Journal Pre-proof

Multivariate tests of independence and their application in correlation analysis between financial markets

Long Feng, Xiaoxu Zhang, Binghui Liu

PII: S0047-259X(20)30233-5
DOI: https://doi.org/10.1016/j.jmva.2020.104652
Reference: YJMVA 104652

To appear in: *Journal of Multivariate Analysis*

Received date : 26 May 2019
Revised date : 16 April 2020
Accepted date : 19 June 2020

Multivariate tests of independence and their application in correlation analysis between financial markets

Long Feng[a], Xiaoxu Zhang[b], Binghui Liu[b,*]

[a]School of Statistics and Data Science, LPMC and KLMDASR, Nankai University
[b]School of Mathematics and Statistics & KLAS, Northeast Normal University

Abstract

We consider the multivariate independence testing problem between pairs of random vectors for high-dimensional data and develop three high-dimensional nonparametric independence tests based on spatial sign and spatial rank, which have greater power than many existing popular tests, especially for heavy-tailed distributions. Under the elliptically symmetric distributions, which are much more general than the widely studied multivariate normal distributions, we establish asymptotic properties of the proposed tests and demonstrate their power superiority via frequently used numerical experiments. To explore the correlation between different financial markets, we first apply the proposed methods to test the dependence between the return rate data of the stocks from US S&P500 index and China CSI300 index, and then apply them to test the dependence between the return rate data of the stocks from the Shanghai Stock Exchange and the Shenzhen Stock Exchange in China.

Keywords: Elliptically symmetric, Heavy-tailed, Spatial rank, Spatial sign, Stock return.
2010 MSC: 62F03 Hypothesis testing

1. Introduction

With the deepening of economic globalization and financial integration, the fluctuation of one country's financial market is usually not only affected by internal factors, but also by the fluctuation of other countries' financial markets. Hence, more and more attention has been paid to the relationships between major international financial markets [7, 12], in order to identify linkages between them and to construct a reasonable portfolio for global investment [18]. A test of independence between the return rate vectors of the assets from two financial markets can help investigate their relationship, which can be considered as the basis of follow-up analysis.

In multivariate data analysis, it is important to determine whether two sets of variables are related [16]. Let $(\mathbb{X}, \mathbb{Y})$ be a random sample of size n from a $(p+q)$-variable distribution, where $\mathbb{X} = (\boldsymbol{X}_1, \cdots, \boldsymbol{X}_n)^\top$ with $\boldsymbol{X}_i \in \mathbb{R}^p$ is an $n \times p$ data matrix of the return rates of p assets from one financial market and $\mathbb{Y} = (\boldsymbol{Y}_1, \cdots, \boldsymbol{Y}_n)^\top$ with $\boldsymbol{Y}_j \in \mathbb{R}^q$ is an $n \times q$ data matrix of the return rates of q assets from another financial market. We consider the null hypothesis of the independence of the $\boldsymbol{X}$- and $\boldsymbol{Y}$-variables, written as

$$H_0 : \boldsymbol{X} \text{ and } \boldsymbol{Y} \text{ are independently distributed.} \tag{1}$$

The classical parametric test for (1) is the likelihood ratio test based on the multivariate normal model [26], whose statistic is $\boldsymbol{W} = |\mathbf{A}|/(|\mathbf{A}_{\boldsymbol{XX}}||\mathbf{A}_{\boldsymbol{YY}}|)$, where $\mathbf{A}_{\boldsymbol{XX}}$, $\mathbf{A}_{\boldsymbol{YY}}$ are the sample covariance matrices of $\boldsymbol{X}$, $\boldsymbol{Y}$ respectively, and $\mathbf{A}$ is the sample covariance matrix of $(\boldsymbol{X}^\top, \boldsymbol{Y}^\top)^\top$. It is optimal under the multivariate normal model when the dimensions p, q are fixed and smaller than the sample size n. However, it fails in the high-dimensional situation when p, q are larger than n.

*Corresponding author. Email address: `liubh100@nenu.edu.cn`

In recent years, researchers have paid more and more attention to the independence test of high-dimensional data. For example, Srivastava and Reid [20] proposed a test based on the Frobenius norm of sample covariance and correlation matrices between $\boldsymbol{X}$ and $\boldsymbol{Y}$. Jiang et al. [10] proposed the corrected likelihood ratio test and large-dimensional trace criterion to test the independence of two large sets of multivariate variables. Yang and Pan [27] extended the classic canonical correlation analysis to high-dimensional cases. Yata and Aoshima [28] modified Srivastava and Reid [20]'s test by using the extended cross-data-matrix methodology. In addition, Bodnar et al. [2] proposed alternative tests that are motivated from a classical multivariate analysis of variance and were defined as linear spectral statistics of a Fisher matrix. Despite the progress in this pursuit, there are many problems, one of which is that these tests are normal or similar-to-normal theory methods. Thus, they may fail to deal with data from heavy-tailed distributions, such as the multivariate t-distribution and the mixture of multivariate normal distribution.

To find robust and efficient alternatives to the multivariate normal theory methods, a large number of nonparametric methods, including multivariate sign or rank-based methods, are being developed. For example, Chen and Qin [3] proposed a two-sample test for the means of high-dimensional data, as the data dimension is much larger than the sample size, which does not require explicit conditions on the relationship between the data dimension and sample size. Li et al. [14] proposed two tests for the equality of covariance matrices between two high-dimensional populations, which do not require parametric distribution assumptions for the two populations. Wang et al. [25] proposed a high-dimensional nonparametric test for the population mean vector for a general class of multivariate distributions for non-normal high-dimensional multivariate data. Leung and Drton [13] considered the problem of testing mutual independence between variables and presented some rank-based tests, constructed as sums or sums of squares of pairwise rank correlations, which have power advantages in the case of non-normal distributions even when the data dimension is larger than the available sample size. Guo and Chen [6] considered testing regression coefficients in high-dimensional generalized linear models and proposed a test applicable for diverging dimensions, which is robust enough to accommodate a wide range of link functions. Feng et al. [5] concerned tests for the two-sample location problem when the data dimension is larger than the sample size, which is scalar-invariant and useful when different components have different scales in high-dimensional data. Zou et al. [29] concerned sign-based tests for sphericity in cases in which the data dimension is larger than the sample size, which is robust with respect to high dimensionality. Feng and Liu [4] proposed two rank-based tests inspired by Spearman's rho and Kendall's tau for testing sphericity in case of high-dimensional data.

For tests of independence between two multivariate random vectors, Taskinen et al. [22] proposed an affine invariant extension of the quadrant test statistics based on spatial signs. Taskinen et al. [23] proposed multivariate extensions of Kendall's tau and Spearman's rho statistics. These statistics performed very well in low-dimensional cases, but are not available in high-dimensional cases, since the sample spatial sign or rank covariance matrices to be inverted in the construction of the statistics are singular. To solve the high-dimensional problem, Paindaveine and Verdebout [17] proposed a high-dimensional sign test for some very special distribution types. To make this more general, in this paper, we propose a more extensive high-dimensional multivariate sign test for independence between two random vectors. In addition, we propose two high-dimensional multivariate rank-based tests for independence between two random vectors. The main difference between the proposed rank-based tests and those in [13] is that we test independence between two groups of variables, while Leung and Drton [13] tested mutual independence between all the involved variables. The common feature of all these rank-based tests is the advantage of power in non-normal situations. The theoretical contribution of this paper is its establishment of asymptotic theories of the three proposed nonparametric tests under the family of elliptically symmetric distributions, which is a very large distribution family, including a large number of well-known heavy tailed distributions, such as t distribution, mixed normal distribution, and power law distribution. We construct the corresponding testing procedures based on asymptotic theories such as these and demonstrate the power gain of the proposed testing procedures in comparison with existing tests through numerical results as well as two real data analyses. In particular, the power gain is especially clear in high-dimensional and heavy-tailed situations.

The remainder of the paper is organized as follows. In Sections 2 and 3, we propose a spatial sign test and two spatial rank tests for the high-dimensional independence testing problem and establish their asymptotic properties, respectively. The simulation performance of the three proposed tests are demonstrated in Section 4, followed by the two empirical applications of the proposed methods in correlation analysis between different financial markets in Section 5. Finally, we conclude the paper with some discussions in Section 6 and relegate the technical proofs to the Appendix.

2. High-dimensional multivariate sign test

Let $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n$ be a sequence of independent and identically distributed (iid) observations of a p-dimensional vector $\boldsymbol{X}$ with an elliptically symmetric density

$$\det(\boldsymbol{\Omega}_{\mathbf{X}})^{-1/2} g_{\mathrm{X}}\{\|\boldsymbol{\Omega}_{\mathrm{X}}^{-1/2}(\mathrm{x} - \boldsymbol{\theta}_{\mathrm{X}})\|\}, \tag{2}$$

where $\|\mathbf{z}\| = (\mathbf{z}^\top \mathbf{z})^{1/2}$ denotes the Euclidean length of a vector $\mathbf{z}$, $\boldsymbol{\theta}_{\boldsymbol{X}}$ is the center of symmetry, and $\boldsymbol{\Omega}_{\boldsymbol{X}}$ is a positive definite symmetric $p \times p$ scatter matrix. Similarly, let $\boldsymbol{Y}_1, \ldots, \boldsymbol{Y}_n$ be a sequence of independent and identically distributed (iid) observations of a q-dimensional vector $\boldsymbol{Y}$ with an elliptically symmetric density

$$\det(\boldsymbol{\Omega}_{\mathbf{Y}})^{-1/2} g_{\mathrm{Y}}\{\|\boldsymbol{\Omega}_{\mathrm{Y}}^{-1/2}(\mathrm{y} - \boldsymbol{\theta}_{\mathrm{Y}})\|\}, \tag{3}$$

where $\boldsymbol{\theta}_{\boldsymbol{Y}}$ is the center of symmetry and $\boldsymbol{\Omega}_{\mathbf{Y}}$ is a positive definite symmetric $q \times q$ scatter matrix. Note that elliptically symmetric distributions are second-order distributions with probability densities whose contours of equal height are ellipses. This class is very general and includes the multivariate normal and sine-wave distributions and others that can be generated from certain first-order distributions. The spatial sign function is defined as $\mathbf{U}(\mathbf{z}) = \|\mathbf{z}\|^{-1}\mathbf{z}I(\mathbf{z} \neq \mathbf{0})$. Let $\boldsymbol{\varepsilon}_i^X \doteq \boldsymbol{\Omega}_X^{-1/2}(\boldsymbol{X}_i - \boldsymbol{\theta}_{\boldsymbol{X}})$ for each $i \in \{1, \cdots, n\}$, where "$\doteq$" denotes "is defined as". Then, 1) the modulus $\|\boldsymbol{\varepsilon}_i^X\|$ and the direction $\mathbf{u}_i^X \doteq \mathbf{U}(\boldsymbol{\varepsilon}_i^X)$ are independent; 2) the direction vector $\mathbf{u}_i^X$ is uniformly distributed on the p-dimensional unit sphere; 3) $\mathrm{E}(\mathbf{u}_i^X) = 0$ and $\mathrm{cov}(\mathbf{u}_i^X) = p^{-1}\mathbf{I}_p$, where $\mathbf{I}_p$ denotes the $p \times p$ identity matrix. Similar conclusions can be derived for $\boldsymbol{\varepsilon}_i^Y \doteq \boldsymbol{\Omega}_{\mathbf{Y}}^{-1/2}(\boldsymbol{Y}_i - \boldsymbol{\theta}_{\boldsymbol{Y}})$ and $\mathbf{u}_i^Y \doteq \mathbf{U}(\boldsymbol{\varepsilon}_i^Y)$.

In a traditional fixed-dimension case, to test the independence between two random vectors, the so-called "inner centering and inner standardization" sign-based statistics as follows is commonly used (see Section 10.3 of [16]): $Q_{\mathrm{S}} = npq\mathrm{tr}(\tilde{\mathbf{A}}^\top \tilde{\mathbf{A}})$, where $\tilde{\mathbf{A}} = n^{-1}\sum_{i=1}^n \tilde{\mathbf{U}}_i^X (\tilde{\mathbf{U}}_i^Y)^\top$, $\tilde{\mathbf{U}}_i^X = \mathbf{U}(\mathbf{S}_X^{-1/2}(\boldsymbol{X}_i - \tilde{\boldsymbol{\theta}}_X))$, $\tilde{\mathbf{U}}_i^Y = \mathbf{U}(\mathbf{S}_{\mathbf{Y}}^{-1/2}(\boldsymbol{Y}_i - \tilde{\boldsymbol{\theta}}_{\boldsymbol{Y}}))$ and $\mathrm{tr}(\cdot)$ denotes the trace function of a matrix. Here $\tilde{\boldsymbol{\theta}}_{\boldsymbol{X}}$ and $\mathbf{S}_{\boldsymbol{X}}$ are the HRE's of the location vector and the scatter matrix for $\boldsymbol{X}$ [9], which satisfy the following conditions: $\sum_{i=1}^n \tilde{\mathbf{U}}_i^X = \mathbf{0}$ and $pn^{-1}\sum_{i=1}^n \tilde{\mathbf{U}}_i^X(\tilde{\mathbf{U}}_i^X)^\top = \mathbf{I}_p$. Similarly, the HRE's $\tilde{\boldsymbol{\theta}}_{\boldsymbol{Y}}$ and $\mathbf{S}_{\boldsymbol{Y}}$ for $\boldsymbol{Y}$ can be obtained. As mentioned in [22], under H_0, $Q_{\mathrm{S}} \xrightarrow{d} \chi^2_{pq}$. However, in case of $p, q > n$, Q_{S} fails because $\mathbf{S}_{\boldsymbol{X}}$ and $\mathbf{S}_{\boldsymbol{Y}}$ are singular, which cannot be inverted in the construction of Q_{S}. A common strategy used to resolve this problem is to replace the scatter matrices $\mathbf{S}_{\boldsymbol{X}}$ and $\mathbf{S}_{\boldsymbol{Y}}$ in Q_{S} with $\mathbf{I}_p$ and $\mathbf{I}_q$, respectively. Moreover, as $\mathbf{S}_{\boldsymbol{X}}$ is not available, $\tilde{\boldsymbol{\theta}}_{\boldsymbol{X}}$ is correspondingly not available; hence, we replace $\tilde{\boldsymbol{\theta}}_{\boldsymbol{X}}$ with a rotation equivariant spatial median $\hat{\boldsymbol{\theta}}_{\boldsymbol{X}}$ inspired by Möttönen and Oja [15], which is a minimizer of the criterion function of $L(\boldsymbol{\theta}) = \sum_{i=1}^n \|\boldsymbol{X}_i - \boldsymbol{\theta}\|$. Similarly, we replace $\tilde{\boldsymbol{\theta}}_{\boldsymbol{Y}}$ with $\hat{\boldsymbol{\theta}}_{\boldsymbol{Y}}$.

Based on the above replacement, we rewrite Q_{S} as follows $Q_{\mathrm{S}}' = npq\mathrm{tr}(\hat{\mathbf{A}}^\top \hat{\mathbf{A}}) = 2pqn^{-1}\sum_{1 \le i < j \le n} (\hat{\mathbf{U}}_i^X)^\top \hat{\mathbf{U}}_j^X (\hat{\mathbf{U}}_i^Y)^\top \hat{\mathbf{U}}_j^Y + pq$, where $\hat{\mathbf{A}} = n^{-1}\sum_{i=1}^n \hat{\mathbf{U}}_i^X (\hat{\mathbf{U}}_i^Y)^\top$, $\hat{\mathbf{U}}_i^X = \mathbf{U}(\boldsymbol{X}_i - \hat{\boldsymbol{\theta}}_X)$, $\hat{\mathbf{U}}_i^Y = \mathbf{U}(\boldsymbol{Y}_i - \hat{\boldsymbol{\theta}}_{\boldsymbol{Y}})$. We can see that $\sum_{1 \le i < j \le n} (\hat{\mathbf{U}}_i^X)^\top \hat{\mathbf{U}}_j^X (\hat{\mathbf{U}}_i^Y)^\top \hat{\mathbf{U}}_j^Y$ is the leading role of Q_{S}'. Because $\mathrm{var}(\hat{\mathbf{U}}_i^X) \neq \mathbf{I}_p/p$ and $\mathrm{var}(\hat{\mathbf{U}}_i^Y) \neq \mathbf{I}_q/q$, we consider using the standardization of $\sum_{1 \le i < j \le n} (\hat{\mathbf{U}}_i^X)^\top \hat{\mathbf{U}}_j^X (\hat{\mathbf{U}}_i^Y)^\top \hat{\mathbf{U}}_j^Y$ and hence propose the following high-dimensional multivariate sign test (abbreviated as HS) for testing independence between vectors $\boldsymbol{X}$ and $\boldsymbol{Y}$:

$$T_{\mathrm{HS}} = \frac{n \sum_{1 \le i < j \le n} (\hat{\mathbf{U}}_i^X)^\top \hat{\mathbf{U}}_j^X (\hat{\mathbf{U}}_i^Y)^\top \hat{\mathbf{U}}_j^Y}{\sqrt{2 \sum_{1 \le i < j \le n} ((\hat{\mathbf{U}}_i^X)^\top \hat{\mathbf{U}}_j^X)^2 \sum_{1 \le i < j \le n} ((\hat{\mathbf{U}}_i^Y)^\top \hat{\mathbf{U}}_j^Y)^2}}. \tag{4}$$

Let $\boldsymbol{\Sigma}_{\boldsymbol{X}} = \mathrm{var}(\boldsymbol{X}) = p^{-1}\mathrm{E}(\|\boldsymbol{\varepsilon}_i^X\|^2)\boldsymbol{\Omega}_{\boldsymbol{X}}$, $\boldsymbol{\Sigma}_{\boldsymbol{Y}} = \mathrm{var}(\boldsymbol{Y}) = q^{-1}\mathrm{E}(\|\boldsymbol{\varepsilon}_i^Y\|^2)\boldsymbol{\Omega}_{\boldsymbol{Y}}$, for any $i \in \{1, \cdots, n\}$. Let $\lambda_{\max}(\cdot)$ denote the largest eigenvalue of a matrix. In deriving the asymptotic properties of T_{HS}, we impose the following two commonly used conditions, which were previously used by [25].

(C1) $\mathrm{tr}(\boldsymbol{\Sigma}_{\boldsymbol{X}}^4)\mathrm{tr}(\boldsymbol{\Sigma}_{\boldsymbol{Y}}^4) = o(\mathrm{tr}^2(\boldsymbol{\Sigma}_{\boldsymbol{X}}^2)\mathrm{tr}^2(\boldsymbol{\Sigma}_{\boldsymbol{Y}}^2))$ as $\max\{p, q\} \to \infty$;

(C2) If $p \to \infty$, then $\frac{\mathrm{tr}^4(\boldsymbol{\Sigma}_X)}{\mathrm{tr}^2(\boldsymbol{\Sigma}_X^2)} \exp\left\{-\frac{\mathrm{tr}^2(\boldsymbol{\Sigma}_X)}{128p\lambda_{\max}^2(\boldsymbol{\Sigma}_X)}\right\} = o(1)$; and if $q \to \infty$, then $\frac{\mathrm{tr}^4(\boldsymbol{\Sigma}_Y)}{\mathrm{tr}^2(\boldsymbol{\Sigma}_Y^2)} \exp\left\{-\frac{\mathrm{tr}^2(\boldsymbol{\Sigma}_Y)}{128q\lambda_{\max}^2(\boldsymbol{\Sigma}_Y)}\right\} = o(1)$.

As mentioned by [25], these two conditions are quite relaxed. In particular, condition (C1) holds trivially if all eigenvalues of $\boldsymbol{\Sigma}_{\boldsymbol{X}}$ and $\boldsymbol{\Sigma}_{\boldsymbol{Y}}$ are bounded away from 0 and ∞. In fact, the bounded eigenvalues assumption is commonly adopted in the literature of estimating high-dimensional covariance matrices (see [1]). It has also been shown that condition (C1) holds under some general conditions if some of the eigenvalues are unbounded (see [3]).

Condition (C2) was first imposed by [25], which also holds if all eigenvalues of $\mathbf{\Sigma}_X$ and $\mathbf{\Sigma}_Y$ are bounded away from 0 and ∞. This permits the eigenvalues to be unbounded, as the exponential term is expected to converge to zero quickly if $\mathrm{tr}(\mathbf{\Sigma}_X)/\{\sqrt{p}\lambda_{\max}(\mathbf{\Sigma}_X)\}$ and $\mathrm{tr}(\mathbf{\Sigma}_Y)/\{\sqrt{q}\lambda_{\max}(\mathbf{\Sigma}_Y)\}$ diverge to ∞. In particular, as mentioned in [25], if $p \to \infty$, let $\lambda_1 \le \lambda_2 \le \cdots \le \lambda_p$ be ordered eigenvalues of $\mathbf{\Sigma}_X$. Assume that as $p \to \infty$, k_1 eigenvalues converge to 0; k_2 eigenvalues diverge to ∞, and $p - k_1 - k_2$ eigenvalues remain bounded with lower bound $c_1 > 0$ and upper bound $c_2 < \infty$. Then,

$$\frac{\mathrm{tr}(\mathbf{\Sigma}_X)}{\sqrt{p}\lambda_{\max}(\mathbf{\Sigma}_X)} \ge \frac{k_1\lambda_1 + c_1(p-k_1-k_2) + k_2\lambda_{p-k_2+1}}{\sqrt{p}\lambda_p}, \quad \frac{\mathrm{tr}^2(\mathbf{\Sigma}_X)}{\mathrm{tr}\left(\mathbf{\Sigma}_X^2\right)} \le \frac{k_2^2\lambda_p^2 + (p-k_2)^2 c_2^2 + 2k_2(p-k_2)c_2\lambda_p}{k_1\lambda_1^2 + (p-k_1)c_1^2}.$$

Assume $\lambda_1 = p^{-b_1}$ and $\lambda_p = p^{b_2}$ for $b_1 > 0$, $b_2 > 0$. If k_1 and k_2 are bounded, then

$$\frac{\mathrm{tr}^4(\mathbf{\Sigma}_X)}{\mathrm{tr}^2(\mathbf{\Sigma}_X^2)} \exp\left\{-\frac{\mathrm{tr}^2(\mathbf{\Sigma}_X)}{128p\lambda_{\max}^2(\mathbf{\Sigma}_X)}\right\} = o(1)$$

in condition (C2) is satisfied if $b_2 < \frac{1}{2}$.

Now, under the above two conditions, we present the asymptotic normality of T_{HS} in (4) under H_0 in (1).

