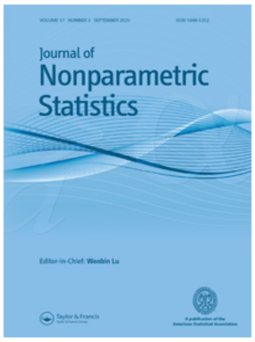

Journal of Nonparametric Statistics

ISSN: 1048-5252 (Print) 1029-0311 (Online) Journal homepage: www.tandfonline.com/journals/gnst20

An adaptive test of the independence of high-dimensional data based on Kendall rank correlation coefficient

Lijuan Han, Yun Ma & Junshan Xie

JOURNAL OF NONPARAMETRIC STATISTICS
2025, VOL. 37, NO. 3, 632–656
https://doi.org/10.1080/10485252.2024.2435852

An adaptive test of the independence of high-dimensional data based on Kendall rank correlation coefficient

Lijuan Han, Yun Ma and Junshan Xie

School of Mathematics and Statistics, Henan University, Kaifeng, People's Republic of China

ABSTRACT

The paper considers the test on the components independence of high-dimensional continuous random vectors based on the Kendall rank correlation coefficient. Under the null hypothesis, it proves that the sum-of-powers type statistics constructed by the Kendall rank correlation coefficient are asymptotically normally distributed and independent when the sum of the orders is odd, and they are also independent with the extreme-value-type test statistics of the Kendall rank correlation coefficient. Based on the asymptotic independence property, it proposes an adaptive test procedure which combines p-values from the sum-of-powers-type test and extreme-value-type test. Some simulations reveal that the proposed test method can significantly enhance the statistical power of the test while keep the type I error well controlled, and it is powerful against a wide range of alternatives (including dense, sparse or moderately sparse signals). To illustrate the use of the proposed adaptive test method, two real data sets are also analysed.

ARTICLE HISTORY
Received 2 June 2024
Accepted 23 November 2024

KEYWORDS
High-dimensional data; adaptive test; independence; Kendall rank correlation coefficient

MSC(2010):
Primary 62E20; Secondary 62G10

1. Introduction

Testing the mutual independence of the components of a random vector is one of the fundamental topics in statistics. As the independence of the components can simplify the modelling and inference tasks of statistical data analysis, the independent component analysis has many applications in causal inference (Chakraborty and Zhang 2019; Pfister et al. 2021), linguistics (Nguyen and Eisenstein 2017), dimension reduction (Sheng and Yin 2016). There have been lots of work on testing for component independence, such as Pearson (1920), Kendall (1938), Hoeffding (1948), Blum et al. (1961), Chen et al. (2010), Fan et al. (2015), Bilodeau and Nangue (2017) and Yao et al. (2018).

Assume that $\boldsymbol{X}$ is a continuous random vector with $\boldsymbol{X} = (\boldsymbol{X}_1, \boldsymbol{X}_2, \dots, \boldsymbol{X}_p)^T \in \mathbb{R}^p$. Let $(X_{1i}, X_{2i}, \dots, X_{pi})^T, i = 1, 2, \dots, n$ be a sample from $\boldsymbol{X}$. We consider the test

$$\begin{aligned} &H_0 : \boldsymbol{X}_1, \boldsymbol{X}_2, \dots, \boldsymbol{X}_p \text{ are mutually independent;} \\ &H_1 : \text{ The negation of } \boldsymbol{H}_0. \end{aligned} \tag{1}$$

CONTACT Junshan Xie junshan@henu.edu.cn School of Mathematics and Statistics, Henan University, Kaifeng 475000, People's Republic of China

Supplemental data for this article can be accessed online at http://dx.doi.org/10.1080/10485252.2024.2435852.

When the population $\boldsymbol{X}$ follows the Gaussian distribution, under traditional asymptotical setup in which the dimension p is fixed and the sample size n tends to infinity, Anderson (2003) established the likelihood ratio test method and proved that the likelihood ratio statistic constructed by the determinant of the sample correlation coefficient matrix asymptotically obeys the chi-squared distribution under the null hypothesis. In high-dimensional cases where the dimension $p = p(n) \to \infty$ as the sample size $n \to \infty$ and $n/p \to \gamma \in (0, \infty)$, Jiang et al. (2013), Jiang and Yang (2013) and Jiang and Qi (2015) studied that likelihood ratio statistic asymptotically obeys normal distribution.

As a simple fact reveals that when the population $\boldsymbol{X}$ is Gaussian, testing H_0 in (1) is equivalent to testing whether the population covariance matrix is a diagonal matrix or the population Pearson correlation matrix is an identity matrix. Denote $\hat{r}_{ij}(1 \le i < j \le p)$ to be the sample Pearson correlation coefficient of the components $\boldsymbol{X}_i$ and $\boldsymbol{X}_j$. Thus a series of test methods based on the entries $\hat{r}_{ij}$ of the sample Pearson correlation matrix have been established, see Schott (2005), Mao (2014), Mao (2020) and He et al. (2021). They roughly fall into two categories according to the functional form of $\hat{r}_{ij}$ in the test statistics.

The first type is using the sum-of-squares-type statistics. For instance, Schott (2005) exploited the sum-of-squares-type statistic $\sum_{i=2}^{p}\sum_{j=1}^{i-1}\hat{r}_{ij}^2$ to test for the independence of the components of high-dimensional random vector with $n/p \to \gamma \in (0, \infty)$, and proved that the test statistic is asymptotically normal under the null hypothesis of (1). Birke and Dette (2005) extended the result to $n/p \to \infty$. Mao (2014) used nonlinear transformation to improve the test statistic of Schott (2005). There are also some other sum-of-squares-type statistics methods including Ledoit and Wolf (2002), Srivastava (2005), Cai and Jiang (2011), He et al. (2021) and Shi et al. (2024). The advantage of the sum-of-squares-type statistic is that the test statistic converges quickly and it is powerful against dense signals in alternatives.

The second type is using the extreme-value-type statistics. When $n/p \to \gamma \ge 1$, Johnstone (2001) established a test statistic based on the largest eigenvalue of the sample covariance matrix, and proved that its limit distribution is Tracy-Widom distribution. Bao et al. (2012) got the similar results on the extreme eigenvalue of the sample Pearson correlation matrix. Jiang (2004) constructed a test statistic $\max_{1\le j<i\le p}\hat{r}_{ij}^2$ and proved that it converges to a Gumbel distribution under the assumption of $p/n \to \gamma \in (0, \infty)$. Several subsequent works weakened the assumptions of moment restrictions on $\boldsymbol{X}_i$ or growth rate restrictions on p and n, one can refer to Li et al. (2010), Li et al. (2012), Liu et al. (2008), Zhou (2007), Cai and Jiang (2012) among others. The extreme-value-type statistic converges slowly but it is powerful against sparse signals in alternatives, which is verified by many researchers, for example, Pillai and Yin (2012), Han et al. (2017), Leung and Drton (2017) and Drton et al. (2020).

When the population no longer follows the Gaussian distribution, the above equivalence test is not valid. Hájek et al. (1999) demonstrated that the advantage of using rank-based tests is that they are not limited by population distribution. Therefore, some scholars consider using rank correlation coefficients to study the mutual independence between variables. The commonly used rank correlation coefficients are Spearman rank correlation coefficient (see Spearman 1904) and Kendall rank correlation coefficient (see Kendall 1938). Bao et al. (2013) established a linear spectral statistic on Spearman rank correlation coefficient. Mao (2016) considered a sum-of-squares-type test on the basis of Spearman rank correlation coefficients. Leung and Drton (2017) studied tests based on sums of rank correlations including Spearman rank correlation coefficient and Kendall rank correlation coefficient. They utilised the asymptotic properties of the U-statistic to establish asymptotic normality of the test statistics. Shi et al. (2023) introduced a max-sum type test based on Spearman rank correlation coefficient. Mao (2018) improved the test statistic on Kendall rank correlation coefficient, and proved that the null limiting distribution of the test statistic is normal when $n, p \to \infty$ or $p \to \infty$, n is fixed.

Recall that $\{X_{\cdot,i} : X_{\cdot,i} = (X_{1i}, X_{2i}, \dots, X_{pi})^T, i = 1, 2, \dots, n\}$ be random sample of the continuous population $\boldsymbol{X} \in \mathbb{R}^p$. For $1 \le k \le p$, denote R_{ki} to be the rank of X_{ki} among $(X_{k1}, X_{k2}, \dots, X_{kn})$. If extracting two sets of random samples from $\boldsymbol{X}_k$ and $\boldsymbol{X}_l$ respectively: $(X_{k1}, X_{k2}, \dots, X_{kn})^T$ and $(X_{l1}, X_{l2}, \dots, X_{ln})^T$, then we can get the corresponding rank vectors $R_k = (R_{k1}, R_{k2}, \dots, R_{kn})^T$ and $R_l = (R_{l1}, R_{l2}, \dots, R_{ln})^T$ respectively. The Kendall rank correlation coefficient of $\boldsymbol{X}_k$ and $\boldsymbol{X}_l$ can be defined as follows:

$$\tau_{kl} = \frac{2}{n(n-1)} S_{kl}, \tag{2}$$

where $S_{kl} = \sum_{i=2}^{n} \sum_{j=1}^{i-1} sgn(R_{ki} - R_{kj}) sgn(R_{li} - R_{lj})$ and $sgn(x)$ is the signum function. We can further define the sample Kendall rank correlation matrix of $\boldsymbol{X}$ by $\hat{R} = (\tau_{kl})_{p \times p}$.

Recently, Shi et al. (2024) proved the sum-of-squares-type statistic and the extreme-value-type statistic based on Kendall rank correlation coefficient are asymptotically independent for high-dimensional population, and they proposed an adaptive high-dimensional independence test method. Wang et al. (2024) established the same adaptive test procedure constructed by the more general three classes of rank-based statistics including Kendall rank statistics as a special case, they also proved the asymptotic independence between the rank-based max-type and sum-type statistics for each of the three classes under the null hypothesis, and they demonstrated the application of the adaptive procedure to the cross-sectional independence test under the panel data regression models.

