$$\hat{\tau}_{i(j+1)}\{1 + (j + 1)\kappa(n_i)\} > 1 \quad (j \geq 0). \tag{S2.2}$$

Then, it holds that $P(\hat{k}_{oi} = k_i) \to 1$ as $m \to \infty$. Note that $\widehat{\Psi}_{i(n_{i(2)})} = 0$ from the fact that $\text{rank}(\boldsymbol{S}_{iD(1)}) \leq n_{i(2)} - 1$. Thus one may choose k_i as $\hat{k}_i = \min\{\hat{k}_{oi}, n_{i(2)} - 2\}$ in actual data analyses. For $\kappa(n_i) = (n_i^{-1} \log n_i)^{1/2}$ in (S2.2), the test procedure by (5.5) with $k_i = \hat{k}_i$, $i = 1, 2$, gave preferable performances throughout our simulations in Sections 6 and S4.2. If $\hat{k}_i = 0$ (that is, (S2.2) holds when $j = 0$), one may consider the test with $\boldsymbol{A}_{i(k_i)} = \boldsymbol{I}_p$. In addition, if $\hat{k}_i = 0$ for $i = 1, 2$, we recommend to use the test by (3.1) with $\boldsymbol{A} = \boldsymbol{I}_p$.

S3 Demonstration

In this section, we introduce two high-dimensional data sets that have the SSE model. We demonstrate the proposed test procedure by (5.5) by using the microarray data sets. We set $\alpha = 0.05$.

We first analyzed leukemia data with 7129 $(= p)$ genes consisting of π_1 : acute lymphoblastic leukemia ($n_1 = 47$ samples) and π_2 : acute myeloid leukemia ($n_2 = 25$ samples) given by Golub et al. (1999). We transformed each sample by $\boldsymbol{x}_{ij} - (\overline{\boldsymbol{x}}_{1n_1} + \overline{\boldsymbol{x}}_{2n_2})/2$ for all i, j, so that $\boldsymbol{\mu}_1 = \boldsymbol{\mu}_2 = 0$ under $H_0 : \ \boldsymbol{\mu}_1 = \boldsymbol{\mu}_2$. Then, (A-vii) and (A-x) hold under H_0. We calculated that $\hat{\eta}_1 = 0.697$ and $\hat{\eta}_2 = 0.602$. Since $\hat{\eta}_i$s are larger than $(n_1^{-1} \log n_1)^{1/2} = 0.286$ or $(n_2^{-1} \log n_2)^{1/2} = 0.359$, we concluded from Proposition S2.2 that (1.6) holds for $i = 1, 2$. We used the test procedure by (5.5). We set $\kappa(n_i) = (n_i^{-1} \log n_i)^{1/2}$ in (S2.2). Let $\tilde{\tau}_{i(j)} = \hat{\tau}_{i(j)}\{1 + j\kappa(n_i)\}$ for all i, j. We calculated that $(\tilde{\tau}_{1(1)}, \tilde{\tau}_{1(2)}, \tilde{\tau}_{1(3)}) = (0.407, 0.993, 1.302)$ and $(\tilde{\tau}_{2(1)}, \tilde{\tau}_{2(2)}, \tilde{\tau}_{2(3)}, \tilde{\tau}_{2(4)}) = (0.579, 0.7, 0.902, 1.307)$, so that $\hat{k}_1 = 2$ and $\hat{k}_1 = 3$. Thus, we chose $k_1 = 2$ and $k_2 = 3$. We calculated that $\widehat{T}_*/\widehat{K}_{1*}^{1/2} = 46.866$. By using (5.5), we rejected H_0 with size 0.05 according to the arguments in Section 5.2.

Next, we analyzed prostate cancer data with 12625 ($= p$) genes consisting of π_1 : normal prostate ($n_1 = 50$ samples) and π_2 : prostate tumor ($n_2 = 52$ samples) given by Singh et al. (2002). We transformed each sample as before. We calculated that $(\hat{\eta}_1, \hat{\eta}_2) = (1.01, 1.009)$ and $(\hat{k}_1, \hat{k}_2) = (4, 3)$ from (S2.2) with $\kappa(n_i) = (n_i^{-1} \log n_i)^{1/2}$. Hence, we used the test procedure by (5.5) with $k_1 = 4$ and $k_2 = 3$. Then, we calculated that $\widehat{T}_*/\widehat{K}_{1*}^{1/2} = 27.497$. Hence, we rejected H_0 by using (5.5). In addition, we considered two cases: (a) π_1 : the first 25 samples ($n_1 = 25$) and π_2 : the last 25 samples ($n_2 = 25$) from the normal prostate; and (b) π_1 : the first 26 samples ($n_1 = 26$) and π_2 : the last 26 samples ($n_2 = 26$) from the prostate tumor. Note that H_0 is true for (a) and (b). We applied the test procedure by (5.5) to the cases. Then, we accepted H_0 both for (a) and (b). We also applied the test procedures by (3.1) with $\boldsymbol{A} = \boldsymbol{I}_p$ and (4.2) to the cases. Then, H_0 was rejected by them both for (a) and (b).

S4 Additional Simulations

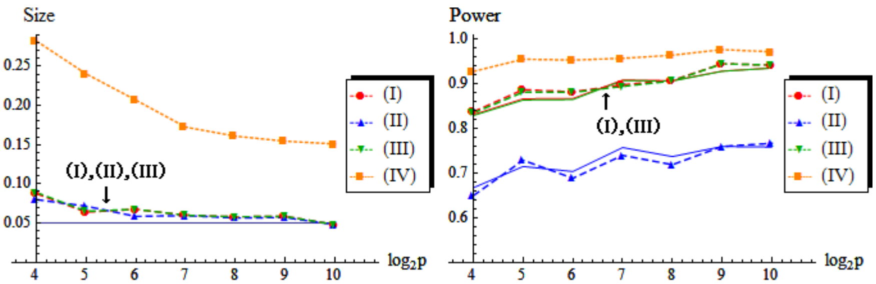
In this section, we give additional simulations for Sections 3.3 and 6 of the main work in Aoshima and Yata (2016).

S4.1 Simulations for NSSE Model

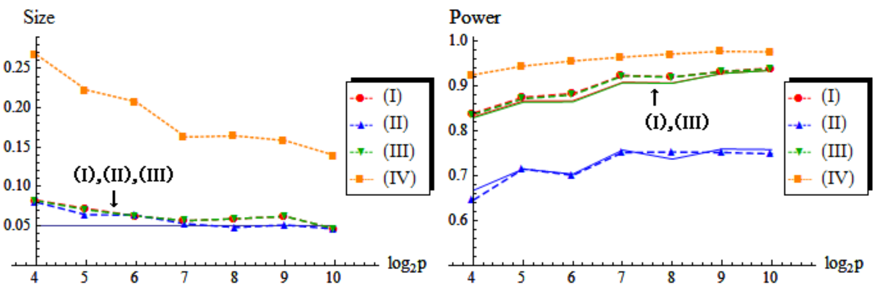
In this section, we give additional simulations for Section 3.3 under the NSSE model.

We set $\alpha = 0.05$, $p = 2^s$, $s = 4, ..., 10$, $n_1 = \lceil p^{1/2} \rceil$, $n_2 = 2n_1$ and $\boldsymbol{\mu}_1 = \boldsymbol{0}$. When considering the alternative hypothesis, we set $\boldsymbol{\mu}_2 = (1, ..., 1, 0, ..., 0, -1, ..., -1)^T$ whose first 5 elements are 1 and last 5 elements are -1. We generated $\breve{\boldsymbol{x}}_{ij}$, $j = 1, 2, ...,$ $(i = 1, 2)$ independently from a multivariate skew normal (MSN) distribution, $\mathrm{SN}_p(\boldsymbol{\Omega}, \boldsymbol{\alpha})$, with correlation matrix $\boldsymbol{\Omega} = (0.3^{|i-j|^{1/2}})$ and shape parameter vector $\boldsymbol{\alpha}$. Note that $E(\breve{\boldsymbol{x}}_{ij}) = (2/\pi)^{1/2} \boldsymbol{\Omega\alpha}/(1 + \boldsymbol{\alpha}^T \boldsymbol{\Omega\alpha})^{1/2}$ ($= \breve{\boldsymbol{\mu}}$, say) and $\mathrm{Var}(\breve{\boldsymbol{x}}_{ij}) = \boldsymbol{\Omega} - \breve{\boldsymbol{\mu}}\breve{\boldsymbol{\mu}}^T$ ($= \breve{\boldsymbol{\Sigma}}$, say). We set $\boldsymbol{x}_{ij} = c_i^{1/2}(\breve{\boldsymbol{x}}_{ij} - \breve{\boldsymbol{\mu}}) + \boldsymbol{\mu}_i$ for all i, j, where $(c_1, c_2) = (1, 1.5)$. Note that $\boldsymbol{\Sigma}_1 = \breve{\boldsymbol{\Sigma}}$ and $\boldsymbol{\Sigma}_2 = 1.5\breve{\boldsymbol{\Sigma}}$. We considered three cases: (a) $\boldsymbol{\alpha} = \boldsymbol{1}_p$; (b) $\boldsymbol{\alpha} = 4\boldsymbol{1}_p$; and (c) $\boldsymbol{\alpha} = 16\boldsymbol{1}_p$, where $\boldsymbol{1}_p = (1, ..., 1)^T$. See Azzalini and Dalla Valle (1996) and Azzalini and Capitanio (1999) for the details of the MSN distribution. Note that (1.4) is met. Also, note that (A-i) is met. See Remark S4.1. Similar to Section 3.3, we calculated $\overline{\alpha}$ and $1 - \overline{\beta}$ with 2000 replications for the test procedures given by (3.1) with (I) $\boldsymbol{A} = \boldsymbol{I}_p$, (II) $\boldsymbol{A} = \boldsymbol{A}_\star$, (III) $\boldsymbol{A} = \boldsymbol{A}_{\star(d)}$ and (IV) $\boldsymbol{A} = \widehat{\boldsymbol{A}}_{\star(d)}$. Note that (A-iv) is met for (I) to (III). In Fig. S4.1, for (a) to (c), we plotted $\overline{\alpha}$ in the left panel and $1 - \overline{\beta}$ in the right panel. We also plotted the asymptotic power, $\Phi(\Delta(\boldsymbol{A})/\{K(\boldsymbol{A})\}^{1/2} - z_\alpha\{K_1(\boldsymbol{A})/K(\boldsymbol{A})\}^{1/2})$, for (I) to (III) by using Theorem 3.

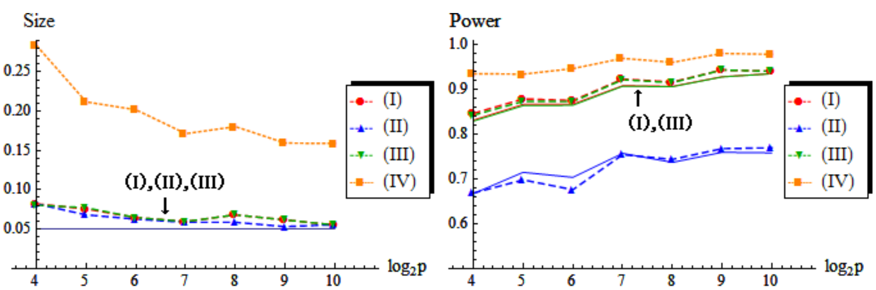
We observed that the plots become close to the theoretical value even for the skewed distributions. The tests with (I) and (III) gave similar performances for (a) to (c). This is probably because $\sigma_{i(j)} \to c_i$ as $p \to \infty$ for all i, j in those settings. Similar to Fig. 1, the test with (I) gave better performances compared to (II) for (a) to (c). See Sections 3.2 and 3.3 for the details.



(a) $\mathrm{SN}_p(\boldsymbol{\Omega}, \boldsymbol{\alpha})$ with $\boldsymbol{\alpha} = \mathbf{1}_p$.



(b) $\mathrm{SN}_p(\boldsymbol{\Omega}, \boldsymbol{\alpha})$ with $\boldsymbol{\alpha} = 4\mathbf{1}_p$.



(c) $\mathrm{SN}_p(\boldsymbol{\Omega}, \boldsymbol{\alpha})$ with $\boldsymbol{\alpha} = 16\mathbf{1}_p$.

Figure S4.1: Test procedures by (3.1) when (I) $\boldsymbol{A} = \boldsymbol{I}_p$, (II) $\boldsymbol{A} = \boldsymbol{A}_\star$, (III) $\boldsymbol{A} = \boldsymbol{A}_{\star(d)}$ and (IV) $\boldsymbol{A} = \widehat{\boldsymbol{A}}_{\star(d)}$ for $p = 2^s$, $s = 4, ..., 10$, $n_1 = \lceil p^{1/2} \rceil$ and $n_2 = 2n_1$. For (a) to (c), the values of $\overline{\alpha}$ are denoted by the dashed lines in the left panel and the values of $1 - \overline{\beta}$ are denoted by the dashed lines in the right panel. The asymptotic powers were given by $\Phi(\Delta(\boldsymbol{A})/\{K(\boldsymbol{A})\}^{1/2} - z_\alpha\{K_1(\boldsymbol{A})/K(\boldsymbol{A})\}^{1/2})$ for (I) to (III) which are denoted by the solid lines in the right panels.

Remark S4.1. Let $\boldsymbol{b}_1 = \boldsymbol{\Omega}^{1/2}\boldsymbol{\alpha}/||\boldsymbol{\Omega}^{1/2}\boldsymbol{\alpha}||$ and $\boldsymbol{b}_2, ..., \boldsymbol{b}_p$ be p-dimensional vectors such that $||\boldsymbol{b}_s|| = 1$, $\boldsymbol{b}_1^T\boldsymbol{b}_s = 0$ for $s = 2, ..., p$, and $\sum_{s=1}^p \boldsymbol{b}_s\boldsymbol{b}_s^T = \boldsymbol{I}_p$. Then, from Propositions 3 and 6 in Azzalini and Capitanio (1999), $\boldsymbol{b}_1^T\boldsymbol{\Omega}^{-1/2}\breve{\boldsymbol{x}}_{ij}, ..., \boldsymbol{b}_p^T\boldsymbol{\Omega}^{-1/2}\breve{\boldsymbol{x}}_{ij}$ are independent. Hence, (A-i) is met from the fact that $\boldsymbol{x}_{ij} - \boldsymbol{\mu}_i = c_i^{1/2}\sum_{s=1}^p \boldsymbol{\Omega}^{1/2}\boldsymbol{b}_s\{\boldsymbol{b}_s^T\boldsymbol{\Omega}^{-1/2}(\breve{\boldsymbol{x}}_{ij} - \breve{\boldsymbol{\mu}})\}$.

S4.2 Simulations for SSE Model

In this section, we give additional simulations for Section 6 under the SSE model.

We set $\alpha = 0.05$, $\boldsymbol{\mu}_1 = \mathbf{0}$ and

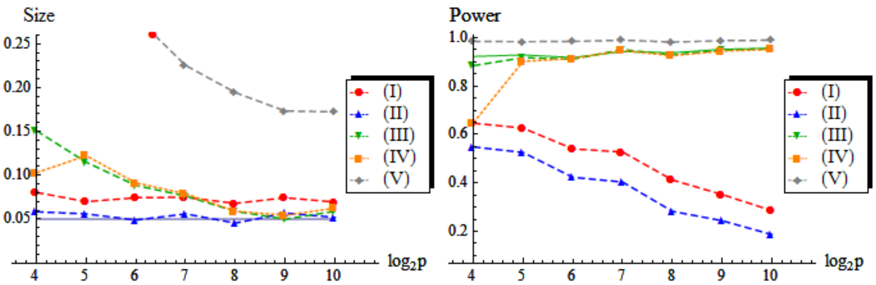
$$\boldsymbol{\Sigma}_i = \begin{pmatrix} \boldsymbol{\Sigma}_{(1)} & \boldsymbol{O}_{2,p-2} \\ \boldsymbol{O}_{p-2,2} & \boldsymbol{\Sigma}_{i(2)} \end{pmatrix} \quad \text{with } \boldsymbol{\Sigma}_{(1)} = \text{diag}(p^{2/3}, p^{1/2}) \tag{S4.1}$$

for $i = 1, 2$. When considering the alternative hypothesis, we set $\boldsymbol{\mu}_2 = (0, ..., 0, 1, 1, 1, 1)^T$ whose last 4 elements are 1. We set $\kappa(n_i) = (n_i^{-1} \log n_i)^{1/2}$ in (S2.2). We checked the performance of five tests: (I) from (3.1) with $\boldsymbol{A} = \boldsymbol{I}_p$, (II) from (4.2), (III) from (5.5), (IV) from (5.5) with $k_i = \hat{k}_i$, $i = 1, 2$, and (V) from (6.1). Let us write that $\boldsymbol{x}_{ij} = (x_{i1(j)}, ..., x_{ip(j)})^T$, $\boldsymbol{\mu}_i = (\mu_{i1}, ..., \mu_{ip})^T$, $\boldsymbol{x}_{ij(2)} = (x_{i3(j)}, ..., x_{ip(j)})^T$ and $\boldsymbol{\mu}_{i(2)} = (\mu_{i3}, ..., \mu_{ip})^T$ for all i, j. We supposed that $(x_{i1(j)}, x_{i2(j)})^T$s are i.i.d. as $N_2(\mathbf{0}, \boldsymbol{\Sigma}_{(1)})$.