Theorem 1. *Under conditions (C1), (C2) and H_0 in (1), if $(p,q) = O(n^2)$, $T_{HS} \xrightarrow{d} \mathcal{N}(0,1)$, where T_{Hs} is given in (4).*

To illustrate and compare the efficiencies of different test statistics for independence, we derive the limiting distribution of the test statistic under specific contiguous alternative sequences (see Section 10.4 in [16]). Let

$$\begin{pmatrix} \boldsymbol{X}_i - \boldsymbol{\theta}_X \\ \boldsymbol{Y}_i - \boldsymbol{\theta}_Y \end{pmatrix} = \begin{pmatrix} \mathbf{I}_p & \mathbf{M}_1 \\ \mathbf{M}_2 & \mathbf{I}_q \end{pmatrix} \begin{pmatrix} \boldsymbol{X}_i^* - \boldsymbol{\theta}_{X^*} \\ \boldsymbol{Y}_i^* - \boldsymbol{\theta}_{Y^*} \end{pmatrix},$$

where $\mathbf{M}_1 \in \mathbb{R}^{p\times q}$, $\mathbf{M}_2 \in \mathbb{R}^{q\times p}$, $\boldsymbol{X}_i^*$ and $\boldsymbol{Y}_i^*$ are independent, with density functions (2) and (3), respectively. Define

$$\mathbf{A}_X^* = \mathrm{E}(\mathbf{U}(\boldsymbol{X}_i^* - \boldsymbol{\theta}_{X^*})\mathbf{U}(\boldsymbol{X}_i^* - \boldsymbol{\theta}_{X^*})^\top), \quad \mathbf{A}_Y^* = \mathrm{E}(\mathbf{U}(\boldsymbol{Y}_i^* - \boldsymbol{\theta}_{Y^*})\mathbf{U}(\boldsymbol{Y}_i^* - \boldsymbol{\theta}_{Y^*})^\top), \tag{5}$$

$$\mathbf{\Lambda} = \mathrm{E}((r_i^{X^*})^{-1} r_i^{Y^*})\mathbf{M}_1\mathbf{A}_Y^* + \mathrm{E}((r_i^{Y^*})^{-1} r_i^{X^*})\mathbf{A}_X^*\mathbf{M}_2^\top, \tag{6}$$

where $r_i^{X^*} = \|\boldsymbol{X}_i^* - \boldsymbol{\theta}_{X^*}\|$, $r_i^{Y^*} = \|\boldsymbol{Y}_i^* - \boldsymbol{\theta}_{Y^*}\|$. Define the covariance matrix of $\boldsymbol{X}_i^*$ and $\boldsymbol{Y}_i^*$ as $\mathbf{\Sigma}_X^*$ and $\mathbf{\Sigma}_Y^*$, respectively. We impose the following conditions for an alternative hypothesis:

(C1′) $\mathrm{tr}(\mathbf{\Sigma}_X^{*4})\mathrm{tr}(\mathbf{\Sigma}_Y^{*4}) = o(\mathrm{tr}^2(\mathbf{\Sigma}_X^{*2})\mathrm{tr}^2(\mathbf{\Sigma}_Y^{*2}))$ as $\max\{p,q\} \to \infty$;

(C2′) If $p \to \infty$, then $\frac{\mathrm{tr}^4(\mathbf{\Sigma}_X^*)}{\mathrm{tr}^2(\mathbf{\Sigma}_X^{*2})} \exp\left\{-\frac{\mathrm{tr}^2(\mathbf{\Sigma}_X^*)}{128p\lambda_{\max}^2(\mathbf{\Sigma}_X^*)}\right\} = o(1)$; and if $q \to \infty$, then $\frac{\mathrm{tr}^4(\mathbf{\Sigma}_Y^*)}{\mathrm{tr}^2(\mathbf{\Sigma}_Y^{*2})} \exp\left\{-\frac{\mathrm{tr}^2(\mathbf{\Sigma}_Y^*)}{128q\lambda_{\max}^2(\mathbf{\Sigma}_Y^*)}\right\} = o(1)$;

(C3′) $n\mathrm{tr}(\mathbf{\Lambda}^\top\mathbf{\Lambda}) = O(\sigma_1^*)$, $\{\mathrm{E}((r_i^{X^*})^{-1}r_i^{Y^*})\}^2\mathrm{tr}(\mathbf{M}_1\mathbf{A}_Y^2\mathbf{M}_1^\top\mathbf{A}_X) = o(\sigma_1^{*2})$, $\{\mathrm{E}((r_i^{Y^*})^{-1}r_i^{X^*})\}^2\mathrm{tr}(\mathbf{M}_2\mathbf{A}_X^2\mathbf{M}_2^\top\mathbf{A}_Y) = o(\sigma_1^{*2})$, where $\sigma_1^{*2} = n(2(n-1))^{-1}\mathrm{tr}(\mathbf{A}_X^{*2})\mathrm{tr}(\mathbf{A}_Y^{*2})$.

Under the above sequence of alternatives, we obtain the following limiting distribution of T_{HS} in (4).

Theorem 2. *Under conditions (C1′)-(C3′), if $(p,q) = O(n^2)$, $T_{HS} \xrightarrow{d} \mathcal{N}\left(n\mathrm{tr}(\mathbf{\Lambda}^\top\mathbf{\Lambda})/\sqrt{2\mathrm{tr}(\mathbf{A}_X^{*2})\mathrm{tr}(\mathbf{A}_Y^{*2})}, 1\right)$, where T_{HS} is given in (4), $\mathbf{A}_X^*$, $\mathbf{A}_Y^*$ are given in (5) and $\mathbf{\Lambda}$ is given in (6).*

3. High-dimensional multivariate rank test

Next, we propose two spatial rank tests for independence that are essentially high-dimensional multivariate extensions of Spearman's rho and Kendall's tau tests for independence testing problems.

3.1. High-dimensional Spearman's rho test

For the independence testing problem in the traditional fixed dimension case, the multivariate Spearman's rho test statistic is proposed [23]: $Q_{\mathrm{R}} = npq\mathrm{tr}((\tilde{\boldsymbol{\Sigma}}^{\mathrm{R}}_{XY})^\top \tilde{\boldsymbol{\Sigma}}^{\mathrm{R}}_{XY})/\{\mathrm{tr}((\tilde{\boldsymbol{\Sigma}}^{\mathrm{R}}_{X})^2)\mathrm{tr}((\tilde{\boldsymbol{\Sigma}}^{\mathrm{R}}_{Y})^2)\}$, where

$$\tilde{\boldsymbol{\Sigma}}^{\mathrm{R}}_{XY} = n^{-1}\sum_{i=1}^n \tilde{\mathbf{R}}_i^X(\tilde{\mathbf{R}}_i^Y)^\top,\ \ \tilde{\boldsymbol{\Sigma}}^{\mathrm{R}}_{X} = n^{-1}\sum_{i=1}^n \tilde{\mathbf{R}}_i^X(\tilde{\mathbf{R}}_i^X)^\top,\ \ \tilde{\boldsymbol{\Sigma}}^{\mathrm{R}}_{Y} = n^{-1}\sum_{i=1}^n \tilde{\mathbf{R}}_i^Y(\tilde{\mathbf{R}}_i^Y)^\top,$$

$$\tilde{\mathbf{R}}_i^X = n^{-1}\sum_{j=1}^n \mathbf{U}((\mathbf{S}^{\mathrm{R}}_X)^{-1/2}(\boldsymbol{X}_i - \boldsymbol{X}_j)),\ \ \tilde{\mathbf{R}}_i^Y = n^{-1}\sum_{j=1}^n \mathbf{U}((\mathbf{S}^{\mathrm{R}}_Y)^{-1/2}(\boldsymbol{Y}_i - \boldsymbol{Y}_j)).$$

$\mathbf{S}^{\mathrm{R}}_X$ and $\mathbf{S}^{\mathrm{R}}_Y$ are full-rank transformation matrices that satisfy $\tilde{\boldsymbol{\Sigma}}^{\mathrm{R}}_X \propto \mathbf{I}_p$ and $\tilde{\boldsymbol{\Sigma}}^{\mathrm{R}}_Y \propto \mathbf{I}_q$, respectively. Recall that under the null hypothesis as well as some general assumptions, it can be concluded that $Q_{\mathrm{R}} \xrightarrow{d} \chi^2_{pq}$ [23]. However, in high-dimensional cases when $p > n$, Q_{R} is not available, as the matrices $\mathbf{S}^{\mathrm{R}}_X$ and $\mathbf{S}^{\mathrm{R}}_Y$ are singular, which cannot be inverted in the construction of Q_{R}. To tackle this problem, we can use a similar strategy to that used in the previous section. We can simply replace $\mathbf{S}^{\mathrm{R}}_X$ and $\mathbf{S}^{\mathrm{R}}_Y$ with $\mathbf{I}_p$ and $\mathbf{I}_q$ in Q_{R}, respectively. The resulting test statistic is $Q_{\mathrm{R}}' = npq\mathrm{tr}((\hat{\boldsymbol{\Sigma}}^{\mathrm{R}}_{XY})^\top \hat{\boldsymbol{\Sigma}}^{\mathrm{R}}_{XY})/\{\mathrm{tr}((\hat{\boldsymbol{\Sigma}}^{\mathrm{R}}_{X})^2)\mathrm{tr}((\hat{\boldsymbol{\Sigma}}^{\mathrm{R}}_{Y})^2)\}$, where

$$\hat{\boldsymbol{\Sigma}}^{\mathrm{R}}_{XY} = n^{-1}\sum_{i=1}^n \hat{\mathbf{R}}_i^X(\hat{\mathbf{R}}_i^Y)^\top,\ \ \hat{\boldsymbol{\Sigma}}^{\mathrm{R}}_{X} = n^{-1}\sum_{i=1}^n \hat{\mathbf{R}}_i^X(\hat{\mathbf{R}}_i^X)^\top,\ \ \hat{\boldsymbol{\Sigma}}^{\mathrm{R}}_{Y} = n^{-1}\sum_{i=1}^n \hat{\mathbf{R}}_i^Y(\hat{\mathbf{R}}_i^Y)^\top,$$

$$\hat{\mathbf{R}}_i^X = n^{-1}\sum_{j=1}^n \mathbf{U}(\boldsymbol{X}_i - \boldsymbol{X}_j),\ \ \hat{\mathbf{R}}_i^Y = n^{-1}\sum_{j=1}^n \mathbf{U}(\boldsymbol{Y}_i - \boldsymbol{Y}_j).$$

By using the commonly used leave-out strategy for Q_{R}', we develop a high-dimensional version of Spearman's rho test (abbreviated as HR) for testing the independence between $\boldsymbol{X}$ and $\boldsymbol{Y}$:

$$T_{\mathrm{HR}} = \frac{\sqrt{2}n\sum^* \mathbf{U}(\boldsymbol{X}_i - \boldsymbol{X}_j)^\top \mathbf{U}(\boldsymbol{X}_k - \boldsymbol{X}_\ell)\mathbf{U}(\boldsymbol{Y}_i - \boldsymbol{Y}_\ell)^\top \mathbf{U}(\boldsymbol{Y}_k - \boldsymbol{Y}_j)}{\sqrt{\sum^*[\mathbf{U}(\boldsymbol{X}_i - \boldsymbol{X}_j)^\top \mathbf{U}(\boldsymbol{X}_k - \boldsymbol{X}_\ell)]^2 \sum^*[\mathbf{U}(\boldsymbol{Y}_i - \boldsymbol{Y}_j)^\top \mathbf{U}(\boldsymbol{Y}_k - \boldsymbol{Y}_\ell)]^2}}, \tag{7}$$

where $\sum^*$ denotes summation over distinct indexes. Here "leave-out" means that we remove the items with some common indices, $\mathbf{U}(\boldsymbol{X}_i - \boldsymbol{X}_j)^\top \mathbf{U}(\boldsymbol{X}_k - \boldsymbol{X}_\ell)\mathbf{U}(\boldsymbol{Y}_i - \boldsymbol{Y}_\ell)^\top \mathbf{U}(\boldsymbol{Y}_k - \boldsymbol{Y}_j)$'s, whose indices i, j, k, l are not mutually different, from $\sum_{i,j,k,l} \mathbf{U}(\boldsymbol{X}_i - \boldsymbol{X}_j)^\top \mathbf{U}(\boldsymbol{X}_k - \boldsymbol{X}_\ell)\mathbf{U}(\boldsymbol{Y}_i - \boldsymbol{Y}_\ell)^\top \mathbf{U}(\boldsymbol{Y}_k - \boldsymbol{Y}_j)$. As mentioned in [3], such items with common indices will generally lead to additional bias and stronger demands on the dimensionality.

Below, we present the asymptotic normality of T_{HR}. Define

$$\mathbf{B}^*_X = \mathrm{E}(\boldsymbol{V}_i^{X^*}(\boldsymbol{V}_i^{X^*})^\top),\ \mathbf{B}^*_Y = \mathrm{E}(\boldsymbol{V}_i^{Y^*}(\boldsymbol{V}_i^{Y^*})^\top), \tag{8}$$

where $\boldsymbol{V}_i^{X^*} = \mathrm{E}(\mathbf{U}(\boldsymbol{X}_i^* - \boldsymbol{X}_j^*)|\boldsymbol{X}_i^*)$, $\boldsymbol{V}_i^{Y^*} = \mathrm{E}(\mathbf{U}(\boldsymbol{Y}_i^* - \boldsymbol{Y}_j^*)|\boldsymbol{Y}_i^*)$; and then define

$$\tilde{\boldsymbol{\Lambda}} = \mathrm{E}\left((\tilde{r}_{ij}^{X^*})^{-1}\tilde{r}_{ij}^{Y^*}\right)\mathbf{M}_1\mathbf{B}^*_Y + \mathrm{E}\left((\tilde{r}_{ij}^{Y^*})^{-1}\tilde{r}_{ij}^{X^*}\right)\mathbf{B}^*_X\mathbf{M}_2^\top, \tag{9}$$

where $\tilde{r}_{ij}^{X^*} = \|\boldsymbol{X}_i^* - \boldsymbol{X}_j^*\|$ and $\tilde{r}_{ij}^{Y^*} = \|\boldsymbol{Y}_i^* - \boldsymbol{Y}_j^*\|$. To derive the limiting distribution of T_{HR} under the alternative hypothesis, we impose the following condition to replace condition (C3$'$):

(C4$'$) $n\mathrm{tr}(\tilde{\boldsymbol{\Lambda}}^\top\tilde{\boldsymbol{\Lambda}}) = O(\sigma_2^*)$, $\left\{\mathrm{E}\left((r_{ij}^{X^*})^{-1}\tilde{r}_{ij}^{Y^*}\right)\right\}^2 \mathrm{tr}(\mathbf{M}_1\mathbf{B}_Y^{*2}\mathbf{M}_1^\top\mathbf{B}^*_X) = o(\sigma_2^{*2})$, $\left\{\mathrm{E}\left((r_{ij}^{Y^*})^{-1}\tilde{r}_{ij}^{X^*}\right)\right\}^2 \mathrm{tr}(\mathbf{M}_2\mathbf{B}_X^{*2}\mathbf{M}_2^\top\mathbf{B}^*_Y) = o(\sigma_2^{*2})$, where $\sigma_2^{*2} = n\{2(n-1)\}^{-1}\mathrm{tr}(\mathbf{B}_X^{*2})\mathrm{tr}(\mathbf{B}_Y^{*2})$.

Theorem 3. (i) *Under conditions (C1), (C2) and H_0 in (1), $T_{HR} \xrightarrow{d} \mathcal{N}(0, 1)$, where T_{HR} is given in (7).*

(ii) *Under conditions (C1$'$), (C2$'$) and (C4$'$), $T_{HR} \xrightarrow{d} \mathcal{N}\left(n\mathrm{tr}(\tilde{\boldsymbol{\Lambda}}^\top\tilde{\boldsymbol{\Lambda}})/\sqrt{2\mathrm{tr}(\mathbf{B}_X^{*2})\mathrm{tr}(\mathbf{B}_Y^{*2})}, 1\right)$, where T_{HR} is given in (7), $\mathbf{B}_X^{*2}$, $\mathbf{B}_Y^{*2}$ are given in (8) and $\tilde{\boldsymbol{\Lambda}}$ is given in (9).*

3.2. High-dimensional Kendall's tau test

Taskinen et al. [23] also proposed the multivariate Kendall's tau test statistic for the independence problem in traditional fixed-dimension cases: $Q_{\mathrm{T}} = npq\mathrm{tr}((\tilde{\boldsymbol{\Sigma}}^{\mathrm{T}}_{\boldsymbol{XY}})^\top \tilde{\boldsymbol{\Sigma}}^{\mathrm{T}}_{\boldsymbol{XY}})/\{4(n-1)^2\mathrm{tr}((\tilde{\boldsymbol{\Sigma}}^{\mathrm{R}}_{\boldsymbol{X}})^2)\mathrm{tr}((\tilde{\boldsymbol{\Sigma}}^{\mathrm{R}}_{\boldsymbol{Y}})^2)\}$, where

$$\tilde{\boldsymbol{\Sigma}}^{\mathrm{T}}_{\boldsymbol{XY}} = \frac{1}{n^2}\sum_{i=1}^n\sum_{j=1}^n \mathbf{U}((\mathbf{S}^{\mathrm{R}}_{\boldsymbol{X}})^{-1/2}(\boldsymbol{X}_i - \boldsymbol{X}_j))\mathbf{U}((\mathbf{S}^{\mathrm{R}}_{\boldsymbol{Y}})^{-1/2}(\boldsymbol{Y}_i - \boldsymbol{Y}_j))^\top.$$

Q_{T} is also asymptotically chi-square distributed with pq degrees of freedom, under the null distribution and some general assumptions. Like Q_{R}, Q_{T} is also not available in the high-dimensional case; hence, in Q_{T}, we can similarly replace the scatter matrix $\mathbf{S}^{\mathrm{R}}_{\boldsymbol{X}}$ and $\mathbf{S}^{\mathrm{R}}_{\boldsymbol{Y}}$ with $\mathbf{I}_p$ and $\mathbf{I}_q$ respectively, and accordingly consider the following statistics $Q'_{\mathrm{T}} = npq\mathrm{tr}((\hat{\boldsymbol{\Sigma}}^{\mathrm{T}}_{\boldsymbol{XY}})^\top \hat{\boldsymbol{\Sigma}}^{\mathrm{T}}_{\boldsymbol{XY}})/\{\mathrm{tr}((\hat{\boldsymbol{\Sigma}}^{\mathrm{R}}_{\boldsymbol{X}})^2)\mathrm{tr}((\hat{\boldsymbol{\Sigma}}^{\mathrm{R}}_{\boldsymbol{Y}})^2)\}$, where $\hat{\boldsymbol{\Sigma}}^{\mathrm{T}}_{\boldsymbol{XY}} = n^{-2}\sum_{i=1}^n\sum_{j=1}^n \mathbf{U}(\boldsymbol{X}_i - \boldsymbol{X}_j)\mathbf{U}(\boldsymbol{Y}_i - \boldsymbol{Y}_j)^\top$. By using the leave-out strategy for Q'_{T}, we develop a high-dimensional version of Kendall's tau test (abbreviated as HT) for testing the independence between $\boldsymbol{X}$ and $\boldsymbol{Y}$:

$$T_{\mathrm{HT}} = \frac{n\sum^* \mathbf{U}(\boldsymbol{X}_i - \boldsymbol{X}_j)^\top \mathbf{U}(\boldsymbol{X}_k - \boldsymbol{X}_\ell)\mathbf{U}(\boldsymbol{Y}_i - \boldsymbol{Y}_j)^\top \mathbf{U}(\boldsymbol{Y}_k - \boldsymbol{Y}_\ell)}{\sqrt{2\sum^*[\mathbf{U}(\boldsymbol{X}_i - \boldsymbol{X}_j)^\top \mathbf{U}(\boldsymbol{X}_k - \boldsymbol{X}_\ell)]^2\sum^*[\mathbf{U}(\boldsymbol{Y}_i - \boldsymbol{Y}_j)^\top \mathbf{U}(\boldsymbol{Y}_k - \boldsymbol{Y}_\ell)]^2}}. \tag{10}$$

Theorem 4. (i) *Under conditions (C1), (C2) and H_0 in* (1), $T_{HT} \xrightarrow{d} \mathcal{N}(0,1)$, *where T_{HT} is given in* (10).

(ii) *Under conditions (C1$'$), (C2$'$) and (C4$'$), $T_{HT} \xrightarrow{d} \mathcal{N}\left(n\mathrm{tr}(\tilde{\boldsymbol{\Lambda}}^\top\tilde{\boldsymbol{\Lambda}})/\sqrt{2\mathrm{tr}(\mathbf{B}^{*2}_{\boldsymbol{X}})\mathrm{tr}(\mathbf{B}^{*2}_{\boldsymbol{Y}})}, 1\right)$, where T_{HT} is given in* (10), *$\mathbf{B}^{*2}_{\boldsymbol{X}}$, $\mathbf{B}^{*2}_{\boldsymbol{Y}}$ are given in* (8) *and $\tilde{\boldsymbol{\Lambda}}$ is given in* (9).

3.3. Power comparison

According to Theorems 1-4, the power functions of T_{HS}, T_{HR}, T_{HT} in (4), (7), (10), are

$$\beta_{\mathrm{HS}}(\mathbf{M}_1,\mathbf{M}_2) = \Phi\left(-z_\alpha + \frac{n\mathrm{tr}(\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda})}{\sqrt{2\mathrm{tr}(\mathbf{A}^{*2}_{\boldsymbol{X}})\mathrm{tr}(\mathbf{A}^{*2}_{\boldsymbol{Y}})}}\right),\quad \beta_{\mathrm{HR}}(\mathbf{M}_1,\mathbf{M}_2) = \Phi\left(-z_\alpha + \frac{n\mathrm{tr}(\tilde{\boldsymbol{\Lambda}}^\top\tilde{\boldsymbol{\Lambda}})}{\sqrt{2\mathrm{tr}(\mathbf{B}^{*2}_{\boldsymbol{X}})\mathrm{tr}(\mathbf{B}^{*2}_{\boldsymbol{Y}})}}\right),$$

$$\beta_{\mathrm{HT}}(\mathbf{M}_1,\mathbf{M}_2) = \Phi\left(-z_\alpha + \frac{n\mathrm{tr}(\tilde{\boldsymbol{\Lambda}}^\top\tilde{\boldsymbol{\Lambda}})}{\sqrt{2\mathrm{tr}(\mathbf{B}^{*2}_{\boldsymbol{X}})\mathrm{tr}(\mathbf{B}^{*2}_{\boldsymbol{Y}})}}\right),$$

respectively, where z_α is α-quantile of the standard normal distribution. In addition, the power function of the testing method proposed in [28] (abbreviated as EC) is

$$\beta_{\mathrm{EC}}(\mathbf{M}_1,\mathbf{M}_2) = \Phi\left(-z_\alpha + \frac{n\mathrm{tr}(\boldsymbol{\Sigma}^\top_{\boldsymbol{XY}}\boldsymbol{\Sigma}_{\boldsymbol{XY}})}{\sqrt{2\mathrm{tr}(\boldsymbol{\Sigma}^{*2}_{\boldsymbol{X}})\mathrm{tr}(\boldsymbol{\Sigma}^{*2}_{\boldsymbol{Y}})}}\right),$$

where $\boldsymbol{\Sigma}_{\boldsymbol{XY}}$ is the covariance matrix between $\boldsymbol{X}$ and $\boldsymbol{Y}$. Therefore, the asymptotic relative efficiencies (AREs) of the proposed tests with respect to EC are

$$\mathrm{ARE(HS, EC)} = \frac{\mathrm{tr}(\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda})}{\mathrm{tr}(\boldsymbol{\Sigma}^\top_{\boldsymbol{XY}}\boldsymbol{\Sigma}_{\boldsymbol{XY}})}\sqrt{\frac{\mathrm{tr}(\boldsymbol{\Sigma}^{*2}_{\boldsymbol{X}})\mathrm{tr}(\boldsymbol{\Sigma}^{*2}_{\boldsymbol{Y}})}{\mathrm{tr}(\mathbf{A}^{*2}_{\boldsymbol{X}})\mathrm{tr}(\mathbf{A}^{*2}_{\boldsymbol{Y}})}},\quad \mathrm{ARE(HR, EC)} = \mathrm{ARE(HT, EC)} = \frac{\mathrm{tr}(\tilde{\boldsymbol{\Lambda}}^\top\tilde{\boldsymbol{\Lambda}})}{\mathrm{tr}(\boldsymbol{\Sigma}^\top_{\boldsymbol{XY}}\boldsymbol{\Sigma}_{\boldsymbol{XY}})}\sqrt{\frac{\mathrm{tr}(\boldsymbol{\Sigma}^{*2}_{\boldsymbol{X}})\mathrm{tr}(\boldsymbol{\Sigma}^{*2}_{\boldsymbol{Y}})}{\mathrm{tr}(\mathbf{B}^{*2}_{\boldsymbol{X}})\mathrm{tr}(\mathbf{B}^{*2}_{\boldsymbol{Y}})}},$$

and

$$\mathrm{ARE(HS, HR)} = \frac{\mathrm{tr}(\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda})}{\mathrm{tr}(\tilde{\boldsymbol{\Lambda}}^\top\tilde{\boldsymbol{\Lambda}})}\sqrt{\frac{\mathrm{tr}(\mathbf{B}^{*2}_{\boldsymbol{X}})\mathrm{tr}(\mathbf{B}^{*2}_{\boldsymbol{Y}})}{\mathrm{tr}(\mathbf{A}^{*2}_{\boldsymbol{X}})\mathrm{tr}(\mathbf{A}^{*2}_{\boldsymbol{Y}})}},\quad \mathrm{ARE(HT, HR)} = 1.$$