In this paper, we mainly construct a more general sum-of-powers-type test statistics based on Kendall rank correlation coefficient and prove its asymptotic normality. We indicate the sum-of-powers-type test statistics with different orders are asymptotically independent, and they are all asymptotically independent with the extreme-value-type test statistic in Han et al. (2017). Based on the asymptotical properties, we propose an adaptive test procedure for independence testing of high-dimensional continuous data, and demonstrate that the proposed adaptive testing method can significantly enhance empirical power without distortion of empirical size when dealing with both sparse and non-sparse data testing.

2. Theoretical results

In this paper, we will develop an adaptive test method to revisit the test (1) by constructing the sum-of-powers-type test statistic. To the end, we will first establish the statistical asymptotic properties of a family of sum-of-powers-type test statistic with index $1 \le \gamma \le \infty$ as

$$t(\gamma) = \sum_{k=2}^{p} \sum_{l=1}^{k-1} \tau_{kl}^{\gamma}. \tag{3}$$

In fact, when $\gamma = 2$, it is just the sum-of-squares-type statistic studied in Mao (2018). When $\gamma \to \infty$, we have $[t(\gamma)]^{1/\gamma} \to \max_{1\le i<j\le p} |\tau_{ij}|$, which is equivalent to the extreme-value-type statistic

$$t(\infty) = \max_{1\le j<i\le p} \tau_{ij}^2 \tag{4}$$

studied in Han et al. (2017). Thus the statistics $t(\gamma), 1 \le \gamma \le \infty$ includes both the sum-of-squares-type statistic and the extreme-type statistic as special cases.

In order to give the characteristics of the statistic $t(\gamma)$ with $\gamma \in \mathbb{N}$, we first calculate the mean, variance and covariance of the test statistic $t(\gamma)$. To the end, the cumulants and moments of S_{kl} or τ_{kl} need to be examined. Let κ_i and m_i be the ith cumulant and moment of S_{kl} in (2), respectively. Silverstone (1950) obtained that if X_k is independent of X_l, the odd cumulants are zero and the even cumulants of S_{kl} can be obtained through the following formula

$$(-1)^{i+1}\kappa_{2i} = \frac{2^{2i-1}}{i}B'_{2i}\left\{\frac{B_{2i+1}(n)}{2i+1} + n^{2i} - n\right\}, \quad i \in \mathbb{N}, \tag{5}$$

where $B_{2i+1}(n)$ is the $(2i+1)$th Bernoulli polynomial in n, $B'_{2i} = |B_{2i}|$, B_{2i} is the numerical value of the $(2i)$th Bernoulli number. Thus we can obtain that

$$\begin{aligned}
\kappa_2 &= \frac{1}{18}n(n-1)(2n+5),\\
\kappa_4 &= -\frac{1}{225}n(6n^4+15n^3+10n^2-31),\\
\kappa_6 &= \frac{1}{1323}n(6n^6+21n^5+21n^4-7n^2-41),\\
\kappa_8 &= -\frac{8}{675}n(10n^8+45n^7+60n^6-42n^4+20n^2-93).
\end{aligned}$$

Using the formula in Lucaks (1955), Mao (2018) presented the specific expressions of m_2, m_4, m_6, m_8 as follows.

$$\begin{aligned}
m_2 &= \kappa_2,\\
m_4 &= \kappa_4 + 3\kappa_2^2,\\
m_6 &= \kappa_6 + 15\kappa_4\kappa_2 + 15\kappa_2^3,\\
m_8 &= \kappa_8 + 35\kappa_4^2 + 28\kappa_6\kappa_2 + 210\kappa_4\kappa_2^2 + 105\kappa_2^4.
\end{aligned} \tag{6}$$

For $\gamma \in \Gamma = \{1,2,3,4\}$, denote

$$\mu(\gamma) = E[t(\gamma)] = \sum_{k=2}^{p}\sum_{l=1}^{k-1}\mu_{kl}^{\gamma}, \quad \sigma^2(\gamma) = E[t(\gamma)-\mu(\gamma)]^2,$$

where $\mu_{kl}^{\gamma} = E(\tau_{kl}^{\gamma})$.

By the results in (6), we can prove that

Proposition 2.1: *Under H_0, for $\gamma \in \Gamma = \{1,2,3,4\}$, we have*

$$\mu(\gamma) = \begin{cases} \dfrac{2^{\gamma-1}p(p-1)}{n^{\gamma}(n-1)^{\gamma}} m_{\gamma}, & \text{if } \gamma \text{ is even,} \\ 0, & \text{if } \gamma \text{ is odd.} \end{cases}$$

where m_γ ($\gamma = 2, 4$) are listed in (6).

Proposition 2.2: *Under H_0, for $\gamma \in \Gamma = \{1,2,3,4\}$, we have*

$$\sigma^2(\gamma) = \begin{cases} \dfrac{2^{2\gamma-1}p(p-1)}{n^{2\gamma}(n-1)^{2\gamma}}(m_{2\gamma} - m_{\gamma}^2), & \text{if } \gamma \text{ is even,} \\ \frac{2^{2\gamma-1}p(p-1)}{n^{2\gamma}(n-1)^{2\gamma}} m_{2\gamma}, & \text{if } \gamma \text{ is odd.} \end{cases}$$

where m_γ ($\gamma = 2, 4, 6, 8$) are listed in (6).

Proposition 2.3: *Under H_0, for $\beta, \gamma \in \Gamma = \{1,2,3,4\}$, we have*

$$Cov[t(\beta), t(\gamma)] = \begin{cases} \dfrac{2^{\beta+\gamma-1}p(p-1)}{n^{\beta+\gamma}(n-1)^{\beta+\gamma}}(m_{\beta+\gamma} - m_{\beta}m_{\gamma}), & \text{if } \beta \text{ and } \gamma \text{ are both even,} \\ 0, & \text{if } \beta+\gamma \text{ is odd,} \\ \frac{2^{\beta+\gamma-1}p(p-1)}{n^{\beta+\gamma}(n-1)^{\beta+\gamma}} m_{\beta+\gamma}, & \text{if } \beta \text{ and } \gamma \text{ are both odd.} \end{cases}$$

where m_γ ($\gamma = 2, 4, 6, 8$) are listed in (6).

Remark 2.1: From Propositions 2.2 and 2.3, we can easily derive the Pearson correlation coefficient of $t(\beta)$ and $t(\gamma)$ that

$$Corr[t(\beta), t(\gamma)] = \frac{Cov[t(\beta), t(\gamma)]}{\sigma(\beta)\sigma(\gamma)}.$$

We can next establish the asymptotic distribution of $\{t(\gamma), \gamma \in \Gamma\}$ as follows:

Theorem 2.1: *For $\gamma \in \Gamma = \{1,2,3,4\}$, under H_0, when $p \to \infty$, $n \to \infty$ or n is fixed,*

$$\{[t(\gamma) - \mu(\gamma)]/\sigma(\gamma)\}_{\gamma\in\Gamma}^{T} \xrightarrow{d} N(\mathbf{0}, \mathbf{\Omega}),$$

where $\mathbf{\Omega} = (\omega_{\gamma_1,\gamma_2})$ and

$$\omega_{\gamma_1,\gamma_2} = \begin{cases} 1, & \text{if } \gamma_1 = \gamma_2 \in \Gamma, \\ Corr[t(\gamma_1), t(\gamma_2)], & \text{if } \gamma_1 \neq \gamma_2 \in \Gamma. \end{cases}$$

When $\gamma \to \infty$, we will also consider the asymptotic distribution of $t(\gamma)$ as $\gamma \to \infty$. In fact, for the statistic $t(\infty)$ defined in (4), Han et al. (2017) proved that $t(\infty)$, under

H_0, asymptotically obeys the Gumbel distribution, that is, when $n, p \to \infty$ and $\log p = o(n^{1/3})$, for $x \in R$,

$$P\left\{\frac{9n(n-1)}{2(2n+5)}t(\infty) - 4\log p + \log\log p \leq x\right\} \to exp\left\{-(8\pi)^{-\frac{1}{2}} exp\left(-\frac{x}{2}\right)\right\}.$$

We can further present the statistical asymptotical properties of $t(\gamma), \gamma \in \Gamma = \{1,2,3,4\}$ and $t(\infty)$, and they will play an important role in establishing the proposed adaptive test procedure.

Theorem 2.2: *For $\gamma \in \Gamma = \{1,2,3,4\}$, under H_0,*

(1) *if $\beta + \gamma$ is odd, $t(\beta)$ is asymptotically independent of $t(\gamma)$ when $p \to \infty$, $n \to \infty$ or n is fixed.*
(2) *$\{[t(\gamma) - \mu(\gamma)]/\sigma(\gamma)\}^T_{\gamma \in \Gamma}$ is asymptotically independent of $[\frac{9n(n-1)}{2(2n+5)}t(\infty) - a_p]$ as $n, p \to \infty$ and $\log p = o(n^{1/3})$, where $a_p = 4\log p - \log\log p$.*

3. An adaptive test procedure

In this section, we will establish an adaptive test procedure for the test (1). When considering the index $1 \leq \gamma < \infty$, we know that the statistic $t(\gamma)$ defined in (3) is the sum-of-powers-type statistic, and it is powerful against dense signals in alternatives. When taking $\gamma = \infty$, the statistic $t(\infty)$ is the extreme-value-type statistic, it is powerful against sparse signals in alternatives. However, since the truth is often unknown in practice, it is unclear which types of statistics should be chosen. In this paper, we adopt an adaptive testing procedure by combining the information from statistics $t(\gamma), 1 \leq \gamma < \infty$ of different orders, which would yield high power against various alternatives.

