Cross Projection Test for High-Dimensional Mean Vectors

Guanpeng Wang and Hengjian Cui*

Capital Normal University and Weifang University

Supplementary Material

This Supplement contains proofs of the theorems in the paper and other contributed results. Subsections S1–S5 contain proofs of theorems 1, 2, 3, 4 and 5 respectively. Subsection S6 contains some additional simulation results for the performance of the test statistic T^2_{CP}.

S1 Proof of Theorem 1

First, we restate the asymptotic distribution for linear quadratic forms (see Theorem 2.1 in Srivastava (2009)).

Lemma 1. *We assume that z_{ij} are i.i.d. random variables with $\mathrm{E}(z_{ij}) = 0$, var$(z_{ij}) = 1$, fourth moment κ, and $\bar{\mathbf{z}} = (\bar{z}_1, \ldots, \bar{z}_p)^T$, where $\bar{z}_i = \frac{1}{n}\sum_{j=1}^n z_{ij}$, $i = 1, \ldots, p$ $j = 1, \ldots, n$. Then for any $p \times p$ symmetric matrix, $\mathbf{A} = (a_{ij})$, suppose the following assumptions hold: (i):$\lim\limits_{p\to\infty} \max\limits_{1\le j\le p} \left(\frac{a_{jj}^2}{p}\right) = 0$ and (ii): $\lim\limits_{p\to\infty} \left(\mathrm{tr}\mathbf{A}_+^i/p\right) < \infty$, $i = 1, 2, 4$, where $\mathbf{A}_+ = (a_{ij+})$ is a $p \times p$ symmetric matrix defined by $a_{ii+} = a_{ii}$,* *and* $a_{ij+} = |a_{ij}|$*. As* $n, p \to \infty$*, then the following result holds:*

$$P\Big[\Big(\frac{n\bar{\mathbf{z}}^T\mathbf{A}\bar{\mathbf{z}} - \mathrm{tr}(\mathbf{A})}{\sqrt{2p\tau_2}}\Big) \le x\Big] = \Phi(x),$$

where $\Phi(x)$ *is the cumulative distribution function of a standard normal random variable, and* $\tau_2 = \frac{\mathrm{tr}(\mathbf{A}^2)}{p}$.

Proof of Theorem 1: Recall that the definition of t_{o}^2 is

$$t_{\mathrm{o}}^2 = n_1\bar{\mathbf{x}}_1^T U_2 \mathbf{W}_1^{-1} U_2^T \bar{\mathbf{x}}_1 + n_2\bar{\mathbf{x}}_2^T U_1 \mathbf{W}_2^{-1} U_1^T \bar{\mathbf{x}}_2,$$

where $\mathbf{W}_1 = \mathrm{diag}(\mathbf{u}_{21}^T\mathbf{\Sigma}\mathbf{u}_{21}, \dots, \mathbf{u}_{2p}^T\mathbf{\Sigma}\mathbf{u}_{2p})$ and $\mathbf{W}_2 = \mathrm{diag}(\mathbf{u}_{11}^T\mathbf{\Sigma}\mathbf{u}_{11}, \dots, \mathbf{u}_{1p}^T\mathbf{\Sigma}\mathbf{u}_{1p})$. In this proof, we will prove the asymptotic normality of t_{o}^2 in two steps. The first step is to prove that

$$\frac{n_1\bar{\mathbf{x}}_1^T U_2 \mathbf{W}_1^{-1} U_2^T \bar{\mathbf{x}}_1 - p}{\sqrt{2\mathrm{tr}(\mathbf{R}_1^2)}} \xrightarrow{d.} N(0,1) \quad \text{and} \quad \frac{n_2\bar{\mathbf{x}}_2^T U_1 \mathbf{W}_2^{-1} U_1^T \bar{\mathbf{x}}_2 - p}{\sqrt{2\mathrm{tr}(\mathbf{R}_2^2)}} \xrightarrow{d.} N(0,1). \tag{S.1}$$

The second step is to prove that the two aforementioned parts (S.1) are uncorrelated terms. Now, from the independent components structure, we can see that

$$\begin{aligned} n_1\bar{\mathbf{x}}_1^T U_2 \mathbf{W}_1^{-1} U_2^T \bar{\mathbf{x}}_1 &= n_1\bar{\mathbf{z}}_1^T\big(\Gamma^T U_2 \mathbf{W}_1^{-1} U_2^T \Gamma\big)\bar{\mathbf{z}}_1 \\ &=: n_1\bar{\mathbf{z}}_1^T\mathbf{B}\bar{\mathbf{z}}_1, \end{aligned}$$

where $\mathbf{B} = \Gamma^T U_2 \mathbf{W}_1^{-1} U_2^T \Gamma$. It follows that

$$\begin{aligned} \mathrm{tr}(\mathbf{B}) &= \mathrm{tr}(\Gamma^T U_2 \mathbf{W}_1^{-1} U_2^T \Gamma) = \mathrm{tr}(\mathbf{W}_1^{-1} U_2^T \Gamma\Gamma^T U_2) \\ &= \mathrm{tr}(\mathbf{W}_1^{-1} U_2^T \mathbf{\Sigma} U_2) = \mathrm{tr}(\mathbf{W}_1^{-1/2} U_2^T \mathbf{\Sigma} U_2 \mathbf{W}_1^{-1/2}) = \mathrm{tr}(\mathbf{R}_1) = p, \end{aligned}$$

and

$$\begin{aligned}\mathrm{tr}(\mathbf{B}^2) &= \mathrm{tr}(\Gamma^T U_2 \mathbf{W}_1^{-1} U_2^T \Gamma \Gamma^T U_2 \mathbf{W}_1^{-1} U_2^T \Gamma) = \mathrm{tr}(\mathbf{W}_1^{-1} U_2^T \mathbf{\Sigma} U_2 \mathbf{W}_1^{-1} U_2^T \Gamma \Gamma^T U_2) \\ &= \mathrm{tr}(\mathbf{W}_1^{-1} U_2^T \mathbf{\Sigma} U_2 \mathbf{W}_1^{-1} U_2^T \mathbf{\Sigma} U_2) = \mathrm{tr}(\mathbf{W}_1^{-1/2} U_2^T \mathbf{\Sigma} U_2 \mathbf{W}_1^{-1} U_2^T \mathbf{\Sigma} U_2 \mathbf{W}_1^{-1/2}) \\ &= \mathrm{tr}(\mathbf{R}_1^2).\end{aligned}$$

In the framework of our projection test, as long as Assumption 3 holds, the following conclusions can also be naturally established

$$\lim_{p\to\infty} \varrho_i = \lim_{p\to\infty} \left(\frac{\mathrm{tr}((\mathbf{R}_1)^i)}{p} \right) = \varrho_{i0} < \infty, \quad i = 1, \ldots, 4, \tag{S.2}$$

where $\mathbf{R}_1 = \mathbf{D}_2^{-1/2}(U_2^T \mathbf{\Sigma} U_2)\mathbf{D}_2^{-1/2}$ and $\mathbf{D}_2 = \mathrm{diag}(\mathbf{u}_{21}^T \mathbf{\Sigma} \mathbf{u}_{21}, \ldots, \mathbf{u}_{2p}^T \mathbf{\Sigma} \mathbf{u}_{2p})$ for given projection matrix U_2. Let the other correlation coefficient matrix $\mathbf{R}_2 = \mathbf{D}_1^{-1/2}(U_1^T \mathbf{\Sigma} U_1)\mathbf{D}_1^{-1/2}$, where $\mathbf{D}_1 = \mathrm{diag}(\mathbf{u}_{11}^T \mathbf{\Sigma} \mathbf{u}_{11}, \ldots, \mathbf{u}_{1p}^T \mathbf{\Sigma} \mathbf{u}_{1p})$. Similarly, for given projection matrix U_1, projection correlation matrix $\mathbf{R}_2$ still has the conclusion of (S.2). Particularly, overcoming the correlation between two variables in the covariance matrix by using the projection technique holds for many covariance matrix models, for example, when the covariance matrix is diagonal as well as the band structure, autoregressive, and factor models. Hence, the result of (S.2) holds for two assumptions in Lemma 1, and combining this with Assumption 2 completes the asymptotically standard normality distribution in (S.1). Then, we split expression t_{o}^2 into two terms, writing $n_1 \bar{\mathbf{x}}_1^T U_2 \mathbf{W}_1^{-1} U_2^T \bar{\mathbf{x}}_1$ and $n_2 \bar{\mathbf{x}}_2^T U_1 \mathbf{W}_2^{-1} U_1^T \bar{\mathbf{x}}_2$ as I_1 and I_2, respectively. For the sake of calculation simplicity, the main calculation formulas involved in terms I_1 and I_2 are expressed with simple symbols, which are respectively defined as follows:

$$\begin{aligned} I_1 + I_2 &= \sum_{i=1}^{p}\left(\frac{\sqrt{n_1}\mathbf{u}_{2i}^T\bar{\mathbf{x}}_1}{\sqrt{\mathbf{u}_{2i}^T\boldsymbol{\Sigma}\mathbf{u}_{2i}}}\right)^2 + \sum_{j=1}^{p}\left(\frac{\sqrt{n_2}\mathbf{u}_{1j}^T\bar{\mathbf{x}}_2}{\sqrt{\mathbf{u}_{1j}^T\boldsymbol{\Sigma}\mathbf{u}_{1j}}}\right)^2 \\ &=: \sum_{i=1}^{p}\left(\frac{\sqrt{n_1}\mathbf{u}_{i}^T\bar{\mathbf{x}}_1}{\sqrt{\mathbf{u}_{2i}^T\boldsymbol{\Sigma}\mathbf{u}_{2i}}}\right)^2 + \sum_{j=1}^{p}\left(\frac{\sqrt{n_2}\mathbf{v}_{j}^T\bar{\mathbf{x}}_2}{\sqrt{\mathbf{v}_{j}^T\boldsymbol{\Sigma}\mathbf{v}_{j}}}\right)^2. \end{aligned}$$