To clearly show the relations among HS, HR, HT, and EC, we consider the special case of $\boldsymbol{X}$ and $\boldsymbol{Y}$: $p = q$, $g_{X^*} = g_{Y^*}$, $\boldsymbol{\Omega}_{X^*} = \mathbf{I}_p$, $\boldsymbol{\Omega}_{Y^*} = \mathbf{I}_q$. Now, the power function of T_{HS}, T_{HR}, T_{HT} and T_{EC} becomes

$$\beta_{\mathrm{HS}}(\mathbf{M}_1, \mathbf{M}_2) = \Phi\left(-z_\alpha + \frac{n\{\mathrm{E}((r_i^{X^*})^{-1})\mathrm{E}(r_i^{X^*})\}^2 \mathrm{tr}\{(\mathbf{M}_1 + \mathbf{M}_2^\top)^\top(\mathbf{M}_1 + \mathbf{M}_2^\top)\}}{\sqrt{2}p}\right),$$
$$\beta_{\mathrm{HR}}(\mathbf{M}_1, \mathbf{M}_2) = \Phi\left(-z_\alpha + \frac{n\{\mathrm{E}((\tilde{r}_i^{X^*})^{-1})\mathrm{E}(\tilde{r}_i^{X^*})\}^2 \mathrm{tr}\{(\mathbf{M}_1 + \mathbf{M}_2^\top)^\top(\mathbf{M}_1 + \mathbf{M}_2^\top)\}}{\sqrt{2}p}\right),$$
$$\beta_{\mathrm{HT}}(\mathbf{M}_1, \mathbf{M}_2) = \Phi\left(-z_\alpha + \frac{n\{\mathrm{E}((\tilde{r}_i^{X^*})^{-1})\mathrm{E}(\tilde{r}_i^{X^*})\}^2 \mathrm{tr}\{(\mathbf{M}_1 + \mathbf{M}_2^\top)^\top(\mathbf{M}_1 + \mathbf{M}_2^\top)\}}{\sqrt{2}p}\right),$$
$$\beta_{\mathrm{EC}}(\mathbf{M}_1, \mathbf{M}_2) = \Phi\left(-z_\alpha + \frac{n\mathrm{tr}\{(\mathbf{M}_1 + \mathbf{M}_2^\top)^\top(\mathbf{M}_1 + \mathbf{M}_2^\top)\}}{\sqrt{2}p}\right).$$

Accordingly,

$$\mathrm{ARE}(\mathrm{HS}, \mathrm{EC}) = \{\mathrm{E}((r_i^{X^*})^{-1})\mathrm{E}(r_i^{X^*})\}^2 \geq 1, \quad \mathrm{ARE}(\mathrm{HR}, \mathrm{EC}) = \{\mathrm{E}((\tilde{r}_{ij}^{X^*})^{-1})\mathrm{E}(\tilde{r}_{ij}^{X^*})\}^2 \geq 1,$$
$$\mathrm{ARE}(\mathrm{HT}, \mathrm{EC}) = \{\mathrm{E}((\tilde{r}_{ij}^{X^*})^{-1})\mathrm{E}(\tilde{r}_{ij}^{X^*})\}^2 \geq 1,$$
$$\mathrm{ARE}(\mathrm{HS}, \mathrm{HR}) = \{\mathrm{E}((r_i^{X^*})^{-1})\mathrm{E}(r_i^{X^*})\}^2\{\mathrm{E}((\tilde{r}_{ij}^{X^*})^{-1})\mathrm{E}(\tilde{r}_{ij}^{X^*})\}^{-2} \to 1, \quad \mathrm{ARE}(\mathrm{HT}, \mathrm{HR}) = 1,$$

where the three inequalities are followed by the Cauchy inequality and the convergence is followed by Lemma 1 in [4].

4. Simulation study

We now present simulation results to demonstrate the performance of the proposed tests HS, HR, HT, and compare them with four existing tests proposed by [20], [28], [10], [2], abbreviated as CS, EC, TJ, LH, respectively. Note that all simulation results are obtained based on 2,500 replications. We consider the following three commonly studied simulation settings:

(I) Multivariate normal distribution, $\boldsymbol{X}^* \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}_{X^*})$ and $\boldsymbol{Y} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}_{Y})$;

(II) Multivariate t-distribution, $\boldsymbol{X}^* \sim t_p(\mathbf{0}, \boldsymbol{\Sigma}_{X^*}, 3)$ and $\boldsymbol{Y} \sim t_q(\mathbf{0}, \boldsymbol{\Sigma}_{Y}, 3)$;

(III) Multivariate mixture normal distribution, X_i^*'s are generated from $\mathcal{MN}_{p,\gamma,9}(\mathbf{0}, \boldsymbol{\Sigma}_{X^*}) \doteq \gamma\mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}_{X^*}) + (1 - \gamma)\mathcal{N}(\mathbf{0}, 9\boldsymbol{\Sigma}_{X^*})$, where γ is chosen to be 0.8. Similarly, $\boldsymbol{Y}_i$'s are generated from $\mathcal{MN}_{q,\gamma,9}(\mathbf{0}, \boldsymbol{\Sigma}_{Y})$.

For these three settings, $\boldsymbol{X}^*$ and $\boldsymbol{Y}$ are independent, $\boldsymbol{\Sigma}_{X^*} = (0.5^{|i-j|})_{1\leq i,j\leq p}$ and $\boldsymbol{\Sigma}_{Y} = (0.5^{|i-j|})_{1\leq i,j\leq q}$.

First, we consider the low-dimension case in which $p \leq q < n$. Let $n = 100$ and $p = q \in \{10, 20\}$. For power comparison, we consider four alternative settings. The first is as follows.

(i) $\boldsymbol{X}_i = \boldsymbol{X}_i^* + n^{-1/2}\nu\boldsymbol{W}_i$, where W_i is composed of the first few p variables of $\boldsymbol{Y}_i$. When $\boldsymbol{X}^*$ and $\boldsymbol{Y}$ are generated from settings (I)-(III), we labeled these settings under the alternative hypothesis as (I-i), (II-i), and (III-i), respectively.

If $\nu = 0$, then $\boldsymbol{X}_i$ is independent of $\boldsymbol{Y}_i$, while if ν is large, $\boldsymbol{X}_i$ would be strongly correlated with $\boldsymbol{Y}_i$. Let $\nu \in \{0, 1.5, 2\}$. **Table 1** reports the empirical sizes and power of these seven methods for testing independence between $\boldsymbol{X}$ and $\boldsymbol{Y}$. For setting (I), under the normal model, all seven methods have similar performances. For settings (II) and (III), as non-normal models are used to generate data, HS, HR, HT, and EC have better performance in controlling the empirical sizes.

Let $(p, q) \in \{(80, 100), (160, 200), (320, 400), (640, 800), (800, 1000)\}$ and $n \in \{30, 50, 100\}$. Then, we consider a high-dimension case in which $n < p \leq q$. Data for this case are generated in a manner similar to that described above. Since LH is not designed for data with particularly large dimensions and TJ fails to control the size in non-normal situations, we exclude them from comparison. **Tables 2-4** summarize the empirical sizes and power of the methods for settings (I)-(III), respectively. These tables suggest that for setting (I), under high-dimensional normal models, all

Table 1: The empirical sizes and power of the involved tests for independence between $\boldsymbol{X}$ and $\boldsymbol{Y}$ in low dimensional cases, where the data are generated via setting (i) of the alternative and settings (I), (II), (III) of the distribution. $n = 100$, $p = q \in \{10, 20\}$ and $\nu \in \{0, 1.5, 2\}$.

	$(p,q) = (10,10)$							$(p,q) = (20,20)$						
(p,q)	HR	HT	HS	CS	EC	LH	TJ	HR	HT	HS	CS	EC	LH	TJ
(I) Multivariate normal distribution, $n = 100$														
$\nu = 0$	4.7	5.3	6.7	5.4	7.2	5.6	4.1	5.8	4.6	6.1	6.2	4.8	7.1	4.8
$\nu = 1.5$	38	37	35	38	42	41	28	39	40	39	42	44	36	25
$\nu = 2$	65	66	62	68	80	74	63	68	69	68	73	80	66	55
(II) Multivariate t-distribution, $n = 100$														
$\nu = 0$	5.8	5.4	6.5	12	6.1	15	16	4.3	4.5	5.5	14	5.5	25	22
$\nu = 1.5$	54	52	53	43	48	54	49	57	56	57	43	43	65	54
$\nu = 2$	81	82	82	66	77	85	76	88	87	88	63	75	87	77
(III) Multivariate mixture normal distribution, $n = 100$														
$\nu = 0$	6.2	6.1	6.7	11	5.0	13	100	5.8	4.2	5.5	16	6.0	26	100
$\nu = 1.5$	51	51	50	42	44	54	100	57	56	56	45	48	62	100
$\nu = 2$	79	79	80	69	78	79	100	86	87	87	69	78	82	100

five methods have similar performance; for non-normal settings (II) and (III), HS, HR, HT and EC perform better in controlling the empirical sizes than CS. Furthermore, HS, HR and HT outperform EC in the power comparison.

Furthermore, we investigate the performance of the proposed tests in a situation in which one of the two sets of variables has a low dimension while the other has a high dimension. Specifically, we let $p = 5$ and $q \in \{100, 200, 400\}$ with $n = 100$. The corresponding results are summarized in **Table 5**, where the proposed tests have similar performance to the EC test and the size of the CS test is also out of control in non-normal situations.

Next, we consider the second setting of the alternative.

(ii) This setting is the same as setting (i) except for the construction of $\boldsymbol{X}_i$. Specifically, $\boldsymbol{X}_i = \boldsymbol{X}_i^* + n^{-1/2}\nu\boldsymbol{W}_i$, where $\boldsymbol{W}_i$ is composed of the first few $p/2$ variables of $\boldsymbol{Y}_i$. When $\boldsymbol{X}^*$ and $\boldsymbol{Y}$ are generated from settings (I)-(III), we labeled these settings under the alternative hypothesis as (I-ii), (II-ii), (III-ii), respectively.

For this setting, we let $n = 100$ and $(p,q) \in \{(80,100), (160,200), (320,400), (640,800), (800,1000)\}$. As suggested by the above results of setting (i), the size of the CS test is often out of control, especially for non-normal distributions, and the size performance of the remaining tests are very similar. Hence, we exclude the CS test and the size results in the following comparison. The corresponding results are summarized in **Table 6**, which are very similar to the above results for setting (i).

Finally, we consider the remaining two settings of the alternative as follows.

(iii) Multivariate t-distribution, $\boldsymbol{Z}_i = (\boldsymbol{X}_i^\top, \boldsymbol{Y}_i^\top)^\top$, where $\boldsymbol{Z}_i \sim t_{p+q}(\boldsymbol{0}, \boldsymbol{\Sigma}_{\boldsymbol{Z}}, 3)$.

(iv) Multivariate mixture normal distribution, $\boldsymbol{Z}_i = (\boldsymbol{X}_i^\top, \boldsymbol{Y}_i^\top)^\top$, where $\boldsymbol{Z}_i \sim \mathcal{MN}_{p+q,\gamma,9}(\boldsymbol{0}, \boldsymbol{\Sigma}_{\boldsymbol{Z}})$.

Here, $\boldsymbol{\Sigma}_{\boldsymbol{Z}} = (a_{ij})_{1\le i,j\le p+q}$, $a_{ii} = 1$, $a_{ij} = \rho = n^{-1}$ for $i \neq j$. The difference between settings (iii), (iv) and settings (i), (ii) is whether the joint distribution of $\boldsymbol{X}_i$ and $\boldsymbol{Y}_i$ is considered. In particular, in settings (iii) and (iv), the joint distributions of $\boldsymbol{X}_i$ and $\boldsymbol{Y}_i$ are set to be multivariate t-distribution and multivariate mixture normal distribution, respectively, which are members of the family of elliptically symmetric distributions. The corresponding results are summarized in **Table 7** and suggest that HS is the most powerful of these involved tests. On the other hand, HR and HT still perform similarly to each other and outperform EC in most cases.

In summary, the simulation results show that the three proposed methods are more powerful than existing popular testing procedures, especially for high-dimensional and heavy-tailed data.

Table 2: The empirical sizes and power of the involved tests for testing independence between $\boldsymbol{X}$ and $\boldsymbol{Y}$ in high dimensional cases, where the data are generated via setting (i) of the alternative and setting (I) of the distribution. $n \in \{30, 50, 100\}$, $(p, q) \in \{(80, 100), \cdots, (800, 1000)\}$ and $\nu \in \{0, 1.5, 2\}$.

	$\nu = 0$					$\nu = 1.5$					$\nu = 2$				
(p, q)	HR	HT	HS	CS	EC	HR	HT	HS	CS	EC	HR	HT	HS	CS	EC
							$n = 30$								
(80,100)	6.2	6.0	5.8	5.9	6.1	37	37	37	37	38	70	70	67	68	70
(160,200)	4.8	4.9	5.2	3.8	5.2	33	33	33	34	36	68	67	62	67	71
(320,400)	5.4	5.1	7.5	6.6	5.8	37	37	28	35	38	72	72	48	65	68
(640,800)	5.3	5.6	9.2	6.7	5.1	35	35	29	34	37	71	71	50	64	67
(800,1000)	5.2	5.4	9.8	5.7	5.4	35	35	30	33	36	72	71	51	65	66
							$n = 50$								
(80,100)	5.7	5.4	5.3	5.2	4.8	38	38	38	38	40	71	72	71	73	71
(160,200)	5.3	4.9	5.0	4.7	5.7	37	37	38	36	36	73	72	73	73	76
(320,400)	5.7	5.4	5.1	5.6	5.3	43	42	43	43	38	72	73	73	73	77
(640,800)	5.3	5.8	6.3	5.7	5.1	42	42	42	41	37	72	72	72	71	76
(800,1000)	5.4	5.6	6.8	5.4	5.3	43	42	43	41	38	72	73	72	71	77
							$n = 100$								
(80,100)	5.4	5.2	4.9	5.4	4.9	39	40	40	40	41	74	74	74	74	76
(160,200)	4.9	5.1	5.2	5.1	6.2	43	43	44	43	36	75	75	75	76	76
(320,400)	6.3	5.8	5.7	4.9	6.8	41	40	40	41	40	82	81	82	82	78
(640,800)	5.8	5.7	5.9	4.8	6.3	40	40	40	40	39	81	81	81	82	79
(800,1000)	5.7	5.8	6.3	4.9	6.3	42	40	41	41	39	83	81	82	82	80

5. Empirical application

5.1. Dependence between US and Chinese stock markets

Much research has analyzed correlations between global financial markets, especially in some special periods such as financial crisis. For example, Sunil and Nivedita [21] used correlation and network methods to investigate the effect of important financial indices on the organization structure; Vodenska et al. [24] used network theory and community analysis to understand the structure of the coupled financial network formed by global stock market indices and currencies; Junior and Franca [11] and Sensoy et al. [19] analyzed the cross-correlation matrix of index returns of the main financial markets after the 2007-2009 crisis using random matrix theory methods.

In this section, we conduct a correlation study of the stocks from the S&P500 index and the CSI 300 index as an example to investigate the relationship between US and Chinese financial markets. Specifically, we use the proposed HS method to test independence between the weekly return rate vector of the stocks from the S&P500 index, denoted as $\boldsymbol{X}$, and that of the stocks from the CSI300 index, denoted as $\boldsymbol{Y}$, where the weekly return rate vectors of the stocks from the two indices at different weeks are considered to be i.i.d. observations of $\boldsymbol{X}$ and $\boldsymbol{Y}$, respectively.

We test the independence between $\boldsymbol{X}$ and $\boldsymbol{Y}$ using observations from January 2005 to November 2018. Considering the timeliness of stock analysis, we use one consecutive year as a sliding window, take one week as a step, and then successively test the independence between $\boldsymbol{X}$ and $\boldsymbol{Y}$ using the observations within the sliding window.

In **Fig. 1**, we present the resulting p-value sequence of the independence test between $\boldsymbol{X}$ and $\boldsymbol{Y}$ for all the sliding windows, where each p-value in the sequence at week t corresponds to a one-consecutive-year sliding window from week $t - 52$ to week t. From **Fig. 1**, it can be seen that for most of the time, $\boldsymbol{X}$ and $\boldsymbol{Y}$ are judged as independent under both significance levels of 0.01 and 0.05.

In **Fig. 2**, we present two time series of the prices of the CSI300 index and the S&P500 index, respectively. To build the connection between **Fig. 2** and **Fig. 1**, we draw a vertical red line at each week that corresponds to a p-value smaller than 0.01, that is, the p-value of the independence test between $\boldsymbol{X}$ and $\boldsymbol{Y}$ using observations during the week as well as in the first 52 weeks. The vertical red lines are mainly divided into three parts: 2010-2012, 2016-2017, and 2018-present. In the first part, the two indices show similar trends within some sub-parts; in the second part, they grew

Table 3: The empirical sizes and power of the involved tests for testing independence between $\boldsymbol{X}$ and $\boldsymbol{Y}$ in high dimensional cases, where the data are generated via setting (i) of the alternative and setting (II) of the distribution. $n \in \{30, 50, 100\}$, $(p, q) \in \{(80, 100), \cdots, (800, 1000)\}$ and $\nu \in \{0, 1.5, 2\}$.

	$\nu = 0$					$\nu = 1.5$					$\nu = 2$				
(p, q)	HR	HT	HS	CS	EC	HR	HT	HS	CS	EC	HR	HT	HS	CS	EC
								$n = 30$							
(80,100)	5.2	5.4	5.4	22	6.8	53	54	54	46	45	83	82	82	62	67
(160,200)	4.9	4.9	4.2	22	8.1	54	54	54	48	37	81	82	82	62	67
(320,400)	5.2	5.0	8.6	22	5.8	55	55	57	45	41	80	80	82	61	65
(640,800)	4.9	5.1	9.7	22	5.9	55	55	56	43	40	81	80	82	62	64
(800,1000)	5.6	5.4	10	23	5.8	55	56	57	43	40	80	82	81	64	63
								$n = 50$							
(80,100)	5.6	6.4	5.3	20	5.8	59	59	58	43	39	87	88	87	58	65
(160,200)	5.3	5.3	5.0	21	6.8	62	62	61	43	39	87	87	87	59	69
(320,400)	5.8	5.2	7.8	21	5.7	67	67	66	44	44	87	88	86	58	67
(640,800)	5.7	5.1	8.3	21	5.2	66	66	66	43	40	87	88	85	57	68
(800,1000)	5.4	5.4	8.9	22	5.4	67	66	67	42	41	88	88	86	58	68
								$n = 100$							
(80,100)	5.2	4.5	4.9	23	5.0	62	62	61	43	46	90	90	91	59	71
(160,200)	4.8	5.8	5.2	25	5.1	63	64	64	42	44	92	93	92	54	72
(320,400)	6.0	4.8	5.7	24	4.8	66	67	66	38	44	97	98	97	50	71
(640,800)	5.8	5.1	5.4	24	5.3	66	66	66	39	44	96	96	96	51	70
(800,1000)	5.9	5.5	5.5	25	5.2	65	65	66	39	43	96	96	97	50	70

simultaneously; and in the last part, their trends are just the opposite. **Fig. 3** and **Fig. 4** suggest that in the above three parts with p-values smaller than 0.01, the time series of the return rates of the two indices as well as the differences between the two time series have relatively small fluctuations.

Based on the above results, the stock markets of the two countries are considered independent for most of the time, except for some special periods. With the financial crisis in 2007-2009, the global economic recovery brought about the growth of both the US and Chinese stock markets. At that time, the financial markets of the two countries showed a strong correlation. This may be the reason for the correlation in 2010-2012. On the other hand, after the establishment of the Shanghai-Hong Kong Stock Exchange in 2015, international capital was able to enter China's stock market in large quantities, which will had a significant impact on the Chinese stock market. This may be the reason for the correlation in 2016-2017 and 2018-present.

Finally, as the remaining two testing methods proposed in this paper obtain very similar conclusions, they are not presented in this paper. Moreover, we use a nonparametric method in this study to analyze the stock return rate data because most of the stocks involved have non-normal distributions for their weekly return rates, especially for the stocks from CSI300 index, which is suggested by **Fig. 5**.

5.2. Dependence between the Shanghai and Shenzhen Stock Exchanges

We compiled monthly returns on all the securities in Chinese stock markets that have been listed from June 2005 to May 2019. Because the securities listed in Chinese stock markets change over time, we only consider $p + q = 1340$ securities that were listed throughout the entire period. There are $p = 559$ securities in the Shenzhen Stock Market and $q = 781$ securities in the Shanghai Stock Exchange. From June 2005 to May 2019, a total of $n = 144$ consecutive observations were obtained.

First, we test whether the stocks in the Shenzhen Stock Exchange are independent from the stocks in the Shanghai Stock Exchange. Since HS, HT and HR have very similar performance, below, we only compare EC with HS for this real data analysis. The obtained test statistics of EC and HS are 100.46 and 98.56, respectively, based on which both methods reject the null hypothesis. Thus, we believe that the stock return vectors of the two markets are deeply dependent on each other.

Table 4: The empirical sizes and power of the involved tests for testing independence between $\boldsymbol{X}$ and $\boldsymbol{Y}$ in high dimensional cases, where the data are generated via setting (i) of the alternative and setting (III) of the distribution. $n \in \{30, 50, 100\}$, $(p, q) \in \{(80, 100), \cdots, (800, 1000)\}$ and $v \in \{0, 1.5, 2\}$.