Inspired by the idea from Mosteller and Fisher (1948), Yu et al. (2009), Pan et al. (2014), Xu et al. (2016) and Shi et al. (2023), we propose to combine the statistic $t(\gamma)$ through their p-values. In particular, we employ the minimum p-value to approximate the maximum power. The adaptive test is as follows:

$$T_{asy} = \min_{\gamma \in \Gamma \cup \{\infty\}} P_{t(\gamma)},$$

where $P_{t(\gamma)}$ represents the p-value of $t(\gamma)$.

We can classify the test statistics $t(\gamma), \gamma \in \Gamma = \{1,2,3,4\}$ by judging γ is even or odd. Define $\breve{T}_o$ and $\breve{T}_e$ as the collections of random variables,

$$\breve{T}_o = \{[t(\gamma) - \mu(\gamma)]/\sigma(\gamma) : \gamma \in \Gamma \text{ is odd}\}, \breve{T}_e = \{[t(\gamma) - \mu(\gamma)]/\sigma(\gamma) : \gamma \in \Gamma \text{ is even}\}.$$

Denote the following test statistics T_O and T_E,

$$T_O = \max_{\gamma \in \Gamma, \gamma \text{ is odd}} |[t(\gamma) - \mu(\gamma)]/\sigma(\gamma)|, T_E = \max_{\gamma \in \Gamma, \gamma \text{ is even}} |[t(\gamma) - \mu(\gamma)]/\sigma(\gamma)|.$$

We can get the corresponding p-value statistics of T_O, T_E and $t(\infty)$ which are denoted as P_O, P_E, and P_∞ respectively. Take the minimum p-value of the above test T_O, T_E and $t(\infty)$ as

$$P_{\min} = \min\{P_O, P_E, P_\infty\}.$$

By Theorem 2.2, we know that the statistics in $\breve{T}_o$, $\breve{T}_e$ and $t(\infty)$ are asymptotically independent with $\log p = o(n^{1/3})$ under H_0, which means that P_O, P_E, and P_∞ are asymptotically independent and all obey the uniform distribution $U[0,1]$. After some elementary calculation, we can see that the distribution function of $P_{\min} = \min\{P_O, P_E, P_\infty\}$ is

$$F_0(x) \equiv P(P_{\min} \le x) = 1 - (1-x)^3.$$

Then

$$F_0(P_{\min}) = 1 - (1 - P_{\min})^3,$$

which means that the asymptotic p-value of the adaptive test is

$$P_{asy} = 1 - (1 - P_{\min})^3. \tag{7}$$

Setting the nominal significance level α, we will reject H_0 if $P_{asy} < \alpha$, otherwise, we will accept H_0.

For the adaptive test method in (7), we can see that the type I error is asymptotically controlled by

$$P(P_{asy} < \alpha) = P(P_{\min} < \alpha^*) = 1 - P(P_{\min} \ge \alpha^*),$$

where $\alpha^* = 1 - (1-\alpha)^{1/3}$. Recall that P_O, P_E, and P_∞ are asymptotically independent and all obey the uniform distribution $U[0,1]$ under H_0. Then we have

$$P(P_{asy} < \alpha) \to 1 - (1 - \alpha^*)^3 = \alpha$$

as $p \to \infty$. Therefore, we can obtain the the type I error of the adaptive test method can be well controlled for high-dimensional data test.

For the power analysis of the adaptive test, since

$$P\left(P_{\min} < \alpha^*\right) \ge \min\left\{P\left(P_O < \alpha^*\right), P\left(P_E < \alpha^*\right), P\left(P_\infty < \alpha^*\right)\right\},$$

we can see that the power of adaptive test goes to 1 if there exists $\gamma \in \Gamma \cup \{\infty\}$, such that the power of t_γ goes to 1. Thus the adaptive test method can be powerful against a wide spectrum of alternatives. It should be noted that the power of the adaptive test is not necessarily higher than that of all the $t(\gamma)$ statistics. This is because the power of $t(\gamma)$ is $P(P_{t_\gamma} < \alpha)$, and it is different from $P(P_{t_\gamma} < \alpha^*)$ since $\alpha^* = 1 - (1-\alpha)^{1/3} < \alpha$.

4. Simulations

In this section, we evaluate the finite sample performance of the proposed adaptive test method under various types of alternatives, and also investigate the relationship between the power and sparsity levels. At the same time, we also compare the performance of the proposed adaptive test method with several existing methods.

4.1. Simulation 1: adaptiveness in different situations

We can exhibit the empirical size and power values of the proposed adaptive test under the following seven continuous populations.

(1) Gaussian distribution: The population X obeys p-variate Gaussian distribution, $X \sim N_p(0, I_p)$ for the null and $X \sim N_p(0, \Sigma)$ for the alternative.
(2) Light-tailed Gaussian copula family: The population X belongs to light-tailed Gaussian copula family, $X = Z^{1/3}$, where $Z \sim N_p(0, I_p)$ for the null and $Z \sim N_p(0, \Sigma)$ for the alternative.
(3) Heavy-tailed Gaussian copula family: The population X belongs to heavy-tailed Gaussian copula family, $X = Z^3$, where $Z \sim N_p(0, I_p)$ for the null and $Z \sim N_p(0, \Sigma)$ for the alternative.
(4) Multivariate t distribution: The population X obeys the p-variate t distribution with freedom 3, $X = (X_1, \ldots, X_p)^T$ has independent components for the null and X has covariance matrix Σ for the alternative.
(5) Multivariate exponential distribution: The population X obeys the p-variate exponential distribution. We assume that $X = (X_1, \ldots, X_p)^T$ has independent Exp(0.25) distributed components under H_0. When H_1 holds, we assume that $X = (X_1, \ldots, X_p)^T$ have the dependent structure as follows: for each $j = 1, \ldots, p$, X_j conditioned on X_{-j} follows an exponential distribution of rate $0.25 + \Sigma_{j,-j}X_{-j}$ where $\Sigma_{j,-j}$ denotes the jth of Σ without the diagonal element which is a $p-1$ dimensional row vector, X_{-j} denotes the vector X without the jth entry which is a $p-1$ dimensional column vector.
(6) Multivariate Lognormal distribution: The population $X = (X_1, \ldots, X_p)^T$ has i.i.d. Lognorm(0,1) random variables under H_0. And under a type of sparse alternative H_1: $(X_1, \ldots, X_p)^T = 0.6(Z_1, \ldots, Z_p)^T + 0.8(Z_2, \ldots, Z_{p+1})^T$, where $Z_k(k = 1, \ldots, p)$ are i.i.d. Lognorm(0,1) random variables.
(7) Multivariate Cauchy distribution: The population $X = (X_1, \ldots, X_p)^T$ has i.i.d. Cauchy(0,1) distributed random variables under H_0. And under a type of non-sparse alternative H_1: $(X_1, \ldots, X_p)^T = \Sigma_p^{1/2}(Z_1, \ldots, Z_p)^T$, $k = 1, \ldots, p$, where $Z_k(k = 1, \ldots, p)$ are i.i.d. Cauchy(0,1) random variables, $\Sigma_p = 0.95I_p + 0.051_p1_p^T$.

We choose the sample size $n = \{20, 40, 80, 160\}$ and the dimension $p = \{16, 32, 64, 128, 256\}$ respectively. Under the nominal significance $\alpha = 0.05$, the empirical sizes of the seven test methods related to statistics $t(1), t(2), t(3), t(4), t(\infty)$ and the proposed adaptive test T_{asy} are listed in Table 3-1 to Table 3-7 in the Supplementary Material.

We can see that the empirical sizes of each tests have the similar performance under the seven distributions. Hence, we take the Gaussian distribution case as an example, when both n and p are small (for example $n = 20, p = 16$), the empirical sizes of the tests $t(1)$, $t(2)$, $t(3)$ and $t(4)$ are better than that of $t(\infty)$. As n and p increase, the empirical sizes of the tests $t(1), t(2), t(3), t(4), t(\infty)$ are all quite close to the nominal significance level 0.05. But the growth rate of the empirical sizes of $t(\infty)$ is much slower than other test methods, and it is consistent with the fact that the Gumbel extreme value distribution converges slowly. For the proposed adaptive test T_{asy}, when p is small (for example $p = 16, 32$), the empirical sizes are obviously larger than 0.05. However, as the dimension p increases, the empirical sizes decrease and steadily approach the significance level 0.05, which means the type I errors of the proposed adaptive test T_{asy} and $t(\gamma), \gamma \in \Gamma = \{1, 2, 3, 4\}$ test are all well controlled under H_0 for high-dimensional data.

Next, we consider the empirical powers of the tests $t(1), t(2), t(3), t(4)$, $t(\infty)$ and the adaptive test under sparse and non-sparse alternatives. And we also investigate the relationship between the empirical powers and sparsity levels.

To the end, we will first introduce some notations. For two matrices $A = (a_{ij})_{1\le i,j\le p}$ and $B = diag(b_1, b_2, \ldots, b_p)$, we will define $A \circ B = (a_{ij}b_{jj})_{1\le i,j\le p}$. Denote $\triangle \in \mathcal{R}^{p\times p}$ be a symmetric matrix with eight nonzero entries. We choose four nonzero entries randomly from the upper triangle of $\triangle \in \mathcal{R}^{p\times p}$, each with a magnitude randomly drawn from the uniform distribution in $[0, 1]$, the other four nonzero entries in the lower triangle are determined by symmetry. Let $\lambda_{\min}(A)$ be the smallest eigenvalue of the matrix A. Write $\phi = \{-\lambda_{\min}(I_p + \triangle) + 0.05\}I\{\lambda_{\min}(I_p + \triangle) \le 0\}$.