Hence, it is shown that

$$\begin{aligned} \operatorname{cov}(I_1, I_2) &= \operatorname{cov}\Big(\sum_{i=1}^{p}\Big(\frac{\sqrt{n_1}\mathbf{u}_i^T\bar{\mathbf{x}}_1}{\sqrt{\mathbf{u}_i^T\boldsymbol{\Sigma}\mathbf{u}_i}}\Big)^2, \sum_{j=1}^{p}\Big(\frac{\sqrt{n_2}\mathbf{v}_j^T\bar{\mathbf{x}}_2}{\sqrt{\mathbf{v}_j^T\boldsymbol{\Sigma}\mathbf{v}_j}}\Big)^2\Big) \\ &= \sum_{i=1}^{p}\sum_{j=1}^{p}\Big\{\mathrm{E}\Big(\Big(\frac{\sqrt{n_1}\mathbf{u}_i^T\bar{\mathbf{x}}_1}{\sqrt{\mathbf{u}_i^T\boldsymbol{\Sigma}\mathbf{u}_i}}\Big)^2\Big(\frac{\sqrt{n_2}\mathbf{v}_j^T\bar{\mathbf{x}}_2}{\sqrt{\mathbf{v}_j^T\boldsymbol{\Sigma}\mathbf{v}_j}}\Big)^2\Big) - \mathrm{E}\Big(\frac{\sqrt{n_1}\mathbf{u}_i^T\bar{\mathbf{x}}_1}{\sqrt{\mathbf{u}_i^T\boldsymbol{\Sigma}\mathbf{u}_i}}\Big)^2\mathrm{E}\Big(\frac{\sqrt{n_2}\mathbf{v}_j^T\bar{\mathbf{x}}_2}{\sqrt{\mathbf{v}_j^T\boldsymbol{\Sigma}\mathbf{v}_j}}\Big)^2\Big\} \\ &= \sum_{i=1}^{p}\sum_{j=1}^{p}\mathrm{E}\Big(\Big(\frac{\sqrt{n_1}\mathbf{u}_i^T\bar{\mathbf{x}}_1}{\sqrt{\mathbf{u}_i^T\boldsymbol{\Sigma}\mathbf{u}_i}}\Big)^2\Big(\frac{\sqrt{n_2}\mathbf{v}_j^T\bar{\mathbf{x}}_2}{\sqrt{\mathbf{v}_j^T\boldsymbol{\Sigma}\mathbf{v}_j}}\Big)^2\Big) - p \times p, \end{aligned} \tag{S.3}$$

where

$$\begin{aligned} &\sum_{i,j=1}^{p}\mathrm{E}\Big(\Big(\frac{\sqrt{n_1}\mathbf{u}_i^T\bar{\mathbf{x}}_1}{\sqrt{\mathbf{u}_i^T\boldsymbol{\Sigma}\mathbf{u}_i}}\Big)^2\Big(\frac{\sqrt{n_2}\mathbf{v}_j^T\bar{\mathbf{x}}_2}{\sqrt{\mathbf{v}_j^T\boldsymbol{\Sigma}\mathbf{v}_j}}\Big)^2\Big) \\ =& \sum_{i,j=1}^{p}\mathrm{E}\Big(\frac{n_1\mathbf{u}_i^T\bar{\mathbf{x}}_1\bar{\mathbf{x}}_1^T\mathbf{u}_i}{\mathbf{u}_i^T\boldsymbol{\Sigma}\mathbf{u}_i}\cdot\frac{n_2\mathbf{v}_j^T\bar{\mathbf{x}}_2\bar{\mathbf{x}}_2^T\mathbf{v}_j}{\mathbf{v}_j^T\boldsymbol{\Sigma}\mathbf{v}_j}\Big) = \sum_{i,j=1}^{p}\mathrm{E}\Big(\frac{n_1\mathbf{u}_i^T\Gamma\bar{\mathbf{z}}_1\bar{\mathbf{z}}_1^T\Gamma^T\mathbf{u}_i}{\mathbf{u}_i^T\boldsymbol{\Sigma}\mathbf{u}_i}\cdot\frac{n_2\mathbf{v}_j^T\bar{\mathbf{x}}_2\bar{\mathbf{x}}_2^T\mathbf{v}_j}{\mathbf{v}_j^T\boldsymbol{\Sigma}\mathbf{v}_j}\Big) \\ =& \sum_{i,j=1}^{p}\mathrm{E}\Big(\mathrm{E}\Big(\frac{\mathrm{tr}(n_1\bar{\mathbf{z}}_1\bar{\mathbf{z}}_1^T\Gamma^T\mathbf{u}_i\mathbf{u}_i^T\Gamma)}{\mathbf{u}_i^T\boldsymbol{\Sigma}\mathbf{u}_i}\cdot\frac{\mathrm{tr}(n_2\bar{\mathbf{z}}_2\bar{\mathbf{z}}_2^T\Gamma^T\mathbf{v}_i\mathbf{v}_i^T\Gamma)}{\mathbf{v}_i^T\boldsymbol{\Sigma}\mathbf{v}_i}|\mathbf{S}_{\mathbf{z}_1}, \mathbf{S}_{\mathbf{z}_2}\Big)\Big) \\ =& \sum_{i,j=1}^{p}\mathrm{E}\Big(\frac{1}{\mathbf{u}_i^T\boldsymbol{\Sigma}\mathbf{u}_i}\frac{1}{\mathbf{v}_j^T\boldsymbol{\Sigma}\mathbf{v}_j}\mathrm{E}\Big(\mathrm{tr}(n_1\bar{\mathbf{z}}_1\bar{\mathbf{z}}_1^T\Gamma^T\mathbf{u}_i\mathbf{u}_i^T\Gamma)|\mathbf{S}_{\mathbf{z}_2}\Big)\mathrm{E}\Big(\mathrm{tr}(n_2\bar{\mathbf{z}}_2\bar{\mathbf{z}}_2^T\Gamma^T\mathbf{v}_j\mathbf{v}_j^T\Gamma)|\mathbf{S}_{\mathbf{z}_1}\Big)\Big) \\ =& \sum_{i,j=1}^{p}\mathrm{E}\Big(\frac{1}{\mathbf{u}_i^T\boldsymbol{\Sigma}\mathbf{u}_i}\frac{1}{\mathbf{v}_j^T\boldsymbol{\Sigma}\mathbf{v}_j}\mathrm{tr}\Big(\mathrm{E}(n_1\bar{\mathbf{z}}_1\bar{\mathbf{z}}_1^T|\mathbf{S}_{\mathbf{z}_2})\Gamma^T\mathbf{u}_i\mathbf{u}_i^T\Gamma\Big)\mathrm{tr}\Big(\mathrm{E}(n_2\bar{\mathbf{z}}_2\bar{\mathbf{z}}_2^T|\mathbf{S}_{\mathbf{z}_1})\Gamma^T\mathbf{v}_j\mathbf{v}_j^T\Gamma\Big)\Big), \end{aligned}$$

where $\bar{\mathbf{z}}_1$ and $\bar{\mathbf{z}}_2$ represent the sample means of $\mathbf{z}_i$ in terms of the two partitioned samples in structure (3.6). They correspond to $\mathbf{S}_{\mathbf{z}_1}$ and $\mathbf{S}_{\mathbf{z}_2}$, which are the sample covariance matrices for the two split samples. The two conditional expectations in the last line above are defined as

$$\mathrm{E}(n_1\bar{\mathbf{z}}_1\bar{\mathbf{z}}_1^T|\mathbf{S}_{\mathbf{z}_2}) = c_{11}(\mathbf{S}_{\mathbf{z}_2})\mathbf{I}_p + c_{12}(\mathbf{S}_{\mathbf{z}_2})(\mathbf{1}\mathbf{1}^T - \mathbf{I}_p),$$

$$\mathrm{E}(n_2\bar{\mathbf{z}}_1\bar{\mathbf{z}}_1^T|\mathbf{S}_{\mathbf{z}_2}) = c_{21}(\mathbf{S}_{\mathbf{z}_1})\mathbf{I}_p + c_{22}(\mathbf{S}_{\mathbf{z}_1})(\mathbf{1}\mathbf{1}^T - \mathbf{I}_p),$$

where $\mathbf{1}$ denotes a column vector whose p-dimensional elements are all one: $c_{11}(\mathbf{S}_{\mathbf{z}_2}) = \mathrm{E}(n_1\bar{\mathbf{z}}_{11}^2|\mathbf{S}_{\mathbf{z}_2})$ and $c_{12}(\mathbf{S}_{\mathbf{z}_2}) = \mathrm{E}(n_1\bar{\mathbf{z}}_{11}\bar{\mathbf{z}}_{12}|\mathbf{S}_{\mathbf{z}_2})$ and $c_{21}(\mathbf{S}_{\mathbf{z}_1}) = \mathrm{E}(n_2\bar{\mathbf{z}}_{21}^2|\mathbf{S}_{\mathbf{z}_2})$ and $c_{22}(\mathbf{S}_{\mathbf{z}_1}) = \mathrm{E}(n_2\bar{\mathbf{z}}_{21}\bar{\mathbf{z}}_{22}|\mathbf{S}_{\mathbf{z}_1})$. Among these, $\bar{\mathbf{z}}_{11}$ and $\bar{\mathbf{z}}_{12}$ respectively denote the first and second component elements of $\bar{\mathbf{z}}_1$. Therefore,

$$\begin{aligned}
& \sum_{i,j=1}^{p} \mathrm{E}\Big(\frac{1}{\mathbf{u}_i^T\mathbf{\Sigma}\mathbf{u}_i}\frac{1}{\mathbf{v}_j^T\mathbf{\Sigma}\mathbf{v}_j}\mathrm{tr}\Big(\mathrm{E}(n_1\bar{\mathbf{z}}_1\bar{\mathbf{z}}_1^T|\mathbf{S}_{\mathbf{z}_2})\Gamma^T\mathbf{u}_i\mathbf{u}_i^T\Gamma\Big)\mathrm{tr}\Big(\mathrm{E}(n_2\bar{\mathbf{z}}_2\bar{\mathbf{z}}_2^T|\mathbf{S}_{\mathbf{z}_1})\Gamma^T\mathbf{v}_j\mathbf{v}_j^T\Gamma\Big)\Big) \\
= & \sum_{i,j=1}^{p} \mathrm{E}\Big(\frac{\mathbf{u}_i^T\Gamma\big(c_{11}(\mathbf{S}_{\mathbf{z}_2})\mathbf{I}_p + c_{12}(\mathbf{S}_{\mathbf{z}_2})(\mathbf{1}\mathbf{1}^T - \mathbf{I}_p)\big)\Gamma^T\mathbf{u}_i}{\mathbf{u}_i^T\mathbf{\Sigma}\mathbf{u}_i} \\
& \qquad\qquad \times \frac{\mathbf{v}_j^T\Gamma\big(c_{21}(\mathbf{S}_{\mathbf{z}_1})\mathbf{I}_p + c_{21}(\mathbf{S}_{\mathbf{z}_1})(\mathbf{1}\mathbf{1}^T - \mathbf{I}_p)\big)\Gamma^T\mathbf{v}_j}{\mathbf{v}_j^T\mathbf{\Sigma}\mathbf{v}_j}\Big) \\
= & \sum_{i,j=1}^{p} \mathrm{E}\Big(\frac{\mathbf{u}_i^T\Gamma c_{11}(\mathbf{S}_{\mathbf{z}_2})\mathbf{I}_p\Gamma^T\mathbf{u}_i}{\mathbf{u}_i^T\mathbf{\Sigma}\mathbf{u}_i}\cdot\frac{\mathbf{v}_j^T\Gamma c_{21}(\mathbf{S}_{\mathbf{z}_1})\mathbf{I}_p\Gamma^T\mathbf{v}_j}{\mathbf{v}_j^T\mathbf{\Sigma}\mathbf{v}_j}\Big) \\
= & \sum_{i=1}^{p}\sum_{j=1}^{p}\mathrm{E}\Big(c_{11}(\mathbf{S}_{\mathbf{z}_2})c_{21}(\mathbf{S}_{\mathbf{z}_1})\Big) = \sum_{i=1}^{p}\sum_{j=1}^{p}\mathrm{E}\Big(\mathrm{E}(n_1\bar{\mathbf{z}}_{11}^2|\mathbf{S}_{\mathbf{z}_2})\mathrm{E}(n_2\bar{\mathbf{z}}_{21}^2|\mathbf{S}_{\mathbf{z}_1})\Big) \\
= & \sum_{i=1}^{p}\sum_{j=1}^{p}\mathrm{E}\Big(\mathrm{E}(n_1\bar{\mathbf{z}}_{11}^2|\mathbf{S}_{\mathbf{z}_2})\Big)\mathrm{E}\Big(\mathrm{E}(n_2\bar{\mathbf{z}}_{21}^2|\mathbf{S}_{\mathbf{z}_1})\Big) = p\times p, \qquad \text{(S.4)}
\end{aligned}$$