	$v = 0$					$v = 1.5$					$v = 2$				
(p, q)	HR	HT	HS	CS	EC	HR	HT	HS	CS	EC	HR	HT	HS	CS	EC
							$n = 30$								
(80,100)	5.7	5.0	6.3	26	7.3	54	55	55	46	40	82	82	82	60	68
(160,200)	4.3	4.4	5.6	31	7.5	57	58	57	45	41	87	86	86	61	68
(320,400)	5.8	4.9	6.5	34	8.0	61	61	60	49	39	83	83	82	61	66
(640,800)	5.4	5.1	7.3	34	8.2	60	61	60	48	40	82	82	83	60	65
(800,1000)	5.3	5.4	7.6	33	8.1	60	60	60	47	41	83	82	82	61	66
							$n = 50$								
(80,100)	4.7	4.7	5.9	29	5.8	58	57	57	46	45	87	88	86	62	69
(160,200)	5.4	5.9	4.4	30	4.5	59	59	59	45	40	90	90	89	58	72
(320,400)	4.7	5.4	6.5	32	5.1	66	65	65	44	40	92	91	92	57	72
(640,800)	5.5	5.3	6.9	31	4.9	65	65	65	43	41	91	91	92	56	70
(800,1000)	5.1	5.6	7.1	31	5.8	66	65	67	43	42	90	90	91	57	71
							$n = 100$								
(80,100)	6.1	4.9	4.9	29	5.3	60	60	59	44	37	93	93	92	58	75
(160,200)	4.9	5.3	6.4	35	5.4	62	62	63	45	42	94	93	92	54	77
(320,400)	5.3	4.8	6.0	36	6.8	67	67	66	44	45	92	94	93	54	75
(640,800)	5.1	4.9	5.7	35	6.2	66	66	65	45	45	93	93	93	53	76
(800,1000)	5.0	5.6	5.4	33	5.9	65	65	65	44	46	92	92	94	53	75

It is well-known that the stocks are correlated because they have many common factors. Hence, to remove the influence of these common factors, we consider the following Fama-French three-factor model

$$Z_{ij} = r_{ij} - rf_i = \alpha_j + \beta_{j1}(rm_i - rf_i) + \beta_{j2}SMB_i + \beta_{j3}HML_i + \epsilon_{ij},$$

for $j \in \{1, \cdots, p+q\}$ and $i \in \{1, \cdots, n\}$, where $\{1, \cdots, p\}$ corresponds to the stocks in the Shenzhen Stock Market and $\{p+1, \cdots, p+q\}$ corresponds to the stocks in the Shanghai Stock Exchange. The rate of 10-year Chinese Treasuries is chosen as the risk-free rate (rf_j) for each stock j. The value-weighted return on all the stocks of the Shanghai Stock Exchange and the Shenzhen Stock Exchange is used as a proxy for the market return (rm_i). The average return on the three small portfolios minus the average return on the three big portfolios (SMB_i), and the average return on two value portfolios minus the average return on two growth portfolios (HML_i) are calculated based on the stocks listed on the Shanghai Stock Exchange and the Shenzhen Stock Exchange. We use r_{ij} to denote the return rate of security j on time i. All data are measured in percent per month.

We remove the common factors as follows. Let

$$X_{ij} \doteq Z_{ij} - \left(\hat{\alpha}_j + \hat{\beta}_{j1}(rm_i - rf_i) + \hat{\beta}_{j2}SMB_i + \hat{\beta}_{j3}HML_i\right)$$

for $i \in \{1, \cdots, n\}$ and $j \in \{1, \cdots, p\}$; and for $i \in \{1, \cdots, n\}$ and $j' \in \{1, \cdots, q\}$, let

$$Y_{ij'} \doteq Z_{i,j'+p} - \left(\hat{\alpha}_{j'+p} + \hat{\beta}_{j'+p,1}(rm_i - rf_i) + \hat{\beta}_{j'+p,2}SMB_i + \hat{\beta}_{j'+p,3}HML_i\right).$$

Here $\hat{\alpha}_j, \hat{\beta}_{j1}, \hat{\beta}_{j2}$ and $\hat{\beta}_{j3}$ are the estimations of $\alpha_j, \beta_{j1}, \beta_{j2}$ and β_{j3} under the Fama-French three-factor model. We then consider the null hypothesis of the independence of the $\boldsymbol{X}$- and $\boldsymbol{Y}$-variables: the stocks in the Shenzhen stock market are independent from the stocks in the Shanghai Stock Exchange.

We apply EC and HS to the data of X_{ij}'s and $Y_{ij'}$'s, and the test statistics are 56.64 and 52.90, respectively, based on which the null hypothesis is still rejected. Thus, we still believe that the two markets are deeply dependent on each other. To make the advantage of HS explicit, we adopt a random sampling procedure. In particular, we randomly

Table 5: The empirical sizes and power of the involved tests for testing independence between $\boldsymbol{X}$ and $\boldsymbol{Y}$ in cases of low p with large q, where the data are generated via setting (i) of the alternative and settings (I), (II), (III) of the distribution. $n = 100$, $(p, q) \in \{(5, 100), (5, 200), (5, 400)\}$ and $\nu \in \{0, 3, 6\}$.

	$\nu = 0$					$\nu = 3$					$\nu = 6$				
(p, q)	HR	HT	HS	CS	EC	HR	HT	HS	CS	EC	HR	HT	HS	CS	EC
Multivariate normal distribution, setting (I-i)															
(5,100)	5.3	5.6	5.0	3.8	6.6	29	28	30	33	36	92	90	93	97	97
(5,200)	5.2	5.7	5.0	5.0	4.8	18	18	19	21	25	75	75	77	83	85
(5,400)	4.7	5.3	5.4	3.8	8.1	14	15	15	15	16	60	60	62	58	66
Multivariate t distribution, setting (II-i)															
(5,100)	4.9	5.8	5.6	13	5.3	38	38	40	54	38	93	94	95	93	86
(5,200)	5.3	4.6	4.8	14	6.0	26	26	27	52	30	79	80	80	92	68
(5,400)	5.6	5.4	4.6	14	6.2	21	20	22	50	24	57	56	58	87	54
Multivariate mixture normal distribution, setting (III-i)															
(5,100)	5.9	5.8	5.4	14	5.7	37	38	39	48	39	94	94	95	98	90
(5,200)	4.2	5.6	5.2	16	5.5	25	25	26	41	26	78	76	78	96	76
(5,400)	5.6	4.6	4.8	15	5.6	20	20	21	38	20	57	58	58	94	60

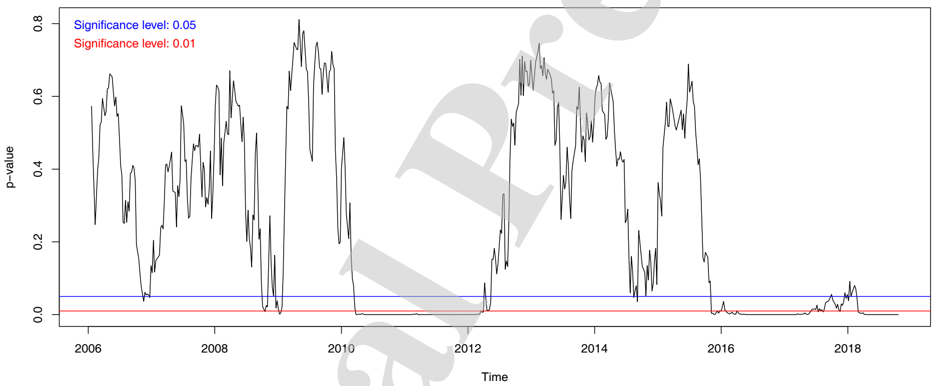

Fig. 1: p-value series of sliding annual independence test between return rates of two groups of stocks in CSI500 index and the S&P500 index, respectively.

select n' months from the total $n = 144$ months, p' stocks from the $p = 559$ stocks in the Shenzhen Stock Exchange, and q' stocks from the $q = 781$ stocks in the Shanghai Stock Exchange. **Table 8** reports the power of these two tests in a different setting of n', p' and q', where for each setting of n', p' and q' we perform random sampling 1,000 times. We observe that HS is more powerful than EC for each setting, which may be due to the heavy-tailed distributions of the return data, presented in **Fig. 6**, as well as the high dimensionality.

6. Conclusions

We have proposed three high-dimensional nonparametric independence tests based on the spatial sign and ranks that provide more powerful alternatives to the widely studied multivariate normal theory methods. The power superiority of the three proposed tests in comparison with existing test procedures is especially clear for high-dimensional and heavy-tailed data, as shown by numerical evidence as well as two real data analyses on stock return rate data.

Choosing the appropriate test in practical applications depends on the distribution of the practical data and the dependence between the two high-dimensional random vectors. The proposed tests are advantageous in non-normal

Table 6: The empirical power of the involved tests for testing independence between $\boldsymbol{X}$ and $\boldsymbol{Y}$, where the data are generated via setting (ii) of the alternative and settings (I), (II), (III) of the distribution. $n = 100$, $(p, q) \in \{(80, 100), (160, 200), (320, 400)\}$ and $\nu \in \{2, 3\}$.

	$\nu = 2$				$\nu = 3$			
(p, q)	HR	HT	HS	EC	HR	HT	HS	EC
setting (I-ii)								
(80,100)	33	33	33	37	81	81	81	82
(160,200)	34	34	35	37	84	84	85	81
(320,400)	34	34	35	35	85	84	85	85
setting (II-ii)								
(80,100)	55	55	56	35	92	93	94	76
(160,200)	55	56	57	39	95	95	96	78
(320,400)	57	58	60	35	95	96	96	78
setting (III-ii)								
(80,100)	50	50	51	35	93	94	94	79
(160,200)	53	53	54	33	96	95	96	81
(320,400)	56	56	57	35	97	96	97	81

Table 7: The empirical power of the involved tests for testing independence between $\boldsymbol{X}$ and $\boldsymbol{Y}$, where the data are generated via settings (iii) and (iv). $n = 100$ and $(p, q) \in \{(160, 200), (320, 400), (640, 800)\}$.

	$n = 30$				$n = 50$				$n = 100$			
(p, q)	HR	HT	HS	EC	HR	HT	HS	EC	HR	HT	HS	EC
(iii)												
(160,200)	78	77	100	49	60	60	100	58	36	36	93	60
(320,400)	96	95	100	51	89	88	100	58	73	72	100	57
(640,800)	100	100	100	52	100	100	95	56	98	98	100	59
(iv)												
(160,200)	81	82	95	50	60	60	100	57	35	34	72	60
(320,400)	96	96	83	51	91	90	98	61	70	70	99	62
(640,800)	100	100	80	49	100	100	98	58	99	99	100	62

situations when testing whether linear dependence exists between two high-dimensional random vectors. In comparing these proposed tests, when the dimension is larger than the square of sample sizes, the spatial rank-based tests generally have better performance in controlling the size than the spatial sign-based tests, which are, however, much more time consuming due to the more complex statistics. Hence, the spatial sign-based test is preferable unless the dimensionality is very large.

Acknowledgments

We thank the Editor, Associate Editor and referees for their insightful and helpful comments, as well as our financial sponsors: NSFC grants 11571068, 11501092, 11631003, the Fundamental Research Funds for the Central Universities grant 2412017BJ002, the Key Laboratory of Applied Statistics of MOE (KLAS) grants 130026507 and 130028612.

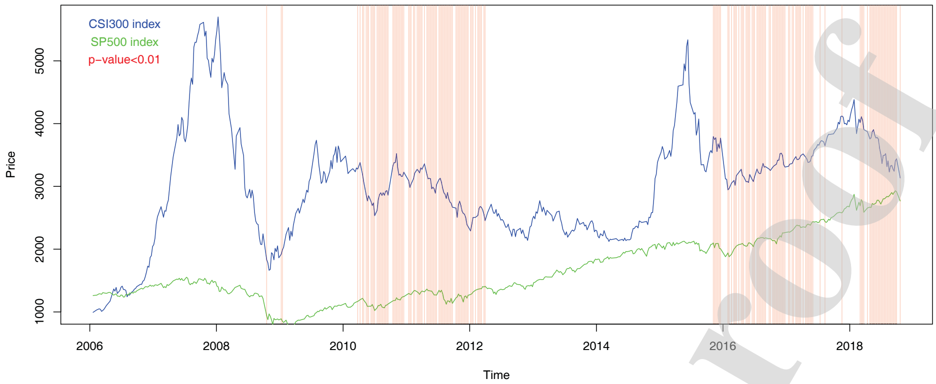

Fig. 2: Time series of the prices of the CSI300 index and the S&P500 index, respectively.

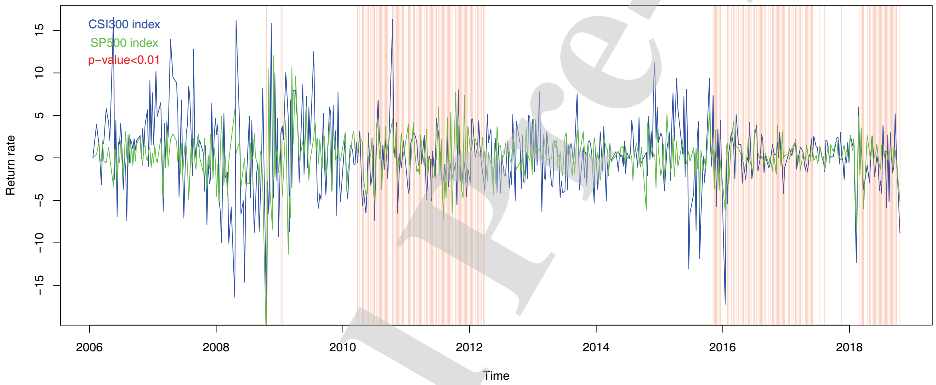

Fig. 3: Time series of the return rates of the CSI300 index and the S&P500 index, respectively.

Appendix

Proof of Theorem 1

Let $\mathbf{U}_i^X = \mathbf{U}(\boldsymbol{X}_i - \boldsymbol{\theta}_X)$, $\mathbf{U}_i^Y = \mathbf{U}(Y_i - \boldsymbol{\theta}_Y)$, $r_i^X = \|\boldsymbol{X}_i - \hat{\boldsymbol{\theta}}_X\|$, $r_i^Y = \|\boldsymbol{Y}_i - \hat{\boldsymbol{\theta}}_Y\|$, $\hat{\boldsymbol{\mu}}_X = \hat{\boldsymbol{\theta}}_X - \boldsymbol{\theta}_X$, $\hat{\boldsymbol{\mu}}_Y = \hat{\boldsymbol{\theta}}_Y - \boldsymbol{\theta}_Y$, $\mathbf{A}_X = \mathrm{E}(\mathbf{U}_i^X(\mathbf{U}_i^X)^\top)$ and $\mathbf{A}_Y = \mathrm{E}(\mathbf{U}_i^Y(\mathbf{U}_i^Y)^\top)$. Before presenting the proof of the main theorems, we propose some necessary lemmas. First, we recall Lemma 1 in [25] as follows.

Lemma 1. *Under conditions (C1) and (C2) given in Section 2,* $\mathrm{E}\left((\mathbf{U}_i^X)^\top\mathbf{U}_j^X\right)^4 = O(1)\left[\mathrm{E}((\mathbf{U}_i^X)^\top\mathbf{U}_j^X)^2\right]^2$, $\mathrm{E}((\mathbf{U}_i^X)^\top\mathbf{A}_X\mathbf{U}_i^X)^2 = O(1)\left[\mathrm{E}((\mathbf{U}_i^X)^\top\mathbf{A}_X\mathbf{U}_i^X)\right]^2$ *and* $\mathrm{E}((\mathbf{U}_i^X)^\top\mathbf{A}_X\mathbf{U}_j^X)^2 = O(1)\left[\mathrm{E}((\mathbf{U}_i^X)^\top\mathbf{A}_X\mathbf{U}_j^X)\right]^2$.

Let $T_r = (n-1)^{-1}\sum_{1\le i<j\le n}(\hat{\mathbf{U}}_i^X)^\top\hat{\mathbf{U}}_j^X(\hat{\mathbf{U}}_i^Y)^\top\hat{\mathbf{U}}_j^Y$ and $\hat{\sigma}_1^2 = 2\{n^2(n-1)^2\}^{-1}\sum_{1\le i<j\le n}((\hat{\mathbf{U}}_i^X)^\top\hat{\mathbf{U}}_j^X)^2\sum_{1\le i<j\le n}((\hat{\mathbf{U}}_i^Y)^\top\hat{\mathbf{U}}_j^Y)^2$. Then $T_{\mathrm{HS}} = T_r/\hat{\sigma}_1$. Let $\sigma_1^2 = n\{2(n-1)\}^{-1}\mathrm{tr}(\mathbf{A}_X^2)\mathrm{tr}(\mathbf{A}_Y^2)$. To prove Theorem 1, we only need to prove the following two propositions.

Proposition 1. *Under conditions (C1), (C2) given in Section 2 and H_0 in (1),* $T_r/\sigma_1 \xrightarrow{d} \mathcal{N}(0,1)$.

Proposition 2. *Under conditions (C1) and (C2) given in Section 2, as* $n\to\infty$, $\hat{\sigma}_1/\sigma_1 \xrightarrow{p} 1$.

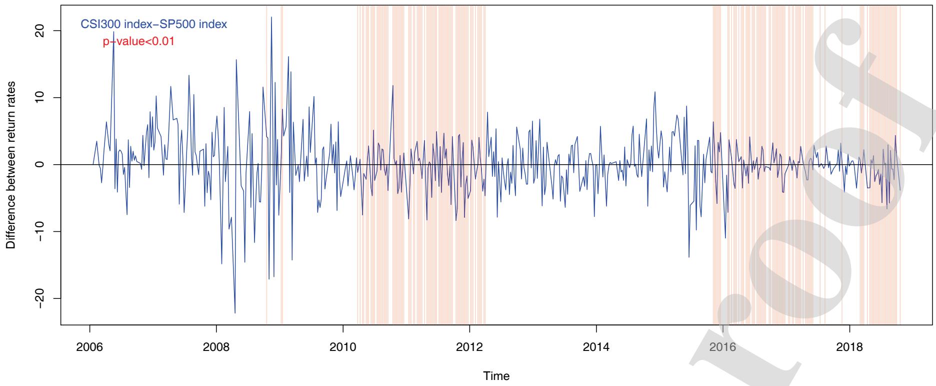

Fig. 4: Time series of the difference of the return rates of the CSI300 index and the S&P500 index, respectively.

Table 8: Empirical power comparison at the 5% level for independence between the Shanghai Stock Exchange and the Shenzhen Stock Exchange.

	$n=12$		$n=15$		$n=20$	
(p,q)	HS	EC	HS	EC	HS	EC
(200,160)	66	61	81	76	94	89
(240,200)	74	69	88	84	98	94
(280,240)	81	77	91	85	99	96

Lemma 2.

$$\mathbf{U}(\boldsymbol{X}_i-\hat{\boldsymbol{\theta}}_X)=\mathbf{U}(\boldsymbol{X}_i-\boldsymbol{\theta}_X)-\frac{1}{r_i^X}(\mathbf{I}_p-\mathbf{U}(\boldsymbol{X}_i-\boldsymbol{\theta}_X)\mathbf{U}(\boldsymbol{X}_i-\boldsymbol{\theta}_X)^\top)(\hat{\boldsymbol{\theta}}_X-\boldsymbol{\theta}_X)-\frac{1}{2(r_i^X)^2}\|(\hat{\boldsymbol{\theta}}_X-\boldsymbol{\theta}_X)\|^2\mathbf{U}(\boldsymbol{X}_i-\boldsymbol{\theta}_X)+o_p(n^{-1}).$$

Note that the proof of Lemma 2 can be found in Lemma 2 in Appendix of [29].

Lemma 3. *$\hat{\boldsymbol{\mu}}_X$ admits the following asymptotic representation: $\hat{\boldsymbol{\mu}}_X=(nc_X)^{-1}\sum_{i=1}^{n}\mathbf{U}(\boldsymbol{X}_i-\boldsymbol{\theta}_X)+o_p(b_{n,p})$, where $c_X=\mathrm{E}((r_i^X)^{-1})$ and $b_{n,p}=c_X^{-1}n^{-1/2}$.*

Note that the proof of Lemma 3 can be found in Lemma 1 in Appendix of [29].

Lemma 4. *Suppose all the conditions imposed in Theorem 1 hold. Let $T_1=(n-1)^{-1}\sum_{1\le i<j\le n}(\mathbf{U}_i^X)^\top\mathbf{U}_j^X(\mathbf{U}_i^Y)^\top\mathbf{U}_j^Y$, then $T_r=T_1+o_p(\sigma_1)$.*

Proof. Using the Taylor expansion, T_r can be written as

$$\frac{n}{n(n-1)}\sum_{1\le i<j\le n}\{\mathbf{U}(\boldsymbol{X}_i-\hat{\boldsymbol{\theta}}_X)^\top\mathbf{U}(\boldsymbol{X}_j-\hat{\boldsymbol{\theta}}_X)\mathbf{U}(\boldsymbol{Y}_i-\hat{\boldsymbol{\theta}}_Y)^\top\mathbf{U}(\boldsymbol{Y}_j-\hat{\boldsymbol{\theta}}_Y)\}$$

$$=\frac{1}{(n-1)}\sum_{1\le i<j\le n}(\mathbf{U}_i^X)^\top\mathbf{U}_j^X(\mathbf{U}_i^Y)^\top\mathbf{U}_j^Y-\frac{1}{(n-1)}\sum_{1\le i<j\le n}\{\frac{1}{r_i^X}\hat{\boldsymbol{\mu}}_X^\top[\mathbf{I}_p-\mathbf{U}_i^X(\mathbf{U}_i^X)^\top]\mathbf{U}_j^X(\mathbf{U}_i^Y)^\top\mathbf{U}_j^Y$$

$$+\frac{1}{r_j^X}\hat{\boldsymbol{\mu}}_X^\top[\mathbf{I}_p-\mathbf{U}_j^X(\mathbf{U}_j^X)^\top]\mathbf{U}_i^X(\mathbf{U}_i^Y)^\top\mathbf{U}_j^Y+\frac{1}{r_i^Y}\hat{\boldsymbol{\mu}}_Y^\top[\mathbf{I}_q-\mathbf{U}_i^Y(\mathbf{U}_i^Y)^\top]\mathbf{U}_j^Y(\mathbf{U}_i^X)^\top\mathbf{U}_j^X$$

$$+\frac{1}{r_j^Y}\hat{\boldsymbol{\mu}}_Y^\top[\mathbf{I}_q-\mathbf{U}_j^Y(\mathbf{U}_j^Y)^\top]\mathbf{U}_i^Y(\mathbf{U}_i^X)^\top\mathbf{U}_j^X\}+\frac{1}{(n-1)}\sum_{1\le i<j\le n}\{\frac{1}{r_i^Xr_i^Y}\hat{\boldsymbol{\mu}}_X^\top[\mathbf{I}_p-\mathbf{U}_i^X(\mathbf{U}_i^X)^\top]\mathbf{U}_j^X$$

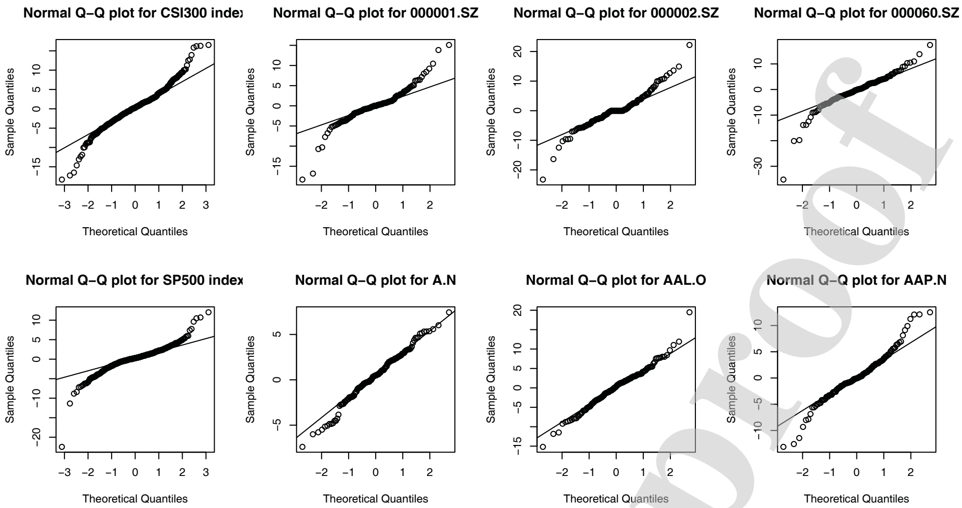

Fig. 5: Q-Q plots of the CSI300 index, the S&P500 index and the first three stocks (arranged in alphabetical order) in each index respectively, which suggest that most stocks have heavy-tailed distributions for their weekly return rates, especially for stocks from the CSI300 index.