In order to study the effectiveness of the proposed adaptive test method under sparse and non-sparse alternatives, we consider six types of covariance matrix in alternatives, the forms of the covariance matrix under the alternatives for the first four distributions are as follows:

(a) $\Sigma_1 = I_p + \triangle + \phi I_p$,
(b) $\Sigma_2 = \{(1-\rho)I_p + \rho 1_{p,k}1_{p,k}^T\} \circ diag\{1^\delta, \ldots, p^\delta\}$,
(c) $\Sigma_3 = \{(1-\rho)I_p + \rho 1_p 1_p^T\} \circ diag\{1^\delta, \ldots, p^\delta\}$,
(d) $\Sigma_4 = \{(\rho^{|i-j|})_{1\le i,j\le p}\} \circ diag\{1^\delta, \ldots, p^\delta\}$.

Here Σ_1 is a sparse covariance matrix whose structure is to ensure positivity. Σ_2 is also a sparse covariance matrix, where $diag\{1^\delta, \ldots, p^\delta\}$ is a $p \times p$ diagonal matrix whose diagonal elements are i^δ, I_n is a $n \times n$ identity matrix, 1_n is a $n \times 1$ vector with all elements 1, $1_{n,k}$ represents a $n \times 1$ vector with first k elements being 1 and the rest 0. Here k represents the sparse level of the matrix, ρ represents the signal of part non-diagonal elements and $\delta \ge 0$ is the scale parameter. We can adjust the sparse alternative through selecting different values of the above parameter. Σ_3 and Σ_4 are both non-sparse covariance matrix. The parameter ρ represents the signal of all non-diagonal elements and the parameter $\delta \ge 0$ is the scale parameter. When $\rho = 0$, the non-diagonal elements of this two matrix are 0; When $\delta = 0$, the non-diagonal elements of Σ_3 are all ρ, the non-diagonal elements of Σ_4 are all the power of ρ. We can adjust the alternative through selecting different values of k. We can also adjust the non-sparse alternative through selecting different values of the above parameters.

We will list the empirical powers of the tests under the alternative with sparse covariance matrix $\Sigma_1 = I_p + \triangle + \phi I_p$. The empirical powers of $t(1), t(2), t(3), t(4)$, $t(\infty)$ and the adaptive test T_{asy} under first five different distributions are listed in Table 3-8 to Table 3-12 in the Supplementary Material. We also investigate the empirical powers of the last two populations under the sparse alternative and the non-sparse alternative in Table 3-13 and 3-14 in the Supplementary Material.

We can see that under the first five distributions, when the alternative data is sparse, the extreme-value-type test method $t(\infty)$ is more powerful than that of other $t(\gamma), 1 \le \gamma <$

∞ test methods, and the reason is that the extreme-value-type test method $t(\infty)$ is powerful against sparse signals. For the proposed adaptive test, the empirical power is higher than $t(1), t(2), t(3), t(4)$ and $t(\infty)$. Under the non-sparse alternative, the empirical power of the adaptive test is the highest among all tests. All the above results demonstrate the performance of the proposed adaptive test is the best among all six tests above.

To investigate the performance of empirical powers under different sparse levels, we will next select sparse covariance matrix $\Sigma_2 = \{(1-\rho)I_p + \rho 1_{p,k} 1_{p,k}^T\} \circ diag\{1^\delta, \ldots, p^\delta\}$ under the light-tailed Gaussian copula family. We will select the parameters of Σ_2 in four situations.

(a) $\delta = 1, \rho = 0.8, k = 2$,
(b) $\delta = 1, \rho = 0.4, k = 5$,
(c) $\delta = 1, \rho = 0.4, k = 7$,
(d) $\delta = 1, \rho = 0.4, k = 13$.

When $\rho = 0.8, k = 2$, the sparsity of Σ_2 is high relatively. When $\rho = 0.4$, we set $k = 5$, $k = 7, k = 13$ respectively, then the sparsity of Σ_2 is decreasing successively. Then we can analyse the influence of different parameters ρ and k on the performance of the empirical powers by Table 3-15 to Table 3-18 in the Supplementary Material.

Table 3-15 in the Supplementary Material shows that among all the tests $t(1), t(2), t(3)$, $t(4)$ and $t(\infty)$, the test $t(\infty)$ performs best under the high sparsity level case (a) for the alternative. And the proposed adaptive test T_{asy} performs better than most single $t(\gamma)$ statistic, its empirical powers are usually close or even higher than the best single $t(\gamma)$ statistic.

By Table 3-16 to Table 3-18 in the Supplementary Material, we can see that, when the sparsity level increases from $k = 5$ to $k = 7$ and then to $k = 13$, which means that when the alternative gets denser, the test $t(\infty)$'s advantage over the single test $t(\gamma)$ is getting weaker and weaker, even the single test $t(\gamma)$ performs better than the test $t(\infty)$. However, the proposed adaptive test T_{asy} still has a significant advantage among all the test methods.

We will also consider the performance of the empirical powers of these tests under non-sparse alternatives, we will choose the alternative covariance matrixes $\Sigma = \Sigma_3$ with $\delta = 1, \rho = 0.05$ and $\Sigma = \Sigma_4$ with $\delta = 1, \rho = 0.2$ respectively. The Gaussian distribution population, the light-tailed Gaussian copula family and the heavy-tailed Gaussian copula family will be investigated. The empirical powers of the tests will be listed in Table 3-19 to Table 3-24 in the Supplementary Material.

We can see that when the alternatives are non-sparse, the tests $t(1)$, $t(2)$, $t(3)$ and $t(4)$ perform well when n and p are large and they are much better than $t(\infty)$, and it is consistent with the fact that the statistic $t(\infty)$ converges slowly. For the proposed adaptive test T_{asy}, under the two non-sparse alternatives, the empirical powers of the adaptive test are close to or even better than the optimal empirical powers of $t(1), t(2), t(3), t(4)$ and $t(\infty)$. Then we can say that the proposed adaptive test is also powerful when the alternatives are dense.

Then we can conclude that under the premise of keeping type I error, the proposed adaptive test T_{asy} can significantly enhance empirical power of the test. It is powerful against a wide range of alternatives, and thus advantageous in practice when the true alternative is unknown.

4.2. Simulation 2: comparison with existing methods

We also compare the performance of the proposed adaptive test method T_{asy} with the existing methods proposed by Schott (2005), Cai and Jiang (2011), Mao (2016) and Han et al. (2017), denoted as S_r, M_r, S_ρ and M_ρ respectively. We also compare T_{asy} with the method in Shi et al. (2024), Shi et al. (2023), which are denoted as S_ρ^* and C_ρ^* respectively. The empirical sizes and the empirical powers are listed in Table 3-25 to Table 3-30 in the Supplementary Material.

Table 3-25 and Table 3-26 in the Supplementary Material presents the empirical sizes of each tests under Gaussian distribution and heavy-tailed Gaussian copula family. It is observed that as the size n and the dimension p increase, the empirical sizes of S_r, S_ρ and S_ρ^* tend to the nominal significance level 0.05 quickly, while the empirical sizes of M_r and M_ρ tend to the nominal significance level 0.05 slowly, the empirical sizes of C_ρ^* tend to the nominal significance level 0.05 with moderate growth rate. For the proposed adaptive test T_{asy}, when p is small, the empirical sizes are obviously larger than 0.05. However, as the dimension p increases, the empirical sizes decrease and steadily approach the nominal significance level 0.05, which means the type I errors of the proposed adaptive test T_{asy} and other tests S_r, S_ρ, S_ρ^*, M_r, M_ρ and C_ρ^* are all well controlled under H_0 for high-dimensional data.

We consider the dense alternatives and spare alternatives under Gaussian distribution and heavy-tailed Gaussian copula family which are shown in Table 3-27 and Table 3-30 in the Supplementary Material.

For the dense alternatives, Table 3-27 and Table 3-28 in the Supplementary Material show that the empirical powers of the proposed adaptive test T_{asy}, S_r, S_ρ, S_ρ^* and C_ρ^* are approaching to one as the dimension p increases. The empirical powers of M_r and M_ρ are low but increasing slowly as the size n increases. These results illustrate that the sum-of-powers-type tests perform well in the dense alternatives. In terms of distributions, we can see that the empirical powers of M_r under heavy-tailed Gaussian copula family are higher than Gaussian copula family. The empirical powers of the proposed adaptive test T_{asy} perform best under the two distributions.

For the sparse alternatives, Table 3-29 and Table 3-30 in the Supplementary Material present that the empirical powers of all tests are approaching to one as the size n increases. The growth rates of M_r and M_ρ are higher than S_r, S_ρ and S_ρ^*. The growth rates of C_ρ^* are higher than M_r and M_ρ. These results illustrate that the extreme-value-type tests perform well in the sparse alternatives and the max-sum test proposed by Shi et al. (2023) is better than S_r, S_ρ, S_ρ^*, M_r and M_ρ. And we can see that the empirical powers of the proposed adaptive test T_{asy} perform best among all tests under the two distributions.

In summary, the proposed adaptive test method T_{asy} has good performances under all above settings. These simulations reveal that the proposed test method can significantly enhance the empirical powers of the test while keep the type I error well controlled, and it has consistently good performances across a wide range of scenarios.

5. Real data example

We will apply the proposed adaptive test to the analysis of the Leaf Dataset and the Parkinson's Disease Dataset.

Table 1. The p-values of the test statistics.

$t(1)$	$t(2)$	$t(3)$	$t(4)$	$t(\infty)$	T_{asy}
4.253375×10^{-5}	2.220446×10^{-16}	2.442491×10^{-15}	0	0.05791024	0

Table 2. The p-values of the test statistics.