where the second equation is followed by one fact:

$$
\begin{aligned}
\frac{c_{12}(\mathbf{S}_{\mathbf{z}_2})\mathbf{u}_i^T\Gamma(\mathbf{1}\mathbf{1}^T-\mathbf{I}_p)\Gamma^T\mathbf{u}_i}{\mathbf{u}_i^T\boldsymbol{\Sigma}\mathbf{u}_i} &= \frac{\mathrm{E}(n_1\bar{\mathbf{z}}_{11}\bar{\mathbf{z}}_{12}|\mathbf{S}_{\mathbf{z}_2})\mathbf{u}_i^T\Gamma(\mathbf{1}\mathbf{1}^T-\mathbf{I}_p)\Gamma^T\mathbf{u}_i}{\mathbf{u}_i^T\boldsymbol{\Sigma}\mathbf{u}_i} \\
&= \frac{n_1\mathrm{E}(\bar{\mathbf{z}}_{11})\mathrm{E}(\bar{\mathbf{z}}_{12})\mathbf{u}_i^T\Gamma(\mathbf{1}\mathbf{1}^T-\mathbf{I}_p)\Gamma^T\mathbf{u}_i}{\mathbf{u}_i^T\boldsymbol{\Sigma}\mathbf{u}_i} = 0,
\end{aligned}
$$

and similarly,

$$
\frac{c_{21}(\mathbf{S}_{\mathbf{z}_1})\mathbf{v}_j^T\Gamma(\mathbf{1}\mathbf{1}^T-\mathbf{I}_p)\Gamma^T\mathbf{v}_j}{\mathbf{v}_j^T\boldsymbol{\Sigma}\mathbf{v}_i} = 0.
$$

In summary, putting the result shown in (S.4) into formula (S.3), we obtain

$$
\mathrm{cov}(I_1, I_2) = \sum_{i=1}^{p}\sum_{j=1}^{p}\mathrm{E}\Big(\Big(\frac{\sqrt{n_1}\mathbf{u}_i^T\bar{\mathbf{x}}_1}{\sqrt{\mathbf{u}_i^T\boldsymbol{\Sigma}\mathbf{u}_i}}\Big)^2\Big(\frac{\sqrt{n_2}\mathbf{v}_j^T\bar{\mathbf{x}}_2}{\sqrt{\mathbf{v}_j^T\boldsymbol{\Sigma}\mathbf{v}_j}}\Big)^2\Big) - p\times p = p^2 - p^2 = 0.
$$

Here, we can declare that I_1 and I_2 are uncorrelated terms. The asymptotic variance of the sum of I_1 and I_2 is equal to $\{2\big(\mathrm{tr}(\mathbf{R}_1^2)+\mathrm{tr}(\mathbf{R}_2^2)\big)\}^{1/2}$. Combining the results in equation (S.1), it can be straightforwardly shown that

$$
\frac{t_{\mathrm{o}}^2 - 2p}{\Big\{2\big(\mathrm{tr}(\mathbf{R}_1^2)+\mathrm{tr}(\mathbf{R}_2^2)\big)\Big\}^{1/2}} \xrightarrow{d.} N(0,1).
$$

This completes the proof of Theorem 1. □

S2 Proof of Theorem 2

In this section, we first show the consistency of Lemma 2 to assist the derivation of Theorem 2.

Lemma 2. *Under the structure of random variables in (3.6), when Assumptions 2 and 4 hold, the given p-dimensional projection directions for* $\mathbf{u}_{2j}$ *from sample* *covariance matrix* $\mathbf{S}_2$*, as* n *goes to infinity for* $j = 1, \ldots, p$ *are*

$$\left|\frac{\mathbf{u}_{2j}^T \mathbf{S}_1 \mathbf{u}_{2j}}{\mathbf{u}_{2j}^T \boldsymbol{\Sigma} \mathbf{u}_{2j}} - 1\right| \xrightarrow{p.} 0,$$

In addition, matrix $\mathbf{S}_2$ *also holds this result for given projection direction matrix* U_1*.*

Proof of Lemma 2: First, the spectral decomposition of covariance matrix $\boldsymbol{\Sigma}$ can be written as

$$\boldsymbol{\Sigma} = \mathbf{V}\Lambda\mathbf{V}^T,$$

where $\mathbf{V}$ and Λ are composed of eigenvectors and eigenvalues of $\boldsymbol{\Sigma}$. Define $\bar{\mathbf{x}}_1^* = \mathbf{V}^T\bar{\mathbf{x}}_1$ and$\bar{\mathbf{x}}_2^* = \mathbf{V}^T\bar{\mathbf{x}}_2$, and $U_2^* = \mathbf{V}^T U_2$ and $U_1^* = \mathbf{V}^T U_1$. Based on the test statistic, T_{CP}^2 is constructed as in equation (2.5). Now, it can be expressed in matrix form as

$$\begin{aligned} T_{\text{CP}}^2 &= n_1\bar{\mathbf{x}}_1^T U_2 \widehat{\mathbf{W}}_1^{-1} U_2^T \bar{\mathbf{x}}_1 + n_2\bar{\mathbf{x}}_2^T U_1 \widehat{\mathbf{W}}_2^{-1} U_1^T \bar{\mathbf{x}}_2 \\ &= n_1\bar{\mathbf{x}}_1^T \mathbf{V}\mathbf{V}^T U_2 \widehat{\mathbf{W}}_1^{-1} U_2^T \mathbf{V}\mathbf{V}^T \bar{\mathbf{x}}_1 + n_2\bar{\mathbf{x}}_2^T \mathbf{V}\mathbf{V}^T U_1 \widehat{\mathbf{W}}_2^{-1} U_1^T \mathbf{V}\mathbf{V}^T \bar{\mathbf{x}}_2 \\ &\triangleq n_1\bar{\mathbf{x}}_1^{*T} U_2^* (\widehat{\mathbf{W}}_1^*)^{-1} U_2^{*T} \bar{\mathbf{x}}_1^* + n_2\bar{\mathbf{x}}_2^{*T} U_1^* (\widehat{\mathbf{W}}_2^*)^{-1} U_1^{*T} \bar{\mathbf{x}}_2^*, \end{aligned} \tag{S.5}$$

where $\widehat{\mathbf{W}}_1^* = \text{diag}(\mathbf{u}_{21}^{*T}\mathbf{S}_1^*\mathbf{u}_{21}^*, \ldots, \mathbf{u}_{2p}^{*T}\mathbf{S}_1^*\mathbf{u}_{2p}^*)$ and $\widehat{\mathbf{W}}_2^* = \text{diag}(\mathbf{u}_{11}^{*T}\mathbf{S}_2^*\mathbf{u}_{11}^*, \ldots, \mathbf{u}_{1p}^{*T}\mathbf{S}_2^*\mathbf{u}_{1p}^*)$, in which $\mathbf{u}_{2i} = \mathbf{V}\mathbf{u}_{2i}^*$ and $\mathbf{S}_1^* = \mathbf{V}^T\mathbf{S}_1\mathbf{V}$. Diagonal matrices $\widehat{\mathbf{W}}_1$ and $\widehat{\mathbf{W}}_2$ include the variance components of vectors $\bar{\mathbf{x}}_1^T U_2$ and $\bar{\mathbf{x}}_2^T U_1$, respectively. Thus, vectors $\bar{\mathbf{x}}_1^T U_2$ and $\bar{\mathbf{x}}_2^T U_1$ have reached the standardization effect. At the beginning of the proof, given the projection direction on each data split, the projection variances have the following consistency: $\widehat{\mathbf{W}}^*_{i,(jj)} \xrightarrow{p.} \mathbf{W}^*_{i,(jj)}$, where $\mathbf{W}^*_{i,(jj)} = \operatorname{diag}(\mathbf{u}_{21}^{*T}\Lambda\mathbf{u}^*_{21}, \ldots, \mathbf{u}_{2p}^{*T}\Lambda\mathbf{u}^*_{2p})$ for $j = 1, 2, \ldots, p$, and $i = 1$ and 2. Under Assumption 4, given projection vectors $\mathbf{u}^*_{2i}$ for $i = 1, 2, \ldots, p$, so that $\mathbf{u}_{2i}^{*T}\Lambda\mathbf{u}^*_{2i} > c_0$, it is easy to find that

$$\left|\frac{\mathbf{u}_{2i}^{*T}\mathbf{S}_1^*\mathbf{u}^*_{2i}}{\mathbf{u}_{2i}^{*T}\Lambda\mathbf{u}^*_{2i}} - 1\right| = \left|\frac{\mathbf{u}_{2i}^{*T}\mathbf{S}_1^*\mathbf{u}^*_{2i} - \mathbf{u}_{2i}^{*T}\Lambda\mathbf{u}^*_{2i}}{\mathbf{u}_{2i}^{*T}\Lambda\mathbf{u}^*_{2i}}\right| = \left|\frac{\mathbf{u}_{2i}^{*T}\mathbf{V}^T\mathbf{S}_1\mathbf{V}\mathbf{u}^*_{2i} - \mathbf{u}_{2i}^{*T}\mathbf{V}^T\boldsymbol{\Sigma}\mathbf{V}\mathbf{u}^*_{2i}}{\mathbf{u}_{2i}^{*T}\mathbf{V}^T\boldsymbol{\Sigma}\mathbf{V}\mathbf{u}^*_{2i}}\right|.$$