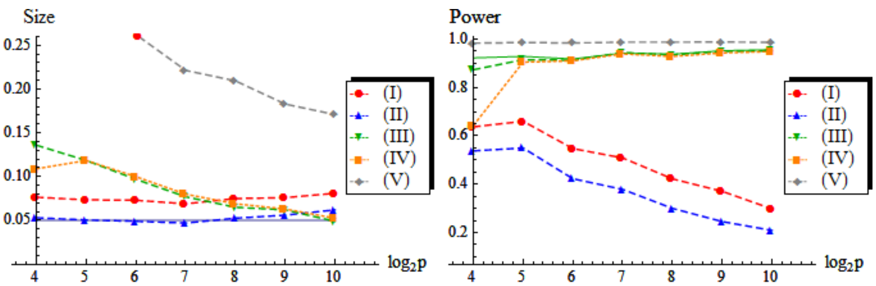
First, we checked the performance of the test procedures for the MSN distribution. We set $p = 2^s$, $n_1 = 3\lceil p^{1/2} \rceil$ and $n_2 = 4\lceil p^{1/2} \rceil$ for $s = 4, ..., 10$. We generated $\breve{\boldsymbol{x}}_{ij(2)}$, $j = 1, 2, ...,$ $(i = 1, 2)$ independently from $\text{SN}_{p-2}(\boldsymbol{\Omega}_i, \boldsymbol{\alpha})$ with $\boldsymbol{\Omega}_1 = (0.3^{|i-j|^{1/2}})$ and $\boldsymbol{\Omega}_2 = (0.5^{|i-j|^{1/2}})$, where $(x_{i1(j)}, x_{i2(j)})^T$ and $\breve{\boldsymbol{x}}_{ij(2)}$ are independent for each j. We considered two cases: (a) $\boldsymbol{\alpha} = 4\mathbf{1}_{p-2}$; and (b) $\boldsymbol{\alpha} = 16\mathbf{1}_{p-2}$. Similar to Section S4.1, we set $\boldsymbol{x}_{ij(2)} = \breve{\boldsymbol{x}}_{ij(2)} - \breve{\boldsymbol{\mu}}_i + \boldsymbol{\mu}_{i(2)}$ for all i, j, where $\breve{\boldsymbol{\mu}}_i = E(\breve{\boldsymbol{x}}_{ij(2)}) = (2/\pi)^{1/2} \boldsymbol{\Omega}_i \boldsymbol{\alpha} / (1 + \boldsymbol{\alpha}^T \boldsymbol{\Omega}_i \boldsymbol{\alpha})^{1/2}$, $i = 1, 2$. Then, we had $\boldsymbol{\Sigma}_{i(2)} = \boldsymbol{\Omega}_i - \breve{\boldsymbol{\mu}}_i \breve{\boldsymbol{\mu}}_i^T$, $i = 1, 2$, in (S4.1). Note that (4.1) and (A-vi) with $k_1 = k_2 = 2$ are met. Similar to Remark S4.1, we note that (A-i) is met. However, (A-viii) is not met. Similar to Section 6, we calculated $\overline{\alpha}$ and $1 - \overline{\beta}$ with 2000 replications for the five test procedures. In Fig. S4.2, for (a) and (b), we plotted $\overline{\alpha}$ in the left panel and $1 - \overline{\beta}$ in the right panel. We observed the performances similar to those in Fig. 2 (a).

Next, we checked the performance of the test procedures for the multivariate skew t (MST) distribution. See Azzalini and Capitanio (2003) and Gupta (2003) for the details of the MST distribution. We considered two cases: (i) $(n_1, n_2) = (40, 60)$ and $p = 50 + 100(s - 1)$ for $s = 1, ..., 7$; and (ii) $p = 500$, $n_1 = 10s$ and $n_2 = 1.5n_1$ for $s = 2, ..., 8$. We generated $\breve{\boldsymbol{x}}_{ij(2)}$, $j = 1, 2, ...,$ $(i = 1, 2)$ independently from a MST distribution, $\text{ST}_{p-2}(\boldsymbol{\Omega}_i, \boldsymbol{\alpha}, \nu)$, with correlation matrix $\boldsymbol{\Omega}_i$, shape parameter vector $\boldsymbol{\alpha}$ and degrees of freedom ν, where $(x_{i1(j)}, x_{i2(j)})^T$ and $\breve{\boldsymbol{x}}_{ij(2)}$ are independent for each j. We set $\boldsymbol{\Omega}_1 = (0.3^{|i-j|^{1/2}})$, $\boldsymbol{\Omega}_2 = (0.5^{|i-j|^{1/2}})$ and $\boldsymbol{\alpha} = 10\mathbf{1}_{p-2}$. We considered two cases: (a) $\nu = 10$ and (b) $\nu = 20$. Note that $E(\breve{\boldsymbol{x}}_{ij(2)}) = (\nu/\pi)^{1/2} \{\Gamma(\nu/2 - 1/2)/\Gamma(\nu/2)\} \boldsymbol{\Omega}_i \boldsymbol{\alpha} / (1 + \boldsymbol{\alpha}^T \boldsymbol{\Omega}_i \boldsymbol{\alpha})^{1/2}$ $(= \breve{\boldsymbol{\mu}}_i$, say) and $\text{Var}(\breve{\boldsymbol{x}}_{ij(2)}) = \nu \boldsymbol{\Omega}_i / (\nu - 2) - \breve{\boldsymbol{\mu}}_i \breve{\boldsymbol{\mu}}_i^T$ $(= \breve{\boldsymbol{\Sigma}}_i$, say), where $\Gamma(\cdot)$ denotes the gamma function. We set $\boldsymbol{x}_{ij(2)} = \breve{\boldsymbol{x}}_{ij(2)} - \breve{\boldsymbol{\mu}}_i + \boldsymbol{\mu}_{i(2)}$ for all i, j. Then, we had $\boldsymbol{\Sigma}_{i(2)} = \breve{\boldsymbol{\Sigma}}_i$, $i = 1, 2$, in (S4.1). Note that (4.1) and (A-vi) with $k_1 = k_2 = 2$ are met. However, (A-i) and (A-viii) are not met. Similar to Fig. S4.2, we plotted $\overline{\alpha}$ in the left panel and $1 - \overline{\beta}$ in the right panel for (i) in Fig. S4.3 and for (ii) in Fig. S4.4. We observed the performances similar to those in Fig. 2 (b) and (c).

Throughout, the test procedure by (5.5) with $k_i = \hat{k}_i$, $i = 1, 2$, gave adequate performances for high-dimensional cases even for the skewed and heavy tailed distributions.



(a) $\mathrm{SN}_{p-2}(\boldsymbol{\Omega}_i, \boldsymbol{\alpha})$ with $\boldsymbol{\alpha} = 4\mathbf{1}_{p-2}$.



(b) $\mathrm{SN}_{p-2}(\boldsymbol{\Omega}_i, \boldsymbol{\alpha})$ with $\boldsymbol{\alpha} = 16\mathbf{1}_{p-2}$.

Figure S4.2: When $p = 2^s$, $n_1 = 3\lceil p^{1/2}\rceil$ and $n_2 = 4\lceil p^{1/2}\rceil$ for $s = 4, ..., 10$, the performances of five tests: (I) from (3.1) with $\boldsymbol{A} = \boldsymbol{I}_p$, (II) from (4.2), (III) from (5.5), (IV) from (5.5) with $k_i = \hat{k}_i$, $i = 1, 2$, and (V) from (6.1). For (a) and (b), the values of $\overline{\alpha}$ are denoted by the dashed lines in the left panel and the values of $1 - \overline{\beta}$ are denoted by the dashed lines in the right panel. The asymptotic power of (III) was given by $\Phi(\Delta_*/K_*^{1/2} - z_\alpha(K_{1*}/K_*)^{1/2})$ which is denoted by the solid line in the right panels. When p is small, $\overline{\alpha}$ for (V) was too high to describe in the left panels.

S5 Appendix A

In this appendix, we give proofs of the theoretical results in Sections 2 and 3 of the main work in Aoshima and Yata (2016).

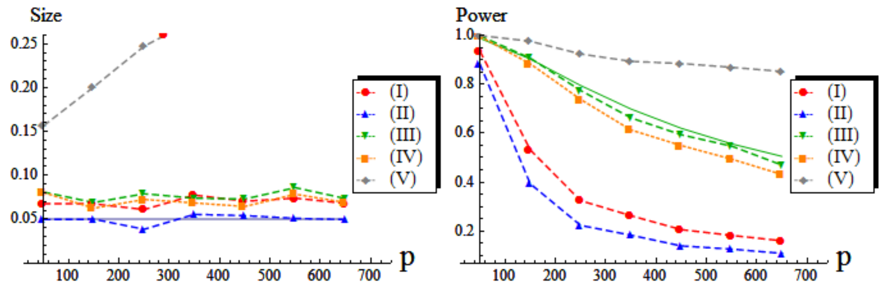
We simply write $T = T(\boldsymbol{A})$, $\Delta = \Delta(\boldsymbol{A})$, $K = K(\boldsymbol{A})$, $K_1 = K_1(\boldsymbol{A})$, $\widehat{K}_1 = \widehat{K}_1(\boldsymbol{A})$ and $K_2 = K_2(\boldsymbol{A})$.

Proof of Theorem 1. We note that for $i = 1, 2$

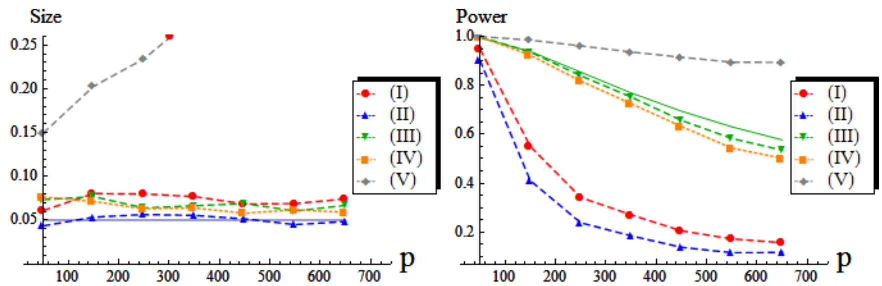
$$\boldsymbol{\mu}_A^T \boldsymbol{\Sigma}_{i,A} \boldsymbol{\mu}_A \le \Delta \lambda_{\max}(\boldsymbol{\Sigma}_{i,A}) \le \Delta \mathrm{tr}(\boldsymbol{\Sigma}_{i,A}^2)^{1/2}. \tag{S5.1}$$

Hence, from the fact that $\mathrm{tr}(\boldsymbol{\Sigma}_{i,A}^2)/n_i^2 \le K_1$ for $i = 1, 2$, it holds that $K_2 = O(\Delta K_1^{1/2})$, so that

$$\mathrm{Var}(T/\Delta) = (K_1 + K_2)/\Delta^2 = K_1/\Delta^2 + O(K_1^{1/2}/\Delta). \tag{S5.2}$$



(a) $\mathrm{ST}_{p-2}(\boldsymbol{\Omega}_i, \boldsymbol{\alpha}, \nu)$ with $\nu = 10$.



(b) $\mathrm{ST}_{p-2}(\boldsymbol{\Omega}_i, \boldsymbol{\alpha}, \nu)$ with $\nu = 20$.

Figure S4.3: When (i) $(n_1, n_2) = (40, 60)$ and $p = 50+100(s-1)$ for $s = 1, ..., 7$, the performances of five tests: (I) from (3.1) with $\boldsymbol{A} = \boldsymbol{I}_p$, (II) from (4.2), (III) from (5.5), (IV) from (5.5) with $k_i = \hat{k}_i$, $i = 1, 2$, and (V) from (6.1).

Thus, under (A-iii), from Chebyshev's inequality, we can claim the result. □

Proof of Theorem 2. We first consider the case when (A-iv) is met. From (S5.1), under (A-ii), it holds that $\boldsymbol{\mu}_A^T \boldsymbol{\Sigma}_{i,A} \boldsymbol{\mu}_A / n_i = o(\Delta \mathrm{tr}(\boldsymbol{\Sigma}_{i,A}^2)^{1/2}/n_i) = o(\Delta K_1^{1/2})$ as $m \to \infty$, so that

$$K_2/K_1 = O\{K_2/(\Delta K_1^{1/2})\} \to 0 \tag{S5.3}$$

under (A-ii) and (A-iv). Let $\boldsymbol{x}_{ij,A} = \boldsymbol{A}^{1/2}\boldsymbol{x}_{ij}$ $(j = 1, ..., n_i)$, $\boldsymbol{\mu}_{i,A} = \boldsymbol{A}^{1/2}\boldsymbol{\mu}_i$ and $\boldsymbol{\Gamma}_{i,A} = \boldsymbol{A}^{1/2}\boldsymbol{\Gamma}_i$ for $i = 1, 2$. We write that

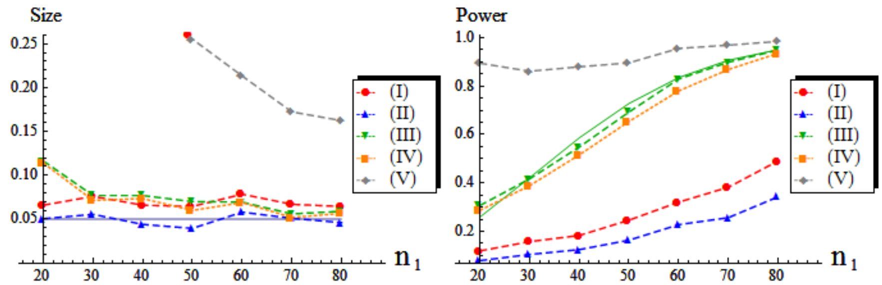
$$\boldsymbol{x}_{ij,A} = \boldsymbol{\Gamma}_{i,A}\boldsymbol{w}_{ij} + \boldsymbol{\mu}_{i,A} \quad \text{for all } i, j. \tag{S5.4}$$

Note that $\mathrm{Var}(\boldsymbol{x}_{ij,A}) = \boldsymbol{\Sigma}_{i,A}$ for $i = 1, 2$. Then, from (S5.3), by using Theorem 5 given in Aoshima and Yata (2015), we can obtain the result when (A-iv) is met.

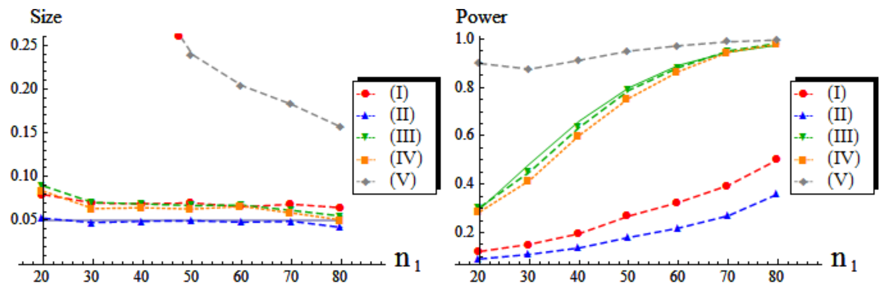
Next, we consider the case when (A-v) is met. Let $\boldsymbol{\mu}_{12} = \boldsymbol{\mu}_1 - \boldsymbol{\mu}_2$. Under (A-v), it holds that

$$T - \Delta = 2\boldsymbol{\mu}_{12}^T \boldsymbol{A}(\overline{\boldsymbol{x}}_{1n_1} - \overline{\boldsymbol{x}}_{2n_2} - \boldsymbol{\mu}_{12}) + o_P(K_2^{1/2}) \tag{S5.5}$$

from the fact that $\mathrm{Var}\{(\overline{\boldsymbol{x}}_{1n_1} - \overline{\boldsymbol{x}}_{2n_2} - \boldsymbol{\mu}_{12})^T \boldsymbol{A}(\overline{\boldsymbol{x}}_{1n_1} - \overline{\boldsymbol{x}}_{2n_2} - \boldsymbol{\mu}_{12}) - \mathrm{tr}(\boldsymbol{S}_{1n_1}\boldsymbol{A})/n_1 - \mathrm{tr}(\boldsymbol{S}_{2n_2}\boldsymbol{A})/n_2\} = K_1$. Let $\omega_j = 2\boldsymbol{\mu}_{12}^T \boldsymbol{A}(\boldsymbol{x}_{1j} - \boldsymbol{\mu}_1)/n_1$ for $j = 1, ..., n_1$, and $\omega_{j+n_1} = -2\boldsymbol{\mu}_{12}^T \boldsymbol{A}(\boldsymbol{x}_{2j} - \boldsymbol{\mu}_2)/n_2$ for



(a) $\mathrm{ST}_{p-2}(\boldsymbol{\Omega}_i, \boldsymbol{\alpha}, \nu)$ with $\nu = 10$.



(b) $\mathrm{ST}_{p-2}(\boldsymbol{\Omega}_i, \boldsymbol{\alpha}, \nu)$ with $\nu = 20$.

Figure S4.4: When (ii) $p = 500$, $n_1 = 10s$ and $n_2 = 1.5n_1$ for $s = 2, ..., 8$, the performances of five tests: (I) from (3.1) with $\boldsymbol{A} = \boldsymbol{I}_p$, (II) from (4.2), (III) from (5.5), (IV) from (5.5) with $k_i = \hat{k}_i$, $i = 1, 2$, and (V) from (6.1).