$t(1)$	$t(2)$	$t(3)$	$t(4)$	$t(\infty)$	T_{asy}
0	0	0	0	0	0

5.1. Leaf dataset

When studying the shape and texture features extracted from digital images of leaf specimens of plant species, an automatic plant recognition system requires a set of discriminating variables and a structured database to train statistical models, thus it is necessary to test the complete independence of the shape and texture features variables. We will adopt the data collected in Silva et al. (2013) and it was available on http://archive.ics.uci.edu/ml/datasets/Leaf. We only examine the characteristics of the sixth plant species here, and the corresponding data matrix has a total of $n = 8$ rows and $p = 14$ columns.

We will calculate the p-values of the test statistics $t(1)$, $t(2)$, $t(3)$, $t(4)$ and $t(\infty)$ respectively. The p-values of $t(1)$, $t(2)$, $t(3)$, $t(4)$, $t(\infty)$ and the proposed adaptive test T_{asy} are shown by Table 1 above, which indicates that the six tests above share the same test conclusion, that is, we should reject the null hypothesis of mutual independence. Consequently, we obtain the features considered here are correlated evidently. This result is consistent with conclusion of Shi et al. (2023).

5.2. Parkinson's disease dataset

Degenerative voice performance is a common symptom in the most majority of Parkinson's disease subjects, who require regular one-on-one rehabilitation with a speech specialist over a long period. A computer programme named LSVT Companion can be used to determine 'acceptable' or 'unacceptable' rehabilitation by analyzing the biomedical speech signal dataset. Tsanas et al. (2014) collected biomedical speech signal processing algorithms from 14 LSVT patients diagnosed with Parkinson's disease. The initial study used 310 algorithms to characterise 126 speech samples, resulting in a design matrix with 126×310. The data can be downloaded from http://archive.ics.uci.edu/ml/datasets/LSVT+Voice+Rehabilitation. An important topic is to test whether all the algorithms are mutual independent.

We will use six different test methods including $t(1)$, $t(2)$, $t(3)$, $t(4)$, $t(\infty)$ and the proposed adaptive method T_{asy} to consider the algorithmic dependence test. The p-values of all the test statistics are listed in Table 2. We can see that all the p-values of the six tests are much less than the nominal significance level $\alpha = 0.05$ and they are all very close to 0, which shows that the six tests have the same test conclusion, that is, we should reject the null hypothesis of mutual independence. Therefore, we conclude that the Parkinson's disease data set has a definite algorithmic dependence. And it is also shows no difference with results of Shi et al. (2023).

6. Proofs

Recall that $t(\gamma) = \sum_{k=2}^{p}\sum_{l=1}^{k-1} \tau_{kl}^{\gamma}$ with $\tau_{kl} = \frac{2}{n(n-1)}S_{kl}$. Then $m_{2i} = E(S_{kl}^{2i})$ will be needed in the following proofs. Mao (2018) presented the specific expressions of m_2, m_4, m_6 and m_8. Before the proof of the main results, we will first calculate the expression of $m_{2i} = E(S_{kl}^{2i})$ with $i = 5, 6, 7, 8$.

Using Faà di Bruno's formula, Lucaks (1955) proved that the pth moment m_p can be expressed by the cumulants $\kappa_i's$ as follows:

$$m_p = \sum \frac{p!}{j_1!(l_1)^{j_1}\cdots j_s!(l_s)^{j_s}}\kappa_{l_1}^{j_1}\cdots\kappa_{l_s}^{j_s},$$

where the summation is extended over all partitions of p satisfying

$$\begin{cases} j_1 + j_2 + \cdots + j_s = k, \\ j_1 l_1 + j_2 l_2 + \cdots + j_s l_s = p. \end{cases}$$

Silverstone (1950) indicated that if X_k is independent of X_l, the $(2i)$th cumulant of S_{kl}, κ_{2i}, can be shown through the following formula

$$(-1)^{i+1}\kappa_{2i} = \frac{2^{2i-1}}{i}B_{2i}'\left\{\frac{B_{2i+1}(n)}{2i+1} + n^{2i} - n\right\}, \tag{8}$$

where $B_{2i+1}(n)$ is the $(2i+1)$th Bernoulli polynomial in n, $B_{2i}' = |B_{2i}|$, B_{2i} is the numerical value of the $(2i)$th Bernoulli number.

Then under H_0, we can obtain the specific expressions of $\kappa_{10}, \kappa_{12}, \kappa_{14}, \kappa_{16}$ by (8) that

$$\begin{aligned}
\kappa_{10} &= \frac{128}{1089}n(6n^{10} + 33n^9 + 55n^8 - 66n^6 + 66n^4 - 33n^2 - 61),\\
\kappa_{12} &= -\frac{353792}{11179350}n(210n^{12} + 1365n^{11} + 2730n^{10} - 5005n^8\\
&\quad + 8580n^6 - 9009n^4 + 4550n^2 - 3421),\\
\kappa_{14} &= \frac{2048}{135}n(6n^{14} + 45n^{13} + 105n^{12} - 273n^{10} + 715n^8\\
&\quad - 1287n^6 + 1365n^4 - 691n^2 + 15),\\
\kappa_{16} &= -\frac{3703808}{5917275}n(2730n^{16} + 23205n^{15} + 61880n^{14} - 216580n^{12} + 804440n^{10}\\
&\quad - 2212210n^8 + 4022200n^6 - 4275908n^4 + 2165800n^2 - 375557).
\end{aligned}$$

Thus we can further get $m_{10}, m_{12}, m_{14}, m_{16}$ as follows.

Lemma 6.1: *Under H_0, we have*

$$\begin{aligned}
m_{10} &= \kappa_{10} + 45\kappa_2\kappa_8 + 210\kappa_4\kappa_6 + 1575\kappa_2\kappa_4^2 + 630\kappa_6\kappa_2^2 + 3150\kappa_4\kappa_2^3 + 945\kappa_2^5,\\
m_{12} &= \kappa_{12} + 66\kappa_2\kappa_{10} + 495\kappa_4\kappa_8 + 462\kappa_6^2 + 12870\kappa_2\kappa_4\kappa_6\\
&\quad + 5775\kappa_4^3 + 1485\kappa_8\kappa_2^2 + 51975\kappa_2^2\kappa_4^2
\end{aligned}$$

$$
\begin{aligned}
&\quad + 13860\kappa_6\kappa_2^3 + 51975\kappa_4\kappa_2^4 + 10395\kappa_2^6,\\
m_{14} &= \kappa_{14} + 91\kappa_2\kappa_{12} + 1001\kappa_4\kappa_{10} + 3003\kappa_6\kappa_8 + 45045\kappa_2\kappa_4\kappa_8\\
&\quad + 42042\kappa_2\kappa_6^2 + 105105\kappa_6\kappa_4^2\\
&\quad + 3003\kappa_{10}\kappa_2^2 + 525525\kappa_2\kappa_4^3 + 407760\kappa_4\kappa_6\kappa_2^2\\
&\quad + 45045\kappa_8\kappa_2^3 + 1576575\kappa_4^2\kappa_2^3 + 315315\kappa_6\kappa_2^4\\
&\quad + 945945\kappa_4\kappa_2^5 + 135135\kappa_2^7,\\
m_{16} &= \kappa_{16} + 120\kappa_2\kappa_{14} + 1820\kappa_4\kappa_{12} + 8008\kappa_6\kappa_{10}\\
&\quad + 6435\kappa_8^2 + 106605\kappa_2\kappa_4\kappa_{10} + 360360\kappa_2\kappa_6\kappa_8\\
&\quad + 840840\kappa_4\kappa_6^2 + 75675600\kappa_6\kappa_2^5 + 450450\kappa_8\kappa_4^2 + 5460\kappa_{12}\kappa_2^2 + 12162150\kappa_2\kappa_6\kappa_4^2\\
&\quad + 2702700\kappa_4\kappa_8\kappa_2^2 + 2522520\kappa_2^2\kappa_6^2 + 2627625\kappa_4^4 + 120120\kappa_{10}\kappa_2^3\\
&\quad + 21882150\kappa_4\kappa_6\kappa_2^3\\
&\quad + 31531500\kappa_2^2\kappa_4^3 + 1351350\kappa_8\kappa_2^4 + 7567560\kappa_6\kappa_2^5\\
&\quad + 47297250\kappa_4^2\kappa_2^4 + 18918900\kappa_4\kappa_2^6\\
&\quad + 135135\kappa_2^8.
\end{aligned}
$$

Proof of Proposition 2.1: Under H_0, we have

$$
\mu(\gamma) = E\left(\sum_{k=2}^{p}\sum_{l=1}^{k-1}\tau_{kl}^{\gamma}\right) = \frac{p(p-1)}{2}\frac{2^{\gamma}}{n^{\gamma}(n-1)^{\gamma}}E(S_{kl}^{\gamma}) = \frac{2^{\gamma-1}p(p-1)}{n^{\gamma}(n-1)^{\gamma}}m_{\gamma},
$$

where $m_{\gamma} = E(S_{kl}^{\gamma})$. By the fact that under H_0, when γ is odd, $m_{\gamma} = 0$, we can easily complete the proof. ■

Proof of Proposition 2.2: By Brown and Eagleson (1984), we know that τ_{kl} for $2 \le k \le p$ and $1 \le l \le k-1$ are of pairwise independence under H_0. Thus we have

$$
\begin{aligned}
\sigma^2(\gamma) &= Var\left(\sum_{k=2}^{p}\sum_{l=1}^{k-1}\tau_{kl}^{\gamma}\right) = \sum_{k=2}^{p}\sum_{l=1}^{k-1}Var(\tau_{kl}^{\gamma})\\
&= \frac{p(p-1)}{2}\left\{E(\tau_{kl}^{\gamma})^2 - [E(\tau_{kl}^{\gamma})]^2\right\}\\
&= \frac{p(p-1)}{2}\left[\frac{2^{2\gamma}m_{2\gamma}}{n^{2\gamma}(n-1)^{2\gamma}} - \left(\frac{2^{\gamma}m_{\gamma}}{n^{\gamma}(n-1)^{\gamma}}\right)^2\right]\\
&= \frac{2^{2\gamma-1}p(p-1)}{n^{2\gamma}(n-1)^{2\gamma}}(m_{2\gamma} - m_{\gamma}^2).
\end{aligned}
$$

As $m_{\gamma} = 0$ when γ is odd, then we can easily get the conclusion. ■

Proof of Proposition 2.3: For $\beta, \gamma \in \Gamma = \{1, 2, 3, 4\}$, $Cov[t(\beta), t(\gamma)]$ can be calculated under the following three situations.