According to this fact,

$$\begin{aligned}
\mathrm{E}(\mathbf{u}_{2i}^{*T}\mathbf{S}_1^*\mathbf{u}_{2i}) &= \mathbf{u}_{2i}^{*T}\mathbf{V}^T\mathrm{E}\Big(\frac{1}{n_1-1}\sum_{j=1}^{n_1}(\mathbf{x}_j - \bar{\mathbf{x}}_1)(\mathbf{x}_j - \bar{\mathbf{x}}_1)^T\Big)\mathbf{V}\mathbf{u}^*_{2i} \\
&= \mathbf{u}_{2i}^{*T}\mathbf{V}^T\boldsymbol{\Sigma}\mathbf{V}\mathbf{u}^*_{2i} = \mathbf{u}_{2i}^{*T}\Lambda\mathbf{u}^*_{2i},
\end{aligned}$$

so that $\mathrm{E}\left(\left|\frac{\mathbf{u}_{2i}^{*T}\mathbf{S}_1^*\mathbf{u}^*_{2i}}{\mathbf{u}_{2i}^{*T}\Lambda\mathbf{u}^*_{2i}} - 1\right|\right) = 0$. Let $\mathbf{y}_j = \mathbf{V}^T\mathbf{x}_j$. Given $\mathbf{u}_{2i}^{*T}$ for $i = 1, 2, \ldots, p$, $\operatorname{var}\Big(\frac{\mathbf{u}_{2i}^{*T}\mathbf{y}_j}{\sqrt{\mathbf{u}_{2i}^{*T}\Lambda\mathbf{u}^*_{2i}}}\Big) = 1$ and

$$\begin{aligned}
\frac{\mathbf{u}_{2i}^{*T}\mathbf{S}_1^*\mathbf{u}^*_{2i}}{\mathbf{u}_{2i}^{*T}\Lambda\mathbf{u}^*_{2i}} &= \frac{1}{n_1-1}\sum_{j=1}^{n_1}\left(\frac{\mathbf{u}_{2i}^{*T}(\mathbf{y}_j - \bar{\mathbf{y}}_j)}{\sqrt{\mathbf{u}_{2i}^{*T}\Lambda\mathbf{u}^*_{2i}}}\right)^2 \\
&= \frac{1}{n_1-1}\sum_{j=1}^{n_1}\left(\frac{\mathbf{u}_{2i}^{*T}\mathbf{y}_j}{\sqrt{\mathbf{u}_{2i}^{*T}\Lambda\mathbf{u}^*_{2i}}}\right)^2 - \frac{n_1}{n_1-1}\left(\frac{\mathbf{u}_{2i}^{*T}\bar{\mathbf{y}}}{\mathbf{u}_{2i}^{*T}\Lambda\mathbf{u}^*_{2i}}\right)^2,
\end{aligned}$$

where $\bar{\mathbf{y}} = \frac{1}{n_1}\sum_{j=1}^{n_1}\mathbf{y}_j$. Suppose that Assumption 2 holds, we have by the law of large numbers that

$$\frac{\mathbf{u}_{2i}^{*T}\mathbf{S}_1^*\mathbf{u}^*_{2i}}{\mathbf{u}_{2i}^{*T}\Lambda\mathbf{u}^*_{2i}} \xrightarrow{p.} 1$$

as $n \to \infty$. Similarly, the consistency can be easily obtained as follows:

$$\frac{\mathbf{u}_{1i}^{*T}\mathbf{S}_2^*\mathbf{u}^*_{2i}}{\mathbf{u}_{1i}^{*T}\Lambda\mathbf{u}^*_{1i}} \xrightarrow{p.} 1$$

as $n \to \infty$, for given projection directions $\mathbf{u}_{1i}$'s, $1 \le i \le p$. Even though the eigenvalues of Λ vary with n_i in a sequence, as long as a n_i is given, our consistency can be obtained through the law of large numbers. Combining the transformation results of (S.5) and the proof conclusions of the consistency property mentioned above, we obtain the following consistency properties:

$$\frac{\mathbf{u}_{2i}^T \mathbf{S}_1 \mathbf{u}_{2i}}{\mathbf{u}_{2i}^T \boldsymbol{\Sigma} \mathbf{u}_{2i}} \xrightarrow{p.} 1 \quad \text{and} \quad \frac{\mathbf{u}_{1i}^T \mathbf{S}_2 \mathbf{u}_{2i}}{\mathbf{u}_{1i}^T \boldsymbol{\Sigma} \mathbf{u}_{1i}} \xrightarrow{p.} 1.$$

This completes the proof of Lemma 2. □

Proof of Theorem2: It can be seen from the expression of equation (S.5) that test statistic T_{CP}^2 can also be written as

$$\begin{aligned} T_{\mathrm{CP}}^2 &= \sum_{i=1}^{p} \left(\frac{\sqrt{n_1} \mathbf{u}_{2i}^T \bar{\mathbf{x}}_1}{\sqrt{\mathbf{u}_{2i}^T \mathbf{S}_1 \mathbf{u}_{2i}}} \right)^2 + \sum_{j=1}^{p} \left(\frac{\sqrt{n_2} \mathbf{u}_{1j}^T \bar{\mathbf{x}}_2}{\sqrt{\mathbf{u}_{1j}^T \mathbf{S}_2 \mathbf{u}_{1j}}} \right)^2 \\ &= \sum_{i=1}^{p} \left(\frac{\sqrt{n_1} \mathbf{u}_{2i}^{*T} \bar{\mathbf{x}}_1^*}{\sqrt{\mathbf{u}_{2i}^{*T} \mathbf{S}_1^* \mathbf{u}_{2i}^*}} \right)^2 + \sum_{j=1}^{p} \left(\frac{\sqrt{n_2} \mathbf{u}_{1j}^{*T} \bar{\mathbf{x}}_2^*}{\sqrt{\mathbf{u}_{1j}^{*T} \mathbf{S}_2^* \mathbf{u}_{1j}^*}} \right)^2 \\ &=: I_3 + I_4. \end{aligned} \tag{S.6}$$

For independent components structure (3.6), covariance matrix $\boldsymbol{\Sigma} = \Gamma\Gamma^T$ is assumed to be positive definite. Define $\delta_j = \frac{1}{p}\mathrm{tr}(\boldsymbol{\Sigma}^j)$, for $j = 1, 2, 3, 4$, when the limitation of δ_j's exists. That is,

$$0 < \lim_{p\to\infty} \delta_j = \delta_{j0} < \infty. \tag{S.7}$$

Let $\hat{\delta}_1 = \frac{1}{p}\mathrm{tr}(\mathbf{S}_n)$ and $\hat{\delta}_2 = \frac{1}{p}[\mathrm{tr}(\mathbf{S}_n^2) - \frac{1}{n-1}(\mathrm{tr}(\mathbf{S}_n))^2]$. Srivastava (2009) proved that $\hat{\delta}_1$ and $\hat{\delta}_2$ are consistent estimators of δ_1 and δ_2 as $(n, p) \to \infty$ in Theorem 2.2. In our framework, when Assumptions 1–3 hold, it implies that the result of equation (S.2) holds. It can be shown that for $i = 1$ and 2, $\frac{1}{p}(\mathrm{tr}(\widehat{\mathbf{R}}_i^2) - p^2/(n_i - 1))$ is a consistent estimator of $\frac{1}{p}\mathrm{tr}(\mathbf{R}_i^2)$ as $(n, p) \to \infty$ for the case $(n-1) = O(p^\tau)$, $0 < \tau \leq 1$, which was proved by Srivastava and Du (2008) in Lemma 3.2. According to the asymptotic normality of Theorem 1, and if Assumptions 1–3 hold, it can be seen that the expression of I_3 and I_4 have the following asymptotic normality distribution:

$$\frac{n_1\bar{\mathbf{x}}_1^T U_2 \widehat{\mathbf{W}}_1^{-1} U_2^T \bar{\mathbf{x}}_1 - p(\frac{n_1-1}{n_1-3})}{\{2\big(\mathrm{tr}(\widehat{\mathbf{R}}_1^2) - \frac{p^2}{(n_1-1)}\big)\}^{1/2}} \xrightarrow{d.} N(0,1), \tag{S.8}$$

and

$$\frac{n_2\bar{\mathbf{x}}_2^T U_1 \widehat{\mathbf{W}}_2^{-1} U_1^T \bar{\mathbf{x}}_2 - p(\frac{n_2-1}{n_2-3})}{\{2\big(\mathrm{tr}(\widehat{\mathbf{R}}_2^2) - \frac{p^2}{(n_2-1)}\big)\}^{1/2}} \xrightarrow{d.} N(0,1). \tag{S.9}$$

It should be pointed out that the expectations of both I_3 and I_4 are respectively obtained by obeying p independent $F(1, n_i - 1)$ distribution with 1 and $n_i - 1$ degrees of freedom under normal distribution. However, regardless of distribution, both $p(\frac{n_1-1}{n_1-3})$ and $p(\frac{n_2-1}{n_2-3})$ converge to p when Assumption 2 holds. Therefore, their limit distributions are unchanged. In real world application, we still adopt the expression of Theorem 2 to further correct the bias of test statistics I_3 and I_4.

To obtain the asymptotic normality property of test statistic T^2_{CP}, based on equations (S.6)–(S.8), we only need to prove that terms I_3 and I_4 are asymptotically irrelevant or irrelevant. According to the conclusions of Lemma 2, it follows that

$$\frac{\mathbf{u}_{2i}^T\mathbf{S}_1\mathbf{u}_{2i}}{\mathbf{u}_{2i}^T\boldsymbol{\Sigma}\mathbf{u}_{2i}} \xrightarrow{p.} 1 \quad \text{and} \quad \frac{\mathbf{u}_{1i}^T\mathbf{S}_2\mathbf{u}_{2i}}{\mathbf{u}_{1i}^T\boldsymbol{\Sigma}\mathbf{u}_{1i}} \xrightarrow{p.} 1.$$

Furthermore, the cross test statistic, T_{CP}^2, yields the following result:

$$\begin{aligned} T_{\mathrm{CP}}^2 &= \sum_{i=1}^{p}\left\{\left(\frac{\sqrt{n_1}\mathbf{u}_{2i}^T\bar{\mathbf{x}}_1}{\sqrt{\mathbf{u}_{2i}^T\mathbf{S}_1\mathbf{u}_{2i}}}\right)^2+\left(\frac{\sqrt{n_2}\mathbf{u}_{1i}^T\bar{\mathbf{x}}_2}{\sqrt{\mathbf{u}_{1i}^T\mathbf{S}_2\mathbf{u}_{1i}}}\right)^2\right\} \\ &\xrightarrow{p.} \sum_{i=1}^{p}\left\{\left(\frac{\sqrt{n_1}\mathbf{u}_{2i}^T\bar{\mathbf{x}}_1}{\sqrt{\mathbf{u}_{2i}^T\boldsymbol{\Sigma}\mathbf{u}_{2i}}}\right)^2+\left(\frac{\sqrt{n_2}\mathbf{u}_{1i}^T\bar{\mathbf{x}}_2}{\sqrt{\mathbf{u}_{1i}^T\boldsymbol{\Sigma}\mathbf{u}_{1i}}}\right)^2\right\} \\ &= I_1 + I_2. \end{aligned} \tag{S.10}$$

To obtain the asymptotic normality of test statistic T_{CP}^2, when I_3 and I_4 have asymptotic normality, they can be translated using the uncorrelated property between I_1 and I_2 because I_3 and I_4 converge with probabilities I_1 and I_2, respectively. According to the consistency property of T_{CP}^2 in (S.10), and because I_1 and I_2 are uncorrelated in Theorem 1, it is easy to find that from the asymptotic normality shown in (S.8) and (S.9), the cross test statistic T_{CP}^2 follows the asymptotic normality

$$\frac{T_{\mathrm{CP}}^2 - p(\frac{n_1-1}{n_1-3}) - p(\frac{n_2-1}{n_2-3})}{\{2\big(\mathrm{tr}(\widehat{\mathbf{R}}_1^2) + \mathrm{tr}(\widehat{\mathbf{R}}_2^2) - \frac{p^2}{n_1-1} - \frac{p^2}{n_2-1}\big)\}^{1/2}} \xrightarrow{d.} N(0,1),$$

where $\widehat{\mathbf{R}}_1$ and $\widehat{\mathbf{R}}_2$ are the sample correlation matrix of projection samples $U_2^T\boldsymbol{X}_1$ and $U_1^T\boldsymbol{X}_2$ with $\boldsymbol{X}_1 = (\mathbf{x}_1, \ldots, \mathbf{x}_{n_1})$ and $\boldsymbol{X}_2 = (\mathbf{x}_{n_1+1}, \ldots, \mathbf{x}_n)$, respectively. This completes the proof of Theorem 2. □