$j = 1, ..., n_2$. Note that $\sum_{j=1}^{n_1+n_2} \omega_j = 2\boldsymbol{\mu}_{12}^T \boldsymbol{A}(\overline{\boldsymbol{x}}_{1n_1} - \overline{\boldsymbol{x}}_{2n_2} - \boldsymbol{\mu}_{12})$ and $\mathrm{Var}(\sum_{j=1}^{n_1+n_2} \omega_j) = K_2$. Note that $E(w_j^4) = O\{(\boldsymbol{\mu}_A^T \boldsymbol{\Sigma}_{1,A} \boldsymbol{\mu}_A)^2 / n_1^4\}$ for $j = 1, ..., n_1$, and $E(w_j^4) = O\{(\boldsymbol{\mu}_A^T \boldsymbol{\Sigma}_{2,A} \boldsymbol{\mu}_A)^2 / n_2^4\}$ for $j = n_1 + 1, ..., n_1 + n_2$, under (A-i). Then, for Lyapunov's condition, it holds that as $n_{\min} \to \infty$

$$\frac{\sum_{j=1}^{n_1+n_2} E(w_j^4)}{K_2^2} = \frac{O\{(\boldsymbol{\mu}_A^T \boldsymbol{\Sigma}_{1,A} \boldsymbol{\mu}_A)^2 / n_1^3 + (\boldsymbol{\mu}_A^T \boldsymbol{\Sigma}_{2,A} \boldsymbol{\mu}_A)^2 / n_2^3\}}{K_2^2} = O(n_{\min}^{-1}) \to 0.$$

Hence, by using Lyapunov's central limit theorem, we have that $\sum_{j=1}^{n_1+n_2} \omega_j / K_2^{1/2} \Rightarrow N(0, 1)$. In view of (S5.5) and $K_2/K = 1 + o(1)$ as $m \to \infty$ under (A-v), we can obtain the result when (A-v) is met. □

Proof of Proposition 1. From (S5.1) and the fact that $\mathrm{tr}(\boldsymbol{\Sigma}_{i,A}^2)/n_i^2 \le K_1$, $i = 1, 2$, it holds that $K_1/K_2 \ge K_1^{1/2}/(8\Delta)$. Thus, (A-v) implies (A-iii). It concludes the result. □

Proof of Lemma 1. From (S5.3), the result is obtained straightforwardly. □

Proofs of Lemma 2 and Corollary 1. From (2.3), (S5.4) and the equation (23) given in Aoshima and Yata (2015), we have that $\widehat{K}_1/K_1 = 1 + o_P(1)$ as $m \to \infty$ under (A-i). It concludes the result of Lemma 2. By using Lemmas 1 and 2, it holds that $\widehat{K}_1/K = 1 + o_P(1)$ under (A-i), so that the result of Corollary 1 is obtained from Theorem 2. $\square$

Proofs of Theorem 3 and Corollary 2. First, we consider Corollary 2. From Theorem 1, under (A-i) and (A-iii), we have that as $m \to \infty$

$$P(T/\widehat{K}_1^{1/2} > z_\alpha) = P(T/\Delta > z_\alpha \widehat{K}_1^{1/2}/\Delta) = P\{1 + o_P(1) > o_P(1)\} \to 1$$

from the fact that $\widehat{K}_1^{1/2}/\Delta = K_1^{1/2}\{1 + o_P(1)\}/\Delta = o_P(1)$ under (A-i) and (A-iii). It concludes the result of Corollary 2 when (A-iii) is met. From Theorem 2, Lemmas 1 and 2, under (A-i), (A-ii) and (A-iv), we have that

$$\begin{aligned} P(T/\widehat{K}_1^{1/2} > z_\alpha) &= P\{(T-\Delta)/K^{1/2} > (z_\alpha K_1^{1/2} - \Delta)/K^{1/2} + o_P(1)\} \\ &= \Phi\{(\Delta - z_\alpha K_1^{1/2})/K^{1/2}\} + o(1) = \Phi(\Delta/K_1^{1/2} - z_\alpha) + o(1). \end{aligned} \tag{S5.6}$$

It concludes the result of Corollary 2 when (A-ii) and (A-iv) are met. We note that $K/K_2 \to 1$ as $m \to \infty$ under (A-v). Then, by combining (S5.6) and Theorem 2, we can conclude the result of Corollary 2 when (A-v) is met.

Next, we consider Theorem 3. By combining (S5.6) and Theorem 2, we can conclude the results about size and power in Theorem 3 when (A-iv) is met. From (S5.2) we note that $K/\Delta^2 \to 0$ under (A-iii). It holds that $\Phi\{(\Delta - z_\alpha K_1^{1/2})/K^{1/2}\} \to 1$ under (A-iii), so that from Corollary 2 we obtain the result about power when (A-iii) is met. Hence, by considering a convergent subsequence of $\Delta/K_1^{1/2}$, we can conclude the result about power in Theorem 3. $\square$

S6 Appendix B

In this appendix, we give proofs of the theoretical results in Sections 4 and 5 of the main work in Aoshima and Yata (2016). Also, we give two lemmas and proofs of the lemmas.

Let $\bar{z}_{ij} = \sum_{l=1}^{n_i} z_{ijl}/n_i$ and $v_{i(j)} = \sum_{l=1}^{n_i}(z_{ijl} - \bar{z}_{ij})^2/(n_i - 1)$ for all i, j. Let $\boldsymbol{u}_{ij} = (z_{ij1}, ..., z_{ijn_i})^T/(n_i-1)^{1/2}$, $\boldsymbol{u}_{oij} = \boldsymbol{P}_{n_i}\boldsymbol{u}_{ij} = (z_{ij1} - \bar{z}_{ij}, ..., z_{ijn_i} - \bar{z}_{ij})^T/(n_i-1)^{1/2}$ and $\dot{\boldsymbol{u}}_{ij} = ||\boldsymbol{u}_{ij}||^{-1}\boldsymbol{u}_{ij}$ for all i, j. Let $\boldsymbol{\zeta}_i$ be an arbitrary unit random n_i-dimensional vector for $i = 1, 2$. Let $\boldsymbol{y}_{ij} = \sum_{s=1}^{k_i} \lambda_{is}^{1/2}\boldsymbol{h}_{is} z_{isj}$ and $\boldsymbol{v}_{ij} = \sum_{s=k_i+1}^{p} \lambda_{is}^{1/2}\boldsymbol{h}_{is} z_{isj}$ for all i, j. Note that $\boldsymbol{x}_{ij} = \boldsymbol{y}_{ij} + \boldsymbol{v}_{ij} + \boldsymbol{\mu}_i$ for all i, j. Let $\psi_{ij} = \mathrm{tr}(\boldsymbol{\Sigma}_i^2)/\lambda_{ij} + n_i\boldsymbol{\mu}_i^T\boldsymbol{\Sigma}_i\boldsymbol{\mu}_i/\lambda_{ij}$ for $i = 1, 2;\ j = 1, ..., k_i$. Let $h_{st} = \boldsymbol{h}_{1s}^T\boldsymbol{h}_{2t}$ for all s, t. We also let $\boldsymbol{M}_i = \boldsymbol{\mu}_i\boldsymbol{1}_{n_i}^T$ for $i = 1, 2$.

Proof of Theorem 4. We assume $\boldsymbol{\mu}_1 = \boldsymbol{\mu}_2 = \boldsymbol{0}$ and $\boldsymbol{h}_{11}^T\boldsymbol{h}_{21} \geq 0$ without loss of generality. Let $\boldsymbol{H}_{i1} = \boldsymbol{h}_{i1}\boldsymbol{h}_{i1}^T$, $\boldsymbol{H}_{i2} = \boldsymbol{I}_p - \boldsymbol{H}_{i1}$, $\boldsymbol{\Sigma}_{i1} = \lambda_{i1}\boldsymbol{H}_{i1}$ and $\boldsymbol{\Sigma}_{i2} = \sum_{j=2}^{p}\lambda_{ij}\boldsymbol{h}_{ij}\boldsymbol{h}_{ij}^T$ for $i = 1, 2$. Note that $\boldsymbol{\Sigma}_i = \boldsymbol{\Sigma}_{i1} + \boldsymbol{\Sigma}_{i2}$ for $i = 1, 2$. We write that

$$T_I = T(\boldsymbol{H}_{11}, \boldsymbol{H}_{21}) + T(\boldsymbol{H}_{12}, \boldsymbol{H}_{22}) - 2\overline{\boldsymbol{x}}_{1n_1}^T(\boldsymbol{H}_{11}\boldsymbol{H}_{22} + \boldsymbol{H}_{12}\boldsymbol{H}_{21})\overline{\boldsymbol{x}}_{2n_2}.$$

We have that $\mathrm{Var}\{T(\boldsymbol{H}_{11}, \boldsymbol{H}_{21})\} = K_1(\boldsymbol{H}_{11}, \boldsymbol{H}_{21}) = 2\sum_{i=1}^{2}\lambda_{i1}^2/\{n_i(n_i-1)\} + 4\lambda_{11}\lambda_{21}(\boldsymbol{h}_{11}^T\boldsymbol{h}_{21})^2/(n_1n_2)$ and $\mathrm{Var}\{T(\boldsymbol{H}_{12}, \boldsymbol{H}_{22})\} = K_1(\boldsymbol{H}_{12}, \boldsymbol{H}_{22}) = 2\sum_{i=1}^{2}\mathrm{tr}(\boldsymbol{\Sigma}_{i2}^2)/\{n_i(n_i-1)\} + 4\mathrm{tr}(\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22})$

$/(n_1n_2)$, where $K_1(\cdot,\cdot)$ is defined in Section S1.2. Let $\psi=(\lambda_{11}/n_1+\lambda_{21}/n_2)$. Then, under (4.1) it holds that as $m\to\infty$

$$K_1(\boldsymbol{H}_{11},\boldsymbol{H}_{21})=2\psi^2\{1+o(1)\}\quad\text{and}\quad K_1(\boldsymbol{H}_{12},\boldsymbol{H}_{22})=o(\psi^2)$$

because $\mathrm{tr}(\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22})\le\{\mathrm{tr}(\boldsymbol{\Sigma}_{12}^2)\mathrm{tr}(\boldsymbol{\Sigma}_{22}^2)\}^{1/2}=(\Psi_{1(2)}\Psi_{2(2)})^{1/2}$. Also, under (4.1) it follows that

$$\mathrm{Var}\{\overline{\boldsymbol{x}}_{1n_1}^T(\boldsymbol{H}_{11}\boldsymbol{H}_{22}+\boldsymbol{H}_{12}\boldsymbol{H}_{21})\overline{\boldsymbol{x}}_{2n_2}\}=\frac{\mathrm{tr}(\boldsymbol{\Sigma}_{11}\boldsymbol{\Sigma}_{22})+\mathrm{tr}(\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{21})}{n_1n_2}=o(\psi^2)$$

because $\mathrm{tr}(\boldsymbol{\Sigma}_{11}\boldsymbol{\Sigma}_{22})\le\lambda_{11}\mathrm{tr}(\boldsymbol{\Sigma}_{22}^2)^{1/2}$ and $\mathrm{tr}(\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{21})\le\lambda_{21}\mathrm{tr}(\boldsymbol{\Sigma}_{12}^2)^{1/2}$. Hence, under (4.1) we have that $K_{1(I)}=2\psi^2\{1+o(1)\}$ and

$$\begin{aligned}T_I&=\sum_{i=1}^2\lambda_{i1}(\bar{z}_{i1}^2-v_{i(1)}/n_i)-2(\lambda_{11}\lambda_{21})^{1/2}\bar{z}_{11}\bar{z}_{21}(\boldsymbol{h}_{11}^T\boldsymbol{h}_{21})+o_P(\psi)\\&=(\lambda_{11}^{1/2}\bar{z}_{11}-\lambda_{21}^{1/2}\bar{z}_{21})^2-\psi+o_P(\psi)\end{aligned}$$

from the fact that $v_{i(1)}=1+o_P(1)$, $i=1,2$. By noting that $E(z_{i1l}^4)$'s are bounded, for Lyapunov's condition, it holds that $\sum_{i=1}^2\sum_{l=1}^{n_i}(\lambda_{i1}^{1/2}z_{i1l}/n_i)^4=o(\psi^2)$. Hence, by using Lyapunov's central limit theorem, we have that $\psi^{-1/2}(\lambda_{11}^{1/2}\bar{z}_{11}-\lambda_{21}^{1/2}\bar{z}_{21})\Rightarrow N(0,1)$. Thus, from $\psi^{-1}T_I=\psi^{-1}(\lambda_{11}^{1/2}\bar{z}_{11}-\lambda_{21}^{1/2}\bar{z}_{21})^2-1+o_P(1)$ and $K_{1(I)}=2\psi^2\{1+o(1)\}$ under (4.1), we have that $T_I/(K_{1(I)}/2)^{1/2}+1\Rightarrow\chi_1^2$. From Lemma 2, it concludes the result. □

Proof of Corollary 3. From Theorem 2, the result is obtained straightforwardly. □

Throughout the proofs of Propositions 2 to 5, Lemmas B.1, B.2, 3 and Theorem 5, we assume (A-vi) and (A-viii). Throughout the proofs of Propositions 2 to 5 and Lemma B.1, we omit the subscript with regard to the population.

Proof of Proposition 2. Let us write that $\boldsymbol{U}_1=\sum_{s=1}^k\lambda_s\boldsymbol{u}_{os}\boldsymbol{u}_{os}^T$ and $\boldsymbol{U}_2=\sum_{s=k+1}^p\lambda_s\boldsymbol{u}_s\boldsymbol{u}_s^T$. Note that $\boldsymbol{S}_D=\boldsymbol{U}_1+\boldsymbol{P}_n\boldsymbol{U}_2\boldsymbol{P}_n$. Also, note that $\boldsymbol{P}_n\hat{\boldsymbol{u}}_j=\hat{\boldsymbol{u}}_j$ and $\hat{\lambda}_j=\hat{\boldsymbol{u}}_j^T\boldsymbol{S}_D\hat{\boldsymbol{u}}_j=\hat{\boldsymbol{u}}_j^T(\boldsymbol{U}_1+\boldsymbol{U}_2)\hat{\boldsymbol{u}}_j$ when $\hat{\lambda}_j>0$. From Lemma 5 in Yata and Aoshima (2013) we can claim that as $m_0\to\infty$

$$\hat{\lambda}_j/\lambda_j-\delta_j=(\hat{\boldsymbol{u}}_j^T\boldsymbol{U}_1\hat{\boldsymbol{u}}_j)/\lambda_j+o_P(1)\ \ \text{for } j=1,...,k.$$

Also, similar to the proofs of Lemmas 3 and 4 in Yata and Aoshima (2012), we have that $\boldsymbol{u}_j^T(\boldsymbol{U}_2-\delta\boldsymbol{I}_n)\boldsymbol{u}_{j'}=O_P(\Psi_{(k+1)}^{1/2}/n)$ and $\boldsymbol{u}_j^T(\boldsymbol{U}_2-\delta\boldsymbol{I}_n)\boldsymbol{\zeta}=O_P(\Psi_{(k+1)}^{1/2}/n^{1/2})$ for $j,j'=1,...,k$, where $\delta=\sum_{s=k+1}^p\lambda_s/(n-1)$. Then, by noting that $\boldsymbol{u}_{oj}^T\boldsymbol{u}_{oj'}=O_P(n^{-1/2})$ $(j\ne j')$ and $||\boldsymbol{u}_{oj}||^2=||\boldsymbol{u}_j||^2+O_P(n^{-1})=1+O_P(n^{-1/2})$ as $n\to\infty$, we can claim that

$$\hat{\lambda}_j/\lambda_j=||\boldsymbol{u}_j||^2+\delta_j+O_P(n^{-1})=1+\delta_j+O_P(n^{-1/2})$$

$$\text{and}\quad\hat{\boldsymbol{u}}_j^T\dot{\boldsymbol{u}}_j=1+O_P(n^{-1})\quad\text{for } j=1,...,k; \tag{S6.1}$$

$$\hat{\boldsymbol{u}}_{j'}^T\boldsymbol{u}_j=O_P(n^{-1/2}\lambda_{j'}/\lambda_j)\quad\text{for } j<j'\le k \tag{S6.2}$$

in a way similar to the proof of Lemma 5 in Yata and Aoshima (2012) and the proof of Lemma 9 in Yata and Aoshima (2013). By noting that $(\boldsymbol{X}-\overline{\boldsymbol{X}})\hat{\boldsymbol{u}}_j=(\boldsymbol{X}-\boldsymbol{M})\hat{\boldsymbol{u}}_j$ when $\hat{\lambda}_j>0$, we write that

$$(\boldsymbol{h}_j^T\hat{\boldsymbol{h}}_j)^2=\{\boldsymbol{h}_j^T(\boldsymbol{X}-\overline{\boldsymbol{X}})\hat{\boldsymbol{u}}_j\}^2/\{(n-1)\hat{\lambda}_j\}=||\boldsymbol{u}_j||^2(\hat{\boldsymbol{u}}_j^T\dot{\boldsymbol{u}}_j)^2(\lambda_j/\hat{\lambda}_j)$$

when $\hat{\lambda}_j>0$. Thus, from (S6.1) we can conclude the results. □

Proof of Proposition 3. We can claim that as $m_0\to\infty$

$$\{\lambda_j(n-1-j)\}^{-1}\Big(\mathrm{tr}(\boldsymbol{S}_D)-\sum_{l=1}^{j}\hat{\lambda}_l\Big)-\delta_j=O_P(n^{-1})\quad\text{for } j=1,...,k$$

in a way similar to the proof of Lemma 11 in Yata and Aoshima (2013). Then, it follows from (S6.1) that

$$\tilde{\lambda}_j/\lambda_j=||\boldsymbol{u}_j||^2+O_P(n^{-1})=1+O_P(n^{-1/2})\quad\text{for } j=1,...,k. \tag{S6.3}$$

Note that $(\boldsymbol{h}_j^T\tilde{\boldsymbol{h}}_j)^2=||\boldsymbol{u}_j||^2(\hat{\boldsymbol{u}}_j^T\dot{\boldsymbol{u}}_j)^2(\lambda_j/\tilde{\lambda}_j)$. Then, from (S6.1) and (S6.3) we can conclude the results. □

Proofs of Propositions 4 and 5. First, we consider Proposition 4. From (S6.1) there exists a unit random vector $\boldsymbol{\varepsilon}_j=(\varepsilon_{j1},...,\varepsilon_{jn})^T$ such that $\dot{\boldsymbol{u}}_j^T\boldsymbol{\varepsilon}_j=0$ and

$$\hat{\boldsymbol{u}}_j=\{1+O_P(n^{-1})\}\dot{\boldsymbol{u}}_j+\boldsymbol{\varepsilon}_j\times O_P(n^{-1/2})\quad\text{for } j=1,...,k \tag{S6.4}$$

as $m_0\to\infty$. By noting that $\dot{\boldsymbol{u}}_j=\boldsymbol{u}_j\{1+o_P(1)\}$ and $\boldsymbol{u}_j^T\boldsymbol{u}_{j'}=O_p(n^{-1/2})$ $(j\neq j')$ as $n\to\infty$, it follows from (S6.4) that

$$\hat{\boldsymbol{u}}_{j'}^T\boldsymbol{u}_j=O_P(n^{-1/2})\quad\text{for } j'<j\le k. \tag{S6.5}$$

Then, from (S6.1) to (S6.3) and (S6.5) it holds that for $j=1,...,k$ $(l=1,...,n)$

$$\frac{\tilde{\boldsymbol{h}}_j^T\boldsymbol{y}_l}{\lambda_j^{1/2}}=\frac{\hat{\boldsymbol{u}}_j^T(\boldsymbol{X}-\boldsymbol{M})^T\boldsymbol{y}_l}{\{(n-1)\tilde{\lambda}_j\lambda_j\}^{1/2}}=\sum_{s=1}^{k}\frac{\lambda_s z_{sl}\hat{\boldsymbol{u}}_j^T\boldsymbol{u}_s}{(\tilde{\lambda}_j\lambda_j)^{1/2}}=z_{jl}+O_P(n^{-1/2}) \tag{S6.6}$$

because $z_{sl}=O_P(1)$ for $s=1,...,k$. Let us write that

$$\boldsymbol{u}_{j(l)}=(z_{j1},...,z_{jl-1},0,z_{jl+1},...,z_{jn})^T/(n-1)^{1/2}\quad\text{for all } j,l.$$