(1) When both β and γ are even, under H_0, for different positive integers h, i, j, k, we have $E(\tau_{ij}^{\beta}\tau_{ik}^{\gamma}) = E(\tau_{hi}^{\beta}\tau_{jk}^{\gamma})$. Then

$$
\begin{aligned}
&Cov[t(\beta), t(\gamma)] \\
&= Cov\left(\sum_{k=2}^{p}\sum_{l=1}^{k-1}\tau_{kl}^{\beta}, \sum_{k=2}^{p}\sum_{l=1}^{k-1}\tau_{kl}^{\gamma}\right) \\
&= E\left\{\left[\sum_{k=2}^{p}\sum_{l=1}^{k-1}\tau_{kl}^{\beta} - \sum_{k=2}^{p}\sum_{l=1}^{k-1}E(\tau_{kl}^{\beta})\right]\cdot\left[\sum_{k=2}^{p}\sum_{l=1}^{k-1}\tau_{kl}^{\gamma} - \sum_{k=2}^{p}\sum_{l=1}^{k-1}E(\tau_{kl}^{\gamma})\right]\right\} \\
&= E\left\{\left[\sum_{k=2}^{p}\sum_{l=1}^{k-1}\tau_{kl}^{\beta}\right]\cdot\left[\sum_{k=2}^{p}\sum_{l=1}^{k-1}\tau_{kl}^{\gamma}\right]\right\} - \frac{p^2(p-1)^2}{4}E(\tau_{kl}^{\beta})E(\tau_{kl}^{\gamma}) \\
&= \sum_{k=2}^{p}\sum_{l=1}^{k-1}E(\tau_{kl}^{\beta+\gamma}) + \frac{p(p-1)}{2}\left[\frac{p(p-1)}{2} - 1\right]E(\tau_{kl}^{\beta}\tau_{kj}^{\gamma}) \\
&\quad - \frac{p^2(p-1)^2}{4}E(\tau_{kl}^{\beta})E(\tau_{kl}^{\gamma}) \\
&= \frac{p(p-1)}{2}\left[\frac{2^{\beta+\gamma}m_{\beta+\gamma}}{n^{\beta+\gamma}(n-1)^{\beta+\gamma}} - \frac{2^{\beta+\gamma}m_{\beta}m_{\gamma}}{n^{\beta+\gamma}(n-1)^{\beta+\gamma}}\right] \\
&= \frac{2^{\beta+\gamma-1}p(p-1)}{n^{\beta+\gamma}(n-1)^{\beta+\gamma}}(m_{\beta+\gamma} - m_{\beta}m_{\gamma}).
\end{aligned}
$$

(2) When both β and γ are odd,

$$
\begin{aligned}
&Cov[t(\beta), t(\gamma)] \\
&= E\left\{\left[\sum_{k=2}^{p}\sum_{l=1}^{k-1}\tau_{kl}^{\beta}\right]\cdot\left[\sum_{k=2}^{p}\sum_{l=1}^{k-1}\tau_{kl}^{\gamma}\right]\right\} \\
&= \sum_{k=2}^{p}\sum_{l=1}^{k-1}E(\tau_{kl}^{\beta+\gamma}) + \frac{p(p-1)}{2}\left[\frac{p(p-1)}{2} - 1\right]E(\tau_{kl}^{\beta}\tau_{kj}^{\gamma}) \\
&= \frac{2^{\beta+\gamma-1}p(p-1)}{n^{\beta+\gamma}(n-1)^{\beta+\gamma}}m_{\beta+\gamma}.
\end{aligned}
$$

(3) When $\beta+\gamma$ is odd. Without loss of generality, we assume β is even and γ is odd, so we have $E(\tau_{ij}^{\gamma}) = 0$ under H_0, furthermore, $E(\tau_{ij}^{\beta}\tau_{ik}^{\gamma}) = E(\tau_{hi}^{\beta}\tau_{jk}^{\gamma}) = 0$. Then

$$
\begin{aligned}
&Cov[t(\beta), t(\gamma)] \\
&= E\left\{\left[\sum_{k=2}^{p}\sum_{l=1}^{k-1}\tau_{kl}^{\beta} - \sum_{k=2}^{p}\sum_{l=1}^{k-1}E(\tau_{kl}^{\beta})\right]\cdot\left[\sum_{k=2}^{p}\sum_{l=1}^{k-1}\tau_{kl}^{\gamma}\right]\right\} \\
&= E\left\{\left[\sum_{k=2}^{p}\sum_{l=1}^{k-1}\tau_{kl}^{\beta}\right]\cdot\left[\sum_{k=2}^{p}\sum_{l=1}^{k-1}\tau_{kl}^{\gamma}\right]\right\} - \sum_{k=2}^{p}\sum_{l=1}^{k-1}E(\tau_{kl}^{\beta})E\left[\sum_{k=2}^{p}\sum_{l=1}^{k-1}\tau_{kl}^{\gamma}\right]
\end{aligned}
$$

$$
\begin{aligned}
&= \sum_{k=2}^{p}\sum_{l=1}^{k-1} E(\tau_{kl}^{\beta+\gamma}) + \frac{p(p-1)}{2}\left[\frac{p(p-1)}{2} - 1\right] E(\tau_{kl}^{\beta}\tau_{kj}^{\gamma}) \\
&\quad - \frac{p^2(p-1)^2}{4} E(\tau_{kl}^{\beta})E(\tau_{kl}^{\gamma}) \\
&= \sum_{k=2}^{p}\sum_{l=1}^{k-1} E(\tau_{kl}^{\beta+\gamma}) + \frac{p(p-1)}{2} E(\tau_{kl}^{\beta}\tau_{kj}^{\gamma}) \\
&= 0.
\end{aligned}
$$
■

Proof of Theorem 2.1: Firstly, we will prove that the statistic $t(\gamma)$ converges to a normal distribution for any fixed $\gamma \in \Gamma = \{1, 2, 3, 4\}$.

Recall that $t(\gamma) = \sum_{k=2}^{p}\sum_{l=1}^{k-1} \tau_{kl}^{\gamma}$, $\mu(\gamma) = E[t(\gamma)]$ and $\sigma^2(\gamma) = Var[t(\gamma)]$. We can write the standard statistic

$$
\frac{t(\gamma) - \mu(\gamma)}{\sigma(\gamma)} = \sum_{k=2}^{p} Y_{nk},
$$

where $Y_{nk} = \sigma^{-1}(\gamma)\sum_{l=1}^{k-1} \pi_{kl}$ and $\pi_{kl} = \tau_{kl}^{\gamma} - E(\tau_{kl}^{\gamma})$.

Denote $\mathcal{F}_{n,k}$ be a σ-field generated by $\{R_1, \ldots R_k\}$ for $1 \le k \le p$. Thus for each $n \ge 2$, Y_{nk} is $\mathcal{F}_{n,k}$ measurable. Note that τ_{kl} is independent of $\mathcal{F}_{n,l-1}$ under H_0. Then $E(\pi_{kl}|\mathcal{F}_{n,l-1}) = E(\pi_{kl}) = 0$. Therefore, under H_0, $\{Y_{nk}, \mathcal{F}_{n,k}, 2 \le k \le p\}$ is a martingale difference sequence. By the central limit theorem of martingale difference (See Theorem 2.3 in McLeish 1974), it is sufficient to show that under H_0,

(i) $E(\max_{1\le k\le p} |Y_{nk}|^2) \le C < \infty$,
(ii) $\max_{1\le k\le p} |Y_{nk}| \xrightarrow{p} 0$ as $n \to \infty$,
(iii) $\sum_{k=1}^{p} Y_{nk}^2 \xrightarrow{p} 1$ as $n \to \infty$.

Since the conditions (i) and (ii) are consequences of the classical Lindeberg condition, that is, for any $\epsilon > 0$, $\sum_{k=2}^{p} E[Y_{nk}^2 I(|Y_{nk}| \ge \epsilon)] \to 0$ as $n \to \infty$. By the Cauchy-Schwarz inequality and the Chebyshev inequality, we know

$$
\sum_{k=2}^{p} E[Y_{nk}^2 I(|Y_{nk}| \ge \epsilon)] \le \sum_{k=2}^{p} \sqrt{E(Y_{nk}^4)}\sqrt{P(|Y_{nk}| \ge \epsilon)} \le \frac{1}{\epsilon^2}\sum_{k=2}^{p} E(Y_{nk}^4).
$$

Thus, to establish the central limit theorem of $\frac{t(\gamma)-\mu(\gamma)}{\sigma(\gamma)}$, we only need to show

$$
\sum_{k=2}^{p} E(Y_{nk}^4) \to 0 \tag{9}
$$

and

$$
\sum_{k=2}^{p} Y_{nk}^2 \xrightarrow{p} 1 \tag{10}
$$

as $n \to \infty$.

Firstly, we will prove condition (9). When γ is even, note that $\omega_2 = Var(\tau_{kl}^{\gamma}), \omega_4 = E[\tau_{kl}^{\gamma} - E(\tau_{kl}^{\gamma})]^4$. Under H_0, for different l_1, l_2, l_3, l_4, we know $\tau_{kl_1}, \tau_{kl_2}, \tau_{kl_3}, \tau_{kl_4}$ are mutually independent when R_k is given, then

$$E(\pi_{kl_1}\pi_{kl_2}\pi_{kl_3}\pi_{kl_4}) = E[E(\pi_{kl_1}\pi_{kl_2}\pi_{kl_3}\pi_{kl_4}|R_k)] = E\left[\prod_{i=1}^{4} E(\pi_{kl_i}|R_k)\right] = E\left[\prod_{i=1}^{4} E(\pi_{kl_i})\right] = 0,$$

where $E(\pi_{kl_i}|R_k) = E(\pi_{kl_i})$ can be obtained by Brown and Eagleson (1984).