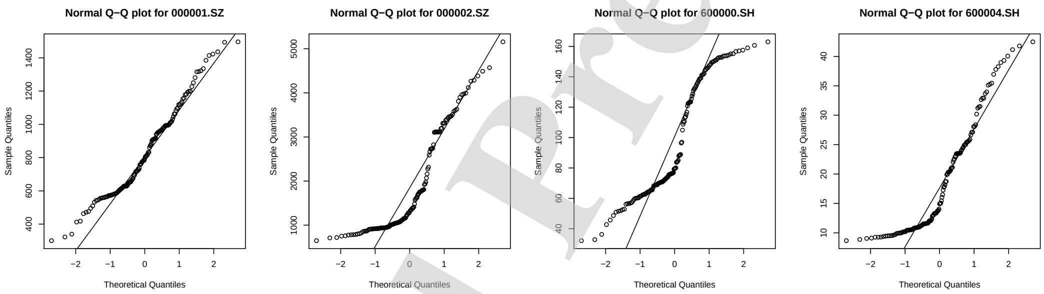

Fig. 6: Q-Q plots for four stocks used for demonstration: 000001.SZ and 000002.SZ in the Shenzhen Stock Exchange, 600000.SH and 600004.SH in the Shanghai Stock Exchange, which suggest that most stocks in the two markets have heavy-tailed distributions for their monthly return rates.

$$
\begin{aligned}
&\times (\mathbf{U}_j^Y)^\top[\mathbf{I}_q - \mathbf{U}_i^Y(\mathbf{U}_i^Y)^\top]\hat{\boldsymbol{\mu}}_{\boldsymbol{Y}} + \frac{1}{r_i^X r_j^Y}\hat{\boldsymbol{\mu}}_{\boldsymbol{X}}^\top[\mathbf{I}_p - \mathbf{U}_i^X(\mathbf{U}_i^X)^\top]\mathbf{U}_j^X(\mathbf{U}_i^Y)^\top[\mathbf{I}_q - \mathbf{U}_j^Y(\mathbf{U}_j^Y)^\top]\hat{\boldsymbol{\mu}}_{\boldsymbol{Y}}\\
&+ \frac{1}{r_j^X r_i^Y}\hat{\boldsymbol{\mu}}_{\boldsymbol{X}}^\top[\mathbf{I}_p - \mathbf{U}_j^X(\mathbf{U}_j^X)^\top]\mathbf{U}_i^X(\mathbf{U}_j^Y)^\top[\mathbf{I}_q - \mathbf{U}_i^Y(\mathbf{U}_i^Y)^\top]\hat{\boldsymbol{\mu}}_{\boldsymbol{Y}} + \frac{1}{r_j^X r_j^Y}\hat{\boldsymbol{\mu}}_{\boldsymbol{X}}^\top[\mathbf{I}_p - \mathbf{U}_j^X(\mathbf{U}_j^X)^\top]\mathbf{U}_i^X\\
&\times (\mathbf{U}_i^Y)^\top[\mathbf{I}_q - \mathbf{U}_j^Y(\mathbf{U}_j^Y)^\top]\hat{\boldsymbol{\mu}}_{\boldsymbol{Y}} + \frac{1}{r_i^X r_j^X}\hat{\boldsymbol{\mu}}_{\boldsymbol{X}}^\top[\mathbf{I}_p - \mathbf{U}_i^X(\mathbf{U}_i^X)^\top][\mathbf{I}_p - \mathbf{U}_j^X(\mathbf{U}_j^X)^\top]\hat{\boldsymbol{\mu}}_{\boldsymbol{X}}\\
&\times \mathbf{U}(\boldsymbol{Y}_i - \hat{\boldsymbol{\theta}}_{\boldsymbol{Y}})^\top\mathbf{U}(\boldsymbol{Y}_j - \hat{\boldsymbol{\theta}}_{\boldsymbol{Y}}) + \frac{1}{r_i^Y r_j^Y}\hat{\boldsymbol{\mu}}_{\boldsymbol{Y}}^\top[\mathbf{I}_q - \mathbf{U}_i^Y(\mathbf{U}_i^Y)^\top][\mathbf{I}_q - \mathbf{U}_j^Y(\mathbf{U}_j^Y)^\top]\hat{\boldsymbol{\mu}}_{\boldsymbol{Y}}\\
&\times \mathbf{U}(\boldsymbol{X}_i - \hat{\boldsymbol{\theta}}_{\boldsymbol{X}})^\top\mathbf{U}(\boldsymbol{X}_j - \hat{\boldsymbol{\theta}}_{\boldsymbol{X}})\} - \frac{1}{(n-1)}\sum_{1\le i<j\le n}\{\frac{1}{r_i^X r_j^X r_i^Y}\hat{\boldsymbol{\mu}}_{\boldsymbol{X}}^\top[\mathbf{I}_p - \mathbf{U}_i^X(\mathbf{U}_i^X)^\top]\\
&\times [\mathbf{I}_p - \mathbf{U}_j^X(\mathbf{U}_j^X)^\top]\hat{\boldsymbol{\mu}}_{\boldsymbol{X}}\hat{\boldsymbol{\mu}}_{\boldsymbol{Y}}[\mathbf{I}_q - \mathbf{U}_i^Y(\mathbf{U}_i^Y)^\top]\mathbf{U}_j^Y + \frac{1}{r_i^X r_j^X r_j^Y}\hat{\boldsymbol{\mu}}_{\boldsymbol{X}}^\top[\mathbf{I}_p - \mathbf{U}_i^X(\mathbf{U}_i^X)^\top]\\
&\times [\mathbf{I}_p - \mathbf{U}_j^X(\mathbf{U}_j^X)^\top]\hat{\boldsymbol{\mu}}_{\boldsymbol{X}}\hat{\boldsymbol{\mu}}_{\boldsymbol{Y}}[\mathbf{I}_q - \mathbf{U}_j^Y(\mathbf{U}_j^Y)^\top]\mathbf{U}_i^Y + \frac{1}{r_i^X r_i^Y r_j^Y}\hat{\boldsymbol{\mu}}_{\boldsymbol{X}}^\top[\mathbf{I}_p - \mathbf{U}_i^X(\mathbf{U}_i^X)^\top]
\end{aligned}
$$

$$\times \mathbf{U}_j^X \hat{\boldsymbol{\mu}}_Y^\top [\mathbf{I}_q - \mathbf{U}_i^Y (\mathbf{U}_i^Y)^\top][\mathbf{I}_q - \mathbf{U}_j^Y (\mathbf{U}_j^Y)^\top]\hat{\boldsymbol{\mu}}_Y + \frac{1}{r_j^X r_i^Y r_j^Y} \hat{\boldsymbol{\mu}}_X^\top [\mathbf{I}_p - \mathbf{U}_j^X (\mathbf{U}_j^X)^\top]$$
$$\times \mathbf{U}_i^X \hat{\boldsymbol{\mu}}_Y^\top [\mathbf{I}_q - \mathbf{U}_i^Y (\mathbf{U}_i^Y)^\top][\mathbf{I}_q - \mathbf{U}_j^Y (\mathbf{U}_j^Y)^\top]]\hat{\boldsymbol{\mu}}_Y\} + R_0 + o_p(n^{-4})$$
$$= \frac{1}{(n-1)} \sum_{1 \le i < j \le n} (\mathbf{U}_i^X)^\top \mathbf{U}_j^X (\mathbf{U}_i^Y)^\top \mathbf{U}_j^Y + o_p(\sigma_1),$$

where R_0 denote the rest part of T_r. For simplicity, we only show $(n-1)^{-1} \sum_{1 \le i < j \le n} (r_i^X)^{-1} \hat{\boldsymbol{\mu}}_X^\top [\mathbf{I}_p - \mathbf{U}_i^X (\mathbf{U}_i^X)^\top] \mathbf{U}_j^X (\mathbf{U}_i^Y)^\top \mathbf{U}_j^Y = o_p(\sigma_1)$, while we can similarly know that the other parts in T_r are all $o_p(\sigma_1)$. Let $G_1 = (n-1)^{-1} \sum_{1 \le i < j \le n} r_i^X \hat{\boldsymbol{\mu}}_X^\top \mathbf{U}_j^X (\mathbf{U}_i^Y)^\top \mathbf{U}_j^Y$, then $\mathrm{E}(G_1^2) = o_p(\sigma_1^2)$, because

$$\mathrm{E}(G_1^2) = \frac{1}{n^2(n-1)^2} \sum_{1 \le i < j \le n} \mathrm{E}\{\frac{1}{r_i^X c_X} (\mathbf{U}_i^X)^\top \mathbf{U}_j^X (\mathbf{U}_i^Y)^\top \mathbf{U}_j^Y\}^2$$
$$= \frac{1}{2n(n-1)} \mathrm{E}((\mathbf{U}_i^X)^\top \mathbf{U}_j^X)^2 \mathrm{E}((\mathbf{U}_i^Y)^\top \mathbf{U}_j^Y)^2 = \frac{1}{2n(n-1)} \mathrm{tr}(\mathbf{A}_X^2)\mathrm{tr}(\mathbf{A}_Y^2) = o_p(\sigma_1^2).$$

Then, we conclude that $T_r = T_1 + o_p(\sigma_1)$. □

Lemma 5. *Suppose that all the conditions imposed in Theorem 1 hold, then* $T_1/\sigma_1 \xrightarrow{d} N(0,1)$.

Proof. Let $Z_j = (n-1)^{-1} \sum_{i=1}^{j-1} (\mathbf{U}_i^X)^\top \mathbf{U}_j^X (\mathbf{U}_i^Y)^\top \mathbf{U}_j^Y$, for $j = 2, \cdots, n$. Let $S_m = \sum_{j=2}^{m} Z_j$, $\boldsymbol{V}_i = (\boldsymbol{X}_i^\top, \boldsymbol{Y}_i^\top)^\top$ and $\mathcal{F}_m = \sigma\{\boldsymbol{V}_1, \cdots \boldsymbol{V}_m\}$, which is the σ-algebra generated by $\{\boldsymbol{V}_1, \cdots \boldsymbol{V}_m\}$. Hence $T_1 = \sum_{j=2}^{n} Z_j$. We can verify that for each n, $\{S_m, \mathcal{F}_m\}_{m=2}^n$ is a sequence of zero mean and square integrable martingale. In order to prove the normality of T_1, according to [8], it suffices to show the following tow results:

$$\frac{\sum_{j=2}^n \mathrm{E}[Z_j^2 | \mathcal{F}_{j-1}]}{\sigma_1^2} \xrightarrow{p} 1,$$

and for any $\epsilon > 0$,

$$\sigma_1^{-2} \sum_{j=2}^{n} \mathrm{E}[Z_j^2 I(|Z_j| > \epsilon\sigma_1) | \mathcal{F}_{j-1}] \xrightarrow{p} 0.$$

Below, we will prove the first result. We see that

$$\sum_{j=2}^{n} \mathrm{E}[Z_j^2 | \mathcal{F}_{j-1}] = \frac{1}{(n-1)^2} \sum_{j=2}^{n} \mathrm{E}[(\sum_{i=1}^{j-1} (\mathbf{U}_i^X)^\top \mathbf{U}_j^X (\mathbf{U}_i^Y)^\top \mathbf{U}_j^Y)^2 | \mathcal{F}_{j-1}]$$
$$= \frac{1}{(n-1)^2} \sum_{j=2}^{n} \mathrm{E}[(\sum_{i_1, i_2 = 1}^{j-1} (\mathbf{U}_{i_1}^X)^\top \mathbf{U}_j^X (\mathbf{U}_j^X)^\top \mathbf{U}_{i_2}^X (\mathbf{U}_{i_1}^Y)^\top \mathbf{U}_j^Y (\mathbf{U}_j^Y)^\top \mathbf{U}_{i_2}^Y) | \mathcal{F}_{j-1}]$$
$$= \frac{1}{(n-1)^2} \sum_{j=2}^{n} \sum_{i_1, i_2 = 1}^{j-1} (\mathbf{U}_{i_1}^X)^\top \mathrm{E}[\mathbf{U}_j^X (\mathbf{U}_j^X)^\top | \mathcal{F}_{j-1}] \mathbf{U}_{i_2}^X (\mathbf{U}_{i_1}^Y)^\top \mathrm{E}[\mathbf{U}_j^Y (\mathbf{U}_j^Y)^\top | \mathcal{F}_{j-1}] \mathbf{U}_{i_2}^Y$$
$$= \frac{1}{(n-1)^2} \sum_{j=2}^{n} \sum_{i_1, i_2 = 1}^{j-1} (\mathbf{U}_{i_1}^X)^\top \mathbf{A}_X \mathbf{U}_{i_2}^X (\mathbf{U}_{i_1}^Y)^\top \mathbf{A}_Y \mathbf{U}_{i_2}^Y$$
$$= \frac{1}{(n-1)^2} \sum_{j=2}^{n} \sum_{i=1}^{j-1} (\mathbf{U}_i^X)^\top \mathbf{A}_X \mathbf{U}_i^X (\mathbf{U}_i^Y)^\top \mathbf{A}_Y \mathbf{U}_i^Y + \frac{1}{(n-1)^2} \sum_{j=2}^{n} \sum_{i_1 \ne i_2}^{j-1} (\mathbf{U}_{i_1}^X)^\top \mathbf{A}_X \mathbf{U}_{i_2}^X (\mathbf{U}_{i_1}^Y)^\top \mathbf{A}_Y \mathbf{U}_{i_2}^Y \doteq C_1 + C_2.$$

By using simple algebra, we can obtain that $\mathrm{E}(C_1) = \sigma_1^2$, $\mathrm{E}(C_2) = 0$ and

$$\mathrm{var}(C_1) = \frac{1}{(n-1)^4}\sum_{j=2}^{n} j^2\{\mathrm{E}((\mathbf{U}_i^X)^\top \mathbf{A}_X \mathbf{U}_i^X)^2 \mathrm{E}((\mathbf{U}_i^Y)^\top \mathbf{A}_Y \mathbf{U}_i^Y)^2 - \mathrm{tr}^2(\mathbf{A}_X^2)\mathrm{tr}^2(\mathbf{A}_Y^2)\},$$
$$\mathrm{var}(C_2) = \frac{1}{(n-1)^4}\sum_{j=3}^{n} \frac{j(n-j+1)(j-1)}{2}\mathrm{tr}(\mathbf{A}_X^4)\mathrm{tr}(\mathbf{A}_Y^4) = o_p(\sigma_1^4).$$

By Lemma 1, we can see that $\mathrm{E}((\mathbf{U}_i^X)^\top \mathbf{A}_X \mathbf{U}_i^X)^2 = O(1)\mathrm{E}^2((\mathbf{U}_i^X)^\top \mathbf{A}_X \mathbf{U}_i^X) = O(\mathrm{tr}^2(\mathbf{A}_X^2))$ and $\mathrm{E}((\mathbf{U}_i^Y)^\top \mathbf{A}_Y \mathbf{U}_i^Y)^2 = O(\mathrm{tr}^2(\mathbf{A}_Y^2))$. Hence $\mathrm{var}(C_1) = o_p(\sigma_1^4)$ and $C_1/\sigma_1^2 \xrightarrow{p} 1$. By using condition C(1), we have $\mathrm{var}(C_2) = o_p(\sigma_1^4)$, which implies that $C_2 = o_p(\sigma_1^2)$.

Next, we will prove the second result. Note that $\sigma_1^{-2}\sum_{j=2}^n \mathrm{E}[Z_j^2 I(|Z_j| > \epsilon\sigma_1|)|\mathcal{F}_{j-1}] \le \sigma_1^{-4}\epsilon^{-2}\sum_{j=2}^n \mathrm{E}[Z_j^4|\mathcal{F}_{j-1}]$. Accordingly, the assertion of this lemma is true if we can show $\mathrm{E}\left\{\sum_{j=2}^n \mathrm{E}[Z_j^4|\mathcal{F}_{j-1}]\right\} = o(\sigma_1^4)$. Note that

$$\mathrm{E}\left\{\sum_{j=2}^{n}\mathrm{E}[Z_j^4|\mathcal{F}_{j-1}]\right\} = \sum_{j=2}^{n}\mathrm{E}(Z_j^4) = O(n^{-4})\sum_{j=2}^{n}\mathrm{E}(\sum_{i=1}^{j-1}(\mathbf{U}_i^X)^\top \mathbf{A}_X \mathbf{U}_j^X \mathbf{U}_i^Y \mathbf{A}_Y \mathbf{U}_j^Y)^4,$$

which can be decomposed as $3Q + P$. Here

$$\begin{aligned} Q =& O(n^{-4})\sum_{j=2}^{n}\sum_{s<t}^{j-1}\sum^{j-1}\mathrm{E}((\mathbf{U}_j^X)^\top \mathbf{A}_X \mathbf{U}_s^X(\mathbf{U}_s^X)^\top \mathbf{A}_X \mathbf{U}_j^X(\mathbf{U}_j^X)^\top \mathbf{A}_X \mathbf{U}_t^X(\mathbf{U}_t^X)^\top \mathbf{A}_X \mathbf{U}_j^X \\ &\times (\mathbf{U}_j^Y)^\top \mathbf{A}_Y \mathbf{U}_s^Y(\mathbf{U}_s^Y)^\top \mathbf{A}_Y \mathbf{U}_j^Y(\mathbf{U}_j^Y)^\top \mathbf{A}_Y \mathbf{U}_t^Y(\mathbf{U}_t^Y)^\top \mathbf{A}_Y \mathbf{U}_j^Y), \\ P =& O(n^{-4})\sum_{j=2}^{n}\sum_{i=1}^{j-1}\mathrm{E}(((U_j^X)^\top \mathbf{A}_X \mathbf{U}_i^X)^4)\mathrm{E}(((\mathbf{U}_j^Y)^\top \mathbf{A}_Y \mathbf{U}_i^Y)^4). \end{aligned}$$

Obviously, $Q = O(n^{-1})\mathrm{E}((\mathbf{U}_j^X)^\top \mathbf{A}_X \mathbf{U}_s^X(\mathbf{U}_s^X)^\top \mathbf{A}_X \mathbf{U}_j^X)^2\mathrm{E}((\mathbf{U}_j^Y)^\top \mathbf{A}_Y \mathbf{U}_s^Y(\mathbf{U}_s^Y)^\top \mathbf{A}_Y \mathbf{U}_j^Y)^2$. By Lemma 1, we have

$$\mathrm{E}((\mathbf{U}_j^X)^\top \mathbf{A}_X \mathbf{U}_s^X(\mathbf{U}_s^X)^\top \mathbf{A}_X \mathbf{U}_j^X)^2 = O(1)\{\mathrm{E}^2((\mathbf{U}_j^X)^\top \mathbf{A}_X \mathbf{U}_s^X(\mathbf{U}_s^X)^\top \mathbf{A}_X \mathbf{U}_j^X)\} = O(\mathrm{tr}^4(\mathbf{A}_X^2))$$

and $\mathrm{E}(((U_j^X)^\top \mathbf{A}_X \mathbf{U}_i^X)^4) = O(\mathrm{tr}^4(\mathbf{A}_X^2))$. Thus, $Q = o(\sigma_1^4)$, $P = o(\sigma_1^4)$ and $T_1/\sqrt{\mathrm{var}(T_1)} \xrightarrow{d} \mathcal{N}(0,1)$, where $\mathrm{var}(T_1) = n\{2(n-1)\}^{-1}\mathrm{tr}(\mathbf{A}_X^2)\mathrm{tr}(A_Y^2)$. This complete the proof of Lemma 5. □

Proof of Proposition 1: Using Lemma 3 and Lemma 4, proof of Proposition 1 can be directly obtained.

Proof of Proposition 2: Taking the same procedure as in the proof of Lemma 4, we can see that

$$\hat{\sigma}_1^2 = \frac{2}{n^2(n-1)^2}\sum_{1\le i<j\le n}((\hat{\mathbf{U}}_i^X)^\top \hat{\mathbf{U}}_j^X)^2\sum_{1\le i<j\le n}((\hat{\mathbf{U}}_i^Y)^\top \hat{\mathbf{U}}_j^Y)^2 = \frac{2}{n^2(n-1)^2}\sum_{1\le i<j\le n}((\mathbf{U}_i^X)^\top \mathbf{U}_j^X)^2\sum_{1\le i<j\le n}((\mathbf{U}_i^Y)^\top \mathbf{U}_j^Y)^2 + o_p(\sigma_1^2).$$

Obviously, $\mathrm{E}(\frac{2}{n^2(n-1)^2}\sum_{1\le i<j\le n}((\mathbf{U}_i^X)^\top \mathbf{U}_j^X)^2\sum_{1\le i<j\le n}((\mathbf{U}_i^Y)^\top \mathbf{U}_j^Y)^2) = \sigma_1^2$, which implies that

$$\begin{aligned} &\mathrm{var}(\frac{2}{n^2(n-1)^2}\sum_{1\le i<j\le n}((\mathbf{U}_i^X)^\top \mathbf{U}_j^X)^2\sum_{1\le i<j\le n}((\mathbf{U}_i^Y)^\top \mathbf{U}_j^Y)^2) \\ =&O(n^{-4})\mathrm{E}((\mathbf{U}_i^X)^\top \mathbf{U}_j^X)^4\mathrm{E}((\mathbf{U}_i^Y)^\top \mathbf{U}_j^Y)^4 + O(n^{-2})\mathrm{E}\{((\mathbf{U}_i^X)^\top \mathbf{U}_j^X)^2((\mathbf{U}_i^X)^\top \mathbf{U}_l^X)^2\}\mathrm{E}\{((\mathbf{U}_i^Y)^\top \mathbf{U}_j^Y)^2((\mathbf{U}_i^Y)^\top \mathbf{U}_l^Y)^2\} \\ =&O(n^{-4}\mathrm{tr}^2(\mathbf{A}_X^2)\mathrm{tr}^2(\mathbf{A}_Y^2)) + O(n^{-2}\mathrm{tr}^2(\mathbf{A}_X^2)\mathrm{tr}^2(\mathbf{A}_Y^2)) = o_p(\sigma_1^4). \end{aligned}$$