We have that $E\{(\sum_{s=k+1}^{p}\lambda_s z_{sl}\boldsymbol{u}_{j(l)}^T\boldsymbol{u}_{s(l)}/\lambda_j)^2\}=O\{\Psi_{(k+1)}/(n\lambda_j^2)\}=O(n^{-1})$ and $E(||\sum_{s=k+1}^{p}\lambda_s z_{sl}\boldsymbol{u}_{s(l)}/\lambda_j||^2)=O(\Psi_{(k+1)}/\lambda_j^2)=O(1)$ for $j=1,...,k$. It follows that

$$\sum_{s=k+1}^{p}\frac{\lambda_s z_{sl}\boldsymbol{u}_{j(l)}^T\boldsymbol{u}_{s(l)}}{\lambda_j}=O_P(n^{-1/2})\quad\text{and}\quad\boldsymbol{\zeta}^T\sum_{s=k+1}^{p}\frac{\lambda_s z_{sl}\boldsymbol{u}_{s(l)}}{\lambda_j}=O_P(1) \tag{S6.7}$$

from the fact that $|\boldsymbol{\zeta}^T\sum_{s=k+1}^{p}\lambda_s z_{sl}\boldsymbol{u}_{s(l)}/\lambda_j|\le||\boldsymbol{\zeta}||\cdot||\sum_{s=k+1}^{p}\lambda_s z_{sl}\boldsymbol{u}_{s(l)}/\lambda_j||$ and Markov's inequality. Let $d_n=(n-1)/(n-2)$. Here, from (S6.4) we write that for $j=1,...,k$

$$d_n\hat{\boldsymbol{u}}_{j(l)}=\{1+O_P(n^{-1})\}\boldsymbol{u}_{j(l)}/||\boldsymbol{u}_j||+\boldsymbol{\varepsilon}_{j(l)}\times O_P(n^{-1/2})+(n-2)^{-1}\hat{u}_{jl}\boldsymbol{1}_{n(l)}, \tag{S6.8}$$

where $\boldsymbol{\varepsilon}_{j(l)} = (\varepsilon_{j1}, ..., \varepsilon_{jl-1}, 0, \varepsilon_{jl+1}, ..., \varepsilon_{jn})^T$. Note that $||(n-2)^{-1}\hat{u}_{jl}\mathbf{1}_{n(l)}|| = O_P(n^{-1/2})$ since $|\hat{u}_{jl}| \le 1$. Then, it follows from (S6.3), (S6.7) and (S6.8) that for $j = 1, ..., k$

$$\frac{\tilde{\boldsymbol{h}}_{jl}^T \boldsymbol{v}_l}{\lambda_j^{1/2}} = d_n \frac{\hat{\boldsymbol{u}}_{j(l)}^T (\boldsymbol{X} - \boldsymbol{M})^T \boldsymbol{v}_l}{\{(n-1)\tilde{\lambda}_j \lambda_j\}^{1/2}} = d_n \sum_{s=k+1}^{p} \frac{\lambda_s z_{sl} \hat{\boldsymbol{u}}_{j(l)}^T \boldsymbol{u}_{s(l)}}{(\tilde{\lambda}_j \lambda_j)^{1/2}} = O_P(n^{-1/2}). \tag{S6.9}$$

We note that $\mathrm{Var}(\sum_{s=k+1}^p \lambda_s z_{sl}^2/\lambda_j) = O(\Psi_{(k+1)}/\lambda_j^2)$, so that $(n-1)^{-1/2} \sum_{s=k+1}^p \lambda_s z_{sl}^2/\lambda_j = (n-1)^{1/2}\delta_j + O_P(n^{-1/2})$ for $j = 1, ..., k$, because $E(\sum_{s=k+1}^p \lambda_s z_{sl}^2/\lambda_j) = (n-1)\delta_j$. Then, it follows from (S6.3) and (S6.7) that for $j = 1, ..., k$

$$\begin{aligned}\frac{(d_n\tilde{\boldsymbol{h}}_j - \tilde{\boldsymbol{h}}_{jl})^T \boldsymbol{v}_l}{\lambda_j^{1/2}} &= d_n \frac{(\hat{\boldsymbol{u}}_j - \hat{\boldsymbol{u}}_{j(l)})^T (\boldsymbol{X} - \boldsymbol{M})^T \boldsymbol{v}_l}{\{(n-1)\tilde{\lambda}_j \lambda_j\}^{1/2}} \\ &= d_n \hat{u}_{jl} \sum_{s=k+1}^{p} \frac{\lambda_s z_{sl}^2}{\{(n-1)\tilde{\lambda}_j \lambda_j)\}^{1/2}} - \frac{\hat{u}_{jl}\mathbf{1}_{n(l)}^T}{n-2} \sum_{s=k+1}^{p} \frac{\lambda_s z_{sl} \boldsymbol{u}_{s(l)}}{(\tilde{\lambda}_j \lambda_j)^{1/2}} \\ &= d_n \hat{u}_{jl}(n-1)^{1/2}\delta_j\{1 + o_P(1)\} + O_P(n^{-1/2}). \end{aligned} \tag{S6.10}$$

By combining (S6.6) and (S6.9) with (S6.10), we can conclude the result of $\tilde{\boldsymbol{h}}_j$ in Proposition 4. As for $\hat{\boldsymbol{h}}_j$, by noting that $\hat{\boldsymbol{h}}_j = (\tilde{\lambda}_j/\hat{\lambda}_j)^{1/2}\tilde{\boldsymbol{h}}_j$, $||\boldsymbol{u}_j||^2 = 1 + O_P(n^{-1/2})$, (S6.1) and (S6.3), we can conclude the result.

Next, we consider Proposition 5. From (S6.3) we have that for $j = 1, ..., k$

$$\begin{aligned}\frac{(d_n\tilde{\boldsymbol{h}}_j - \tilde{\boldsymbol{h}}_{jl})^T \boldsymbol{y}_l}{\lambda_j^{1/2}} &= d_n \hat{u}_{jl} \sum_{s=1}^{k} \frac{\lambda_s z_{sl}^2}{\{(n-1)\tilde{\lambda}_j \lambda_j)\}^{1/2}} - \frac{\hat{u}_{jl}\mathbf{1}_{n(l)}^T}{n-2} \sum_{s=1}^{k} \frac{\lambda_s z_{sl} \boldsymbol{u}_{s(l)}}{(\tilde{\lambda}_j \lambda_j)^{1/2}} \\ &= d_n \hat{u}_{jl} \times O_P\{(n^{1/2}\lambda_j)^{-1}\lambda_1\} \end{aligned} \tag{S6.11}$$

from the fact that $\mathbf{1}_{n(l)}^T \boldsymbol{u}_{s(l)} = O_P(1)$ and $z_{sl} = O_P(1)$, $s = 1, ..., k$. Then, by combining (S6.6) and (S6.9) with (S6.11), we can conclude the result. □

Lemma B.1. *Assume (A-vi) and (A-viii). It holds for $j = 1, ..., k$ that as $m_0 \to \infty$*

$$\sum_{l=1}^{n} \frac{\tilde{x}_{jl} - x_{jl}}{n} = O_P(\psi_j^{1/2}/n) \quad \text{and} \quad \sum_{l=1}^{n} \frac{(\tilde{x}_{jl} - x_{jl})^2}{n} = O_P(\psi_j/n).$$

Proof. First, we consider the first result. Let $\boldsymbol{\eta}_{sj(l)} = \lambda_s z_{sl} \boldsymbol{u}_{s(l)}/\lambda_j^{1/2}$, $\boldsymbol{\xi}_{sj(l)} = \lambda_s^{1/2}\mu_{(s)}\boldsymbol{u}_{s(l)}/\lambda_j^{1/2}$ and $\boldsymbol{\omega}_{sj(l)} = \boldsymbol{\eta}_{sj(l)} + \boldsymbol{\xi}_{sj(l)}$ for all j, l, s, where $\boldsymbol{u}_{s(l)}$ is given in the proofs of Propositions 4 and 5. Then, we write that when $\hat{\lambda}_j > 0$,

$$\tilde{x}_{jl} = d_n \frac{\hat{\boldsymbol{u}}_{j(l)}^T (\boldsymbol{X} - \boldsymbol{M})^T \boldsymbol{x}_l}{\{(n-1)\tilde{\lambda}_j\}^{1/2}} = d_n \frac{\lambda_j^{1/2}}{\tilde{\lambda}_j^{1/2}} \hat{\boldsymbol{u}}_{j(l)}^T \sum_{s=1}^{p} \boldsymbol{\omega}_{sj(l)}, \tag{S6.12}$$

where $d_n = (n-1)/(n-2)$. Let $\boldsymbol{e}_1 = (1, 0, ..., 0)^T$,..., $\boldsymbol{e}_n = (0, ..., 0, 1)^T$ be the standard basis vectors of dimension n. In view of (S6.3) and (S6.8), by noting that $||\boldsymbol{u}_j||^2 = 1 + O_P(n^{-1/2})$,

$||(n-2)^{-1}\hat{u}_{jl}\mathbf{1}_{n(l)}|| = O_P(n^{-1/2})$ as $n\to\infty$ and $\hat{\boldsymbol{u}}_j - \hat{\boldsymbol{u}}_{j(l)} = \hat{u}_{jl}\boldsymbol{e}_l - (n-1)^{-1}\hat{u}_{jl}\mathbf{1}_{n(l)}$ for $l=1,...,n$, we have that as $m_0\to\infty$

$$d_n\hat{\boldsymbol{u}}_{j(l)}(\lambda_j/\tilde{\lambda}_j)^{1/2} = \boldsymbol{u}_{j(l)}/||\boldsymbol{u}_j||^2 + (n-2)^{-1}\hat{u}_{jl}\mathbf{1}_{n(l)} + (\boldsymbol{\varepsilon}_j - \varepsilon_{jl}\boldsymbol{e}_l)\times O_P(n^{-1/2}) + \boldsymbol{\zeta}_{jl}\times O_P(n^{-1}) \quad \text{for all } l \text{ and } j=1,...,k, \tag{S6.13}$$

where $\boldsymbol{\varepsilon}_j$ and ε_{jl} are given in the proofs of Propositions 4 and 5 and $\boldsymbol{\zeta}_{jl}$ is a random unit vector depending on j and l. Note that $O_P(n^{-1/2})$ and $O_P(n^{-1})$ in (S6.13) do not depend on l. In view of (A-viii), we have that for $j=1,...,k$

$$E\Big\{\Big(\sum_{l=1}^n \boldsymbol{u}_{j(l)}^T\sum_{s=1(\neq j)}^p \boldsymbol{\eta}_{sj(l)}\Big)^2\Big\} = \sum_{l\neq l'}^n\sum_{s,s'(\neq j)}^p \frac{\lambda_s\lambda_{s'}E(z_{jl}z_{sl}z_{s'l}z_{jl'}z_{sl'}z_{s'l'})}{(n-1)^2\lambda_j} + O\{\mathrm{tr}(\boldsymbol{\Sigma}^2)/\lambda_j\} = O\{\mathrm{tr}(\boldsymbol{\Sigma}^2)/\lambda_j\}. \tag{S6.14}$$

On the other hand, we have that for $j=1,...,k$

$$E\Big\{\Big(\sum_{l=1}^n \boldsymbol{u}_{j(l)}^T\sum_{s=1(\neq j)}^p \boldsymbol{\xi}_{sj(l)}\Big)^2\Big\} = O\Big(n\sum_{s=1(\neq j)}^p \frac{\lambda_s\mu_{(s)}^2}{\lambda_j}\Big) = O(n\boldsymbol{\mu}^T\boldsymbol{\Sigma}\boldsymbol{\mu}/\lambda_j). \tag{S6.15}$$

Then, by using Markov's inequality, it follows from (S6.14) and (S6.15) that

$$\sum_{l=1}^n \boldsymbol{u}_{j(l)}^T\sum_{s=1(\neq j)}^p \boldsymbol{\omega}_{sj(l)}/||\boldsymbol{u}_j||^2 = O_P(\psi_j^{1/2}). \tag{S6.16}$$

Also, we have that $E\{\sum_{l=1}^n(\mathbf{1}_{n(l)}^T\sum_{s=1}^p\boldsymbol{\omega}_{sj(l)})^2\} = O(n\psi_j)$, $E(\|\sum_{l=1}^n\sum_{s=1}^p\boldsymbol{\omega}_{sj(l)}\|^2) = O(n\psi_j)$ and $E(\sum_{l=1}^n\|\sum_{s=1}^p\boldsymbol{\omega}_{sj(l)}\|^2) = O(n\psi_j)$ for $j=1,...,k$. Thus, it holds that

$$\Big|\sum_{l=1}^n\hat{u}_{jl}\mathbf{1}_{n(l)}^T\sum_{s=1}^p\boldsymbol{\omega}_{sj(l)}\Big| \le \Big(\sum_{l=1}^n\hat{u}_{jl}^2\Big)^{1/2}\Big\{\sum_{l=1}^n\Big(\mathbf{1}_{n(l)}^T\sum_{s=1}^p\boldsymbol{\omega}_{sj(l)}\Big)^2\Big\}^{1/2} = O_P(n^{1/2}\psi_j^{1/2}),$$
$$\Big|\boldsymbol{\varepsilon}_j^T\sum_{l=1}^n\sum_{s=1}^p\boldsymbol{\omega}_{sj(l)}\Big| \le ||\boldsymbol{\varepsilon}_j||\cdot\Big\|\sum_{l=1}^n\sum_{s=1}^p\boldsymbol{\omega}_{sj(l)}\Big\| = O_P(n^{1/2}\psi_j^{1/2}) \quad \text{and}$$
$$\Big|\sum_{l=1}^n\boldsymbol{\zeta}_{jl}^T\sum_{s=1}^p\boldsymbol{\omega}_{sj(l)}\Big| \le \Big(\sum_{l=1}^n||\boldsymbol{\zeta}_{jl}||^2\Big)^{1/2}\Big(\sum_{l=1}^n\Big\|\sum_{s=1}^p\boldsymbol{\omega}_{sj(l)}\Big\|^2\Big)^{1/2} = O_P(n\psi_j^{1/2}) \tag{S6.17}$$

by using Markov's inequality and Schwarz's inequality. Then, by noting that $\boldsymbol{e}_l^T\boldsymbol{\omega}_{sj(l)} = 0$ for all l,s, we have from (S6.12), (S6.13), (S6.16) and (S6.17) that for $j=1,...,k$

$$\sum_{l=1}^n\frac{\tilde{x}_{jl}-x_{jl}}{n} = \sum_{l=1}^n\frac{x_{jl}}{n}\Big(\frac{||\boldsymbol{u}_{j(l)}||^2-||\boldsymbol{u}_j||^2}{||\boldsymbol{u}_j||^2}\Big) + O_P(n^{-1}\psi_j^{1/2}) = -\sum_{l=1}^n\frac{x_{jl}z_{jl}^2}{n(n-1)||\boldsymbol{u}_j||^2} + O_P(n^{-1}\psi_j^{1/2}) = O_P(n^{-1}\psi_j^{1/2}) \tag{S6.18}$$

because it holds that $|\sum_{l=1}^n x_{jl}z_{jl}^2| \le (\sum_{l=1}^n x_{jl}^2\sum_{l'=1}^n z_{jl'}^4)^{1/2}$, $E(\sum_{l=1}^n x_{jl}^2) = n(\lambda_j+\mu_{(j)}^2)$, $E(\sum_{l=1}^n z_{jl}^4) = O(n)$, $\lambda_j \le \mathrm{tr}(\boldsymbol{\Sigma}^2)/\lambda_j$ and $\mu_{(j)}^2 \le \boldsymbol{\mu}^T\boldsymbol{\Sigma}\boldsymbol{\mu}/\lambda_j$. Thus, we can conclude the first result.