We will consider the following four cases: $l_1 = l_2 \neq l_3 = l_4$, or $l_1 = l_3 \neq l_2 = l_4$, or $l_1 = l_4 \neq l_2 = l_3$, or $l_1 = l_2 = l_3 = l_4$.

For the first case,

$$E(\pi_{kl_1}\pi_{kl_2}\pi_{kl_3}\pi_{kl_4}) = E(\pi_{kl_1}^2\pi_{kl_3}^2) = E[E(\pi_{kl_1}^2\pi_{kl_3}^2|R_k)] = E[E(\pi_{kl_1}^2|R_k)E(\pi_{kl_3}^2|R_k)] = E(\pi_{kl_1}^2)E(\pi_{kl_3}^2) = \omega_2^2.$$

The second and third case are same to the first one. For the last case, $E(\pi_{kl_1}\pi_{kl_2}\pi_{kl_3}\pi_{kl_4}) = E(\pi_{kl}^4) = \omega_4$. Then we have

$$\begin{aligned}
\sum_{k=2}^{p} E(Y_{nk}^4) &= \sigma^{-4}(\gamma)\sum_{k=2}^{p} E\left(\sum_{l=1}^{k-1}\pi_{kl}\right)^4 \\
&= \sigma^{-4}(\gamma)\sum_{k=2}^{p} E\left(\sum_{l_1=1}^{k-1}\sum_{l_2=1}^{k-1}\sum_{l_3=1}^{k-1}\sum_{l_4=1}^{k-1}\pi_{kl_1}\pi_{kl_2}\pi_{kl_3}\pi_{kl_4}\right) \\
&= \sigma^{-4}(\gamma)\sum_{k=2}^{p} E\left[\sum_{l_1=1}^{k-1}\sum_{l_2=1}^{k-1}\sum_{l_3=1}^{k-1}\sum_{l_4=1}^{k-1}\pi_{kl_1}\pi_{kl_2}\pi_{kl_3}\pi_{kl_4}I(l_1 \neq l_2 \neq l_3 \neq l_4)\right. \\
&\quad +6\sum_{l=1}^{k-1}\sum_{l_1=1}^{k-1}\sum_{l_2=1}^{k-1}\pi_{kl}^2\pi_{kl_1}\pi_{kl_2}I(l \neq l_1 \neq l_2) + 3\sum_{l_1=1}^{k-1}\sum_{l_2=1}^{k-1}\pi_{kl_1}^2\pi_{kl_2}^2 I(l_1 \neq l_2) \\
&\quad \left. +4\sum_{l_1=1}^{k-1}\sum_{l_2=1}^{k-1}\pi_{kl_1}^3\pi_{kl_2}I(l_1 \neq l_2) + \sum_{l_1=1}^{k-1}\pi_{kl_1}^4\right] \\
&= 3\sigma^{-4}(\gamma)\sum_{k=2}^{p}\sum_{l_1=1}^{k-1}\sum_{l_2\neq l_1}^{k-1} E(\pi_{kl_1}^2\pi_{kl_2}^2) + \sigma^{-4}(\gamma)\sum_{k=2}^{p}\sum_{l_1=1}^{k-1} E(\pi_{kl_1}^4) \\
&= 3\sigma^{-4}(\gamma)\omega_2^2\sum_{k=2}^{p}(k-1)(k-2) + \sigma^{-4}(\gamma)\omega_4\sum_{k=2}^{p}(k-1) \\
&= p(p-1)(p-2)\sigma^{-4}(\gamma)\omega_2^2 + \frac{p(p-1)}{2}\sigma^{-4}(\gamma)\omega_4 \\
&= \frac{4}{p^2(p-1)^2}\omega_2^{-2}\omega_2^2 p(p-1)(p-2) + \frac{p(p-1)}{2}\frac{4}{p^2(p-1)^2}\omega_2^{-2}\omega_4
\end{aligned}$$

$$= \frac{4(p-2)}{p(p-1)} + \frac{2\omega_4}{p(p-1)\omega_2^2}.$$

By Lemma 6.1, we can easy to see that for $i = 1, 2, \ldots, 8$, $m_{2i} = O(n^{3i})$. For ω_2 and ω_4, we can obtain

$$\omega_2 = Var(\tau_{kl}^{\gamma}) = O\left(\frac{1}{n^{\gamma}}\right)$$

and

$$\begin{aligned}
\omega_4 &= E(\pi_{kl}^4) = E[\tau_{kl}^{\gamma} - E(\tau_{kl}^{\gamma})]^4 \\
&= E(\tau_{kl}^{4\gamma}) - 4E(\tau_{kl}^{3\gamma})E(\tau_{kl}^{\gamma}) + 6E(\tau_{kl}^{2\gamma})[E(\tau_{kl}^{\gamma})]^2 - 4E(\tau_{kl}^{\gamma})[E(\tau_{kl}^{\gamma})]^3 + [E(\tau_{kl}^{\gamma})]^4 \\
&= \frac{2^{4\gamma}\, m_{4\gamma}}{n^{4\gamma}\,(n-1)^{4\gamma}} - 4\frac{2^{4\gamma}\, m_{3\gamma}\, m_{\gamma}}{n^{4\gamma}\,(n-1)^{4\gamma}} + 6\frac{2^{4\gamma}\, m_{2\gamma}\, {m_{\gamma}}^2}{n^{4\gamma}\,(n-1)^{4\gamma}} - 3\frac{2^{4\gamma}\, {m_{\gamma}}^4}{n^{4\gamma}\,(n-1)^{4\gamma}} \\
&= O\left(\frac{1}{n^{2\gamma}}\right).
\end{aligned}$$

We can see that $\frac{\omega_4}{\omega_2^2}$ only relies on n and it converges to a finite constant as $n \to \infty$. Then $\sum_{k=2}^{p} E(Y_{nk}^4) \to 0$ holds when $p \to \infty$, $n \to \infty$ or n is fixed. Therefore, the condition (9) is proved.

In the sequel, we will prove the condition (10). By Chebyshev inequality, we only need to show

$$(a)E\left(\sum_{k=2}^{p} Y_{nk}^2 - 1\right) = 0, \quad (b)E\left(\sum_{k=2}^{p} Y_{nk}^2 - 1\right)^2 \to 0.$$

Firstly, we will prove condition (a),

$$\begin{aligned}
E\left(\sum_{k=2}^{p} Y_{nk}^2 - 1\right) &= \sigma^{-2}(\gamma)\sum_{k=2}^{p} E\left(\sum_{l=1}^{k-1} \pi_{kl}\right)^2 - 1 \\
&= \sigma^{-2}(\gamma)\sum_{k=2}^{p}\left[\sum_{l_1=1}^{k-1}\sum_{l_2\neq l_1}^{k-1} E(\pi_{kl_1}\pi_{kl_2}) + \sum_{l=1}^{k-1} E(\pi_{kl}^2)\right] - 1 \\
&= \sigma^{-2}(\gamma)\sum_{k=2}^{p}\sum_{l=1}^{k-1} E(\pi_{kl}^2) - 1 \\
&= \frac{p(p-1)}{2}\sigma^{-2}(\gamma)\omega_2 - 1 \\
&= 0,
\end{aligned}$$

where $\sigma^2(\gamma) = \frac{p(p-1)}{2}\omega_2$.

Secondly, we will verify the condition (b). Since

$$E\left(\sum_{k=2}^{p} Y_{nk}^2 - 1\right)^2 = E\left[\left(\sum_{k=2}^{p} Y_{nk}^2\right)^2 - 2\sum_{k=2}^{p} Y_{nk}^2 + 1\right]$$

$$= E\left(\sum_{k=2}^{p} Y_{nk}^2\right)^2 - 1$$

$$= \sum_{k=2}^{p} E(Y_{nk}^4) + \sum_{k_1=2}^{p}\sum_{k_2\neq k_1}^{p} E(Y_{nk_1}^2 Y_{nk_2}^2) - 1.$$

We know that the first item converges to 0. For the second term, we have

$$E(Y_{nk_1}^2 Y_{nk_2}^2) = \sigma^{-4}(\gamma)E\left[\left(\sum_{l_1=1}^{k_1-1}\pi_{k_1l_1}\right)^2\left(\sum_{l_2=1}^{k_2-1}\pi_{k_2l_2}\right)^2\right]$$

$$= \sigma^{-4}(\gamma)\sum_{l_1=1}^{k_1-1}\sum_{l_2=1}^{k_1-1}\sum_{l_3=1}^{k_2-1}\sum_{l_4=1}^{k_2-1} E(\pi_{k_1l_1}\pi_{k_1l_2}\pi_{k_2l_3}\pi_{k_2l_4})$$

when $k_1 \neq k_2$. We can verify that $E(\pi_{k_1l_1}\pi_{k_1l_2}\pi_{k_2l_3}\pi_{k_2l_4}) \neq 0$ only if $l_1 = l_2$ and $l_3 = l_4$. Then

$$E(Y_{nk_1}^2 Y_{nk_2}^2) = \sigma^{-4}(\gamma)\sum_{l_1=1}^{k_1-1}\sum_{l_2=1}^{k_2-1} E(\pi_{k_1l_1}^2\pi_{k_2l_2}^2) = \sigma^{-4}(\gamma)\sum_{l_1=1}^{k_1-1}\sum_{l_2=1}^{k_2-1} E(\pi_{k_1l_1}^2)E(\pi_{k_2l_2}^2)$$

$$= \sigma^{-4}(\gamma)(k_1-1)(k_2-1)\omega_2^2.$$

Therefore,

$$\sum_{k_1=2}^{p}\sum_{k_2\neq k_1}^{p} E(Y_{nk_1}^2 Y_{nk_2}^2) = \sigma^{-4}(\gamma)\omega_2^2\left[\sum_{k_1=2}^{p}\sum_{k_2=2}^{p}(k_1-1)(k_2-1) - \sum_{k=2}^{p}(k-1)^2\right]$$

$$= \frac{4}{p^2(p-1)^2}\left[\frac{p^2(p-1)^2}{4} - \frac{p(p-1)(2p-1)}{6}\right]$$

$$= 1 - \frac{2(2p-1)}{3p(p-1)}.$$

When $p \to \infty$, $\sum_{k_1=2}^{p}\sum_{k_2\neq k_1}^{p} E(Y_{nk_1}^2 Y_{nk_2}^2) \to 1$. Hence, $E(\sum_{k=2}^{p} Y_{nk}^2 - 1)^2 \to 0$, which means that the condition (b) is proved.