Thus, $\hat{\sigma}_1^2 = \sigma_1^2(1 + o_p(1))$. □

Proof of Theorem 2

By the Taylor Expansion, we have

$$\begin{aligned}
\mathbf{U}(\boldsymbol{X}_i-\hat{\boldsymbol{\theta}}_X) =& \mathbf{U}(\boldsymbol{X}_i^*-\boldsymbol{\theta}_{X^*}-(\hat{\boldsymbol{\theta}}_{X^*}-\boldsymbol{\theta}_{X^*})+\mathbf{M}_1(\boldsymbol{Y}_i^*-\boldsymbol{\theta}_{Y^*})-\mathbf{M}_1(\hat{\boldsymbol{\theta}}_{Y^*}-\boldsymbol{\theta}_{Y^*}))\\
=& \mathbf{U}(\boldsymbol{X}_i-\boldsymbol{\theta}_X)-\frac{1}{r_i^X}(\mathbf{I}_p-\mathbf{U}(\boldsymbol{X}_i-\boldsymbol{\theta}_X)\mathbf{U}(\boldsymbol{X}_i-\boldsymbol{\theta}_X)^\top)(\hat{\boldsymbol{\theta}}_X-\boldsymbol{\theta}_X)\\
&+\frac{r_i^{Y^*}}{r_i^{X^*}}(\mathbf{I}_p-\mathbf{U}(\boldsymbol{X}_i-\boldsymbol{\theta}_X)\mathbf{U}(\boldsymbol{X}_i-\boldsymbol{\theta}_X)^\top)\mathbf{M}_1\mathbf{U}(\boldsymbol{Y}_i^*-\boldsymbol{\theta}_{Y^*})-\frac{1}{r_i^{X^*}}(\mathbf{I}_p-\mathbf{U}(\boldsymbol{X}_i-\boldsymbol{\theta}_X)\mathbf{U}(\boldsymbol{X}_i-\boldsymbol{\theta}_X)^\top)\mathbf{M}_1(\hat{\boldsymbol{\theta}}_{Y^*}-\boldsymbol{\theta}_{Y^*})\\
&-\frac{1}{2(r_i^X)^2}\|(\hat{\boldsymbol{\theta}}_X-\boldsymbol{\theta}_X)+\mathbf{M}_1(\boldsymbol{Y}_i^*-\boldsymbol{\theta}_{Y^*})-\mathbf{M}_1(\hat{\boldsymbol{\theta}}_{Y^*}-\boldsymbol{\theta}_{Y^*}))\|^2\mathbf{U}(\boldsymbol{X}_i-\boldsymbol{\theta}_X)+o_p(n^{-1})
\end{aligned}$$

and

$$\begin{aligned}
\mathbf{U}(\boldsymbol{Y}_i-\hat{\boldsymbol{\theta}}_Y) =& \mathbf{U}(\boldsymbol{Y}_i^*-\boldsymbol{\theta}_{Y^*}-(\hat{\boldsymbol{\theta}}_{Y^*}-\boldsymbol{\theta}_{Y^*})+\mathbf{M}_2(\boldsymbol{X}_i^*-\boldsymbol{\theta}_{X^*})-\mathbf{M}_2(\hat{\boldsymbol{\theta}}_{X^*}-\boldsymbol{\theta}_{X^*}))\\
=& \mathbf{U}(\boldsymbol{Y}_i-\boldsymbol{\theta}_Y)-\frac{1}{r_i^Y}(\mathbf{I}_p-\mathbf{U}(\boldsymbol{Y}_i-\boldsymbol{\theta}_Y)\mathbf{U}(\boldsymbol{Y}_i-\boldsymbol{\theta}_Y)^\top)(\hat{\boldsymbol{\theta}}_Y-\boldsymbol{\theta}_Y)\\
&+\frac{r_i^{X^*}}{r_i^{Y^*}}(\mathbf{I}_p-\mathbf{U}(\boldsymbol{Y}_i-\boldsymbol{\theta}_Y)\mathbf{U}(\boldsymbol{Y}_i-\boldsymbol{\theta}_Y)^\top)\mathbf{M}_2\mathbf{U}(\boldsymbol{X}_i^*-\boldsymbol{\theta}_{X^*})-\frac{1}{r_i^{Y^*}}(\mathbf{I}_p-\mathbf{U}(\boldsymbol{Y}_i-\boldsymbol{\theta}_Y)\mathbf{U}(\boldsymbol{Y}_i-\boldsymbol{\theta}_Y)^\top)\mathbf{M}_2(\hat{\boldsymbol{\theta}}_{X^*}-\boldsymbol{\theta}_{X^*})\\
&-\frac{1}{2(r_i^Y)^2}\|(\hat{\boldsymbol{\theta}}_Y-\boldsymbol{\theta}_Y)+\mathbf{M}_2(\boldsymbol{X}_i^*-\boldsymbol{\theta}_{X^*})-\mathbf{M}_2(\hat{\boldsymbol{\theta}}_{X^*}-\boldsymbol{\theta}_{X^*}))\|^2\mathbf{U}(\boldsymbol{Y}_i-\boldsymbol{\theta}_Y)+o_p(n^{-1}).
\end{aligned}$$

Then, taking the same procedure as in Lemma 3, under conditions (C1$'$)-(C3$'$), we have

$$\begin{aligned}
T_r =& T_1+\frac{1}{n-1}\sum_{1\le i<j\le n}(r_i^{X^*})^{-1}r_i^{Y^*}(r_j^{X^*})^{-1}r_j^{Y^*}(\mathbf{U}_i^{Y^*})^\top\mathbf{M}_1^\top\mathbf{M}_1\mathbf{U}_j^{Y^*}(\mathbf{U}_i^{Y^*})^\top\mathbf{U}_j^{Y^*}\\
&+\frac{1}{n-1}\sum_{1\le i<j\le n}(r_i^{Y^*})^{-1}r_i^{X^*}(r_j^{Y^*})^{-1}r_j^{X^*}(\mathbf{U}_i^{X^*})^\top\mathbf{M}_2^\top\mathbf{M}_2\mathbf{U}_j^{X^*}(\mathbf{U}_i^{X^*})^\top\mathbf{U}_j^{X^*}\\
&+\frac{1}{n-1}\sum_{1\le i<j\le n}(r_i^{X^*})^{-1}r_i^{Y^*}(r_j^{Y^*})^{-1}r_j^{X^*}(\mathbf{U}_i^{Y^*})^\top\mathbf{M}_1^\top\mathbf{U}_j^{X^*}(\mathbf{U}_i^{Y^*})^\top\mathbf{M}_2\mathbf{U}_j^{X^*}\\
&+\frac{1}{n-1}\sum_{1\le i<j\le n}(r_i^{Y^*})^{-1}r_i^{X^*}(r_j^{X^*})^{-1}r_j^{Y^*}(\mathbf{U}_i^{X^*})^\top\mathbf{M}_2^\top\mathbf{U}_j^{Y^*}(\mathbf{U}_i^{X^*})^\top\mathbf{M}_1\mathbf{U}_j^{Y^*}\\
&+\frac{1}{n-1}\sum_{1\le i<j\le n}(r_i^{Y^*})^{-1}r_i^{X^*}(\mathbf{U}_i^{X^*})^\top\mathbf{U}_j^{X^*}(\mathbf{U}_i^{Y^*})^\top\mathbf{M}_2\mathbf{U}_j^{X^*}+\frac{1}{n-1}\sum_{1\le i<j\le n}(r_i^{X^*})^{-1}r_i^{Y^*}(\mathbf{U}_i^{Y^*})^\top\mathbf{M}_1^\top\mathbf{U}_j^{X^*}(\mathbf{U}_i^{Y^*})^\top\mathbf{U}_j^{Y^*}+o_p(\sigma_1).
\end{aligned}$$

By condition (C3$'$), we have

$$\mathrm{E}\left\{\frac{1}{n-1}\sum_{1\le i<j\le n}(r_i^{Y^*})^{-1}r_i^{X^*}(\mathbf{U}_i^{X^*})^\top\mathbf{U}_j^{X^*}(\mathbf{U}_i^{Y^*})^\top\mathbf{M}_2\mathbf{U}_j^{X^*}\right\}^2=O[\{\mathrm{E}((r_i^{Y^*})^{-1}r_i^{X^*})\}^2\mathrm{tr}(\mathbf{M}_2\mathbf{A}_X^2\mathbf{M}_2^\top\mathbf{A}_Y)]=o(\sigma_1^2),$$

$$\mathrm{E}\left\{\frac{1}{n-1}\sum_{1\le i<j\le n}(r_i^{X^*})^{-1}r_i^{Y^*}(\mathbf{U}_i^{Y^*})^\top\mathbf{M}_1^\top\mathbf{U}_j^{X^*}(\mathbf{U}_i^{Y^*})^\top\mathbf{U}_j^{Y^*}\right\}^2=O[\{\mathrm{E}((r_i^{X^*})^{-1}r_i^{Y^*})\}^2\mathrm{tr}(\mathbf{M}_1\mathbf{A}_Y^2\mathbf{M}_1^\top\mathbf{A}_X)]=o(\sigma_1^2),$$

$$\frac{1}{n-1}\sum_{1\le i<j\le n}(r_i^{Y^*})^{-1}r_i^{X^*}(r_j^{Y^*})^{-1}r_j^{X^*}(\mathbf{U}_i^{X^*})^\top\mathbf{M}_2^\top\mathbf{M}_2\mathbf{U}_j^{X^*}(\mathbf{U}_i^{X^*})^\top\mathbf{U}_j^{X^*}=\frac{n}{2}\{\mathrm{E}((r_i^{Y^*})^{-1}r_i^{X^*})\}^2\mathrm{tr}(\mathbf{M}_2\mathbf{A}_X^{*2}\mathbf{M}_2^\top)+o_p(\sigma_1),$$

$$\frac{1}{n-1}\sum_{1\le i<j\le n}(r_i^{X^*})^{-1}r_i^{Y^*}(r_j^{X^*})^{-1}r_j^{Y^*}(\mathbf{U}_i^{Y^*})^\top\mathbf{M}_1^\top\mathbf{M}_1\mathbf{U}_j^{Y^*}(\mathbf{U}_i^{Y^*})^\top\mathbf{U}_j^{Y^*}=\frac{n}{2}\{\mathrm{E}((r_i^{X^*})^{-1}r_i^{Y^*})\}^2\mathrm{tr}(\mathbf{M}_1\mathbf{A}_Y^{*2}\mathbf{M}_1^\top)+o_p(\sigma_1),$$

$$\frac{1}{n-1}\sum_{1\le i<j\le n}(r_i^{X^*})^{-1}r_i^{Y^*}(r_j^{Y^*})^{-1}r_j^{X^*}(\mathbf{U}_i^{Y^*})^\top\mathbf{M}_1^\top\mathbf{U}_j^{X^*}(\mathbf{U}_i^{Y^*})^\top\mathbf{M}_2\mathbf{U}_j^{X^*}=\frac{n}{2}\{\mathrm{E}((r_i^{X^*})^{-1}r_i^{Y^*})\mathrm{E}((r_i^{Y^*})^{-1}r_i^{X^*})\}\mathrm{tr}(\mathbf{M}_1\mathbf{A}_Y^*\mathbf{M}_2\mathbf{A}_X^*)+o_p(\sigma_1),$$

$$\frac{1}{n-1}\sum_{1\le i<j\le n}(r_i^{Y^*})^{-1}r_i^{X^*}(r_j^{X^*})^{-1}r_j^{Y^*}(\mathbf{U}_i^{X^*})^\top\mathbf{M}_2^\top\mathbf{U}_j^{Y^*}(\mathbf{U}_i^{X^*})^\top\mathbf{M}_1\mathbf{U}_j^{Y^*}=\frac{n}{2}\{\mathrm{E}((r_i^{X^*})^{-1}r_i^{Y^*})\mathrm{E}((r_i^{Y^*})^{-1}r_i^{X^*})\}\mathrm{tr}(\mathbf{M}_2\mathbf{A}_X^*\mathbf{M}_1\mathbf{A}_Y^*)+o_p(\sigma_1).$$

Thus, $T_r=T_1+n\mathrm{tr}(\mathbf{\Lambda}^\top\mathbf{\Lambda})/2+o_p(\sigma_1)$, where $\mathbf{\Lambda}=\mathrm{E}((r_i^{X^*})^{-1}r_i^{Y^*})\mathbf{M}_1\mathbf{A}_Y^*+\mathrm{E}((r_i^{Y^*})^{-1}r_i^{X^*})\mathbf{A}_X^*\mathbf{M}_2^\top$. According to the results of Theorem 1, we can easily obtain the result. Here we complete the proof.

Proof of Theorem 3

Define $\boldsymbol{V}_i^X=\mathrm{E}(\mathbf{U}(\boldsymbol{X}_i-\boldsymbol{X}_j)|\boldsymbol{X}_i)$, $\boldsymbol{V}_j^X=-\mathrm{E}(\mathbf{U}(\boldsymbol{X}_i-\boldsymbol{X}_j)|\boldsymbol{X}_j)$, $\boldsymbol{V}_i^Y=\mathrm{E}(\mathbf{U}(\boldsymbol{Y}_i-\boldsymbol{Y}_j)|\boldsymbol{Y}_i)$, $\boldsymbol{V}_j^Y=-\mathrm{E}(\mathbf{U}(\boldsymbol{Y}_i-\boldsymbol{Y}_j)|\boldsymbol{Y}_j)$, $\mathbf{U}(\boldsymbol{Y}_i-\boldsymbol{Y}_j)=\boldsymbol{V}_i^Y+\boldsymbol{V}_j^Y+\boldsymbol{W}_{ij}^Y$, $\mathbf{U}(\boldsymbol{Y}_k-\boldsymbol{Y}_\ell)=\boldsymbol{V}_k^Y+\boldsymbol{V}_\ell^Y+\boldsymbol{W}_{k\ell}^Y$, $\mathbf{U}(\boldsymbol{X}_i-\boldsymbol{X}_j)=\boldsymbol{V}_i^X+\boldsymbol{V}_j^X+\boldsymbol{W}_{ij}^X$, $\mathbf{U}(\boldsymbol{X}_k-\boldsymbol{X}_\ell)=\boldsymbol{V}_k^X+\boldsymbol{V}_\ell^X+\boldsymbol{W}_{k\ell}^X$, $\mathbf{B}_X=\mathrm{E}(\boldsymbol{V}_i^X(\boldsymbol{V}_i^X)^\top)$ and $\mathbf{B}_Y=\mathrm{E}(\boldsymbol{V}_i^Y(\boldsymbol{V}_i^Y)^\top)$.

Define

$$T_\rho=\frac{1}{(n-1)(n-2)(n-3)}\sum^*\{\mathbf{U}(\boldsymbol{X}_i-\boldsymbol{X}_j)^\top\mathbf{U}(\boldsymbol{X}_k-\boldsymbol{X}_\ell)\mathbf{U}(\boldsymbol{Y}_i-\boldsymbol{Y}_\ell)^\top\mathbf{U}(\boldsymbol{Y}_k-\boldsymbol{Y}_j)\},$$
$$\hat{\sigma}_2^2=\frac{1}{2n^2(n-1)^2(n-2)^2(n-3)^2}\sum^*(\mathbf{U}(\boldsymbol{X}_i-\boldsymbol{X}_j)^\top\mathbf{U}(\boldsymbol{X}_k-\boldsymbol{X}_\ell))^2\times\sum^*(\mathbf{U}(\boldsymbol{Y}_i-\boldsymbol{Y}_j)^\top\mathbf{U}(\boldsymbol{Y}_k-\boldsymbol{Y}_\ell))^2.$$

Hence, $T_{\mathrm{HR}}=T_\rho/\hat{\sigma}_2$. To prove Theorem 3, we only need to prove the following proposition.

Proposition 3. *Under conditions (C1), (C2) given in Section 2 and H_0 in (1), as $n\to\infty$, $T_\rho/\sigma_2\xrightarrow{d}\mathcal{N}(0,1)$.*

Here $\sigma_2^2=\frac{8n}{n-1}\mathrm{tr}(\mathbf{B}_X^2)\mathrm{tr}(\mathbf{B}_Y^2)$.

Proposition 4. *Under conditions (C1$'$), (C2$'$) and (C4$'$) given in Section 2, as $n\to\infty$, $\hat{\sigma}_2^2/\sigma_2^2\xrightarrow{p}1$.*

Lemma 6. *As $n\to\infty$, $(\tilde{\sigma}_2)^{-1}\left\{(n-1)^{-1}\sum^*(\boldsymbol{V}_i^X)^\top\boldsymbol{V}_k^X(\boldsymbol{V}_i^Y)^\top\boldsymbol{V}_k^Y\right\}\xrightarrow{d}\mathcal{N}(0,1)$, where $\tilde{\sigma}_2^2=2n(n-1)^{-1}\mathrm{tr}(\mathbf{B}_X^2)\mathrm{tr}(\mathbf{B}_Y^2)$.*

Proof. Define $T_2=(n-1)^{-1}\sum^*(\boldsymbol{V}_i^X)^\top\boldsymbol{V}_k^X(\boldsymbol{V}_i^Y)^\top\boldsymbol{V}_k^Y$. Obviously, $\mathrm{E}(T_2)=0$ and

$$\mathrm{var}(T_2)=\frac{1}{(n-1)^2}\mathrm{E}\left\{\sum^*(\boldsymbol{V}_i^X)^\top\boldsymbol{V}_k^X(\boldsymbol{V}_i^Y)^\top\boldsymbol{V}_k^Y\right\}^2=\frac{2n}{n-1}\mathrm{E}((\boldsymbol{V}_i^X)^\top\boldsymbol{V}_k^X)^2\mathrm{E}((\boldsymbol{V}_i^Y)^\top\boldsymbol{V}_k^Y)^2=\frac{2n}{n-1}\mathrm{tr}(\mathbf{B}_X^2)\mathrm{tr}(\mathbf{B}_Y^2).$$

We only need to show the asymptotic normality of T_2. For each $i\in\{2,\cdots,n\}$, define $\tilde{Z}_i=(n-1)^{-1}\sum_{k=1}^{i-1}(\boldsymbol{V}_i^X)^\top\boldsymbol{V}_k^X(\boldsymbol{V}_i^Y)^\top\boldsymbol{V}_k^Y$ and $\boldsymbol{V}_i=(\boldsymbol{X}_i^\top,\boldsymbol{Y}_i^\top)^\top$. Then, for each $m\in\{2,\cdots,n\}$, define $\tilde{S}_m=\sum_{i=2}^m\tilde{Z}_i$ and $\tilde{\mathcal{F}}_m=\sigma\{\boldsymbol{V}_1,\cdots\boldsymbol{V}_m\}$, where $\sigma\{\boldsymbol{V}_1,\cdots\boldsymbol{V}_m\}$ is the σ-algebra generated by $\{\boldsymbol{V}_1,\cdots\boldsymbol{V}_m\}$. Now $T_2=\sum_{i=2}^n\tilde{Z}_i$. We can verify that for each n, $\{\tilde{S}_m,\tilde{\mathcal{F}}_m\}_{m=2}^n$ is a sequence of zero mean and square integrable martingale. In order to prove the normality of $\tilde{Z}_2$, according to [8], it suffices to show the following two results:

$$\frac{\sum_{i=2}^n\mathrm{E}[\tilde{Z}_i^2|\tilde{\mathcal{F}}_{i-1}]}{\tilde{\sigma}_2^2}\xrightarrow{p}1$$

and for any $\epsilon > 0$,

$$\tilde{\sigma}_2^{-2}\sum_{i=2}^{n}\mathrm{E}[\tilde{Z}_i^2 I(|\tilde{Z}_i| > \epsilon\tilde{\sigma}_2|)|\tilde{\mathcal{F}}_{i-1}] \xrightarrow{p} 0.$$

Below we will prove the first result. Note that

$$\begin{aligned}
\sum_{i=2}^{n}\mathrm{E}[\tilde{Z}_i^2|\tilde{\mathcal{F}}_{i-1}] =& \frac{1}{(n-1)^2}\sum_{i=2}^{n}\mathrm{E}[(\sum_{k=1}^{i-1}(\boldsymbol{V}_i^{\boldsymbol{X}})^\top\boldsymbol{V}_k^{\boldsymbol{X}}(\boldsymbol{V}_i^{\boldsymbol{Y}})^\top\boldsymbol{V}_k^{\boldsymbol{Y}})^2|\tilde{\mathcal{F}}_{i-1}]\\
=& \frac{1}{(n-1)^2}\sum_{i=2}^{n}\mathrm{E}[(\sum_{k_1,k_2=1}^{i-1}(\boldsymbol{V}_{k_1}^{\boldsymbol{X}})^\top\boldsymbol{V}_i^{\boldsymbol{X}}(\boldsymbol{V}_i^{\boldsymbol{X}})^\top\boldsymbol{V}_{k_2}^{\boldsymbol{X}}(\boldsymbol{V}_{k_1}^{\boldsymbol{Y}})^\top\boldsymbol{V}_i^{\boldsymbol{Y}}(\boldsymbol{V}_i^{\boldsymbol{Y}})^\top\boldsymbol{V}_{k_2}^{\boldsymbol{Y}})|\tilde{\mathcal{F}}_{i-1}]\\
=& \frac{1}{(n-1)^2}\sum_{i=2}^{n}\sum_{k_1,k_2=1}^{i-1}(\boldsymbol{V}_{k_1}^{\boldsymbol{X}})^\top\mathrm{E}[\boldsymbol{V}_i^{\boldsymbol{X}}(\boldsymbol{V}_i^{\boldsymbol{X}})^\top|\tilde{\mathcal{F}}_{i-1}]\boldsymbol{V}_{k_2}^{\boldsymbol{X}}(\boldsymbol{V}_{k_1}^{\boldsymbol{Y}})^\top\mathrm{E}[\boldsymbol{V}_i^{\boldsymbol{Y}}(\mathbf{R}_i^{\boldsymbol{Y}})^\top|\tilde{\mathcal{F}}_{i-1}]\boldsymbol{V}_{k_2}^{\boldsymbol{Y}}\\
=& \frac{1}{(n-1)^2}\sum_{i=2}^{n}\sum_{k=1}^{i-1}(\boldsymbol{V}_k^{\boldsymbol{X}})^\top\mathbf{B}_{\boldsymbol{X}}\boldsymbol{V}_k^{\boldsymbol{X}}(\boldsymbol{V}_k^{\boldsymbol{Y}})^\top\mathbf{B}_{\boldsymbol{Y}}\boldsymbol{V}_k^{\boldsymbol{Y}} + \frac{1}{(n-1)^2}\sum_{i=2}^{n}\sum_{k_1\neq k_2}^{i-1}(\boldsymbol{V}_{k_1}^{\boldsymbol{X}})^\top\mathbf{B}_{\boldsymbol{X}}\boldsymbol{V}_{k_2}^{\boldsymbol{X}}(\boldsymbol{V}_{k_1}^{\boldsymbol{Y}})^\top\mathbf{B}_{\boldsymbol{Y}}\boldsymbol{V}_{k_2}^{\boldsymbol{Y}} \doteq C_3 + C_4.
\end{aligned}$$

By simple algebra, we can obtain that $\mathrm{E}(C_3) = \tilde{\sigma}_2^2$, $\mathrm{E}(C_4) = 0$ and

$$\begin{aligned}
\mathrm{var}(C_3) &= \frac{1}{(n-1)^4}\sum_{i=2}^{n} j^2\{\mathrm{E}((\boldsymbol{V}_k^{\boldsymbol{X}})^\top\mathbf{B}_{\boldsymbol{X}}\boldsymbol{V}_k^{\boldsymbol{X}}(\boldsymbol{V}_k^{\boldsymbol{Y}})^\top\mathbf{B}_{\boldsymbol{Y}}\boldsymbol{V}_k^{\boldsymbol{Y}})^2 - \mathrm{tr}^2(\mathbf{B}_{\boldsymbol{X}}^2)\mathrm{tr}^2(\mathbf{B}_{\boldsymbol{Y}}^2)\};\\
\mathrm{var}(C_2) &= \frac{1}{(n-1)^4}\sum_{i=3}^{n}\frac{i(n-i+1)(i-1)}{2}\mathrm{tr}(\mathbf{B}_{\boldsymbol{X}}^4)\mathrm{tr}(\mathbf{B}_{\boldsymbol{Y}}^4) = o_p(\tilde{\sigma}_2^4).
\end{aligned}$$

By Lemma 1 we can easily get $\mathrm{E}((\boldsymbol{V}_k^{\boldsymbol{X}})^\top\mathbf{B}_{\boldsymbol{X}}\boldsymbol{V}_k^{\boldsymbol{X}})^2 = O(1)\mathrm{E}^2((\boldsymbol{V}_k^{\boldsymbol{X}})^\top\mathbf{B}_{\boldsymbol{X}}\boldsymbol{V}_k^{\boldsymbol{X}}) = O(\mathrm{tr}^2(\mathbf{B}_{\boldsymbol{X}}^2))$, and similarly, we get $\mathrm{E}((\boldsymbol{V}_k^{\boldsymbol{Y}})^\top\mathbf{B}_{\boldsymbol{Y}}\boldsymbol{V}_k^{\boldsymbol{Y}})^2 = O(\mathrm{tr}^2(\mathbf{B}_{\boldsymbol{Y}}^2))$. Hence $\mathrm{var}(C_3) = o_p(\tilde{\sigma}_2^4)$ and $C_3/\tilde{\sigma}_2^2 \xrightarrow{p} 1$. By using condition (C4$'$), we have $\mathrm{var}(C_4) = o_p(\tilde{\sigma}_2^4)$, which implies that $C_4 = o_p(\tilde{\sigma}_2^2)$.