Next, we consider the second result. From (S6.12) and (S6.13) we have that

$$\begin{aligned}\sum_{l=1}^{n}\frac{(\tilde{x}_{jl}-x_{jl})^2}{n} =&O_P\Big(\sum_{l=1}^{n}\frac{x_{jl}^2 z_{jl}^4}{n^3}\Big)+O_P\Big\{\sum_{l=1}^{n}\Big(\boldsymbol{u}_{j(l)}^T\sum_{s=1(\neq j)}^{p}\boldsymbol{\omega}_{sj(l)}\Big)^2/n\Big\}\\ &+O_P\Big\{\sum_{l=1}^{n}\hat{u}_{jl}^2\Big(\boldsymbol{1}_{n(l)}^T\sum_{s=1}^{p}\boldsymbol{\omega}_{sj(l)}\Big)^2/n^3\Big\}\\ &+O_P\Big\{\boldsymbol{\varepsilon}_j^T\sum_{l=1}^{n}\Big(\sum_{s=1}^{p}\boldsymbol{\omega}_{sj(l)}\Big)\Big(\sum_{s=1}^{p}\boldsymbol{\omega}_{sj(l)}\Big)^T\boldsymbol{\varepsilon}_j/n^2\Big\}\\ &+O_P\Big\{\sum_{l=1}^{n}\boldsymbol{\zeta}_{jl}^T\Big(\sum_{s=1}^{p}\boldsymbol{\omega}_{sj(l)}\Big)\Big(\sum_{s=1}^{p}\boldsymbol{\omega}_{sj(l)}\Big)^T\boldsymbol{\zeta}_{jl}/n^3\Big\}. \qquad \text{(S6.19)}\end{aligned}$$

By using Markov's inequality, for any $\tau>0$, it holds that $P(\sum_{l=1}^{n}x_{jl}^2\geq \tau n\psi_j)=O(\tau^{-1})$ and $\sum_{l=1}^{n}P(z_{jl}^4\geq \tau n)=O(\tau^{-1})$ for $j=1,...,k$, so that

$$\sum_{l=1}^{n}x_{jl}^2 z_{jl}^4=O_P\big(n\psi_j\max_{l=1,\dots,n}z_{jl}^4\big)=O_P(n^2\psi_j). \qquad \text{(S6.20)}$$

We have that for $j=1,...,k$

$$\begin{aligned}\sum_{l=1}^{n}\Big(\boldsymbol{u}_{j(l)}^T\sum_{s=1(\neq j)}^{p}\boldsymbol{\omega}_{sj(l)}\Big)^2&=O_P(\psi_j),\\ \sum_{l=1}^{n}\hat{u}_{jl}^2\Big(\boldsymbol{1}_{n(l)}^T\sum_{s=1}^{p}\boldsymbol{\omega}_{sj(l)}\Big)^2&\leq\sum_{l=1}^{n}\Big(\boldsymbol{1}_{n(l)}^T\sum_{s=1}^{p}\boldsymbol{\omega}_{sj(l)}\Big)^2=O_P(n\psi_j),\\ \boldsymbol{\varepsilon}_j^T\sum_{l=1}^{n}\Big(\sum_{s=1}^{p}\boldsymbol{\omega}_{sj(l)}\Big)\Big(\sum_{s=1}^{p}\boldsymbol{\omega}_{sj(l)}\Big)^T\boldsymbol{\varepsilon}_j&\leq\sum_{l=1}^{n}\Big\|\sum_{s=1}^{p}\boldsymbol{\omega}_{sj(l)}\Big\|^2=O_P(n\psi_j)\ \text{ and}\\ \sum_{l=1}^{n}\boldsymbol{\zeta}_{jl}^T\Big(\sum_{s=1}^{p}\boldsymbol{\omega}_{sj(l)}\Big)\Big(\sum_{s=1}^{p}\boldsymbol{\omega}_{sj(l)}\Big)^T\boldsymbol{\zeta}_{jl}&\leq\sum_{l=1}^{n}\Big\|\sum_{s=1}^{p}\boldsymbol{\omega}_{sj(l)}\Big\|^2=O_P(n\psi_j) \qquad \text{(S6.21)}\end{aligned}$$

because it holds that $E\{\sum_{l=1}^{n}(\boldsymbol{u}_{j(l)}^T\sum_{s=1(\neq j)}^{p}\boldsymbol{\omega}_{sj(l)})^2\}=O(\psi_j)$, $E\{\sum_{l=1}^{n}(\boldsymbol{1}_{n(l)}^T\sum_{s=1}^{p}\boldsymbol{\omega}_{sj(l)})^2\}=O(n\psi_j)$, $E(\sum_{l=1}^{n}\|\sum_{s=1}^{p}\boldsymbol{\omega}_{sj(l)}\|^2)=O(n\psi_j)$ and $\hat{u}_{jl}^2\leq 1$ for all l. Then, by combining (S6.20) and (S6.21) with (S6.19), we can conclude the second result. □

Lemma B.2. *Assume (A-vi) and (A-viii). It holds that as $m\to\infty$*

$$\begin{aligned}&\tilde{\boldsymbol{h}}_{1j}^T\boldsymbol{h}_{2j'}=h_{jj'}+O_P(n_1^{-1/2})\quad\textit{and}\quad \boldsymbol{h}_{1j}^T\tilde{\boldsymbol{h}}_{2j'}=h_{jj'}+O_P(n_2^{-1/2});\\ &\tilde{\boldsymbol{h}}_{1j}^T\tilde{\boldsymbol{h}}_{2j'}-h_{jj'}=\tilde{\boldsymbol{h}}_{1j}^T\boldsymbol{h}_{2j'}-h_{jj'}+\boldsymbol{h}_{1j}^T\tilde{\boldsymbol{h}}_{2j'}-h_{jj'}+O_P\{(n_1n_2)^{-1/2}\}\end{aligned}$$

for $j=1,...,k_1$ and $j'=1,...,k_2$.

Proof. First, we consider the first result. We note that $(\boldsymbol{X}_i-\overline{\boldsymbol{X}}_i)\hat{\boldsymbol{u}}_{ij}=(\boldsymbol{X}_i-\boldsymbol{M}_i)\hat{\boldsymbol{u}}_{ij}$ when

$\hat{\lambda}_{ij} > 0$. From (S6.1) to (S6.5) we have that as $m \to \infty$

$$\tilde{\boldsymbol{h}}_{1j}^T \boldsymbol{h}_{2j'} = \frac{\hat{\boldsymbol{u}}_{1j}^T (\boldsymbol{X}_1 - \boldsymbol{M}_1)^T \boldsymbol{h}_{2j'}}{\{(n_i - 1)\tilde{\lambda}_{1j}\}^{1/2}} = \frac{\hat{\boldsymbol{u}}_{1j}^T \sum_{s=1}^p \lambda_{1s}^{1/2} h_{sj'} \boldsymbol{u}_{1s}}{\tilde{\lambda}_{1j}^{1/2}}$$

$$= h_{jj'} + \frac{\hat{\boldsymbol{u}}_{1j}^T \sum_{s=k_1+1}^p \lambda_{1s}^{1/2} h_{sj'} \boldsymbol{u}_{1s}}{\lambda_{1j}^{1/2}\{1 + o_P(1)\}} + O_P(n_1^{-1/2}) \quad \text{(S6.22)}$$

for $j = 1, ..., k_1$ and $j' = 1, ..., k_2$. It holds that for $j = 1, ..., k_1$ and $j' = 1, ..., k_2$

$$E\Big\{ \Big(\boldsymbol{u}_{1j}^T \sum_{s=k_1+1}^p \lambda_{1s}^{1/2} h_{sj'} \boldsymbol{u}_{1s} \Big)^2 \Big\} = O\Big(\sum_{s=k_1+1}^p \lambda_{1s} h_{sj'}^2 / n_1 \Big) = O(\lambda_{1k_1+1}/n_1);$$

$$E\Big(\Big\| \sum_{s=k_1+1}^p \lambda_{1s}^{1/2} h_{sj'} \boldsymbol{u}_{1s} \Big\|^2 \Big) = O\Big(\sum_{s=k_1+1}^p \lambda_{1s} h_{sj'}^2 \Big) = O(\lambda_{1k_1+1})$$

because $\sum_{s=k_1+1}^p h_{sj'}^2 \le 1$. Then, by using Markov's inequality, it follows from (S6.4) that for $j = 1, ..., k_1$ and $j' = 1, ..., k_2$

$$\frac{\hat{\boldsymbol{u}}_{1j}^T \sum_{s=k_1+1}^p \lambda_{1s}^{1/2} h_{sj'} \boldsymbol{u}_{1s}}{\lambda_{1j}^{1/2}} = O_P\{ n_1^{-1/2} (\lambda_{1k_1+1}/\lambda_{1j})^{1/2} \}. \quad \text{(S6.23)}$$

Thus, by combining (S6.22) with (S6.23), we can conclude the result for $\tilde{\boldsymbol{h}}_{1j}^T \boldsymbol{h}_{2j'}$. As for $\boldsymbol{h}_{1j}^T \tilde{\boldsymbol{h}}_{2j'}$, we obtain the result similarly.

Next, we consider the second result. From (S6.2), (S6.5) and (S6.23) we have that for $j \ne l = 1, ..., k_1$ and $j' \ne l' = 1, ..., k_2$

$$\frac{\hat{\boldsymbol{u}}_{1j}^T (\sum_{s=k_1+1}^p \lambda_{1s}^{1/2} \lambda_{2l'}^{1/2} h_{sl'} \boldsymbol{u}_{1s} \boldsymbol{u}_{2l'}^T) \hat{\boldsymbol{u}}_{2j'}}{\lambda_{1j}^{1/2} \lambda_{2j'}^{1/2}} = O_P\Big(\frac{\hat{\boldsymbol{u}}_{1j}^T \sum_{s=k_1+1}^p \lambda_{1s}^{1/2} h_{sl'} \boldsymbol{u}_{1s}}{n_2^{1/2} \lambda_{1j}^{1/2}} \Big)$$

$$= O_P[\{ \lambda_{1k_1+1}/(n_1 n_2 \lambda_{1j}) \}^{1/2}] \quad \text{and}$$

$$\frac{\hat{\boldsymbol{u}}_{1j}^T (\lambda_{1l}^{1/2} \boldsymbol{u}_{1l} \sum_{s'=k_2+1}^p \lambda_{2s'}^{1/2} h_{ls'} \boldsymbol{u}_{2s'}^T) \hat{\boldsymbol{u}}_{2j'}}{\lambda_{1j}^{1/2} \lambda_{2j'}^{1/2}} = O_P[\{ \lambda_{2k_2+1}/(n_1 n_2 \lambda_{2j'}) \}^{1/2}]. \quad \text{(S6.24)}$$

From (S6.1), (S6.3) and (S6.23) we have that for $j = 1, ..., k_1$ and $j' = 1, ..., k_2$

$$\frac{\hat{\boldsymbol{u}}_{1j}^T (\sum_{s=k_1+1}^p \lambda_{1s}^{1/2} \lambda_{2j'}^{1/2} h_{sj'} \boldsymbol{u}_{1s} \boldsymbol{u}_{2j'}^T) \hat{\boldsymbol{u}}_{2j'}}{\tilde{\lambda}_{1j}^{1/2} \tilde{\lambda}_{2j'}^{1/2}} = \frac{\hat{\boldsymbol{u}}_{1j}^T \sum_{s=k_1+1}^p \lambda_{1s}^{1/2} h_{sj'} \boldsymbol{u}_{1s}}{\tilde{\lambda}_{1j}^{1/2}} \{ 1 + O_P(n_2^{-1}) \}$$

$$= \frac{\hat{\boldsymbol{u}}_{1j}^T \sum_{s=k_1+1}^p \lambda_{1s}^{1/2} h_{sj'} \boldsymbol{u}_{1s}}{\tilde{\lambda}_{1j}^{1/2}} + O_P\{ (n_1 n_2)^{-1/2} \} \quad \text{and}$$

$$\frac{\hat{\boldsymbol{u}}_{1j}^T (\lambda_{1j}^{1/2} \boldsymbol{u}_{1j} \sum_{s'=k_2+1}^p \lambda_{2s'}^{1/2} h_{js'} \boldsymbol{u}_{2s'}^T) \hat{\boldsymbol{u}}_{2j'}}{\tilde{\lambda}_{1j}^{1/2} \tilde{\lambda}_{2j'}^{1/2}} = \frac{\sum_{s'=k_2+1}^p \lambda_{2s'}^{1/2} h_{js'} \boldsymbol{u}_{2s'}^T \hat{\boldsymbol{u}}_{2j'}}{\tilde{\lambda}_{2j}^{1/2}} + O_P\{ (n_1 n_2)^{-1/2} \}.$$

(S6.25)

It holds that for $j=1,...,k_1$ and $j'=1,...,k_2$

$$E\Big\{\Big(\sum_{s=k_1+1}^{p}\sum_{s'=k_2+1}^{p}\lambda_{1s}^{1/2}\lambda_{2s'}^{1/2}h_{ss'}\boldsymbol{u}_{1j}^T\boldsymbol{u}_{1s}\boldsymbol{u}_{2s'}^T\boldsymbol{u}_{2j'}\Big)^2\Big\}=O\Big(\frac{\mathrm{tr}(\boldsymbol{\Sigma}_{1*}\boldsymbol{\Sigma}_{2*})}{n_1n_2}\Big);$$
$$E\Big(\Big\|\sum_{s=k_1+1}^{p}\sum_{s'=k_2+1}^{p}\lambda_{1s}^{1/2}\lambda_{2s'}^{1/2}h_{ss'}\boldsymbol{u}_{1j}^T\boldsymbol{u}_{1s}\boldsymbol{u}_{2s'}\Big\|^2\Big)=O\{\mathrm{tr}(\boldsymbol{\Sigma}_{1*}\boldsymbol{\Sigma}_{2*})/n_1\};$$
$$E\Big(\Big\|\sum_{s=k_1+1}^{p}\sum_{s'=k_2+1}^{p}\lambda_{1s}^{1/2}\lambda_{2s'}^{1/2}h_{ss'}\boldsymbol{u}_{1s}\boldsymbol{u}_{2s'}^T\boldsymbol{u}_{2j'}\Big\|^2\Big)=O\{\mathrm{tr}(\boldsymbol{\Sigma}_{1*}\boldsymbol{\Sigma}_{2*})/n_2\};\quad\text{and}$$
$$E\Big(\Big\|\sum_{s=k_1+1}^{p}\sum_{s'=k_2+1}^{p}\lambda_{1s}^{1/2}\lambda_{2s'}^{1/2}h_{ss'}\boldsymbol{u}_{1s}\boldsymbol{u}_{2s'}^T\Big\|_F^2\Big)=O\{\mathrm{tr}(\boldsymbol{\Sigma}_{1*}\boldsymbol{\Sigma}_{2*})\},$$

where $||\cdot||_F$ is the Frobenius norm. Then, by noting that $\mathrm{tr}(\boldsymbol{\Sigma}_{1*}\boldsymbol{\Sigma}_{2*})\le\{\mathrm{tr}(\boldsymbol{\Sigma}_{1*}^2)\mathrm{tr}(\boldsymbol{\Sigma}_{2*}^2)\}^{1/2}$ and $|\boldsymbol{\zeta}_1^T(\sum_{s=k_1+1}^{p}\sum_{s'=k_2+1}^{p}\lambda_{1s}^{1/2}\lambda_{2s'}^{1/2}h_{ss'}\boldsymbol{u}_{1s}\boldsymbol{u}_{2s'}^T)\boldsymbol{\zeta}_2|\le||\sum_{s=k_1+1}^{p}\sum_{s'=k_2+1}^{p}\lambda_{1s}^{1/2}\lambda_{2s'}^{1/2}h_{ss'}\boldsymbol{u}_{1s}\boldsymbol{u}_{2s'}^T||_F$, it follows from (S6.4) that for $j=1,...,k_1$ and $j'=1,...,k_2$

$$\frac{\hat{\boldsymbol{u}}_{1j}^T(\sum_{s=k_1+1}^{p}\sum_{s'=k_2+1}^{p}\lambda_{1s}^{1/2}\lambda_{2s'}^{1/2}h_{ss'}\boldsymbol{u}_{1s}\boldsymbol{u}_{2s'}^T)\hat{\boldsymbol{u}}_{2j'}}{\lambda_{1j}^{1/2}\lambda_{2j'}^{1/2}}=O_P\Big\{\Big(\frac{\mathrm{tr}(\boldsymbol{\Sigma}_{1*}\boldsymbol{\Sigma}_{2*})}{n_1n_2\lambda_{1j}\lambda_{2j'}}\Big)^{1/2}\Big\}=O_P\{(n_1n_2)^{-1/2}\}. \tag{S6.26}$$

Then, from (S6.1) to (S6.5), (S6.24), (S6.25) and (S6.26) we have that for $j=1,...,k_1$ and $j'=1,...,k_2$

$$\begin{aligned}\tilde{\boldsymbol{h}}_{1j}^T\tilde{\boldsymbol{h}}_{2j'}=&\frac{\hat{\boldsymbol{u}}_{1j}^T(\sum_{s,s'}^{p}\lambda_{1s}^{1/2}\lambda_{2s'}^{1/2}h_{ss'}\boldsymbol{u}_{1s}\boldsymbol{u}_{2s'}^T)\hat{\boldsymbol{u}}_{2j'}}{\tilde{\lambda}_{1j}^{1/2}\tilde{\lambda}_{2j'}^{1/2}}\\
=&\frac{\hat{\boldsymbol{u}}_{1j}^T(\sum_{s=1}^{k_1}\sum_{s'=1}^{k_2}\lambda_{1s}^{1/2}\lambda_{2s'}^{1/2}h_{ss'}\boldsymbol{u}_{1s}\boldsymbol{u}_{2s'}^T)\hat{\boldsymbol{u}}_{2j'}}{\tilde{\lambda}_{1j}^{1/2}\tilde{\lambda}_{2j'}^{1/2}}+O_P\{(n_1n_2)^{-1/2}\}\\
&+\frac{\hat{\boldsymbol{u}}_{1j}^T\sum_{s=k_1+1}^{p}\lambda_{1s}^{1/2}h_{sj'}\boldsymbol{u}_{1s}}{\tilde{\lambda}_{1j}^{1/2}}+\frac{\sum_{s'=k_2+1}^{p}\lambda_{2s'}^{1/2}h_{js'}\boldsymbol{u}_{2s'}^T\hat{\boldsymbol{u}}_{2j'}}{\tilde{\lambda}_{2j'}^{1/2}}\\
=&h_{jj'}\Big(\frac{\lambda_{1j}^{1/2}\lambda_{2j'}^{1/2}\hat{\boldsymbol{u}}_{1j}^T\boldsymbol{u}_{1j}\boldsymbol{u}_{2j'}^T\hat{\boldsymbol{u}}_{2j'}}{\tilde{\lambda}_{1j}^{1/2}\tilde{\lambda}_{2j'}^{1/2}}-\frac{\lambda_{1j}^{1/2}\hat{\boldsymbol{u}}_{1j}^T\boldsymbol{u}_{1j}}{\tilde{\lambda}_{1j}^{1/2}}-\frac{\lambda_{2j'}^{1/2}\hat{\boldsymbol{u}}_{2j'}^T\boldsymbol{u}_{2j'}}{\tilde{\lambda}_{2j'}^{1/2}}\Big)\\
&+\tilde{\boldsymbol{h}}_{1j}^T\boldsymbol{h}_{2j'}+\boldsymbol{h}_{1j}^T\tilde{\boldsymbol{h}}_{2j'}+O_P\{(n_1n_2)^{-1/2}\}\\
=&h_{jj'}\Big(\frac{\lambda_{1j}^{1/2}\hat{\boldsymbol{u}}_{1j}^T\boldsymbol{u}_{1j}}{\tilde{\lambda}_{1j}^{1/2}}-1\Big)\Big(\frac{\lambda_{2j'}^{1/2}\hat{\boldsymbol{u}}_{2j'}^T\boldsymbol{u}_{2j'}}{\tilde{\lambda}_{2j'}^{1/2}}-1\Big)+\tilde{\boldsymbol{h}}_{1j}^T\boldsymbol{h}_{2j'}+\boldsymbol{h}_{1j}^T\tilde{\boldsymbol{h}}_{2j'}-h_{jj'}+O_P\{(n_1n_2)^{-1/2}\}\\
=&\tilde{\boldsymbol{h}}_{1j}^T\boldsymbol{h}_{2j'}+\boldsymbol{h}_{1j}^T\tilde{\boldsymbol{h}}_{2j'}-h_{jj'}+O_P\{(n_1n_2)^{-1/2}\}\end{aligned}$$

from the facts that $\tilde{\boldsymbol{h}}_{1j}^T\boldsymbol{h}_{2j'}=\hat{\boldsymbol{u}}_{1j}^T\sum_{s=1}^{p}\lambda_{1s}^{1/2}h_{sj'}\boldsymbol{u}_{1s}/\tilde{\lambda}_{1j}^{1/2}$ and $\boldsymbol{h}_{1j}^T\tilde{\boldsymbol{h}}_{2j'}=\sum_{s'=1}^{p}\lambda_{2s'}^{1/2}h_{js'}\boldsymbol{u}_{2s'}^T\hat{\boldsymbol{u}}_{2j'}/\tilde{\lambda}_{2j'}^{1/2}$. It concludes the second result. □