S3 Proof of Theorem 3

Proof of Theorem 3: Combining the conclusion of Theorem 2 and the content of Theorem 2.1 in Srivastava (2009), when mean vector $\boldsymbol{\mu} \neq \mathbf{0}$, as $n, p \to \infty$, for random variables $\mathbf{x}_i$, we easily see that

$$\frac{n_1(\bar{\mathbf{x}}_1 - \boldsymbol{\mu})^T U_2 \widehat{\mathbf{W}}_1^{-1} U_2^T (\bar{\mathbf{x}}_1 - \boldsymbol{\mu}) + n_2(\bar{\mathbf{x}}_2 - \boldsymbol{\mu})^T U_1 \widehat{\mathbf{W}}_2^{-1} U_1^T (\bar{\mathbf{x}}_2 - \boldsymbol{\mu}) - 2p}{\sqrt{2(\mathrm{tr}(\mathbf{R}_1^2) + \mathrm{tr}(\mathbf{R}_2^2))}} \quad \text{(S.11)}$$

has a standard normal distribution, N(0,1). For the local alternative setting $\boldsymbol{\mu} = \left(\frac{1}{n(n-1)}\right)^{\frac{1}{2}} \boldsymbol{\delta}$,

$$\begin{aligned}
& \frac{1}{\sqrt{p}} \{ n_1(\bar{\mathbf{x}}_1 - \boldsymbol{\mu})^T U_2 \widehat{\mathbf{W}}_1^{-1} U_2^T (\bar{\mathbf{x}}_1 - \boldsymbol{\mu}) + n_2(\bar{\mathbf{x}}_2 - \boldsymbol{\mu})^T U_1 \widehat{\mathbf{W}}_2^{-1} U_1^T (\bar{\mathbf{x}}_2 - \boldsymbol{\mu}) \} \\
= & \frac{1}{\sqrt{p}} (n_1 \bar{\mathbf{x}}_1^T U_2 \widehat{\mathbf{W}}_1^{-1} U_2^T \bar{\mathbf{x}}_1 + n_2 \bar{\mathbf{x}}_2^T U_1 \widehat{\mathbf{W}}_2^{-1} U_1^T \bar{\mathbf{x}}_2) - \frac{2n_1}{\sqrt{pn(n-1)}} \boldsymbol{\delta}^T U_2 \widehat{\mathbf{W}}_1^{-1} U_2^T \bar{\mathbf{x}}_1 \\
& - \frac{2n_2}{\sqrt{pn(n-1)}} \boldsymbol{\delta}^T U_1 \widehat{\mathbf{W}}_2^{-1} U_1^T \bar{\mathbf{x}}_2 + \frac{n_1}{n(n-1)\sqrt{p}} \boldsymbol{\delta}^T U_2 \widehat{\mathbf{W}}_1^{-1} U_2^T \boldsymbol{\delta} \\
& + \frac{n_2}{n(n-1)\sqrt{p}} \boldsymbol{\delta}^T U_1 \widehat{\mathbf{W}}_2^{-1} U_1^T \boldsymbol{\delta}.
\end{aligned}$$

Combining the conditions of (3.7) and (3.8), because $\bar{\mathbf{x}}_{i,j} \xrightarrow{p.} \mu_j = (1/\{n(n-1)\})^{\frac{1}{2}} \boldsymbol{\delta}_j$ as $n \to \infty$ and $\widehat{\mathbf{W}}_{i(jj)} \xrightarrow{p.} \mathbf{W}_{i(jj)}$ for $i = 1$ and 2, it follows that

$$\frac{n_1}{\sqrt{pn(n-1)}} \boldsymbol{\delta}^T U_2 \widehat{\mathbf{W}}_1^{-1} U_2^T \bar{\mathbf{x}}_1 - n_1/(n(n-1)\sqrt{p}) \boldsymbol{\delta}^T U_2 \mathbf{W}_1^{-1} U_2^T \boldsymbol{\delta} \xrightarrow{p.} 0,$$

and

$$\frac{n_2}{\sqrt{pn(n-1)}} \boldsymbol{\delta}^T U_1 \widehat{\mathbf{W}}_2^{-1} U_1^T \bar{\mathbf{x}}_2 - n_2/(n(n-1)\sqrt{p}) \boldsymbol{\delta}^T U_1 \mathbf{W}_2^{-1} U_1^T \boldsymbol{\delta} \xrightarrow{p.} 0.$$

The above conclusion is true only if the following facts are proved:

$$\frac{n_1}{\sqrt{pn(n-1)}}\boldsymbol{\delta}^T U_2 \widehat{\mathbf{W}}_1^{-1} U_2^T \bar{\mathbf{x}}_1 - \frac{n_1}{\sqrt{pn(n-1)}}\boldsymbol{\delta}^T U_2 \widehat{\mathbf{W}}_1^{-1} U_2^T \boldsymbol{\mu} \xrightarrow{p.} 0. \qquad \text{(S.12)}$$

When finite fourth moments exists for random variables, according to the conclusion of $\widehat{\mathbf{W}}_{i(jj)} \xrightarrow{p.} \mathbf{W}_{i(jj)}$, the convergence rate of variance of $\frac{n_1}{\sqrt{pn(n-1)}}(\boldsymbol{\delta}^T U_2 \widehat{\mathbf{W}}_1^{-1} U_2^T \bar{\mathbf{x}}_1)$ can be obtained by the consistency term $\frac{n_1}{\sqrt{pn(n-1)}}(\boldsymbol{\delta}^T U_2 \mathbf{W}_1^{-1} U_2^T \bar{\mathbf{x}}_1)$. By the condition of (3.7), the variance of

$$\text{var}\Big(\frac{n_1}{\sqrt{pn(n-1)}}(\boldsymbol{\delta}^T U_2 \mathbf{W}_1^{-1} U_2^T \bar{\mathbf{x}}_1)|U_2\Big) = \frac{n_1}{n(n-1)}\frac{1}{p}(\boldsymbol{\delta}^T U_2 \mathbf{W}_1^{-1} U_2^T \boldsymbol{\delta}),$$

and

$$\text{var}\Big(\frac{n_2}{\sqrt{pn(n-1)}}(\boldsymbol{\delta}^T U_1 \mathbf{W}_2^{-1} U_1^T \bar{\mathbf{x}}_2)|U_1\Big) = \frac{n_2}{n(n-1)}\frac{1}{p}(\boldsymbol{\delta}^T U_1 \mathbf{W}_2^{-1} U_1^T \boldsymbol{\delta}).$$

By assumption 2, both of the above variances tend to 0 as $n \to \infty$. Thus, the result in (S.12) is obviously established. Therefore,

$$\begin{aligned} &\frac{1}{\sqrt{p}}\{n_1(\bar{\mathbf{x}}_1 - \boldsymbol{\mu})^T U_2 \widehat{\mathbf{W}}_1^{-1} U_2^T (\bar{\mathbf{x}}_1 - \boldsymbol{\mu}) + n_2(\bar{\mathbf{x}}_2 - \boldsymbol{\mu})^T U_1 \widehat{\mathbf{W}}_2^{-1} U_1^T (\bar{\mathbf{x}}_2 - \boldsymbol{\mu}))\} \\ - &\frac{1}{\sqrt{p}}\{T_{\text{CP}}^2 - \frac{1}{n-1}(k\boldsymbol{\delta}^T U_2 \mathbf{W}_1^{-1} U_2^T \boldsymbol{\delta} + (1-k)\boldsymbol{\delta}^T U_1 \mathbf{W}_2^{-1} U_1^T \boldsymbol{\delta})\} \xrightarrow{p.} 0. \end{aligned}$$

Define $\Delta(\boldsymbol{\delta}; n, p) = \frac{1}{n-1}(k\boldsymbol{\delta}^T U_2 \mathbf{W}_1^{-1} U_2^T \boldsymbol{\delta} + (1-k)\boldsymbol{\delta}^T U_1 \mathbf{W}_2^{-1} U_1^T \boldsymbol{\delta})$. By the asymptotic normality result in equation (S.11), as $(n, p) \to \infty$, we easily have

$$\frac{T_{\text{CP}}^2 - \Delta(\boldsymbol{\delta}; n, p) - 2p}{\sqrt{2(\text{tr}(\mathbf{R}_1^2) + \text{tr}(\mathbf{R}_2^2))}} \xrightarrow{d.} N(0, 1).$$

Thus, we have that under local alterative $\boldsymbol{\mu} = \{\frac{1}{(n(n-1))}\}^{\frac{1}{2}}\boldsymbol{\delta}$, the conditions of Theorem 1, and $\frac{1}{p}(\boldsymbol{\delta}^T U_2 \mathbf{W}_1^{-1} U_2^T \boldsymbol{\delta} + \boldsymbol{\delta}^T U_1 \mathbf{W}_2^{-1} U_1^T \boldsymbol{\delta}) \leq C$,

$$\begin{aligned}
&\lim_{(n,p)\to\infty} P(T^2 > z_{1-\alpha}|U_1, U_2) \\
= &\lim_{(n,p)\to\infty} P\left(\frac{T_{\mathrm{CP}}^2 - 2p}{\sqrt{2(\mathrm{tr}(\mathbf{R}_1^2) + \mathrm{tr}(\mathbf{R}_2^2))}} > z_{1-\alpha}|U_1, U_2\right) \\
= &\lim_{(n,p)\to\infty} P\left(\frac{T_{\mathrm{CP}}^2 - \Delta(\boldsymbol{\delta}; n, p) - 2p}{\sqrt{2(\mathrm{tr}(\mathbf{R}_1^2) + \mathrm{tr}(\mathbf{R}_2^2))}} > z_{1-\alpha} - \frac{\Delta(n,p)}{\sqrt{2(\mathrm{tr}(\mathbf{R}_1^2) + \mathrm{tr}(\mathbf{R}_2^2))}}\right) \\
= &\lim_{(n,p)\to\infty} \Phi\left(-z_{1-\alpha} + \frac{\Delta(\boldsymbol{\delta}; n, p)}{\sqrt{2(\mathrm{tr}(\mathbf{R}_1^2) + \mathrm{tr}(\mathbf{R}_2^2))}}\right).
\end{aligned}$$

This completes the proof of Theorem 3.