Next, we will prove the second result. Note that

$$\tilde{\sigma}_2^{-2}\sum_{i=2}^{n}\mathrm{E}[\tilde{Z}_i^2 I(|\tilde{Z}_i| > \epsilon\tilde{\sigma}_2|)|\tilde{\mathcal{F}}_{i-1}] \le \tilde{\sigma}_2^{-4}\epsilon^{-2}\sum_{i=2}^{n}\mathrm{E}[\tilde{Z}_i^4|\tilde{\mathcal{F}}_{i-1}].$$

Accordingly, the assertion of this lemma is true if we can show $\mathrm{E}\left\{\sum_{i=2}^{n}\mathrm{E}[\tilde{Z}_i^4|\tilde{\mathcal{F}}_{i-1}]\right\} = o(\tilde{\sigma}_2^4)$. Note that

$$\mathrm{E}\left\{\sum_{i=2}^{n}\mathrm{E}[\tilde{Z}_i^4|\tilde{\mathcal{F}}_{i-1}]\right\} = \sum_{i=2}^{n}\mathrm{E}(\tilde{Z}_i^4) = O(n^{-4})\sum_{i=2}^{n}\mathrm{E}(\sum_{k=1}^{i-1}(\boldsymbol{V}_k^{\boldsymbol{X}})^\top\mathbf{B}_{\boldsymbol{X}}\boldsymbol{V}_k^{\boldsymbol{X}}(\boldsymbol{V}_k^{\boldsymbol{Y}})^\top\mathbf{B}_{\boldsymbol{Y}}\boldsymbol{V}_k^{\boldsymbol{Y}})^4,$$

which can be decomposed as $3\tilde{Q} + \tilde{P}$. Here

$$\begin{aligned}
\tilde{Q} =& O(n^{-4})\sum_{i=2}^{n}\sum_{s<t}^{i-1}\sum^{i-1}\mathrm{E}((\boldsymbol{V}_i^{\boldsymbol{X}})^\top\mathbf{B}_{\boldsymbol{X}}\boldsymbol{V}_s^{\boldsymbol{X}}(\boldsymbol{V}_s^{\boldsymbol{X}})^\top\mathbf{B}_{\boldsymbol{X}}\boldsymbol{V}_i^{\boldsymbol{X}}(\boldsymbol{V}_i^{\boldsymbol{X}})^\top\mathbf{B}_{\boldsymbol{X}}\boldsymbol{V}_t^{\boldsymbol{X}}(\boldsymbol{V}_t^{\boldsymbol{X}})^\top\mathbf{B}_{\boldsymbol{X}}\boldsymbol{V}_i^{\boldsymbol{X}}\\
&\times(\boldsymbol{V}_i^{\boldsymbol{Y}})^\top\mathbf{B}_{\boldsymbol{Y}}\boldsymbol{V}_s^{\boldsymbol{Y}}(\boldsymbol{V}_s^{\boldsymbol{Y}})^\top\mathbf{B}_{\boldsymbol{Y}}\boldsymbol{V}_i^{\boldsymbol{Y}}(\boldsymbol{V}_i^{\boldsymbol{Y}})^\top\mathbf{B}_{\boldsymbol{Y}}\mathbf{B}_{\boldsymbol{Y}}\boldsymbol{V}_t^{\boldsymbol{Y}}(\boldsymbol{V}_t^{\boldsymbol{Y}})^\top\mathbf{B}_{\boldsymbol{Y}}\boldsymbol{V}_i^{\boldsymbol{Y}}),\\
\tilde{P} =& O(n^{-4})\sum_{j=2}^{n}\sum_{i=1}^{j-1}\mathrm{E}((\boldsymbol{V}_i^{\boldsymbol{X}})^\top\mathbf{B}_{\boldsymbol{X}}\boldsymbol{V}_k^{\boldsymbol{X}})^4\mathrm{E}((\boldsymbol{V}_i^{\boldsymbol{Y}})^\top\mathbf{B}_{\boldsymbol{Y}}\boldsymbol{V}_k^{\boldsymbol{Y}})^4.
\end{aligned}$$

Obviously, $\tilde{Q} = O(n^{-1})\mathrm{E}((\boldsymbol{V}_i^{\boldsymbol{X}})^\top\mathbf{B}_{\boldsymbol{X}}\boldsymbol{V}_s^{\boldsymbol{X}}(\boldsymbol{V}_s^{\boldsymbol{X}})^\top\mathbf{B}_{\boldsymbol{X}}\boldsymbol{V}_i^{\boldsymbol{X}})^2\mathrm{E}((\boldsymbol{V}_i^{\boldsymbol{Y}})^\top\mathbf{B}_{\boldsymbol{Y}}\boldsymbol{V}_s^{\boldsymbol{Y}}(\boldsymbol{V}_s^{\boldsymbol{Y}})^\top\mathbf{B}_{\boldsymbol{Y}}\boldsymbol{V}_i^{\boldsymbol{Y}})^2$. By Lemma 1,we have

$$\mathrm{E}((\boldsymbol{V}_i^{\boldsymbol{X}})^\top\mathbf{B}_{\boldsymbol{X}}\boldsymbol{V}_s^{\boldsymbol{X}}(\boldsymbol{V}_s^{\boldsymbol{X}})^\top\mathbf{B}_{\boldsymbol{X}}\boldsymbol{V}_i^{\boldsymbol{X}})^2 = O(1)\mathrm{E}^2((\boldsymbol{V}_i^{\boldsymbol{X}})^\top\mathbf{B}_{\boldsymbol{X}}\boldsymbol{V}_s^{\boldsymbol{X}}(\boldsymbol{V}_s^{\boldsymbol{X}})^\top\mathbf{B}_{\boldsymbol{X}}\boldsymbol{V}_i^{\boldsymbol{X}}) = O(\mathrm{tr}^4(\mathbf{B}_{\boldsymbol{X}}^2))$$

and $\mathrm{E}((\boldsymbol{V}_i^X)^\top \mathbf{B}_X \boldsymbol{V}_k^X)^4 = O(\mathrm{tr}^4(\boldsymbol{V}_X^2))$. Then we can obtain that $\tilde{Q} = o(\tilde{\sigma}_2^4)$, $\tilde{P} = o(\tilde{\sigma}_2^4)$ and $T_2/\sqrt{\mathrm{var}(T_2)} \xrightarrow{d} \mathcal{N}(0,1)$, where $\mathrm{var}(T_2) = 2n(n-1)^{-1}\mathrm{tr}(\mathbf{B}_X^2)\mathrm{tr}(\mathbf{B}_Y^2)$. □

Proof of Proposition 3: Under H_0, we can decompose T_ρ as follows,

$$\begin{aligned}
T_\rho =& \frac{1}{(n-1)(n-2)(n-3)} \sum^* \{\mathbf{U}(\boldsymbol{X}_i - \boldsymbol{X}_j)^\top \mathbf{U}(\boldsymbol{X}_k - \boldsymbol{X}_\ell)\mathbf{U}(\boldsymbol{Y}_i - \boldsymbol{Y}_\ell)^\top \mathbf{U}(\boldsymbol{Y}_k - \boldsymbol{Y}_j)\} \\
=& \frac{1}{(n-1)(n-2)(n-3)} \sum^* (\boldsymbol{V}_i^X + \boldsymbol{V}_j^X + \boldsymbol{W}_{ij}^X)^\top (\boldsymbol{V}_k^X + \boldsymbol{V}_\ell^X + \boldsymbol{W}_{k\ell}^X) \times (\boldsymbol{V}_i^Y + \boldsymbol{V}_\ell^Y + \boldsymbol{W}_{i\ell}^Y)^\top (\boldsymbol{V}_k^Y + \boldsymbol{V}_j^Y + \boldsymbol{W}_{kj}^Y) \\
=& \frac{2}{n-1} \sum^* \{(\boldsymbol{V}_i^X)^\top \boldsymbol{V}_k^X (\boldsymbol{V}_i^Y)^\top \boldsymbol{V}_k^Y\} + \frac{12}{(n-1)(n-2)} \sum^* \{(\boldsymbol{V}_i^X)^\top \boldsymbol{V}_k^X (\boldsymbol{V}_i^Y)^\top \boldsymbol{V}_\ell^Y\} \\
&+ \frac{4}{(n-1)(n-2)(n-3)} \sum^* \{(\boldsymbol{V}_i^X)^\top \boldsymbol{V}_k^X (\boldsymbol{V}_j^Y)^\top \boldsymbol{V}_\ell^Y\} \\
&+ O_p(n^{-3}) \sum^* \Big[(\boldsymbol{W}_{ij}^X)^\top \boldsymbol{V}_k^X (\boldsymbol{V}_j^Y)^\top \boldsymbol{V}_\ell^Y + (\boldsymbol{W}_{ij}^X)^\top \boldsymbol{W}_{k\ell}^X \mathbf{U}(\boldsymbol{Y}_i - \boldsymbol{Y}_j)^\top \mathbf{U}(\boldsymbol{Y}_k - \boldsymbol{Y}_\ell) \\
&\quad + (\boldsymbol{W}_{ij}^X)^\top \boldsymbol{V}_k^X (\boldsymbol{V}_i^Y)^\top \boldsymbol{V}_\ell^Y + (\boldsymbol{W}_{ij}^X)^\top \boldsymbol{V}_k^X (\boldsymbol{W}_{i\ell}^Y)^\top \boldsymbol{V}_k^Y + (\boldsymbol{W}_{ij}^X)^\top \boldsymbol{W}_{k\ell}^X (\boldsymbol{W}_{i\ell}^Y)^\top \boldsymbol{W}_{kj}^Y \\
&\quad + \mathbf{U}(\boldsymbol{X}_i - \boldsymbol{X}_j)^\top \mathbf{U}(\boldsymbol{X}_k - \boldsymbol{X}_\ell)(\boldsymbol{W}_{i\ell}^Y)^\top \boldsymbol{W}_{kj}^Y + (\boldsymbol{W}_{ij}^X)^\top \boldsymbol{W}_{k\ell}^X (\boldsymbol{W}_{i\ell}^Y)^\top \boldsymbol{V}_k^Y \Big] \doteq J_1 + J_2 + J_3 + J_4.
\end{aligned}$$

Based on Lemma 6, it can be concluded that $J_1/\sigma_2 \xrightarrow{d} \mathcal{N}(0,1)$, where $\sigma_2^2 = 8n(n-1)^{-1}\mathrm{tr}(\mathbf{B}_X^2)\mathrm{tr}(\mathbf{B}_Y^2)$. Thus we only need to show the other parts are all $o_p(\sigma_2)$. In fact,

$$\mathrm{E}(J_2^2) = O(n^{-1})\mathrm{E}((\boldsymbol{V}_i^X)^\top \boldsymbol{V}_k^X (\boldsymbol{V}_i^Y)^\top \boldsymbol{V}_\ell^Y)^2 = O(n^{-1})\mathrm{tr}(\mathbf{B}_X^2)\mathrm{tr}(\mathbf{B}_Y^2) = o_p(\sigma_2^2),$$

$$\mathrm{E}(J_3^2) = O(n^{-2})\mathrm{E}((\boldsymbol{V}_i^X)^\top \boldsymbol{V}_k^X (\boldsymbol{V}_j^Y)^\top \boldsymbol{V}_\ell^Y)^2 = O(n^{-2})\mathrm{tr}(\mathbf{B}_X^2)\mathrm{tr}(\mathbf{B}_Y^2) = o_p(\sigma_2^2).$$

For J_4, we just consider the first part in J_4, and rest part can be handled in the similar way.

$$\begin{aligned}
\mathrm{E}(O(n^{-6}) \sum^* (\boldsymbol{W}_{ij}^X)^\top \boldsymbol{V}_k^X (\boldsymbol{V}_j^Y)^\top \boldsymbol{V}_\ell^Y)^2 =& O(n^{-3})\mathrm{E}((\boldsymbol{W}_{ij}^X)^\top \boldsymbol{V}_k^X (\boldsymbol{V}_j^Y)^\top \boldsymbol{V}_\ell^Y)^2 \\
=& O(n^{-3})\mathrm{E}((\boldsymbol{W}_{ij}^X)^\top \boldsymbol{V}_k^X (\boldsymbol{V}_k^X)^\top \boldsymbol{W}_{ij}^X)\mathrm{E}((\boldsymbol{V}_j^Y)^\top \boldsymbol{V}_\ell^Y)^2 = O(n^{-3})\mathrm{E}((\boldsymbol{W}_{ij}^X)^\top \mathbf{B}_X \boldsymbol{W}_{ij}^X)\mathrm{tr}(\mathbf{B}_Y^2).
\end{aligned}$$

Next, we will show $\mathrm{E}((\boldsymbol{W}_{ij}^X)^\top \mathbf{B}_X \boldsymbol{W}_{ij}^X) = O_p(\mathrm{tr}(\mathbf{B}_X^2))$. In fact, $\mathrm{E}(\mathbf{U}(\boldsymbol{X}_i - \boldsymbol{X}_j)^\top \mathbf{B}_X \mathbf{U}(\boldsymbol{X}_i - \boldsymbol{X}_j)) = O_p(\mathrm{tr}(\mathbf{B}_X^2))$, because

$$\begin{aligned}
\mathrm{E}(\mathbf{U}(\boldsymbol{X}_i - \boldsymbol{X}_j)^\top \mathbf{B}_X \mathbf{U}(\boldsymbol{X}_i - \boldsymbol{X}_j)) =& \mathbf{U}(\boldsymbol{X}_i - \boldsymbol{X}_j)^\top \mathbf{B}_X \mathbf{U}(\boldsymbol{X}_i - \boldsymbol{X}_j) = \{\frac{\boldsymbol{X}_i - \boldsymbol{X}_j}{||\boldsymbol{X}_i - \boldsymbol{X}_j||}\}^\top \mathbf{B}_X \mathbf{U}(\boldsymbol{X}_i - \boldsymbol{X}_j) \\
=& \{\frac{\boldsymbol{X}_i - \boldsymbol{X}_0 + \boldsymbol{X}_0 - \boldsymbol{X}_j}{||\boldsymbol{X}_i - \boldsymbol{X}_j||}\}^\top \mathbf{B}_X \mathbf{U}(\boldsymbol{X}_i - \boldsymbol{X}_j) \\
=& \{\frac{\boldsymbol{X}_i - \boldsymbol{X}_0}{||\boldsymbol{X}_i - \boldsymbol{X}_j||}\}^\top \mathbf{B}_X \mathbf{U}(\boldsymbol{X}_i - \boldsymbol{X}_j) + \{\frac{\boldsymbol{X}_0 - \boldsymbol{X}_j}{||\boldsymbol{X}_i - \boldsymbol{X}_j||}\}^\top \mathbf{B}_X \mathbf{U}(\boldsymbol{X}_i - \boldsymbol{X}_j) \\
=& \mathbf{U}(\boldsymbol{X}_i - \boldsymbol{X}_0)^\top \mathbf{B}_X \mathbf{U}(\boldsymbol{X}_i - \boldsymbol{X}_j)\frac{||\boldsymbol{X}_i - \boldsymbol{X}_0||}{||\boldsymbol{X}_i - \boldsymbol{X}_j||} + \mathbf{U}(\boldsymbol{X}_0 - \boldsymbol{X}_j)^\top \mathbf{B}_X \mathbf{U}(\boldsymbol{X}_i - \boldsymbol{X}_j)\frac{||\boldsymbol{X}_0 - \boldsymbol{X}_j||}{||\boldsymbol{X}_i - \boldsymbol{X}_j||}.
\end{aligned}$$

Additionally,

$$\begin{aligned}
\mathrm{E}(\mathbf{U}(\boldsymbol{X}_i - \boldsymbol{X}_0)^\top \mathbf{B}_X \mathbf{U}(\boldsymbol{X}_i - \boldsymbol{X}_j)\frac{||\boldsymbol{X}_i - \boldsymbol{X}_0||}{||\boldsymbol{X}_i - \boldsymbol{X}_j||}) =& \mathrm{E}(\mathrm{E}(\mathbf{U}(\boldsymbol{X}_i - \boldsymbol{X}_0)^\top \mathbf{B}_X \mathbf{U}(\boldsymbol{X}_i - \boldsymbol{X}_j)\frac{||\boldsymbol{X}_i - \boldsymbol{X}_0||}{||\boldsymbol{X}_i - \boldsymbol{X}_j||}|\boldsymbol{X}_i)) \\
=& \mathrm{E}((\boldsymbol{V}_i^X)^\top \mathbf{B}_X \boldsymbol{V}_i^X)\mathrm{E}(\frac{||\boldsymbol{X}_i - \boldsymbol{X}_0||}{||\boldsymbol{X}_i - \boldsymbol{X}_j||}) = O_p(\mathrm{tr}(\mathbf{B}_X^2)).
\end{aligned}$$

So, we have $\mathrm{E}(O(n^{-6})\sum^*(\boldsymbol{W}_{ij}^{\boldsymbol{X}})^\top\boldsymbol{V}_k^{\boldsymbol{X}}(\boldsymbol{V}_j^{\boldsymbol{Y}})^\top\boldsymbol{V}_\ell^{\boldsymbol{Y}})^2 = O(n^{-3})\mathrm{tr}(\mathbf{B}_{\boldsymbol{X}}^2)\mathrm{tr}(\mathbf{B}_{\boldsymbol{Y}}^2) = o_p(\sigma_2^2)$, which completes this proof. $\square$

Proof of Proposition 4:

$$\begin{aligned}\hat{\sigma}_2^2 &= \frac{1}{2n^2(n-1)^2(n-2)^2(n-3)^2}\sum^*(\mathbf{U}(\boldsymbol{X}_i-\boldsymbol{X}_j)^\top\mathbf{U}(\boldsymbol{X}_k-\boldsymbol{X}_\ell))^2\times\sum^*(\mathbf{U}(\boldsymbol{Y}_i-\boldsymbol{Y}_j)^\top\mathbf{U}(\boldsymbol{Y}_k-\boldsymbol{Y}_\ell))^2\\ &= \frac{8}{n^2(n-1)^2}\sum^*((\boldsymbol{V}_i^{\boldsymbol{X}})^\top\boldsymbol{V}_j^{\boldsymbol{X}})^2\sum^*((\boldsymbol{V}_i^{\boldsymbol{Y}})^\top\boldsymbol{V}_j^{\boldsymbol{Y}})^2 + o_p(\sigma_2^2).\end{aligned}$$

Obviously, $\mathrm{E}(8\{n^2(n-1)^2\}^{-1}\sum^*((\boldsymbol{V}_i^{\boldsymbol{X}})^\top\boldsymbol{V}_j^{\boldsymbol{X}})^2\sum^*((\boldsymbol{V}_i^{\boldsymbol{Y}})^\top\boldsymbol{V}_j^{\boldsymbol{Y}})^2) = 8n(n-1)^{-1}\mathrm{tr}(\mathbf{B}_{\boldsymbol{X}}^2)\mathrm{tr}(\mathbf{B}_{\boldsymbol{Y}}^2)$. Then $\mathrm{E}(\hat{\sigma}_2^2) = \sigma_2^2(1+o_p(1))$ and

$$\begin{aligned}&\mathrm{var}(\frac{8}{n^2(n-1)^2}\sum^*((\boldsymbol{V}_i^{\boldsymbol{X}})^\top\boldsymbol{V}_j^{\boldsymbol{X}})^2\sum^*((\boldsymbol{V}_i^{\boldsymbol{Y}})^\top\boldsymbol{V}_j^{\boldsymbol{Y}})^2)^2\\ =&O(n^{-4})\mathrm{E}((\boldsymbol{V}_i^{\boldsymbol{X}})^\top\boldsymbol{V}_j^{\boldsymbol{X}})^4\mathrm{E}((\boldsymbol{V}_i^{\boldsymbol{Y}})^\top\boldsymbol{V}_j^{\boldsymbol{Y}})^4 + O(n^{-2})\mathrm{E}\{((\boldsymbol{V}_i^{\boldsymbol{X}})^\top\boldsymbol{V}_j^{\boldsymbol{X}})^2((\boldsymbol{V}_i^{\boldsymbol{X}})^\top\boldsymbol{V}_\ell^{\boldsymbol{X}})^2\}\mathrm{E}\{((\boldsymbol{V}_i^{\boldsymbol{Y}})^\top\boldsymbol{V}_j^{\boldsymbol{Y}})^2((\boldsymbol{V}_i^{\boldsymbol{Y}})^\top\boldsymbol{V}_\ell^{\boldsymbol{Y}})^2\}\\ =&O(n^{-4}\mathrm{tr}^2(\mathbf{B}_{\boldsymbol{X}}^2)\mathrm{tr}^2(\mathbf{B}_{\boldsymbol{Y}}^2)) + O(n^{-2}\mathrm{tr}^2(\mathbf{B}_{\boldsymbol{X}}^2)\mathrm{tr}^2(\mathbf{B}_{\boldsymbol{Y}}^2)) = o_p(\sigma_2^4).\end{aligned}$$

Thus, $\hat{\sigma}_2^2 = \sigma_2^2(1+o_p(1))$. $\square$

Next, we will prove the corresponding results under the alternative hypothesis. Let $\boldsymbol{V}_i^{\boldsymbol{X}^*} = \mathrm{E}(\mathbf{U}(\boldsymbol{X}_i^*-\boldsymbol{X}_j^*)|\boldsymbol{X}_i^*)$, $\boldsymbol{V}_j^{\boldsymbol{X}^*} = -\mathrm{E}(\mathbf{U}(\boldsymbol{X}_i^*-\boldsymbol{X}_j^*)|\boldsymbol{X}_j^*)$, $\boldsymbol{V}_i^{\boldsymbol{Y}^*} = \mathrm{E}(\mathbf{U}(\boldsymbol{Y}_i^*-\boldsymbol{Y}_j^*)|\boldsymbol{Y}_i^*)$, $\boldsymbol{V}_j^{\boldsymbol{Y}^*} = -\mathrm{E}(\mathbf{U}(\boldsymbol{Y}_i^*-\boldsymbol{Y}_j^*)|\boldsymbol{Y}_j^*)$, $\mathbf{U}(\boldsymbol{Y}_i^*-\boldsymbol{Y}_j^*) = \boldsymbol{V}_i^{\boldsymbol{Y}^*}+\boldsymbol{V}_j^{\boldsymbol{Y}^*}+\boldsymbol{W}_{ij}^{\boldsymbol{Y}^*}$, $\mathbf{U}(\boldsymbol{Y}_k^*-\boldsymbol{Y}_\ell^*) = \boldsymbol{V}_k^{\boldsymbol{Y}^*}+\boldsymbol{V}_\ell^{\boldsymbol{Y}^*}+\boldsymbol{W}_{k\ell}^{\boldsymbol{Y}^*}$, $\mathbf{U}(\boldsymbol{X}_i^*-\boldsymbol{X}_j^*) = \boldsymbol{V}_i^{\boldsymbol{X}^*}+\boldsymbol{V}_j^{\boldsymbol{X}^*}+\boldsymbol{W}_{ij}^{\boldsymbol{X}^*}$, $\mathbf{U}(\boldsymbol{X}_k^*-\boldsymbol{X}_\ell^*) = \boldsymbol{V}_k^{\boldsymbol{X}^*}+\boldsymbol{V}_\ell^{\boldsymbol{X}^*}+\boldsymbol{W}_{k\ell}^{\boldsymbol{X}^*}$, $\mathbf{B}_{\boldsymbol{X}}^* = \mathrm{E}(\boldsymbol{V}_i^{\boldsymbol{X}^*}(\boldsymbol{V}_i^{\boldsymbol{X}^*})^\top)$ and $\mathbf{B}_{\boldsymbol{Y}}^* = \mathrm{E}(\boldsymbol{V}_i^{\boldsymbol{Y}^*}(\boldsymbol{V}_i^{\boldsymbol{Y}^*})^\top)$.