Proof of Theorem 5. We assume (A-ix) and (A-x). Let $\bar{x}_{ij}=\sum_{l=1}^{n_i}x_{ijl}/n_i$ for $i=1,2;\ j=$

$1,...,k_i$. For $i=1,2$ and $j=1,...,k_i$, we have that as $m\to\infty$

$$\bar{x}_{ij}-\mu_{i(j)}=O_P(\lambda_{ij}^{1/2}/n_i^{1/2}) \ \text{ and } \ \sum_{l=1}^{n}\frac{(x_{ijl}-\mu_{i(j)})^2}{n_i}=O_P(\lambda_{ij}) \tag{S6.27}$$

from the facts that $\mathrm{Var}(\sum_{l=1}^{n}x_{ijl}/n_i)=O(\lambda_{ij}/n_i)$ and $P(\sum_{l=1}^{n}(x_{ijl}-\mu_{i(j)})^2\ge \tau n_i\lambda_{ij})=O(\tau^{-1})$ for any $\tau>0$. Let $\bar{x}_{ij\star}=\sum_{l=1}^{n_i}(\tilde{x}_{ijl}-x_{ijl})/n_i$ for $i=1,2;\ j=1,...,k_i$. Note that for $i=1,2;\ j=1,...,k_i$

$$\frac{\psi_{ij}}{n_i^2K_{1*}^{1/2}}=O\Big(\frac{\lambda_{i1}^2}{n_i\mathrm{tr}(\boldsymbol{\Sigma}_{i*}^2)}+\frac{\boldsymbol{\mu}_{i*}^T\boldsymbol{\Sigma}_{i*}\boldsymbol{\mu}_{i*}+\sum_{s=1}^{k_i}\lambda_{is}\mu_{i(s)}^2}{\mathrm{tr}(\boldsymbol{\Sigma}_{i*}^2)}\Big)+o(1)\to 0$$

from the facts that $\mathrm{tr}(\boldsymbol{\Sigma}_i^2)\le k_i\lambda_{i1}^2+\mathrm{tr}(\boldsymbol{\Sigma}_{i*}^2)$, $\mathrm{tr}(\boldsymbol{\Sigma}_{i*}^2)^{1/2}/(n_iK_{1*}^{1/2})=O(1)$, $\boldsymbol{\mu}_i^T\boldsymbol{\Sigma}_i\boldsymbol{\mu}_i=\boldsymbol{\mu}_{i*}^T\boldsymbol{\Sigma}_{i*}\boldsymbol{\mu}_{i*}+\sum_{s=1}^{k_i}\lambda_{is}\mu_{i(s)}^2$ and $\mu_{i(s)}^2=O(\lambda_{is}/n_i)$ for $s=1,...,k_i$. Also, note that $\psi_{ij}^{1/2}\lambda_{ij}^{1/2}/(n_i^{3/2}K_{1*}^{1/2})\to 0$ for $i=1,2;\ j=1,...,k_i$. Then, with the help of Lemma B.1, we have that for $i=1,2;\ j=1,...,k_i$

$$\begin{aligned}
2\sum_{l<l'}^{n_i}\frac{\tilde{x}_{ijl}\tilde{x}_{ijl'}-x_{ijl}x_{ijl'}}{n_i^2}&=\sum_{l,l'}^{n_i}\frac{\tilde{x}_{ijl}\tilde{x}_{ijl'}-x_{ijl}x_{ijl'}}{n_i^2}-\sum_{l=1}^{n_i}\frac{\tilde{x}_{ijl}^2-x_{ijl}^2}{n_i^2}\\
&=\bar{x}_{ij\star}\sum_{l=1}^{n_i}\frac{\tilde{x}_{ijl}+x_{ijl}}{n_i}-\sum_{l=1}^{n_i}\frac{(\tilde{x}_{ijl}+x_{ijl})(\tilde{x}_{ijl}-x_{ijl})}{n_i^2}\\
&=\sum_{l=1}^{n_i}\frac{(\tilde{x}_{ijl}-x_{ijl})+2(x_{ijl}-\mu_{i(j)})+2\mu_{i(j)}}{n_i}\Big(\bar{x}_{ij\star}-\frac{\tilde{x}_{ijl}-x_{ijl}}{n_i}\Big)\\
&=O_P\{(\psi_{ij}^{1/2}/n_i)(\psi_{ij}^{1/2}/n_i+\lambda_{ij}^{1/2}/n_i^{1/2}+\mu_{i(j)})\}=o_P(K_{1*}^{1/2})
\end{aligned}\tag{S6.28}$$

from the fact that $\sum_{l=1}^{n_i}|(x_{ijl}-\mu_{i(j)})(\tilde{x}_{ijl}-x_{ijl})|\le\{\sum_{l=1}^{n_i}(x_{ijl}-\mu_{i(j)})^2\}^{1/2}\{\sum_{l=1}^{n_i}(\tilde{x}_{ijl}-x_{ijl})^2\}^{1/2}$. From Lemma B.2 it holds that for $j=1,...,k_1$ and $j'=1,...,k_2$

$$\begin{aligned}
&\tilde{\boldsymbol{h}}_{1j}^T\tilde{\boldsymbol{h}}_{2j'}=h_{jj'}+O_P(n_{\min}^{-1/2}),\quad \tilde{\boldsymbol{h}}_{1j}^T(\tilde{\boldsymbol{h}}_{2j'}-\boldsymbol{h}_{2j'})=O_P(n_2^{-1/2}),\\
&\tilde{\boldsymbol{h}}_{2j'}^T(\tilde{\boldsymbol{h}}_{1j}-\boldsymbol{h}_{1j})=O_P(n_1^{-1/2}) \ \text{ and } \ (\tilde{\boldsymbol{h}}_{1j}-\boldsymbol{h}_{1j})^T(\tilde{\boldsymbol{h}}_{2j'}-\boldsymbol{h}_{2j'})=O_P\{(n_1n_2)^{-1/2}\}.
\end{aligned}\tag{S6.29}$$

Then, it follows from Lemma B.1, (S6.27) and (S6.29) that for $j=1,...,k_1$ and $j'=1,...,k_2$

$$\begin{aligned}
&\frac{\sum_{l=1}^{n_1}(\tilde{x}_{1jl}\tilde{\boldsymbol{h}}_{1j}-x_{1jl}\boldsymbol{h}_{1j})^T\sum_{l'=1}^{n_2}(\tilde{x}_{2j'l'}\tilde{\boldsymbol{h}}_{2j'}-x_{2j'l'}\boldsymbol{h}_{2j'})}{n_1n_2}\\
&=\{\bar{x}_{1j\star}\tilde{\boldsymbol{h}}_{1j}+\bar{x}_{1j}(\tilde{\boldsymbol{h}}_{1j}-\boldsymbol{h}_{1j})\}^T\{\bar{x}_{2j'\star}\tilde{\boldsymbol{h}}_{2j}+\bar{x}_{2j'}(\tilde{\boldsymbol{h}}_{2j'}-\boldsymbol{h}_{2j'})\}\\
&=O_P\{(\psi_{1j}^{1/2}\psi_{2j'}^{1/2}/(n_1n_2)\}+O_P\{(\psi_{1j}^{1/2}/n_1)(\lambda_{2j'}^{1/2}/n_2^{1/2}+\mu_{2(j')})/n_2^{1/2}\}\\
&\quad+O_P\{(\psi_{2j'}^{1/2}/n_2)(\lambda_{1j}^{1/2}/n_1^{1/2}+\mu_{1(j)})/n_1^{1/2}\}\\
&\quad+O_P\{(\lambda_{1j}^{1/2}/n_1^{1/2}+\mu_{1(j)})(\lambda_{2j'}^{1/2}/n_2^{1/2}+\mu_{2(j')})/(n_1n_2)^{1/2}\}\\
&=o_P(K_{1*}^{1/2})
\end{aligned}\tag{S6.30}$$

from the fact that $\lambda_{ij}/(n_i^2K_{1*}^{1/2})=O\{\lambda_{ij}/(n_i\mathrm{tr}(\boldsymbol{\Sigma}_{i*}^2)^{1/2})\}=o(1)$ for $i=1,2;\ j=1,...,k_i$. Note that $\boldsymbol{\mu}_{i*}=\boldsymbol{A}_{i(k_i)}\boldsymbol{\mu}_i=\sum_{s=k_i+1}^p\mu_{i(s)}\boldsymbol{h}_{is}$ for $i=1,2$. We write that when $\hat{\lambda}_{1j}>0$,

$$\tilde{\boldsymbol{h}}_{1j}^T\Big(\sum_{l=1}^{n_2}\frac{\boldsymbol{v}_{2l}}{n_2}+\boldsymbol{\mu}_{2*}\Big)=\frac{\hat{\boldsymbol{u}}_{1j}^T(\boldsymbol{X}_1-\boldsymbol{M}_1)(\sum_{l=1}^{n_2}\boldsymbol{v}_{2l}/n_2+\boldsymbol{\mu}_{2*})}{(n_1-1)^{1/2}\tilde{\lambda}_{1j}^{1/2}}$$
$$=\hat{\boldsymbol{u}}_{1j}^T\sum_{s=1}^p\sum_{s'=k_2+1}^p\frac{\lambda_{1s}^{1/2}h_{ss'}\boldsymbol{u}_{1s}(\lambda_{2s'}^{1/2}\bar{z}_{2s'}+\mu_{2(s')})}{\tilde{\lambda}_{1j}^{1/2}}.\tag{S6.31}$$

It holds that

$$E\Big\{\Big(\boldsymbol{u}_{1j}^T\sum_{s=k_1+1}^p\sum_{s'=k_2+1}^p\lambda_{1s}^{1/2}h_{ss'}\boldsymbol{u}_{1s}(\lambda_{2s'}^{1/2}\bar{z}_{2s'}+\mu_{2(s')})\Big)^2\Big\}$$
$$=O\Big\{(\sum_{s=k_1+1}^p\lambda_{1s}\boldsymbol{h}_{1s}^T(\boldsymbol{\Sigma}_{2*}/n_2+\boldsymbol{\mu}_{2*}\boldsymbol{\mu}_{2*}^T)\boldsymbol{h}_{1s}\Big)/n_1\Big\}$$
$$=O\{\mathrm{tr}(\boldsymbol{\Sigma}_{1*}\boldsymbol{\Sigma}_{2*})/(n_1n_2)+\boldsymbol{\mu}_{2*}^T\boldsymbol{\Sigma}_{1*}\boldsymbol{\mu}_{2*}/n_1\}\ \text{ for } j=1,...,k_1;$$
$$E\Big(\Big\|\sum_{s=k_1+1}^p\sum_{s'=k_2+1}^p\lambda_{1s}^{1/2}h_{ss'}\boldsymbol{u}_{1s}(\lambda_{2s'}^{1/2}\bar{z}_{2s'}+\mu_{2(s')})\Big\|^2\Big)$$
$$=O\{\mathrm{tr}(\boldsymbol{\Sigma}_{1*}\boldsymbol{\Sigma}_{2*})/n_2+\boldsymbol{\mu}_{2*}^T\boldsymbol{\Sigma}_{1*}\boldsymbol{\mu}_{2*}\};\quad\text{and}$$
$$E\Big\{\Big(\sum_{s'=k_2+1}^p h_{js'}(\lambda_{2s'}^{1/2}\bar{z}_{2s'}+\mu_{2(s')})\Big)^2\Big\}$$
$$=O(\boldsymbol{h}_{1j}^T(\boldsymbol{\Sigma}_{2*}/n_2+\boldsymbol{\mu}_{2*}\boldsymbol{\mu}_{2*}^T)\boldsymbol{h}_{1j})=O\{\lambda_{2k_2+1}/n_2+(\boldsymbol{h}_{1j}^T\boldsymbol{\mu}_{2*})^2\}\ \text{ for } j=1,...,k_1.$$

In view of (S6.1) to (S6.5) and (S6.31), by using Markov's inequality, we have that for $j=1,...,k_1$

$$\tilde{\boldsymbol{h}}_{1j}^T\Big(\sum_{l=1}^{n_2}\frac{\boldsymbol{v}_{2l}}{n_2}+\boldsymbol{\mu}_{2*}\Big)=\sum_{s'=k_2+1}^p h_{js'}(\lambda_{2s'}^{1/2}\bar{z}_{2s'}+\mu_{2(s')})$$
$$+O_P[\{\mathrm{tr}(\boldsymbol{\Sigma}_{1*}\boldsymbol{\Sigma}_{2*})/(\lambda_{1j}n_1n_2)\}^{1/2}+\lambda_{2k_2+1}^{1/2}/(n_1n_2)^{1/2}]$$
$$+O_P\Big[\Big\{\boldsymbol{\mu}_{2*}^T\boldsymbol{\Sigma}_{1*}\boldsymbol{\mu}_{2*}/(\lambda_{1j}n_1)+\sum_{j'=1}^{k_1}(\boldsymbol{h}_{1j'}^T\boldsymbol{\mu}_{2*})^2/n_1\Big\}^{1/2}\Big].\tag{S6.32}$$

Note that $\boldsymbol{h}_{1j}^T(\sum_{l=1}^{n_2}\boldsymbol{v}_{2l}/n_2+\boldsymbol{\mu}_{2*})=\sum_{s'=k_2+1}^p h_{js'}(\lambda_{2s'}^{1/2}\bar{z}_{2s'}+\mu_{2(s')})$ and $\sum_{s'=k_2+1}^p h_{js'}(\lambda_{2s'}^{1/2}\bar{z}_{2s'}+\mu_{2(s')})=O_P(\lambda_{2k_2+1}^{1/2}/n_2^{1/2}+\boldsymbol{h}_{1j}^T\boldsymbol{\mu}_{2*})$. Also, note that $\lambda_{ik_i+1}=o\{\mathrm{tr}(\boldsymbol{\Sigma}_{i*}^2)^{1/2}\}$ for $i=1,2$. Then, it follows from Lemma B.1, (S6.27) and (S6.32) that for $j=1,...,k_1$

$$\{\bar{x}_{1j\star}\tilde{\boldsymbol{h}}_{1j}+\bar{x}_{1j}(\tilde{\boldsymbol{h}}_{1j}-\boldsymbol{h}_{1j})\}^T\Big(\sum_{l=1}^{n_2}\frac{\boldsymbol{v}_{2l}}{n_2}+\boldsymbol{\mu}_{2*}\Big)=o_P(K_{1*}^{1/2}).\tag{S6.33}$$

Similarly, it follows that for $j=1,...,k_2$

$$\Big(\sum_{l=1}^{n_1}\frac{\boldsymbol{v}_{1l}}{n_1}+\boldsymbol{\mu}_{1*}\Big)^T\{\bar{x}_{2j\star}\tilde{\boldsymbol{h}}_{2j}+\bar{x}_{2j}(\tilde{\boldsymbol{h}}_{2j}-\boldsymbol{h}_{2j})\}=o_P(K_{1*}^{1/2}).\tag{S6.34}$$

In view of (S6.30), (S6.33) and (S6.34), we have that

$$\frac{\sum_{l=1}^{n_1}(\boldsymbol{x}_{1l}-\sum_{j=1}^{k_1}\tilde{x}_{1jl}\tilde{\boldsymbol{h}}_{1j})^T\sum_{l'=1}^{n_2}(\boldsymbol{x}_{2l'}-\sum_{j'=1}^{k_2}\tilde{x}_{2j'l'}\tilde{\boldsymbol{h}}_{2j'})}{n_1n_2} = \frac{\sum_{l=1}^{n_1}(\boldsymbol{v}_{1l}+\boldsymbol{\mu}_{1*})^T\sum_{l'=1}^{n_2}(\boldsymbol{v}_{2l'}+\boldsymbol{\mu}_{2*})}{n_1n_2} + o_P(K_{1*}^{1/2}). \tag{S6.35}$$

Then, by combining (S6.28) with (S6.35), we have that $\widehat{T}_* - T_* = o_P(K_{1*}^{1/2})$, so that from Corollary 3, $(\widehat{T}_* - \Delta_*)/K_*^{1/2} \Rightarrow N(0,1)$ under $\limsup_{m\to\infty}\Delta_*^2/K_{1*} < \infty$. It concludes the results. □

Proof of Lemma 3. We assume (A-ix). Let $\boldsymbol{S}_{i(yy)} = (n_i-1)^{-1}\sum_{j=1}^{n_i}(\boldsymbol{y}_{ij}-\overline{\boldsymbol{y}}_i)(\boldsymbol{y}_{ij}-\overline{\boldsymbol{y}}_i)^T$, $\boldsymbol{S}_{i(yv)} = (n_i-1)^{-1}\sum_{j=1}^{n_i}(\boldsymbol{y}_{ij}-\overline{\boldsymbol{y}}_i)(\boldsymbol{v}_{ij}-\overline{\boldsymbol{v}}_i)^T$, $\boldsymbol{S}_{i(vy)} = \boldsymbol{S}_{i(yv)}^T$ and $\boldsymbol{S}_{i(vv)} = (n_i-1)^{-1}\sum_{j=1}^{n_i}(\boldsymbol{v}_{ij}-\overline{\boldsymbol{v}}_i)(\boldsymbol{v}_{ij}-\overline{\boldsymbol{v}}_i)^T$ for $i=1,2$, where $\overline{\boldsymbol{y}}_i = \sum_{j=1}^{n_i}\boldsymbol{y}_{ij}/n_i$ and $\overline{\boldsymbol{v}}_i = \sum_{j=1}^{n_i}\boldsymbol{v}_{ij}/n_i$. Note that $\boldsymbol{S}_{in_i} = \boldsymbol{S}_{i(yy)} + \boldsymbol{S}_{i(yv)} + \boldsymbol{S}_{i(vy)} + \boldsymbol{S}_{i(vv)}$ for $i=1,2$. Also, note that $\boldsymbol{S}_{i(yy)} = \sum_{j=1}^{k_i}\lambda_{ij}||\boldsymbol{u}_{oij}||^2\boldsymbol{h}_{ij}\boldsymbol{h}_{ij}^T + \sum_{j\neq j'}^{k_i}\lambda_{ij}^{1/2}\lambda_{ij'}^{1/2}\boldsymbol{u}_{oij}^T\boldsymbol{u}_{oij'}\boldsymbol{h}_{ij}\boldsymbol{h}_{ij'}^T$ for $i=1,2$. We write that for $i=1,2$

$$\boldsymbol{S}_{in_i}\sum_{j=1}^{k_i}\hat{\boldsymbol{h}}_{ij}\hat{\boldsymbol{h}}_{ij}^T = \sum_{j=1}^{k_i}\hat{\lambda}_{ij}\hat{\boldsymbol{h}}_{ij}\hat{\boldsymbol{h}}_{ij}^T = \sum_{j=1}^{k_i}\tilde{\lambda}_{ij}\tilde{\boldsymbol{h}}_{ij}\tilde{\boldsymbol{h}}_{ij}^T \ \ (=\widehat{\boldsymbol{S}}_{i(yy)},\text{ say}).$$