S4 Proof of Theorem 4

Proof of Theorem 4: Define two events, ε_1 and ε_2, as

$$\varepsilon_1 = \left\{\max_{1\leq j\leq p} |\bar{x}_j - \mu_j|/s_{jj}^{1/2} < \delta_{n,p}/\sqrt{n}\right\}$$

and

$$\varepsilon_2 = \left\{\frac{4}{9} \leq s_{jj}/\sigma_{jj} \leq \frac{9}{4}, \forall j = 1, 2, \ldots, p\right\}.$$

For any $j \in \mathcal{S}(\boldsymbol{\mu})$, it follows that $|\mu_j| > 3\sigma_{jj}^{1/2}\delta_{n,p}/\sqrt{n}$ by the definition of $\mathcal{S}(\boldsymbol{\mu})$. Then, under event $\varepsilon_1 \cap \varepsilon_2$,

$$\frac{|\bar{x}_j|}{s_{jj}^{1/2}} \geq \frac{|\mu_j| - |\bar{x}_j - \mu_j|}{s_{jj}^{1/2}} \geq \frac{2\mu_j}{3\sigma_{jj}^{1/2}} - \delta_{n,p}/\sqrt{n} > \delta_{n,p}/\sqrt{n}.$$

This implies that $j \in \widehat{\mathcal{S}}$. Hence, $\mathcal{S}(\boldsymbol{\mu}) \subset \widehat{\mathcal{S}}$. In fact, we have proved this statement on event $\varepsilon_1 \cap \varepsilon_2$ uniformly for $\boldsymbol{\mu} \in \mathcal{U}$:

$$\inf_{\boldsymbol{\mu} \in \mathcal{U}} P(\mathcal{S}(\boldsymbol{\mu}) \subset \widehat{\mathcal{S}} | \boldsymbol{\mu}) \to 1.$$

Furthermore, under the null hypothesis (H_0), by Assumption 3,

$$P(J_0 = 0 | H_0) = P(\widehat{\mathcal{S}} = \emptyset | H_0) = P(\max_{1 \le j \le p} \{|\bar{x}_j| / s_{jj}^{1/2}\} < \delta_{n,p} / \sqrt{n} | H_0) \to 1.$$

In addition, by $\inf_{\boldsymbol{\mu} \in \mathcal{U}} P(\mathcal{S}(\boldsymbol{\mu}) \subset \widehat{\mathcal{S}} | \boldsymbol{\mu}) \to 1$,

$$\begin{aligned} & \sup_{\boldsymbol{\mu} \in \mathcal{U}} P(J_0 \le n | \mathcal{S}(\boldsymbol{\mu}) \ne \emptyset) \\ \le \; & \sup_{\boldsymbol{\mu} \in \mathcal{U}} P(J_0 \le n, \widehat{\mathcal{S}} \ne \emptyset | \mathcal{S}(\boldsymbol{\mu}) \ne \emptyset) + \sup_{\boldsymbol{\mu} \in \mathcal{U}} P(\widehat{\mathcal{S}} = \emptyset | \mathcal{S}(\boldsymbol{\mu}) \ne \emptyset) \\ \le \; & \sup_{\boldsymbol{\mu} \in \mathcal{U}} P\big(n \cdot \mathbf{1}\{\max_{1 \le j \le p} (|\bar{x}_j| / s_{jj}^{1/2}) > \delta_{n,p} / \sqrt{n}\} \le n, \widehat{\mathcal{S}} \ne \emptyset | \mathcal{S}(\boldsymbol{\mu}) \ne \emptyset\big) + o(1) \to 0. \end{aligned}$$

Therefore, $\inf_{\boldsymbol{\mu} \in \mathcal{U}} P(J_0 > n | \mathcal{S}(\boldsymbol{\mu}) \ne \emptyset) \to 1$. This completes the proof of Theorem 4. □

S5 Proof of Theorem 5

Proof of Theorem 5: According to the result in Theorem 4, $P(J_0 = 0 | H_0) \to 1$. This implies that $J = J_{\text{CPT}} + J_0 \xrightarrow{d.} N(0,1)$ under the null hypothesis (H_0). Hence, one must only prove that $\inf_{\boldsymbol{\mu} \in \mathcal{U}_s} P(J_1 \ge z_\alpha | \boldsymbol{\mu}) \to 1$. By the definitions of J_0 and $\widehat{\mathcal{S}}$, these two events are equivalent. That is, $\{J_0 < n\} = \{\widehat{\mathcal{S}} = \emptyset\}$. Because

$\inf_{\boldsymbol{\mu}\in\mathcal{U}} P(\mathcal{S}(\boldsymbol{\mu}) \subset \widehat{\mathcal{S}}|\boldsymbol{\mu}) \to 1$ and $\mathcal{U}_s = \{\boldsymbol{\mu} : \mathcal{S}(\boldsymbol{\mu}) \neq \emptyset\}$,

$$\sup_{\boldsymbol{\mu}\in\mathcal{U}_s} P(J_0 < n|\boldsymbol{\mu}) = \sup_{\boldsymbol{\mu}\in\mathcal{U}_s} P(\widehat{\mathcal{S}} = \emptyset|\boldsymbol{\mu}) \leq \sup_{\{\boldsymbol{\mu}:\mathcal{S}(\boldsymbol{\mu})\neq\emptyset\}} P(\widehat{\mathcal{S}} = \emptyset, \mathcal{S}(\boldsymbol{\mu}) \subset \widehat{\mathcal{S}}|\boldsymbol{\mu}) + o(1).$$

It can be obviously found that the first term of the last inequality is zero, so $\sup_{\boldsymbol{\mu}\in\mathcal{U}_s} P(J_0 \geq n|\boldsymbol{\mu}) \to 1$. Hence, as $n \to \infty$,

$$\inf_{\boldsymbol{\mu}\in\mathcal{U}_s} P(J > z_\alpha|\boldsymbol{\mu}) \geq \inf_{\boldsymbol{\mu}\in\mathcal{U}_s} P(n + J_{\mathrm{CPT}} > z_\alpha) \to 1,$$

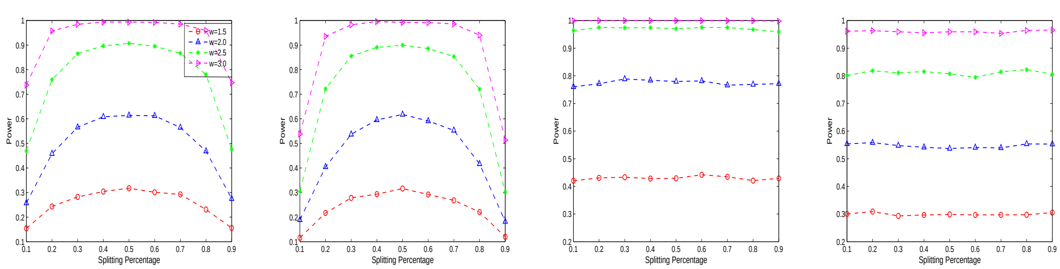
which completes the proof content of Theorem 5. □

S6 Presentation of additional simulation results

Many of the simulation results are listed for reference in this section to avoid redundant text and to help the reader understand the article. These include the exploration of reasonable split percentages for three distributions (Examples (a)–(c)), the comparison of empirical and theoretical power, and the simulation results for both the dense and sparse mean tests.

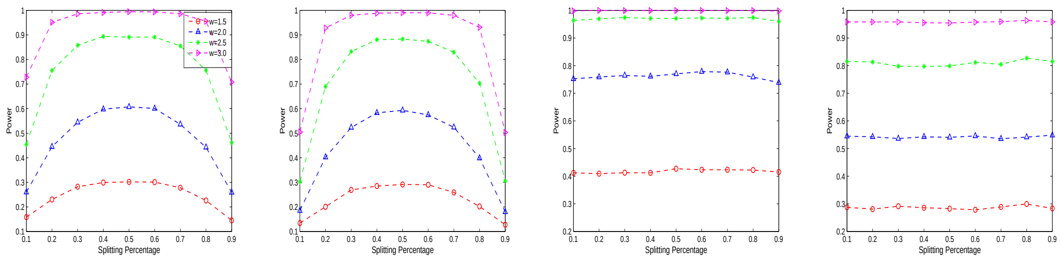
S6.1 Reasonable splitting percentage

Define the splitting percentage for the two-group sample as ς; thus, $n_1 = [n \cdot \varsigma]$ and $n_2 = n - n_1$, where $[x]$ means rounding x to the nearest integer. In this section, we explore this trade-off in simulations by taking a range of ς over (0, 1): $10\%, 20\%, \ldots, 90\%$, and we compare the power of each grid value. It should be noted that when $\varsigma = 10\%$, the empirical power of the CPT is similar to that of $\varsigma = 90\%$ because the CPT is a summation of two statistics, T_1^2 and T_2^2, in equation (2.5), which are obtained by cross projection. The mean vector is set to $\boldsymbol{\mu} = (w/25 * \mathbf{1}_{0.3p}, \mathbf{0}_{0.7p})^T$ throughout in this exploration. Figures 1–3 show the empirical power curves with setting $(n, p) = (150, 300)$, in which w are drawn with the values of 1.5, 2.0, 2.5, and 3.0, respectively. The optimal splitting percentages vary for many simulations, with most peaks occurring at a grid value of $40\% - 60\%$ in factor model structures, $\boldsymbol{\Sigma}_1$ and $\boldsymbol{\Sigma}_2$, while a splitting percentage of the covariance in the range of $20\% - 80\%$ for the remaining two structures is acceptable. It is difficult in practical application to choose an optimal splitting percentage that performs consistently because of the unknown covariance structure. Therefore, we suggest that $40\% - 60\%$ is a reasonable range in our projection framework.



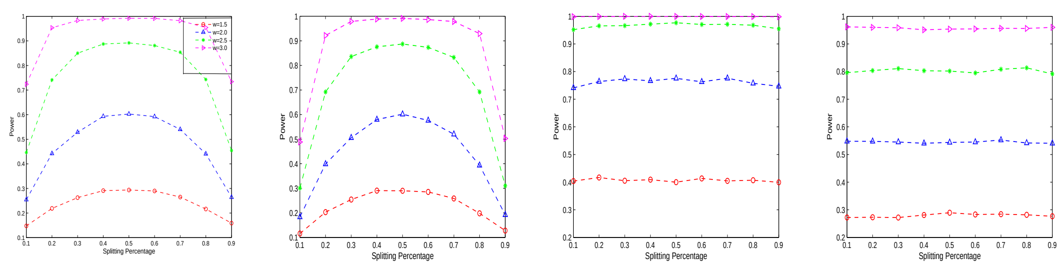
(A) Example(a), Σ_1 (B) Example(a), Σ_2 (C) Example(a), Σ_3 (D) Example(a), Σ_4

Figure 1: Empirical power under multivariate normal data changes with the splitting percentage.



(A) Example(b), Σ_1 (B) Example(b), Σ_2 (C) Example(b), Σ_3 (D) Example(b), Σ_4

Figure 2: Empirical power when multivariate student t data change with the splitting percentage.



(A) Example(c), Σ_1 (B) Example(c), Σ_2 (C) Example(c), Σ_3 (D) Example(c), Σ_4

Figure 3: Empirical power when the multivariate chi-square data change with the splitting percentage.