Taking the same procedure as in Proposition 3, under conditions (C1$'$), (C2$'$) and (C4$'$), we have

$$\begin{aligned}T_\rho/2 =&T_2 + \frac{1}{n-1}\sum_{1\le i<j\le n}(\tilde{r}_{ij}^{\boldsymbol{X}^*})^{-1}\tilde{r}_{ij}^{\boldsymbol{Y}^*}(\tilde{r}_{ij}^{\boldsymbol{X}^*})^{-1}\tilde{r}_{ij}^{\boldsymbol{Y}^*}(\boldsymbol{V}_i^{\boldsymbol{Y}^*})^\top\mathbf{M}_1^\top\mathbf{M}_1\boldsymbol{V}_j^{\boldsymbol{Y}^*}(\boldsymbol{V}_i^{\boldsymbol{Y}^*})^\top\boldsymbol{V}_j^{\boldsymbol{Y}^*}\\ &+\frac{1}{n-1}\sum_{1\le i<j\le n}(\tilde{r}_{ij}^{\boldsymbol{Y}^*})^{-1}\tilde{r}_{ij}^{\boldsymbol{X}^*}(\tilde{r}_{ij}^{\boldsymbol{Y}^*})^{-1}\tilde{r}_{ij}^{\boldsymbol{X}^*}(\boldsymbol{V}_i^{\boldsymbol{X}^*})^\top\mathbf{M}_2^\top\mathbf{M}_2\boldsymbol{V}_j^{\boldsymbol{X}^*}(\boldsymbol{V}_i^{\boldsymbol{X}^*})^\top\boldsymbol{V}_j^{\boldsymbol{X}^*}\\ &+\frac{1}{n-1}\sum_{1\le i<j\le n}(\tilde{r}_{ij}^{\boldsymbol{X}^*})^{-1}\tilde{r}_{ij}^{\boldsymbol{Y}^*}(\tilde{r}_{ij}^{\boldsymbol{Y}^*})^{-1}\tilde{r}_{ij}^{\boldsymbol{X}^*}(\boldsymbol{V}_i^{\boldsymbol{Y}^*})^\top\mathbf{M}_1^\top\boldsymbol{V}_j^{\boldsymbol{X}^*}(\boldsymbol{V}_i^{\boldsymbol{Y}^*})^\top\mathbf{M}_2\boldsymbol{V}_j^{\boldsymbol{X}^*}\\ &+\frac{1}{n-1}\sum_{1\le i<j\le n}(\tilde{r}_{ij}^{\boldsymbol{Y}^*})^{-1}\tilde{r}_{ij}^{\boldsymbol{X}^*}(\tilde{r}_{ij}^{\boldsymbol{X}^*})^{-1}\tilde{r}_{ij}^{\boldsymbol{Y}^*}(\boldsymbol{V}_i^{\boldsymbol{X}^*})^\top\mathbf{M}_2^\top\boldsymbol{V}_j^{\boldsymbol{Y}^*}(\boldsymbol{V}_i^{\boldsymbol{X}^*})^\top\mathbf{M}_1\boldsymbol{V}_j^{\boldsymbol{Y}^*}\\ &+\frac{1}{n-1}\sum_{1\le i<j\le n}(\tilde{r}_{ij}^{\boldsymbol{Y}^*})^{-1}\tilde{r}_{ij}^{\boldsymbol{X}^*}(\boldsymbol{V}_i^{\boldsymbol{X}^*})^\top\boldsymbol{V}_j^{\boldsymbol{X}^*}(\boldsymbol{V}_i^{\boldsymbol{Y}^*})^\top\mathbf{M}_2\boldsymbol{V}_j^{\boldsymbol{X}^*} + \frac{1}{n-1}\sum_{1\le i<j\le n}(\tilde{r}_{ij}^{\boldsymbol{X}^*})^{-1}\tilde{r}_{ij}^{\boldsymbol{Y}^*}(\boldsymbol{V}_i^{\boldsymbol{Y}^*})^\top\mathbf{M}_1^\top\boldsymbol{V}_j^{\boldsymbol{X}^*}(\boldsymbol{V}_i^{\boldsymbol{Y}^*})^\top\boldsymbol{V}_j^{\boldsymbol{Y}^*} + o_p(\sigma_2).\end{aligned}$$

By condition (C4$'$), we have

$$\begin{aligned}\mathrm{E}\left\{\frac{1}{n-1}\sum_{1\le i<j\le n}(\tilde{r}_{ij}^{\boldsymbol{Y}^*})^{-1}\tilde{r}_{ij}^{\boldsymbol{X}^*}(\boldsymbol{V}_i^{\boldsymbol{X}^*})^\top\boldsymbol{V}_j^{\boldsymbol{X}^*}(\boldsymbol{V}_i^{\boldsymbol{Y}^*})^\top\mathbf{M}_2\boldsymbol{V}_j^{\boldsymbol{X}^*}\right\}^2 =&O[\{\mathrm{E}((\tilde{r}_{ij}^{\boldsymbol{Y}^*})^{-1}\tilde{r}_{ij}^{\boldsymbol{X}^*})\}^2\mathrm{tr}(\mathbf{M}_2\mathbf{B}_{\boldsymbol{X}}^2\mathbf{M}_2^\top\mathbf{B}_{\boldsymbol{Y}})] = o(\sigma_2^2),\\ \mathrm{E}\left\{\frac{1}{n-1}\sum_{1\le i<j\le n}(\tilde{r}_{ij}^{\boldsymbol{X}^*})^{-1}\tilde{r}_{ij}^{\boldsymbol{Y}^*}(\boldsymbol{V}_i^{\boldsymbol{Y}^*})^\top\mathbf{M}_1^\top\boldsymbol{V}_j^{\boldsymbol{X}^*}(\boldsymbol{V}_i^{\boldsymbol{Y}^*})^\top\boldsymbol{V}_j^{\boldsymbol{Y}^*}\right\}^2 =&O[\{\mathrm{E}((\tilde{r}_{ij}^{\boldsymbol{X}^*})^{-1}\tilde{r}_{ij}^{\boldsymbol{Y}^*})\}^2\mathrm{tr}(\mathbf{M}_1\mathbf{B}_{\boldsymbol{Y}}^2\mathbf{M}_1^\top\mathbf{B}_{\boldsymbol{X}})] = o(\sigma_2^2),\\ \frac{1}{n-1}\sum_{1\le i<j\le n}(\tilde{r}_{ij}^{\boldsymbol{Y}^*})^{-1}\tilde{r}_{ij}^{\boldsymbol{X}^*}(\tilde{r}_{ij}^{\boldsymbol{Y}^*})^{-1}\tilde{r}_{ij}^{\boldsymbol{X}^*}(\boldsymbol{V}_i^{\boldsymbol{X}^*})^\top\mathbf{M}_2^\top\mathbf{M}_2\boldsymbol{V}_j^{\boldsymbol{X}^*}(\boldsymbol{V}_i^{\boldsymbol{X}^*})^\top\boldsymbol{V}_j^{\boldsymbol{X}^*} =&\frac{n}{2}\{\mathrm{E}((\tilde{r}_{ij}^{\boldsymbol{Y}^*})^{-1}\tilde{r}_{ij}^{\boldsymbol{X}^*})\}^2\mathrm{tr}(\mathbf{M}_2\mathbf{B}_{\boldsymbol{X}}^{*2}\mathbf{M}_2^\top) + o_p(\sigma_2),\\ \frac{1}{n-1}\sum_{1\le i<j\le n}(\tilde{r}_{ij}^{\boldsymbol{X}^*})^{-1}\tilde{r}_{ij}^{\boldsymbol{Y}^*}(\tilde{r}_{ij}^{\boldsymbol{X}^*})^{-1}\tilde{r}_{ij}^{\boldsymbol{Y}^*}(\boldsymbol{V}_i^{\boldsymbol{Y}^*})^\top\mathbf{M}_1^\top\mathbf{M}_1\boldsymbol{V}_j^{\boldsymbol{Y}^*}(\boldsymbol{V}_i^{\boldsymbol{Y}^*})^\top\boldsymbol{V}_j^{\boldsymbol{Y}^*} =&\frac{n}{2}\{\mathrm{E}((\tilde{r}_{ij}^{\boldsymbol{X}^*})^{-1}\tilde{r}_{ij}^{\boldsymbol{Y}^*})\}^2\mathrm{tr}(\mathbf{M}_1\mathbf{B}_{\boldsymbol{Y}}^{*2}\mathbf{M}_1^\top) + o_p(\sigma_2),\\ \frac{1}{n-1}\sum_{1\le i<j\le n}(\tilde{r}_{ij}^{\boldsymbol{X}^*})^{-1}\tilde{r}_{ij}^{\boldsymbol{Y}^*}(\tilde{r}_{ij}^{\boldsymbol{Y}^*})^{-1}\tilde{r}_{ij}^{\boldsymbol{X}^*}(\boldsymbol{V}_i^{\boldsymbol{Y}^*})^\top\mathbf{M}_1^\top\boldsymbol{V}_j^{\boldsymbol{X}^*}(\boldsymbol{V}_i^{\boldsymbol{Y}^*})^\top\mathbf{M}_2\boldsymbol{V}_j^{\boldsymbol{X}^*} =&\frac{n}{2}\{\mathrm{E}((\tilde{r}_{ij}^{\boldsymbol{X}^*})^{-1}\tilde{r}_{ij}^{\boldsymbol{Y}^*})\mathrm{E}((\tilde{r}_{ij}^{\boldsymbol{Y}^*})^{-1}\tilde{r}_{ij}^{\boldsymbol{X}^*})\}\mathrm{tr}(\mathbf{M}_1\mathbf{B}_{\boldsymbol{Y}}^*\mathbf{M}_2\mathbf{B}_{\boldsymbol{X}}^*) + o_p(\sigma_2),\end{aligned}$$

$$\frac{1}{n-1}\sum_{1\le i<j\le n},(\tilde r_{ij}^{Y^*})^{-1}\tilde r_{ij}^{X^*}(\tilde r_{ij}^{X^*})^{-1}\tilde r_{ij}^{Y^*}(\boldsymbol V_i^{X^*})^\top\mathbf M_2^\top\boldsymbol V_j^{Y^*}(\boldsymbol V_i^{X^*})^\top\mathbf M_1\boldsymbol V_j^{Y^*}=\frac{n}{2}\{\mathrm E((\tilde r_{ij}^{X^*})^{-1}\tilde r_{ij}^{Y^*})\mathrm E((\tilde r_{ij}^{Y^*})^{-1}\tilde r_{ij}^{X^*})\}\mathrm{tr}(\mathbf M_2\mathbf B_X^*\mathbf M_1\mathbf B_Y^*)+o_p(\sigma_2).$$

Thus, $T_\rho/2 = T_2 + n\mathrm{tr}(\tilde{\mathbf\Lambda}^\top\tilde{\mathbf\Lambda})/2 + o_p(\sigma_2)$. According to the results of Theorem 3-(i), we can easily obtain the result. Here we complete the proof.

Proof of Theorem 4

First, we will prove the results under the null hypothesis. Define

$$T_\tau = \frac{1}{(n-1)(n-2)(n-3)}\sum^*\{\mathbf U(\boldsymbol X_i-\boldsymbol X_j)^\top\mathbf U(\boldsymbol X_k-\boldsymbol X_\ell)\mathbf U(\boldsymbol Y_i-\boldsymbol Y_j)^\top\mathbf U(\boldsymbol Y_k-\boldsymbol Y_\ell)\},$$
$$\hat\sigma_3^2 = \frac{2}{n^2(n-1)^2(n-2)^2(n-3)^2}\sum^*(\mathbf U(\boldsymbol X_i-\boldsymbol X_j)^\top\mathbf U(\boldsymbol X_k-\boldsymbol X_\ell))^2\times\sum^*(\mathbf U(\boldsymbol Y_i-\boldsymbol Y_j)^\top\mathbf U(\boldsymbol Y_k-\boldsymbol Y_\ell))^2.$$

Hence $T_{\mathrm{HT}} = T_\tau/\hat\sigma_3$. To prove Theorem 4-(i), we only need to prove the following propositions.

Proposition 5. *Under conditions (C1), (C2) given in Section 2 and H_0 in (1) ,as $n\to\infty$, $T_\tau/\sigma_3 \xrightarrow{d} \mathcal N(0,1)$, where $\sigma_3^2 = 32n(n-1)^{-1}\mathrm{tr}(\mathbf B_X^2)\mathrm{tr}(\mathbf B_Y^2)$.*

Proposition 6. *Under conditions (C1) and (C2) given in Section 2, as $n\to\infty$, $\hat\sigma_3^2/\sigma_3^2 \xrightarrow{p} 1$.*

Proof of Proposition 5: Under H_0, similar to T_ρ, we can decompose T_τ as follows:

$$\begin{aligned}
T_\tau =& \frac{1}{(n-1)(n-2)(n-3)}\sum^*\{\mathbf U(\boldsymbol X_i-\boldsymbol X_j)^\top\mathbf U(\boldsymbol X_k-\boldsymbol X_\ell)\mathbf U(\boldsymbol Y_i-\boldsymbol Y_j)^\top\mathbf U(\boldsymbol Y_k-\boldsymbol Y_\ell)\}\\
=& \frac{4}{n-1}\sum^*\{(\boldsymbol V_i^X)^\top\boldsymbol V_k^X(\boldsymbol V_i^Y)^\top\boldsymbol V_k^Y\} + \frac{8}{(n-1)(n-2)}\sum^*\{(\boldsymbol V_i^X)^\top\boldsymbol V_k^X(\boldsymbol V_i^Y)^\top\boldsymbol V_\ell^Y\}\\
&+\frac{4}{(n-1)(n-2)(n-3)}\sum^*\{(\boldsymbol V_i^X)^\top\boldsymbol V_k^X(\boldsymbol V_j^Y)^\top\boldsymbol V_\ell^Y\} + O_p(n^{-3})\sum^*[(\boldsymbol W_{ij}^X)^\top\boldsymbol V_k^X(\boldsymbol V_j^Y)^\top\boldsymbol V_\ell^Y\\
&+(\boldsymbol W_{ij}^X)^\top\boldsymbol V_k^X(\boldsymbol V_i^Y)^\top\boldsymbol V_\ell^Y + (\boldsymbol W_{ij}^X)^\top\boldsymbol V_k^X(\boldsymbol W_{ij}^Y)^\top\boldsymbol V_k^Y + (\boldsymbol W_{ij}^X)^\top\boldsymbol W_{k\ell}^X\mathbf U(\boldsymbol Y_i-\boldsymbol Y_j)^\top\mathbf U(\boldsymbol Y_k-\boldsymbol Y_\ell)\\
&+\mathbf U(\boldsymbol X_i-\boldsymbol X_j)^\top\mathbf U(\boldsymbol X_k-\boldsymbol X_\ell)(\boldsymbol W_{ij}^Y)^\top\boldsymbol W_{k\ell}^Y + (\boldsymbol W_{ij}^X)^\top\boldsymbol W_{k\ell}^X(\boldsymbol W_{ij}^Y)^\top\boldsymbol V_k^Y + (\boldsymbol W_{ij}^X)^\top\boldsymbol W_{k\ell}^X(\boldsymbol W_{ij}^Y)^\top\boldsymbol W_{k\ell}^Y]\\
=& \frac{4}{n-1}\sum^*\{(\boldsymbol V_i^X)^\top\boldsymbol V_k^X(\boldsymbol V_i^Y)^\top\boldsymbol V_k^Y\} + o_p(\sigma_3).
\end{aligned}$$

Based on Lemma 6, it can be concluded that $T_\tau/\sigma_3 \xrightarrow{d} \mathcal N(0,1)$. □

Proof of Proposition 6:

$$\begin{aligned}
\hat\sigma_3^2 =& \frac{2}{n^2(n-1)^2(n-2)^2(n-3)^2}\sum^*(\mathbf U(\boldsymbol X_i-\boldsymbol X_j)^\top\mathbf U(\boldsymbol X_k-\boldsymbol X_\ell))^2\times\sum^*(\mathbf U(\boldsymbol Y_i-\boldsymbol Y_j)^\top\mathbf U(\boldsymbol Y_k-\boldsymbol Y_\ell))^2\\
=& \frac{32}{n^2(n-1)^2}\sum^*((\boldsymbol V_i^X)^\top\boldsymbol V_j^X)^2\sum^*((\boldsymbol V_i^Y)^\top\boldsymbol V_j^Y)^2 + o_p(\sigma_3^2).
\end{aligned}$$

$\mathrm E(32\{n^2(n-1)^2\}^{-1}\sum^*((\boldsymbol V_i^X)^\top\boldsymbol V_j^X)^2\sum^*((\boldsymbol V_i^Y)^\top\boldsymbol V_j^Y)^2) = \sigma_3^2$. Since

$$\begin{aligned}
&\mathrm{var}\left\{\frac{2}{n^2(n-1)^2}\sum^*(\mathbf U(\boldsymbol X_i-\boldsymbol X_j)^\top\mathbf U(\boldsymbol X_k-\boldsymbol X_\ell))^2\sum^*(\mathbf U(\boldsymbol Y_i-\boldsymbol Y_j)^\top\mathbf U(\boldsymbol Y_k-\boldsymbol Y_\ell))^2\right\}\\
=&O(n^{-4})\mathrm E((\boldsymbol V_i^X)^\top\boldsymbol V_j^X)^4\mathrm E((\boldsymbol V_i^Y)^\top\boldsymbol V_j^Y)^4) + O(n^{-2})E\{((\boldsymbol V_i^X)^\top\boldsymbol V_j^X)^2((\boldsymbol V_i^X)^\top\boldsymbol V_\ell^X)^2\}E\{((\boldsymbol V_i^Y)^\top\boldsymbol V_j^Y)^2((\boldsymbol V_i^Y)^\top\boldsymbol V_\ell^Y)^2\}\\
=&O(n^{-4}\mathrm{tr}^2(\mathbf B_X^2)\mathrm{tr}^2(\mathbf B_Y^2)) + O(n^{-2}\mathrm{tr}^2(\mathbf B_X^2)\mathrm{tr}^2(\mathbf B_Y^2)) = o_p(\sigma_3^4),
\end{aligned}$$

we can see that $\hat\sigma_3^2 = \sigma_3^2(1+o_p(1))$. □

The proof of Theorem 4-(ii) are very similar to the proof of Theorem 3-(ii), hence we omit the details here.

References

[1] P. J. Bickel, E. Levina, Covariance regularization by thresholding, Annals of Statistics 36 (2008) 2577–2604.
[2] T. Bodnar, H. Dette, N. Parolya, Testing for independence of large dimensional vectors, The Annals of Statistics 47 (2019) 2977–3008.
[3] S. X. Chen, Y.-L. Qin, A two-sample test for high-dimensional data with applications to gene-set testing, The Annals of Statistics 38 (2010) 808–835.
[4] L. Feng, B. Liu, High-dimensional rank tests for sphericity, Journal of Multivariate Analysis 155 (2017) 217–233.
[5] L. Feng, C. Zou, Z. Wang, Multivariate-sign-based high-dimensional tests for the two-sample location problem, Journal of the American Statistical Association 111 (2016) 721–735.
[6] B. Guo, S. X. Chen, Tests for high dimensional generalized linear models, Journal of the Royal Statistical Society: Series B (Statistical Methodology) 78 (2016) 1079–1102.
[7] X. Guo, H. Zhang, T. Tian, Development of stock correlation networks using mutual information and financial big data, Plos One 13 (2018) e0195941.
[8] P. Hall, C. C. Heyde, Martingale limit theory and its application, Academic press, 2014.
[9] T. P. Hettmansperger, R. H. Randles, A practical affine equivariant multivariate median, Biometrika 89 (2002) 851–860.
[10] D. Jiang, Z. Bai, S. Zheng, Testing the independence of sets of large-dimensional variables, Science China Mathematics 56 (2013) 135–147.
[11] L. S. Junior, I. D. P. Franca, Correlation of financial markets in times of crisis, Physica A-statistical Mechanics & Its Applications 391 (2012) 187–208.
[12] D. Y. Kenett, M. Raddant, L. Zatlavi, T. Lux, E. Ben-Jacob, Correlations and dependencies in the global financial village, in: International Journal of Modern Physics: Conference Series, volume 16, pp. 13–28.
[13] D. Leung, M. Drton, Testing independence in high dimensions with sums of squares of rank correlations, Statistics 46 (2015) 280–307.
[14] J. Li, S. X. Chen, et al., Two sample tests for high-dimensional covariance matrices, The Annals of Statistics 40 (2012) 908–940.
[15] J. Möttönen, H. Oja, Multivariate spatial sign and rank methods, Journaltitle of Nonparametric Statistics 5 (1995) 201–213.
[16] H. Oja, Multivariate nonparametric methods with R: an approach based on spatial signs and ranks, Springer Science & Business Media, 2010.
[17] D. Paindaveine, T. Verdebout, On high-dimensional sign tests, Bernoulli 22 (2016) 1745–1769.
[18] T. Preis, D. Y. Kenett, H. E. Stanley, D. Helbing, E. Benjacob, Quantifying the behavior of stock correlations under market stress, Sci Rep 2 (2012) 752.
[19] A. Sensoy, S. Yuksel, M. Erturka, Analysis of cross-correlations between financial markets after the 2008 crisis, Physica A 392 (2013) 5027–5045.
[20] M. S. Srivastava, N. Reid, Testing the structure of the covariance matrix with fewer observations than the dimension, Journal of Multivariate Analysis 112 (2012) 156–171.
[21] K. Sunil, D. Nivedita, Correlation and network analysis of global financial indices, Physical Review E Statistical Nonlinear & Soft Matter Physics 86 (2012) 026101.
[22] S. Taskinen, A. Kankainen, H. Oja, Sign test of independence between two random vectors, Statistics & Probability Letters 62 (2003) 9–21.
[23] S. Taskinen, H. Oja, R. H. Randles, Multivariate nonparametric tests of independence, Journal of the American Statistical Association 100 (2005) 916–925.
[24] I. Vodenska, A. P. Becker, D. Zhou, D. Y. Kenett, H. E. Stanley, S. Havlin, Community analysis of global financial markets, Risks 4 (2016) 13.
[25] L. Wang, B. Peng, R. Li, A high-dimensional nonparametric multivariate test for mean vector, Journal of the American Statistical Association 110 (2015) 1658.
[26] S. S. Wilks, On the independence of k sets of normally distributed statistical variables, Econometrica 3 (1935) 309–326.
[27] Y. Yang, G. Pan, Independence test for high dimensional data based on regularized canonical correlation coefficients, Annals of Statistics 43 (2015).
[28] K. Yata, M. Aoshima, High-dimensional inference on covariance structures via the extended cross-data-matrix methodology, Journal of Multivariate Analysis 151 (2016) 151–166.
[29] C. Zou, L. Peng, L. Feng, Z. Wang, Multivariate-sign-based high-dimensional tests for sphericity, Biometrika 1 (2014) 229–236.