Then, by noting that $||\boldsymbol{u}_{oij}||^2 = ||\boldsymbol{u}_{ij}||^2 + O_P(n_i^{-1})$ and $\boldsymbol{u}_{oij}^T\boldsymbol{u}_{oij'} = O_P(n_i^{-1/2})$ $(j\neq j')$ as $n_i\to\infty$, it follows from (S6.3) that as $m\to\infty$

$$\begin{aligned}\boldsymbol{S}_{i(yy)} - \widehat{\boldsymbol{S}}_{i(yy)} &= \sum_{j=1}^{k_i}\tilde{\lambda}_{ij}(\boldsymbol{h}_{ij}\boldsymbol{h}_{ij}^T - \tilde{\boldsymbol{h}}_{ij}\tilde{\boldsymbol{h}}_{ij}^T) + \sum_{j\neq j'}^{k_i}\lambda_{ij}^{1/2}\lambda_{ij'}^{1/2}\boldsymbol{u}_{oij}^T\boldsymbol{u}_{oij'}\boldsymbol{h}_{ij}\boldsymbol{h}_{ij'}^T + O_P(n_i^{-1})\sum_{j=1}^{k_i}\lambda_{ij}\boldsymbol{h}_{ij}\boldsymbol{h}_{ij}^T \\ &= \sum_{j=1}^{k_i}\tilde{\lambda}_{ij}\{(\boldsymbol{h}_{ij}-\tilde{\boldsymbol{h}}_{ij})\boldsymbol{h}_{ij}^T - \tilde{\boldsymbol{h}}_{ij}(\tilde{\boldsymbol{h}}_{ij}-\boldsymbol{h}_{ij})^T\} \\ &\quad + O_P(n_i^{-1})\sum_{j=1}^{k_i}\lambda_{ij}\boldsymbol{h}_{ij}\boldsymbol{h}_{ij}^T + O_P(n_i^{-1/2})\sum_{j\neq j'}^{k_i}\lambda_{ij}^{1/2}\lambda_{ij'}^{1/2}\boldsymbol{h}_{ij}\boldsymbol{h}_{ij'}^T \end{aligned} \tag{S6.36}$$

for $i=1,2$. From Lemma B.2, (S6.29) and (S6.36) we have that

$$\mathrm{tr}\{(\boldsymbol{S}_{1(yy)} - \widehat{\boldsymbol{S}}_{1(yy)})(\boldsymbol{S}_{2(yy)} - \widehat{\boldsymbol{S}}_{2(yy)})\} = O_P\{\lambda_{11}\lambda_{21}(n_1n_2)^{-1/2}\},$$

so that

$$\mathrm{tr}\{(\boldsymbol{S}_{1(yy)} - \widehat{\boldsymbol{S}}_{1(yy)})(\boldsymbol{S}_{2(yy)} - \widehat{\boldsymbol{S}}_{2(yy)})\}/(n_1n_2) = o_P(K_{1*}) \tag{S6.37}$$

from the facts that $\lambda_{11}\lambda_{21}(n_1n_2)^{-3/2} \le \lambda_{11}^2/n_1^3 + \lambda_{21}^2/n_1^3$ and $\lambda_{i1} = o(n_i^{1/2}\mathrm{tr}(\boldsymbol{\Sigma}_{i*}^2)^{1/2})$. Note that $\boldsymbol{S}_{i(yv)} = \sum_{j=1}^{k_i}\sum_{s=k_i+1}^{p}\lambda_{ij}^{1/2}\lambda_{is}^{1/2}\boldsymbol{u}_{oij}^T\boldsymbol{u}_{ois}\boldsymbol{h}_{ij}\boldsymbol{h}_{is}^T$ for $i=1,2$. Here, we write that when $\hat{\lambda}_{1j} > 0$,

$$\begin{aligned}\tilde{\boldsymbol{h}}_{1j}^T\boldsymbol{S}_{2(yv)}\boldsymbol{h}_{1j} &= \sum_{j'=1}^{k_2}\lambda_{2j'}^{1/2}\frac{\hat{\boldsymbol{u}}_{1j}^T(\boldsymbol{X}_1-\boldsymbol{M}_1)^T(\sum_{s'=k_2+1}^{p}\lambda_{2s'}^{1/2}\boldsymbol{u}_{o2j'}^T\boldsymbol{u}_{o2s'}\boldsymbol{h}_{2j'}\boldsymbol{h}_{2s'}^T)}{(n_1-1)^{1/2}\tilde{\lambda}_{1j}^{1/2}}\boldsymbol{h}_{1j} \\ &= \sum_{j'=1}^{k_2}\lambda_{2j'}^{1/2}\hat{\boldsymbol{u}}_{1j}^T\sum_{s=1}^{p}\sum_{s'=k_2+1}^{p}\frac{\lambda_{1s}^{1/2}h_{sj'}h_{js'}\boldsymbol{u}_{1s}\lambda_{2s'}^{1/2}\boldsymbol{u}_{o2j'}^T\boldsymbol{u}_{o2s'}}{\tilde{\lambda}_{1j}^{1/2}}.\end{aligned} \tag{S6.38}$$

It holds that for $j'=1,...,k_2$

$$E\Big\{\Big(\boldsymbol{u}_{1j}^T\sum_{s=k_1+1}^{p}\sum_{s'=k_2+1}^{p}\lambda_{1s}^{1/2}h_{sj'}h_{js'}\boldsymbol{u}_{1s}\lambda_{2s'}^{1/2}\boldsymbol{u}_{o2j'}^T\boldsymbol{u}_{o2s'}\Big)^2\Big\}$$
$$=O\Big(\frac{\boldsymbol{h}_{1j}^T\boldsymbol{\Sigma}_{2*}\boldsymbol{h}_{1j}\boldsymbol{h}_{2j'}^T\boldsymbol{\Sigma}_{1*}\boldsymbol{h}_{2j'}}{n_1n_2}\Big)=O\Big(\frac{\lambda_{1k_1+1}\lambda_{2k_2+1}}{n_1n_2}\Big)\ \text{ for } j=1,...,k_1,$$
$$E\Big(\Big\|\sum_{s=k_1+1}^{p}\sum_{s'=k_2+1}^{p}\lambda_{1s}^{1/2}h_{sj'}h_{js'}\boldsymbol{u}_{1s}\lambda_{2s'}^{1/2}\boldsymbol{u}_{o2j'}^T\boldsymbol{u}_{o2s'}\Big\|^2\Big)=O\Big(\frac{\lambda_{1k_1+1}\lambda_{2k_2+1}}{n_2}\Big)$$
$$\text{and}\quad E\Big\{\Big(\sum_{s'=k_2+1}^{p}h_{sj'}h_{js'}\lambda_{2s'}^{1/2}\boldsymbol{u}_{o2j'}^T\boldsymbol{u}_{o2s'}\Big)^2\Big\}=O(\lambda_{2k_2+1}/n_2)\ \text{ for } j,s=1,...,k_1. \tag{S6.39}$$

In view of (S6.1) to (S6.5) and (S6.38), by using Markov's inequality, we have that for $j=1,...,k_1$

$$\lambda_{1j}\tilde{\boldsymbol{h}}_{1j}^T\boldsymbol{S}_{2(yv)}\boldsymbol{h}_{1j}=\lambda_{1j}\sum_{j'=1}^{k_2}\lambda_{2j'}^{1/2}\sum_{s'=k_2+1}^{p}h_{jj'}h_{js'}\lambda_{2s'}^{1/2}\boldsymbol{u}_{o2j'}^T\boldsymbol{u}_{o2s'}+o_P(n_1n_2K_{1*}) \tag{S6.40}$$

because $\lambda_{ik_i+1}=o\{\text{tr}(\boldsymbol{\Sigma}_{i*}^2)^{1/2}\}$ for $i=1,2$. Similarly, it follows that

$$\lambda_{1j}\boldsymbol{h}_{1j}^T\boldsymbol{S}_{2(yv)}\tilde{\boldsymbol{h}}_{1j}=\lambda_{1j}\sum_{j'=1}^{k_2}\lambda_{2j'}^{1/2}\sum_{s'=k_2+1}^{p}h_{jj'}h_{js'}\lambda_{2s'}^{1/2}\boldsymbol{u}_{o2j'}^T\boldsymbol{u}_{o2s'}+o_P(n_1n_2K_{1*}). \tag{S6.41}$$

We write that

$$\tilde{\boldsymbol{h}}_{1j}^T\boldsymbol{S}_{2(yv)}\tilde{\boldsymbol{h}}_{1j}=\sum_{j'=1}^{k_2}\lambda_{2j'}^{1/2}\hat{\boldsymbol{u}}_{1j}^T\sum_{s,t}\sum_{s'=k_2+1}^{p}\frac{\lambda_{1s}^{1/2}\lambda_{1t}^{1/2}h_{sj'}h_{ts'}\boldsymbol{u}_{1s}\lambda_{2s'}^{1/2}\boldsymbol{u}_{o2j'}^T\boldsymbol{u}_{o2s'}\boldsymbol{u}_{1t}^T}{\tilde{\lambda}_{1j}}\hat{\boldsymbol{u}}_{1j}. \tag{S6.42}$$

It holds that for $j'=1,...,k_2$

$$E\Big\{\Big\|\sum_{s,t\ge k_1+1}^{p}\sum_{s'=k_2+1}^{p}\lambda_{1s}^{1/2}\lambda_{1t}^{1/2}h_{sj'}h_{ts'}\boldsymbol{u}_{1s}\lambda_{2s'}^{1/2}\boldsymbol{u}_{o2j'}^T\boldsymbol{u}_{o2s'}\boldsymbol{u}_{1t}^T\Big\|_F^2\Big\}$$
$$=O\{\text{tr}(\boldsymbol{\Sigma}_{1*}\boldsymbol{\Sigma}_{2*})\boldsymbol{h}_{2j'}^T\boldsymbol{\Sigma}_{1*}\boldsymbol{h}_{2j'}/n_2+\boldsymbol{h}_{2j'}^T\boldsymbol{\Sigma}_{1*}\boldsymbol{\Sigma}_{2*}\boldsymbol{\Sigma}_{1*}\boldsymbol{h}_{2j'}/(n_1n_2)\}$$
$$=O\{\text{tr}(\boldsymbol{\Sigma}_{1*}\boldsymbol{\Sigma}_{2*})\lambda_{1k_1+1}/n_2+\lambda_{1k_1+1}^2\lambda_{2k_2+1}/(n_1n_2)\}. \tag{S6.43}$$

Then, in a way similar to (S6.26), by combing (S6.39) and (S6.43) with (S6.42), we have that for $j=1,...,k_1$

$$\lambda_{1j}\tilde{\boldsymbol{h}}_{1j}^T\boldsymbol{S}_{2(yv)}\tilde{\boldsymbol{h}}_{1j}=\lambda_{1j}\sum_{j'=1}^{k_2}\lambda_{2j'}^{1/2}\sum_{s'=k_2+1}^{p}h_{jj'}h_{js'}\lambda_{2s'}^{1/2}\boldsymbol{u}_{o2j'}^T\boldsymbol{u}_{o2s'}+o_P(n_1n_2K_{1*}). \tag{S6.44}$$

Also, from (S6.39) we have that for $j,j'=1,...,k_1$

$$\lambda_{1j}^{1/2}\lambda_{1j'}^{1/2}\boldsymbol{h}_{1j}^T\boldsymbol{S}_{2(yv)}\boldsymbol{h}_{1j'}=\lambda_{1j}^{1/2}\lambda_{1j'}^{1/2}\sum_{s=1}^{k_2}\lambda_{2s}^{1/2}\sum_{s'=k_2+1}^{p}h_{js}h_{j's'}\lambda_{2s'}^{1/2}\boldsymbol{u}_{o2s}^T\boldsymbol{u}_{o2s'}=o_P(n_1^{3/2}n_2K_{1*}). \tag{S6.45}$$

Then, it follows from (S6.36), (S6.40), (S6.41), (S6.44) and (S6.45) that

$$\mathrm{tr}\{(\boldsymbol{S}_{1(yy)}-\widehat{\boldsymbol{S}}_{1(yy)})\boldsymbol{S}_{2(yv)}\}=\mathrm{tr}\{(\boldsymbol{S}_{1(yy)}-\widehat{\boldsymbol{S}}_{1(yy)})\boldsymbol{S}_{2(vy)}\}=o_P(n_1n_2K_{1*}). \qquad \text{(S6.46)}$$

Similarly, it follows that $\mathrm{tr}\{(\boldsymbol{S}_{2(yy)}-\widehat{\boldsymbol{S}}_{2(yy)})\boldsymbol{S}_{1(yv)}\}=\mathrm{tr}\{(\boldsymbol{S}_{2(yy)}-\widehat{\boldsymbol{S}}_{2(yy)})\boldsymbol{S}_{1(vy)}\}=o_P(n_1n_2K_{1*})$. Note that $\boldsymbol{S}_{i(vv)}=\sum_{s,s'\ge k_i+1}^{p}\lambda_{is}^{1/2}\lambda_{is'}^{1/2}\boldsymbol{u}_{ois}^T\boldsymbol{u}_{ois'}\boldsymbol{h}_{is}\boldsymbol{h}_{is'}^T$ for $i=1,2$. Then, in a way similar to $\boldsymbol{S}_{i(yv)}$, we can claim that for $i=1,2$ $(j\ne i)$

$$\mathrm{tr}\{(\boldsymbol{S}_{i(yy)}-\widehat{\boldsymbol{S}}_{i(yy)})\boldsymbol{S}_{j(vv)}\}=o_P(n_1n_2K_{1*}). \qquad \text{(S6.47)}$$

Then, by combining (S6.46) and (S6.47) with (S6.37), we have that

$$\mathrm{tr}(\boldsymbol{S}_{1n_1}\widehat{\boldsymbol{A}}_{1(k_1)}\boldsymbol{S}_{2n_2}\widehat{\boldsymbol{A}}_{2(k_2)})=\mathrm{tr}\{(\boldsymbol{S}_{1n_1}-\boldsymbol{S}_{1(yy)})(\boldsymbol{S}_{2n_1}-\boldsymbol{S}_{2(yy)})\}+o_P(n_1n_2K_{1*}). \qquad \text{(S6.48)}$$

Let $\boldsymbol{\Sigma}_{i\star}=\sum_{j=1}^{k_i}\lambda_{ij}\boldsymbol{h}_{ij}\boldsymbol{h}_{ij}^T$ for $i=1,2$. We can evaluate that

$$\begin{aligned}
E[\{\mathrm{tr}(\boldsymbol{S}_{1(yv)}\boldsymbol{S}_{2(yv)})\}^2]=&O\Big(\frac{\mathrm{tr}(\boldsymbol{\Sigma}_{1\star}\boldsymbol{\Sigma}_{2*})\mathrm{tr}(\boldsymbol{\Sigma}_{1*}\boldsymbol{\Sigma}_{2\star})}{n_1n_2}\Big)=O\Big(\frac{\lambda_{11}\lambda_{21}\mathrm{tr}(\boldsymbol{\Sigma}_{1*}^2)^{1/2}\mathrm{tr}(\boldsymbol{\Sigma}_{2*}^2)^{1/2}}{n_1n_2}\Big);\\
E[\{\mathrm{tr}(\boldsymbol{S}_{1(yv)}\boldsymbol{S}_{2(vy)})\}^2]=&O\Big(\frac{\mathrm{tr}(\boldsymbol{\Sigma}_{1\star}\boldsymbol{\Sigma}_{2*})\mathrm{tr}(\boldsymbol{\Sigma}_{1*}\boldsymbol{\Sigma}_{2\star})}{n_1n_2}\Big)=O\Big(\frac{\lambda_{11}\lambda_{21}\mathrm{tr}(\boldsymbol{\Sigma}_{1*}^2)^{1/2}\mathrm{tr}(\boldsymbol{\Sigma}_{2*}^2)^{1/2}}{n_1n_2}\Big);\\
\text{and}\quad E[\{\mathrm{tr}(\boldsymbol{S}_{i(yv)}\boldsymbol{S}_{j(vv)})\}^2]=&O\Big(\frac{\mathrm{tr}(\boldsymbol{\Sigma}_{i\star}\boldsymbol{\Sigma}_{j*}\boldsymbol{\Sigma}_{i*}\boldsymbol{\Sigma}_{j*})}{n_i}\Big)+O\Big(\frac{\mathrm{tr}(\boldsymbol{\Sigma}_{i\star}\boldsymbol{\Sigma}_{j*})\mathrm{tr}(\boldsymbol{\Sigma}_{1*}\boldsymbol{\Sigma}_{2*})}{n_1n_2}\Big)\\
=&o\Big(\frac{\lambda_{i1}\mathrm{tr}(\boldsymbol{\Sigma}_{i*}^2)^{1/2}\mathrm{tr}(\boldsymbol{\Sigma}_{j*}^2)}{n_i}\Big)
\end{aligned}$$

for $i=1,2$ $(j\ne i)$. Then, we have that

$$\mathrm{tr}\{(\boldsymbol{S}_{1n_1}-\boldsymbol{S}_{1(yy)})(\boldsymbol{S}_{2n_1}-\boldsymbol{S}_{2(yy)})\}-\mathrm{tr}(\boldsymbol{S}_{1(vv)}\boldsymbol{S}_{2(vv)})=o_P(n_1n_2K_{1*}). \qquad \text{(S6.49)}$$

With the help of (23) in Aoshima and Yata (2015), we claim that $\mathrm{tr}(\boldsymbol{S}_{1(vv)}\boldsymbol{S}_{2(vv)})/\mathrm{tr}(\boldsymbol{\Sigma}_{1*}\boldsymbol{\Sigma}_{2*})=1+o_P(1)$. Hence, from (S6.48) and (S6.49) we have that

$$\mathrm{tr}(\boldsymbol{S}_{1n_1}\widehat{\boldsymbol{A}}_{1(k_1)}\boldsymbol{S}_{2n_2}\widehat{\boldsymbol{A}}_{1(k_2)})/(n_1n_2)=\mathrm{tr}(\boldsymbol{\Sigma}_{1*}\boldsymbol{\Sigma}_{2*})/(n_1n_2)+o_P(K_{1*}).$$

By using Lemma S2.1, we can conclude the result. □

Proof of Theorem 6. Similar to the proof of Theorem 3, by combining Theorem 5 and Lemma 3, we can conclude the result. □

S7 Appendix C

In this appendix, we give proofs of the theoretical results in Section S2.