S6.2 Comparison of empirical and theoretical power

In this subsection, we illustrate that the performance of the empirical power is close to the theoretical power calculated by Theorem 3. The mean vector is set to $\boldsymbol{\mu} = (w/30 * \mathbf{1}_{0.5p}, \mathbf{0}_{0.5p})^T$ throughout in this presentation. Figures 4–5 show the curves of the empirical power and the approximated theoretical power with the setting $(n, p) = (250, 300)$, where w is plotted with values from 1.0 to 1.8. According to Theorem 3, the asymptotic power of standardized CPT statistic T^2

as $(n,p)\to\infty$ is given by

$$\boldsymbol{\beta}(T^2|\boldsymbol{\delta}) \simeq \mathrm{E}_{U_1,U_2}\Big(\Phi\Big(-z_{1-\alpha}+\frac{\Delta(\boldsymbol{\delta};n,p)}{\sqrt{2(\mathrm{tr}(\mathbf{R}_1^2)+\mathrm{tr}(\mathbf{R}_2^2))}}\Big)\Big).$$

Let $\tilde{\boldsymbol{\beta}}(T^2|\boldsymbol{\delta})$ be an approximation of $\mathrm{E}_{U_1,U_2}\Big(\Phi\Big(-z_{1-\alpha}+\frac{\Delta(\boldsymbol{\delta};n,p)}{\sqrt{2(\mathrm{tr}(\mathbf{R}_1^2)+\mathrm{tr}(\mathbf{R}_2^2))}}\Big)\Big)$, where

$$\tilde{\boldsymbol{\beta}}(T^2|\boldsymbol{\delta}) = \frac{1}{m}\sum_{i=1}^{m}\Phi\Big(-z_{1-\alpha}+\frac{\Delta^{(i)}(\boldsymbol{\delta};n,p)}{\sqrt{2\big(\mathrm{tr}(\mathbf{R}_1^{(i)})^2+\mathrm{tr}(\mathbf{R}_1^{(i)})^2\big)}}\Big)$$

with $\Delta^{(i)}(\boldsymbol{\delta};n,p) = \frac{1}{n-1}\big(k\boldsymbol{\delta}^T U_2^{(i)}\mathbf{W}_1^{(i)-1}U_2^{(i)T}\boldsymbol{\delta}+(1-k)\boldsymbol{\delta}^T U_1^{(i)}\mathbf{W}_2^{(i)-1}U_1^{(i)T}\boldsymbol{\delta}\big)$ and $\mathbf{W}_1^{(i)}$, $\mathbf{W}_2^{(i)}$, $\mathbf{R}_1^{(i)}$ and $\mathbf{R}_2^{(i)}$ are obtained by replacing U_1 and U_2 with $U_1^{(i)}$ and $U_2^{(i)}$ in the definitions of $\mathbf{W}_1$, $\mathbf{W}_2$, $\mathbf{R}_1$ and $\mathbf{R}_2$, respectively, where the projection directions $U_1^{(i)}$ and $U_2^{(i)}$ are the eigenvectors of the sample covariance matrix of i-th iteration.

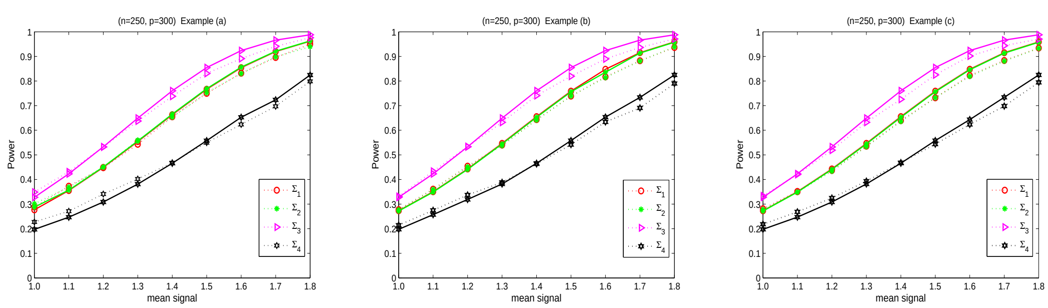


Figure 4: The dashed and solid lines represent the fitted plots of the empirical power and the approximated theoretical power ($\tilde{\boldsymbol{\beta}}(T^2|\boldsymbol{\delta})$) of CPT with increasing signal strength, respectively.

It can observe from Figure 4 that the empirical power is very close to the approximated theoretical power calculated by Theorem 3. Of course, with a sufficiently large sample size and dimensionality, the convergence of our CPT to the asymptotic normality under the local alternative would be better. Therefore, it is normal to have a slight difference on a few points in the fitted plots in Figure 4.

S6.3 Simulation results for dense and sparse mean tests

We first show the performance of CPT in terms of empirical size and power on dense mean settings for the multivariate student t and multivariate normal chi-square distributions in Tables S1–S2, respectively. It can be seen from Tables S1–S2 that the performance of our proposed CPT is similar to that of Table 1 under the normal distribution.

Table S1: Empirical size and power (%) of test statistics (Example(b), nominal $\alpha = 0.05$)

		Size					Dense mean $w = 2$					Dense mean $w = 3$				
	Type	$T^2_{\rm CP}$	$T^2_{\rm OP}$	$T^2_{\rm BS}$	$T^2_{\rm S}$	$T^2_{\rm CQ}$	$T^2_{\rm CP}$	$T^2_{\rm OP}$	$T^2_{\rm BS}$	$T^2_{\rm S}$	$T^2_{\rm CQ}$	$T^2_{\rm CP}$	$T^2_{\rm OP}$	$T^2_{\rm BS}$	$T^2_{\rm S}$	$T^2_{\rm CQ}$
	$\mathbf{\Sigma}_1$	5.7	5.5	7.0	7.2	5.1	**93.3**	74.4	14.2	17.1	9.0	**100.0**	100.0	71.5	78.5	42.9
$n = 200$	$\mathbf{\Sigma}_2$	6.4	5.4	7.1	7.0	5.3	**92.6**	71.6	8.3	9.0	5.3	**100.0**	100.0	11.8	15.4	8.9
$p = 250$	$\mathbf{\Sigma}_3$	5.9	5.4	5.7	5.7	5.1	96.9	79.7	99.9	99.9	99.8	**100.0**	100.0	100.0	100.0	100.0
	$\mathbf{\Sigma}_4$	6.1	5.0	5.2	5.5	4.6	78.5	36.4	95.9	96.4	92.9	**100.0**	99.4	100.0	100.0	100.0
	$\mathbf{\Sigma}_1$	6.0	5.1	7.1	7.5	4.4	**97.5**	91.5	16.6	20.0	11.5	**100.0**	100.0	91.1	99.3	79.2
$n = 200$	$\mathbf{\Sigma}_2$	6.0	4.6	6.8	7.0	4.2	**98.1**	89.7	8.4	9.1	5.9	**100.0**	100.0	11.8	15.4	8.3
$p = 350$	$\mathbf{\Sigma}_3$	6.2	4.9	5.1	5.5	4.3	99.5	96.4	100.0	100.0	100.0	**100.0**	100.0	100.0	100.0	100.0
	$\mathbf{\Sigma}_4$	5.9	4.8	5.9	6.0	5.8	89.1	64.3	98.7	99.0	97.7	**100.0**	100.0	100.0	100.0	100.0

The three distributions have very similar empirical powers across different combinations of sample sizes and dimensions on dense mean settings. Thus, to save space, the power function graph of Example (a) (multivariate normal distribution) is shown as a trend graph that gradually increases with the mean signal. The eight subgraphs in Figure 5 highlight the advantages of our proposed

Table S2: Empirical size and power (%) of test statistics (Example(c), nominal $\alpha = 0.05$)

		Size					Dense mean $w = 2$					Dense mean $w = 3$				
	Type	T^2_{CP}	T^2_{OP}	T^2_{BS}	T^2_{S}	T^2_{CQ}	T^2_{CP}	T^2_{OP}	T^2_{BS}	T^2_{S}	T^2_{CQ}	T^2_{CP}	T^2_{OP}	T^2_{BS}	T^2_{S}	T^2_{CQ}
	$\boldsymbol{\Sigma}_1$	5.7	5.5	6.7	7.2	5.0	**93.0**	57.0	12.8	13.3	9.9	**100.0**	100.0	55.2	69.9	48.9
$n = 200$	$\boldsymbol{\Sigma}_2$	6.2	5.2	7.3	7.1	5.3	**92.6**	61.0	7.7	8.3	5.6	**100.0**	100.0	11.3	14.7	8.1
$p = 250$	$\boldsymbol{\Sigma}_3$	5.8	5.5	5.4	–	5.9	97.3	59.1	99.8	98.7	99.9	**100.0**	100.0	100.0	100.0	100.0
	$\boldsymbol{\Sigma}_4$	5.9	5.5	5.2	7.7	4.8	77.6	23.2	95.7	89.3	93.4	**100.0**	98.4	100.0	100.0	100.0
	$\boldsymbol{\Sigma}_1$	5.8	5.1	7.3	7.8	4.9	**97.5**	73.8	16.5	16.9	11.3	**100.0**	100.0	87.5	95.1	70.9
$n = 200$	$\boldsymbol{\Sigma}_2$	5.8	5.5	7.2	7.6	6.2	**98.1**	76.3	7.9	8.4	5.4	**100.0**	100.0	11.4	14.1	7.9
$p = 350$	$\boldsymbol{\Sigma}_3$	5.0	5.5	5.3	–	4.4	99.4	76.5	100.0	99.8	100.0	**100.0**	100.0	100.0	100.0	100.0
	$\boldsymbol{\Sigma}_4$	6.0	5.2	5.3	7.8	4.3	88.8	36.7	98.9	96.2	98.0	**100.0**	100.0	100.0	100.0	100.0

CPT method over the method of optimal projection direction, T^2_{OP}, which coincide with the empirical size and power shown in Table 1 when the empirical size is controlled at the nominal level of 0.05.

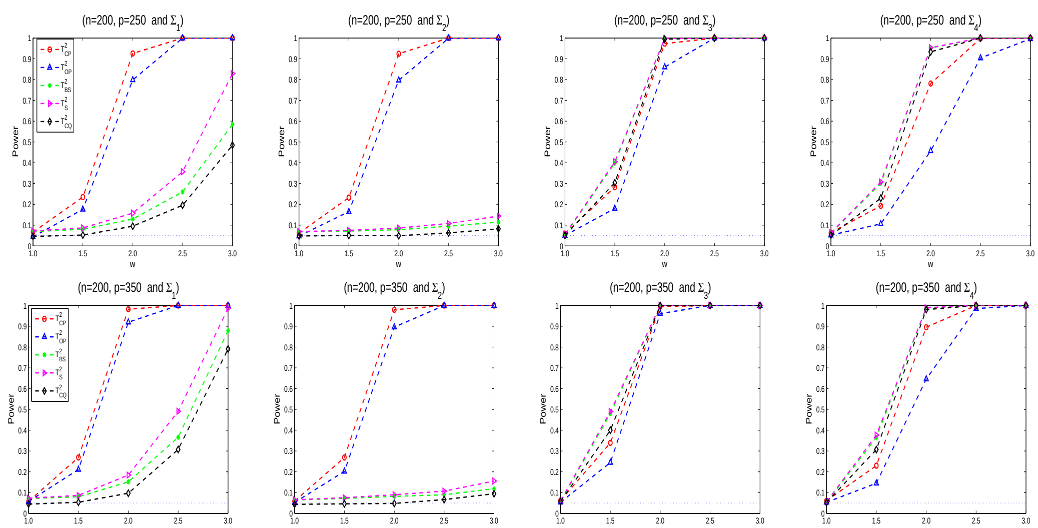


Figure 5: Comparing the empirical power under the settings of $\boldsymbol{\Sigma}_1$–$\boldsymbol{\Sigma}_4$ in Example (a).