Proofs of Propositions S2.1 and S2.2. We omit the subscript with regard to the population for the sake of simplicity. First, we consider Proposition S2.1. By using Lemmas 1 and 5 in Yata and Aoshima (2013), under (A-i) and (1.4), we can obtain $\boldsymbol{\zeta}^T\{\sum_{s=1}^{p}\lambda_s\boldsymbol{u}_s\boldsymbol{u}_s^T-\mathrm{tr}(\boldsymbol{\Sigma})\boldsymbol{I}_n/(n-1)\}\boldsymbol{\zeta}=$

$o_P\{\mathrm{tr}(\boldsymbol{\Sigma}^2)^{1/2}\}$ as $m_0 \to \infty$, so that $\hat{\lambda}_1 = \mathrm{tr}(\boldsymbol{\Sigma})/(n-1) + o_P\{\mathrm{tr}(\boldsymbol{\Sigma}^2)^{1/2}\}$ because $\hat{\lambda}_1 - \mathrm{tr}(\boldsymbol{\Sigma})/(n-1) = \hat{\boldsymbol{u}}_1^T\{\sum_{s=1}^p \lambda_s \boldsymbol{u}_s \boldsymbol{u}_s^T - \boldsymbol{I}_n \mathrm{tr}(\boldsymbol{\Sigma})/(n-1)\}\hat{\boldsymbol{u}}_1$. Then, by noting $\mathrm{tr}(\boldsymbol{S}_D) - \mathrm{tr}(\boldsymbol{\Sigma}) = o_P\{\mathrm{tr}(\boldsymbol{\Sigma}^2)^{1/2}\}$ under (A-i), we can claim $\tilde{\lambda}_1/\mathrm{tr}(\boldsymbol{\Sigma}^2)^{1/2} = o_P(1)$ under (A-i) and (1.4). Thus, from (2.3) we conclude the first result of Proposition S2.1. Under (1,6) there exists a fixed integer $j_\star$ such that $\lambda_{j_\star}/\lambda_1 \to 0$. Note that $(\sum_{i=j_\star}^p \lambda_i^4)/\lambda_1^4 \le \lambda_{j_\star}^2 \mathrm{tr}(\boldsymbol{\Sigma}^2)/\lambda_1^4 = o(1)$. Then, by using Lemma 1 and Corollary 4.1 in Yata and Aoshima (2013), we can claim that $\tilde{\lambda}_1/\lambda_1 = 1 + o_P(1)$ under (A-i) and (1.6). It concludes the results of Proposition S2.1.

Next, we consider Proposition S2.2. Let $\phi(n)$ be any function such that $\phi(n) \to 0$ and $n^{1/4}\phi(n) \to \infty$ as $n \to \infty$. Let $\lambda = \phi(n)\mathrm{tr}(\boldsymbol{\Sigma}^2)^{1/2}$. Assume that $\lambda_1^2/\mathrm{tr}(\boldsymbol{\Sigma}^2) = O(n^{-c})$ as $m_0 \to \infty$ with some fixed constant $c > 1/2$. Then, there is at least one positive integer t (> 2) satisfying $c(t/2-1) > t/4$, so that $\mathrm{tr}(\boldsymbol{\Sigma}^t)/\lambda^t \le \lambda_1^{t-2}/(\phi(n)^t \mathrm{tr}(\boldsymbol{\Sigma}^2)^{t/2-1}) = o(1)$. Then, by using Lemmas 1 and 5 in Yata and Aoshima (2013), we can obtain

$$\hat{\lambda}_1/\lambda = \hat{\boldsymbol{u}}_1^T \boldsymbol{S}_D \hat{\boldsymbol{u}}_1/\lambda = \mathrm{tr}(\boldsymbol{\Sigma})/\{(n-1)\lambda\} + o_P(1)$$

under (A-viii). Then, by noting that $\mathrm{tr}(\boldsymbol{S}_D) - \mathrm{tr}(\boldsymbol{\Sigma}) = o_P\{\mathrm{tr}(\boldsymbol{\Sigma}^2)^{1/2}\}$ under (A-viii), we can claim that $\tilde{\lambda}_1^2/\mathrm{tr}(\boldsymbol{\Sigma}^2) = o_P[\{\phi(n)\}^2]$ under (A-viii). Thus from (2.3) we conclude the result of Proposition S2.2. $\square$

Proofs of Lemma S2.1, Propositions S2.3 and S2.4. We assume (A-i) and (A-vi). We omit the subscript with regard to the population for the sake of simplicity. Let $\boldsymbol{V}_1 = \sum_{j=1}^k \lambda_j \boldsymbol{u}_{j(1)} \boldsymbol{u}_{j(2)}^T$ and $\boldsymbol{V}_2 = \sum_{j=k+1}^p \lambda_j \boldsymbol{u}_{j(1)} \boldsymbol{u}_{j(2)}^T$, where $\boldsymbol{u}_{j(1)} = (z_{j1}, ..., z_{jn_{(1)}})^T/(n_{(1)}-1)^{1/2}$ and $\boldsymbol{u}_{j(2)} = (z_{jn_{(1)}+1}, ..., z_{jn})^T/(n_{(2)}-1)^{1/2}$. Let $\boldsymbol{V}_{o1} = \boldsymbol{P}_{n_{(1)}} \boldsymbol{V}_1 \boldsymbol{P}_{n_{(2)}}$ and $\boldsymbol{V}_{o2} = \boldsymbol{P}_{n_{(1)}} \boldsymbol{V}_2 \boldsymbol{P}_{n_{(2)}}$. Note that $\boldsymbol{S}_{D(1)} = \boldsymbol{P}_{n_{(1)}}(\boldsymbol{V}_1 + \boldsymbol{V}_2)\boldsymbol{P}_{n_{(2)}} = \boldsymbol{V}_{o1} + \boldsymbol{V}_{o2}$. Let us write the singular value decomposition of $\boldsymbol{S}_{D(1)}$ as $\boldsymbol{S}_{D(1)} = \sum_{j=1}^{n_{(2)}-1} \acute{\lambda}_j \acute{\boldsymbol{u}}_{j(1)} \acute{\boldsymbol{u}}_{j(2)}^T$, where $\acute{\boldsymbol{u}}_{j(1)}$ (or $\acute{\boldsymbol{u}}_{j(2)}$) denotes a unit left- (or right-) singular vector corresponding to $\acute{\lambda}_j$. First, we consider Lemma S2.1. By using Lemma 1 and Corollary 5.1 in Yata and Aoshima (2013), we can claim for $j = 1, ..., k$ that as $m_0 \to \infty$

$$\acute{\lambda}_j/\lambda_j = 1 + o_P(1). \tag{S7.1}$$

By noting that $\mathrm{tr}(\boldsymbol{\Sigma}_*^4)/\mathrm{tr}(\boldsymbol{\Sigma}_*^2)^2 \le \lambda_{k+1}^2/\Psi_{(k+1)} = o(1)$ under (A-vi) and by using Lemmas 1 and 4 in Yata and Aoshima (2013), we can claim that $\boldsymbol{\zeta}_{(1)}^T \boldsymbol{P}_{n_{(1)}} \boldsymbol{V}_2 \boldsymbol{P}_{n_{(2)}} \boldsymbol{\zeta}_{(2)}/\Psi_{(k+1)}^{1/2} = o_P(1)$, where $\boldsymbol{\zeta}_{(i)}$ is an arbitrary unit random $n_{(i)}$-dimensional vector for $i = 1, 2$. Hence, we have that

$$\boldsymbol{\zeta}_{(1)}^T \boldsymbol{S}_{D(1)} \boldsymbol{\zeta}_{(2)}/\Psi_{(k+1)}^{1/2} = \boldsymbol{\zeta}_{(1)}^T \boldsymbol{V}_{o1} \boldsymbol{\zeta}_{(2)}/\Psi_{(k+1)}^{1/2} + o_P(1). \tag{S7.2}$$

Then, in a way similar to (A.10) in Yata and Aoshima (2013), we have that $\hat{\Psi}_{(k+1)}/\Psi_{(k+1)} = 1 + o_P(1)$. In view of (S7.1), we can claim that $\hat{\Psi}_{(j)}/\Psi_{(j)} = 1 + o_P(1)$ for $j = 1, ..., k$. It concludes the result of Lemma S2.1.

Next, we consider Proposition S2.3. By noting (S7.2) and $\mathrm{rank}(\boldsymbol{V}_{o1}) \le k$, we have that $\acute{\lambda}_j/\Psi_{(k+1)}^{1/2} = o_P(1)$ for $j > k$. Then, by combining Lemma S2.1 with (S7.1), we can conclude the result of Proposition S2.3.

Finally, we consider Proposition S2.4. We assume (A-viii) and (A-ix). Let $\lambda_* = \phi(n)\Psi_{(k+1)}^{1/2}$, where $\phi(n)$ is defined in the proofs of Propositions S2.1 and S2.2. Assume that $\lambda_{k+1}^2/\Psi_{(k+1)} = O(n^{-c})$ as $m_0 \to \infty$ with some fixed constant $c > 1/2$. Then, there is at least one positive integer t (> 2) satisfying $c(t/2-1) > t/4$, so that $\mathrm{tr}(\boldsymbol{\Sigma}_*^t)/\lambda_*^t \le \lambda_{k+1}^{t-2}/(\phi(n)^t\Psi_{(k+1)}^{t/2-1}) = o(1)$. Hence, similar to (S7.2), we have that

$$\boldsymbol{\zeta}_{(1)}^T\boldsymbol{S}_{D(1)}\boldsymbol{\zeta}_{(2)}/\lambda_* = \boldsymbol{\zeta}_{(1)}^T\boldsymbol{V}_{o1}\boldsymbol{\zeta}_{(2)}/\lambda_* + o_P(1). \tag{S7.3}$$

Let $\acute{\boldsymbol{V}}_{o1} = \boldsymbol{V}_{o1} - \sum_{j=1}^k \acute{\lambda}_j\acute{\boldsymbol{u}}_{j(1)}\acute{\boldsymbol{u}}_{j(2)}^T$. From (S7.3), it holds that $\boldsymbol{\zeta}_{(1)}^T\acute{\boldsymbol{V}}_{o1}\boldsymbol{\zeta}_{(2)}/\lambda_* = o_P(1)$, so that all the singular values of $\acute{\boldsymbol{V}}_1/\lambda_*$ are of the order $o_P(1)$. Then, from the fact that $\mathrm{rank}(\acute{\boldsymbol{V}}_{o1}) \le 2k$, it holds that

$$\mathrm{tr}(\acute{\boldsymbol{V}}_{o1}\acute{\boldsymbol{V}}_{o1}^T)/\Psi_{(k+1)} = k \times o_P[\{\phi(n)\}^2]. \tag{S7.4}$$

Here, in view of (A-viii), we have that $\mathrm{Var}(\boldsymbol{u}_{j(1)}^T\boldsymbol{V}_{o2}\boldsymbol{u}_{j(2)}) = O(\Psi_{(k+1)}/n^2)$ for $j = 1, ..., k$, so that $\boldsymbol{u}_{j(1)}^T\boldsymbol{V}_{o2}\boldsymbol{u}_{j(2)} = O_P(\Psi_{(k+1)}^{1/2}/n)$ for $j = 1, ..., k$. In view of (A-ix), it holds that

$$\mathrm{tr}(\boldsymbol{V}_{o1}\boldsymbol{V}_{o2}^T)/\Psi_{(k+1)} = \mathrm{tr}(\boldsymbol{V}_1\boldsymbol{V}_{o2}^T)/\Psi_{(k+1)} = k \times o_P(n^{-1/2}). \tag{S7.5}$$

On the other hand, we have that $E(||\boldsymbol{u}_{j(1)}^T\boldsymbol{V}_{o2}||^2) = O(\Psi_{(k+1)}/n)$ and $E(||\boldsymbol{u}_{j(2)}^T\boldsymbol{V}_{o2}^T||^2) = O(\Psi_{(k+1)}/n)$ for $j = 1, ..., k$, so that $\boldsymbol{u}_{j(1)}^T\boldsymbol{V}_{o2}\boldsymbol{\zeta}_{(2)} = O_P(\Psi_{(k+1)}^{1/2}/n^{1/2})$ and $\boldsymbol{\zeta}_{(1)}^T\boldsymbol{V}_{o2}\boldsymbol{u}_{j(2)} = O_P(\Psi_{(k+1)}^{1/2}/n^{1/2})$ for $j = 1, ..., k$. Then, in a way similar to the proof of Lemma 12 in Yata and Aoshima (2013), we have that $\acute{\boldsymbol{u}}_{j(l)} = ||\boldsymbol{u}_{j(l)}||^{-1}\boldsymbol{u}_{j(l)}\{1 + O_P(n^{-1/2})\} + \boldsymbol{\varepsilon}_{jl} \times O_P(n^{-1/2})$ with some unit random vector $\boldsymbol{\varepsilon}_{jl}$ for $j = 1, ..., k;\ l = 1, 2$. Hence, from (S7.1) and $\boldsymbol{\zeta}_{(1)}^T\boldsymbol{V}_{o2}\boldsymbol{\zeta}_{(2)} = o_P(\Psi_{(k+1)}^{1/2})$, we have that $\mathrm{tr}(\sum_{j=1}^k \acute{\lambda}_j\acute{\boldsymbol{u}}_{j(1)}\acute{\boldsymbol{u}}_{j(2)}^T\boldsymbol{V}_{o2}^T)/\Psi_{(k+1)} = k \times o_P(n^{-1/2})$. Hence, from (S7.5), it holds that

$$\mathrm{tr}(\acute{\boldsymbol{V}}_{o1}\boldsymbol{V}_{o2}^T)/\Psi_{(k+1)} = k \times o_P(n^{-1/2}). \tag{S7.6}$$

Note that $E\{\mathrm{tr}(\boldsymbol{V}_{o2}\boldsymbol{V}_{o2}^T)\} = \Psi_{(k+1)}$ and $\mathrm{Var}\{\mathrm{tr}(\boldsymbol{V}_{o2}\boldsymbol{V}_{o2}^T)/\Psi_{(k+1)}\} = O(n^{-1})$. Then, by noting that $\widehat{\Psi}_{(k+1)} = \mathrm{tr}\{(\acute{\boldsymbol{V}}_{o1} + \boldsymbol{V}_{o2})(\acute{\boldsymbol{V}}_{o1} + \boldsymbol{V}_{o2})^T\}$, from (S7.4) and (S7.6), we obtain that

$$\widehat{\Psi}_{(k+1)}/\Psi_{(k+1)} = \mathrm{tr}(\boldsymbol{V}_{o2}\boldsymbol{V}_{o2}^T)/\Psi_{(k+1)} + k \times o_P[\{\phi(n)\}^2] = 1 + k \times o_P[\{\phi(n)\}^2]. \tag{S7.7}$$

Similarly, by noting that $\acute{\lambda}_{k+1}/\lambda_* = o_P(1)$ from (S7.3), we can claim that

$$\widehat{\Psi}_{(k+2)}/\Psi_{(k+1)} = \{1 + o(n^{-1/2})\}\widehat{\Psi}_{(k+2)}/\Psi_{(k+2)} = 1 + (k+1) \times o_P[\{\phi(n)\}^2]. \tag{S7.8}$$

By combining (S7.7) and (S7.8), we can conclude the result of Proposition S2.4. □

References

Aoshima, M. and Yata, K. (2011). Two-stage procedures for high-dimensional data. *Sequential Anal. (Editor's special invited paper)* **30**, 356-399.

Aoshima, M. and Yata, K. (2015). Asymptotic normality for inference on multisample, high-dimensional mean vectors under mild conditions. *Methodol. Comput. Appl. Probab.* **17**, 419-439.

Aoshima, M. and Yata, K. (2016). Two-sample tests for high-dimension, strongly spiked eigenvalue models.

Azzalini, A. and Dalla Valle, A. (1996). The multivariate skew-normal distribution. *Biometrika* **83**, 715-726.

Azzalini, A. and Capitanio, A. (1999). Statistical applications of the multivariate skew normal distribution. *J. R. Statist. Soc. Ser. B* **61**, 579-602.

Azzalini, A. and Capitanio, A. (2003). Distributions generated by perturbation of symmetry with emphasis on a multivariate skew t-distribution. *J. R. Statist. Soc. Ser. B* **65**, 367-389.

Golub, T. R., Slonim, D. K., Tamayo, P., Huard, C., Gaasenbeek, M., Mesirov, J. P., Coller, H., Loh, M. L., Downing, J. R., Caligiuri, M. A., Bloomfield, C. D. and Lander, E. S. (1999). Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. *Science* **286**, 531-537.

Gupta, A. K. (2003). Multivariate skew t-distribution. *Statistics* **37**, 359-363.

Singh, D., Febbo, P. G., Ross, K., Jackson, D. G., Manola, J., Ladd, C., Tamayo, P., Renshaw, A. A., D'Amico, A. V., Richie, J. P., Lander, E. S., Loda, M., Kantoff, P. W., Golub, T. R. and Sellers, W. R. (2002). Gene expression correlates of clinical prostate cancer behavior. *Cancer Cell* **1**, 203-209.

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

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

Yata, K. and Aoshima, M. (2013). PCA consistency for the power spiked model in high-dimensional settings. *J. Multivariate Anal.* **122**, 334-354.

Institute of Mathematics, University of Tsukuba, Ibaraki 305-8571, Japan.

E-mail: aoshima@math.tsukuba.ac.jp

Institute of Mathematics, University of Tsukuba, Ibaraki 305-8571, Japan.

E-mail: yata@math.tsukuba.ac.jp