Under the sparse mean settings, Tables S3 and S4 show the empirical size of the CPT approach as well as some tests for random samples generated by the three distributions, including Examples (a)–(c) in the case where the covariance structures are $\mathbf{\Sigma}_2$ and $\mathbf{\Sigma}_3$, respectively. Tables S5 and S6, respectively, describe the empirical power of the multivariate student t and chi-square distributions at novel level $\alpha = 0.05$ under the four types of the covariance matrix. We can observe that the results in these tables have the same performance as in Tables 2 and 3.

Table S3: Empirical size (%) of tests with $\alpha = 0.05$, sparse mean, $\mathbf{\Sigma}_2$

n	p	$T^2_{\rm BS}$	$T^2_{\rm S}$	$T^2_{\rm CQ}$	$T^2_{\rm OP}$	$T^2_{\rm CP}$	$J_{\rm OP}+J_0$	$J_{\rm CPT}+J_0$	$P(\widehat{\mathcal{S}}=\emptyset)$
Example(a): Multivariate Gaussian									
150	150	7.20	7.60	5.04	5.27	6.33	5.29	6.33	99.98
	200	7.07	7.28	4.98	5.01	6.30	5.04	6.32	99.93
	300	6.98	7.32	4.62	4.83	6.52	4.84	6.53	99.99
200	150	6.47	6.96	4.73	5.16	5.98	5.18	5.99	99.97
	200	7.06	7.34	4.93	4.84	6.28	4.86	6.30	99.98
	300	7.67	7.53	5.52	5.42	6.18	5.43	6.19	99.98
Example(b): Multivariate Student t									
150	150	7.09	7.42	4.97	5.39	6.19	5.46	6.24	99.92
	200	6.58	6.91	4.55	5.11	5.82	5.16	5.84	99.93
	300	6.91	7.30	4.74	5.21	5.74	5.23	5.75	99.97
200	150	7.58	7.64	5.21	4.84	5.62	4.85	5.63	99.99
	200	7.22	7.51	4.89	5.69	5.79	5.70	5.79	99.99
	300	7.13	7.46	4.96	5.20	5.90	5.20	5.90	100.00
Example(c): Multivariate Chi-square									
150	150	7.01	7.34	4.91	5.26	6.20	5.31	6.25	99.93
	200	6.65	6.80	4.51	5.31	6.33	5.35	6.36	99.95
	300	6.96	7.13	4.90	5.37	5.87	5.39	5.88	99.97
200	150	7.27	7.43	4.97	5.37	6.05	5.39	6.06	99.98
	200	6.42	6.62	4.38	5.56	5.34	5.57	5.35	99.99
	300	6.72	7.03	4.68	5.13	6.10	5.13	6.10	100.00

Table S4: Empirical size (%) of tests with $\alpha = 0.05$, sparse mean, $\mathbf{\Sigma}_3$

n	p	T^2_{BS}	T^2_{S}	T^2_{CQ}	T^2_{OP}	T^2_{CP}	$J_{\mathrm{OP}} + J_0$	$J_{\mathrm{CPT}} + J_0$	$P(\widehat{\mathcal{S}} = \emptyset)$
Example(a): Multivariate Gaussian									
150	150	5.47	5.89	4.70	4.56	5.77	4.68	5.90	99.86
	200	5.31	5.99	4.60	4.83	6.12	4.88	6.16	99.93
	300	5.00	5.65	5.33	4.51	5.71	4.53	5.73	99.96
200	100	5.58	6.00	5.00	4.78	5.96	4.80	5.97	99.92
	200	5.75	6.16	5.57	4.99	6.25	5.04	6.28	99.98
	300	4.88	5.35	4.53	4.98	5.13	5.00	5.15	99.98
Example(b): Multivariate Student t									
100	150	5.40	5.94	5.13	5.54	5.46	5.62	5.50	99.92
	200	5.97	6.51	5.10	5.75	5.67	5.79	5.72	99.95
	300	5.95	6.30	4.53	5.57	5.90	5.59	5.92	99.97
200	150	5.11	5.38	4.47	5.59	5.52	5.59	5.52	99.99
	200	5.78	5.90	4.67	5.43	5.93	5.45	5.95	99.98
	300	5.37	5.55	4.93	5.11	5.83	5.13	5.85	99.98
Example(c): Multivariate Chi-square									
150	150	5.50	-	4.43	5.15	5.34	5.92	6.04	99.20
	200	5.38	-	5.20	5.13	5.61	5.90	6.37	99.11
	300	5.74	-	4.97	5.48	5.78	6.08	6.34	99.32
200	150	5.59	8.97	4.47	5.43	5.76	5.69	6.01	99.70
	200	5.46	7.16	4.67	5.21	5.47	5.59	5.78	99.62
	300	4.79	7.07	4.93	5.16	5.51	5.35	5.69	99.79

Table S5: Empirical power (%) of tests with $\alpha = 0.05$ for Example(b)

Type	n	p	T^2_{BS}	T^2_{S}	T^2_{CQ}	T^2_{OP}	T^2_{CP}	$J_{\mathrm{OP}} + J_0$	$J_{\mathrm{CPT}} + J_0$	$P(\widehat{\mathcal{S}} = \emptyset)$
$\mathbf{\Sigma}_1$	150	150	11.87	12.83	8.17	46.77	74.23	49.51	75.00	92.07
		200	10.13	12.60	7.00	43.25	63.85	52.50	68.86	78.19
		300	10.26	12.88	6.57	40.41	49.79	50.39	57.94	77.20
	200	100	12.38	18.09	8.33	77.15	93.20	83.33	94.32	55.76
		200	11.37	12.42	7.37	58.41	86.62	62.51	87.38	84.72
		300	11.48	13.15	8.17	50.33	72.52	58.52	75.78	77.40
$\mathbf{\Sigma}_2$	150	150	7.90	8.44	4.70	34.70	71.94	34.86	71.97	99.36
		200	7.70	8.11	4.43	36.08	64.30	36.18	64.33	99.54
		300	7.87	8.63	5.50	37.45	46.68	37.70	48.90	98.98
	200	150	8.38	10.76	6.30	81.60	93.48	82.52	93.72	87.34
		200	8.12	8.36	5.30	55.01	86.43	55.04	86.43	99.67
		300	8.27	8.35	5.87	28.96	74.02	29.90	74.03	99.90
$\mathbf{\Sigma}_3$	150	150	97.41	97.65	95.00	55.35	86.42	89.43	95.16	16.18
		200	94.32	94.91	90.90	55.20	79.64	85.41	92.28	21.30
		300	86.66	87.45	80.77	42.76	64.55	79.31	85.50	28.90
	200	150	99.73	99.71	99.60	67.77	96.86	97.20	99.19	4.49
		200	99.25	99.27	98.57	64.01	93.17	96.09	98.42	6.61
		300	96.99	97.08	95.33	58.79	84.11	93.07	96.36	10.67
$\mathbf{\Sigma}_4$	150	150	87.00	87.81	81.50	33.30	66.31	77.58	84.07	27.73
		200	80.09	81.19	73.20	29.71	58.15	71.97	79.32	33.81
		300	68.57	70.31	60.00	26.26	46.64	65.18	71.53	41.31
	200	150	96.34	96.39	94.40	42.80	83.32	90.78	94.38	11.61
		200	93.38	93.42	88.34	39.83	75.74	87.53	92.19	15.51
		300	86.14	86.28	78.70	36.04	63.66	82.43	86.87	21.85

Table S6: Empirical power (%) of tests with $\alpha = 0.05$ for Example(c)

Type	n	p	T^2_{BS}	T^2_{S}	T^2_{CQ}	T^2_{OP}	T^2_{CP}	$J_{\text{OP}}+J_0$	$J_{\text{CPT}}+J_0$	$P(\widehat{\mathcal{S}}=\emptyset)$
$\mathbf{\Sigma}_1$	150	150	10.34	13.43	8.07	55.94	73.99	60.57	75.40	86.20
		200	10.68	13.71	7.13	48.98	63.29	53.69	65.57	86.51
		300	9.56	9.97	6.17	22.51	47.84	24.44	48.50	96.65
	200	150	12.83	12.34	8.70	56.29	93.29	57.47	93.34	95.37
		200	12.48	15.26	7.57	60.16	85.58	66.69	87.02	77.91
		300	11.83	14.13	7.10	49.74	73.85	58.00	76.86	79.38
$\mathbf{\Sigma}_2$	150	150	8.20	9.67	4.83	62.87	73.26	63.23	73.52	97.56
		200	7.41	7.91	5.87	32.42	62.82	32.55	62.85	99.62
		300	7.77	8.43	5.20	37.46	47.22	37.77	47.91	99.00
	200	150	8.37	8.81	5.50	69.66	93.51	69.77	93.51	99.45
		200	7.57	8.03	5.83	53.96	85.58	54.23	85.60	98.92
		300	7.55	8.72	5.17	57.93	72.09	58.57	72.42	96.83
$\mathbf{\Sigma}_3$	150	150	97.29	98.41	95.50	49.00	86.11	85.40	93.70	22.52
		200	94.49	96.19	90.73	44.12	77.99	80.07	89.29	30.01
		300	86.00	90.99	79.73	35.61	64.08	68.48	79.97	43.58
	200	150	99.77	99.94	99.60	63.58	97.16	97.73	99.38	3.82
		200	99.35	99.70	98.45	60.43	93.23	95.32	98.21	8.04
		300	97.17	98.29	94.57	53.93	84.21	91.86	95.55	13.20
$\mathbf{\Sigma}_4$	150	150	86.86	89.98	79.27	29.83	66.40	71.43	80.93	35.12
		200	80.36	83.26	74.13	26.94	57.58	65.57	74.53	42.20
		300	68.06	72.88	59.87	24.37	45.13	57.09	65.36	52.72
	200	150	96.38	97.22	94.00	41.23	83.89	89.12	94.10	14.23
		200	93.47	94.91	90.10	37.86	76.26	85.93	90.81	18.11
		300	86.68	89.01	79.47	33.54	63.95	78.24	84.53	27.39

References

Srivastava, M. and Du, M. (2008). "A test for the mean vector with fewer observations than the dimension." *Journal of Multivariate Analysis*, **99(3)**, 386–402.

Srivastava, M. (2009). "A test for the mean vector with fewer observations than the dimension under non-normality." *Journal of Multivariate Analysis*, **100**, 518–532.

Guanpeng Wang

School of Mathematical Sciences, Capital Normal University, Beijing 100048, China.

School of Mathematics and Information Science, Weifang University, Weifang, Shandong 261061, China.

E-mail: (wguanpeng@163.com)

Hengjian Cui

School of Mathematical Sciences, Capital Normal University, Beijing 100048, China.

E-mail: (hjcui@bnu.edu.cn)