Therefore, we have $\sum_{k=2}^{P} Y_{nk}^2 \xrightarrow{p} 1$, which means that (10) is proved.

When γ is odd. Under H_0, we know $E(\tau_{kl}^{\gamma}) = 0$, $Y_{nk} = \sigma^{-1}(\gamma)\sum_{l=1}^{k-1}\pi_{kl}$, where $\pi_{kl} = \tau_{kl}^{\gamma}$. Then

$$\omega_2 = E(\pi_{kl}^2) = E(\tau_{kl}^{2\gamma}) = O\left(\frac{1}{n^{\gamma}}\right),$$

$$\omega_4 = E(\pi_{kl}^4) = E(\tau_{kl}^{4\gamma}) = O\left(\frac{1}{n^{2\gamma}}\right). \tag{11}$$

By (11), $\frac{\omega_4}{\omega_2^2}$ also converges to a finite constant. By the similar arguments as the case that γ is even, we can also prove the asymptotic normality of $t(\gamma)$.

Based on the above arguments, we know that $[t(\gamma) - \mu(\gamma)]/\sigma(\gamma) \xrightarrow{d} N(0,1)$ for any fixed $\gamma \in \Gamma = \{1,2,3,4\}$. By the similar arguments, we can also obtain the asymptotic normality of any linear combination of $\sum_{\gamma=1}^{4} a_\gamma t(\gamma)$ with $a_\gamma \in \mathbb{R}, \gamma = 1,2,3,4$. According to Cramer-Wold theorem, we will get $\{t(\gamma)\}_{\gamma\in\Gamma}^{T}$ converges to a multivariate normal distribution.

Thus the proof of Theorem 2.1 is complete. ■

Proof of Theorem 2.2: When $\beta + \gamma$ is odd, we know that $Cov[t(\beta), t(\gamma)] = 0$ by Proposition 2.3. Thus to show $t(\beta)$ is asymptotically independent of $t(\gamma)$, we only need to show $(t(\beta), t(\gamma))'$ is asymptotically normal distributed. And this can be obtained by proving any fixed linear combination of $t(\beta)$ and $t(\gamma)$ converges to a normal random variable. Therefore, for any real numbers a and b, we need to show the asymptotic normality of $t(\beta,\gamma) \equiv at(\beta) + bt(\gamma)$, where $t(\beta) = \sum_{k=2}^{p}\sum_{l=1}^{k-1}\tau_{kl}^{\beta}, t(\gamma) = \sum_{k=2}^{p}\sum_{l=1}^{k-1}\tau_{kl}^{\gamma}$.

Note that $t(\beta,\gamma) = \sum_{k=2}^{p}\sum_{l=1}^{k-1}(a\tau_{kl}^{\beta} + b\tau_{kl}^{\gamma})$. Without loss of generality, we assume that β is even and γ is odd.

By Propositions 2.1–2.2, we have

$$\begin{aligned} E(a\tau_{kl}^{\beta} + b\tau_{kl}^{\gamma}) &= \frac{a2^{\beta}}{n^{\beta}(n-1)^{\beta}} m_{\beta}, \\ Var(a\tau_{kl}^{\beta} + b\tau_{kl}^{\gamma}) &= a^2 Var(\tau_{kl}^{\beta}) + b^2 Var(\tau_{kl}^{\gamma}) + 2abCov(\tau_{kl}^{\beta}, \tau_{kl}^{\gamma}) \\ &= \frac{a^2 2^{2\beta}}{n^{2\beta}(n-1)^{2\beta}}(m_{2\beta} - {m_{\beta}}^2) + \frac{b^2 2^{2\gamma}}{n^{2\gamma}(n-1)^{2\gamma}} m_{2\gamma}. \end{aligned}$$

Denote

$$Y_{nk}{}^{*} = \sigma^{-1}(\beta,\gamma)\sum_{l=1}^{k-1}\pi_{kl}{}^{*},$$

where

$$\pi_{kl}{}^{*} = a\tau_{kl}^{\beta} + b\tau_{kl}^{\gamma} - E(a\tau_{kl}^{\beta} + b\tau_{kl}^{\gamma}) = a\tau_{kl}^{\beta} + b\tau_{kl}^{\gamma} - \frac{a2^{\beta}}{n^{\beta}(n-1)^{\beta}} m_{\beta}$$

and $\sigma^2(\beta,\gamma) = Var[t(\beta,\gamma)]$. Then

$$\frac{t(\beta,\gamma) - E[t(\beta,\gamma)]}{\sigma(\beta,\gamma)} = \sum_{k=2}^{p} Y_{nk}{}^{*}.$$

Denote $\mathcal{F}_{nk}$ to be a σ-algebra which is generated by $\{R_1, \ldots R_k\}$. Then we can easily get $\{Y_{nk}^{*}, \mathcal{F}_{nk}, 2 \le k \le p, n \ge 2\}$ is a martingale difference sequence. In order to get the central limit theorem of $\sum_{k=2}^{p} Y_{nk}^{*}$, we will also need to verify the similar conditions as the condition (9) and the condition (10).

Denote

$$\omega_2^{*} = Var(a\tau_{kl}^{\beta} + b\tau_{kl}^{\gamma}) = \frac{a^2 2^{2\beta}}{n^{2\beta}(n-1)^{2\beta}}(m_{2\beta} - {m_{\beta}}^2) + \frac{b^2 2^{2\gamma}}{n^{2\gamma}(n-1)^{2\gamma}} m_{2\gamma},$$

$$\omega_4^* = E(\pi_{kl}{}^{*4}) = E[(a\tau_{kl}^{\beta} + b\tau_{kl}^{\gamma}) - E(a\tau_{kl}^{\beta} + b\tau_{kl}^{\gamma})]^4.$$

By Lemma 6.1, we know $\frac{a^2 2^{2\beta}}{n^{2\beta}(n-1)^{2\beta}}(m_{2\beta} - m_{\beta}{}^2) = O(\frac{1}{n^{\beta}})$ and $\frac{b^2 2^{2\gamma}}{n^{2\gamma}(n-1)^{2\gamma}} m_{2\gamma} = O(\frac{1}{n^{\gamma}})$, which means

$$\omega_2^* = O\left(\frac{1}{n^{\min(\beta,\gamma)}}\right). \tag{12}$$

We also have

$$\begin{aligned}
\omega_4^* &= E(\pi_{kl}{}^{*4}) \\
&= E\left[(a\tau_{kl}^{\beta} + b\tau_{kl}^{\gamma}) - \frac{a2^{\beta} m_{\beta}}{n^{\beta}(n-1)^{\beta}}\right]^4 \\
&= E(a\tau_{kl}^{\beta} + b\tau_{kl}^{\gamma})^4 - 4\frac{a2^{\beta} m_{\beta}}{n^{\beta}(n-1)^{\beta}} E(a\tau_{kl}^{\beta} + b\tau_{kl}^{\gamma})^3 \\
&\quad + 6\left[\frac{a2^{\beta} m_{\beta}}{n^{\beta}(n-1)^{\beta}}\right]^2 E(a\tau_{kl}^{\beta} + b\tau_{kl}^{\gamma})^2 - 3\left[\frac{a2^{\beta} m_{\beta}}{n^{\beta}(n-1)^{\beta}}\right]^4 \\
&=: \sum_{i=1}^{4} R_i,
\end{aligned}$$

where

$$\begin{aligned}
R_1 &= E(a\tau_{kl}^{\beta} + b\tau_{kl}^{\gamma})^4 \\
&= a^4 E(\tau_{kl}^{4\beta}) + 4a^3 b E(\tau_{kl}^{3\beta+\gamma}) + 6a^2 b^2 E(\tau_{kl}^{2\beta+2\gamma}) + 4ab^3 E(\tau_{kl}^{\beta+3\gamma}) + b^4 E(\tau_{kl}^{4\gamma}) \\
&= \frac{a^4 2^{4\beta} m_{4\beta}}{n^{4\beta}(n-1)^{4\beta}} + \frac{6a^2 b^2 2^{2\beta+2\gamma} m_{2\beta+2\gamma}}{n^{2\beta+2\gamma}(n-1)^{2\beta+2\gamma}} + \frac{b^4 2^{4\gamma} m_{4\gamma}}{n^{4\gamma}(n-1)^{4\gamma}} \\
&= O\left(\frac{1}{n^{2\beta}}\right) + O\left(\frac{1}{n^{\beta+\gamma}}\right) + O\left(\frac{1}{n^{2\gamma}}\right), \\
R_2 &= -4\frac{a2^{\beta} m_{\beta}}{n^{\beta}(n-1)^{\beta}} E(a\tau_{kl}^{\beta} + b\tau_{kl}^{\gamma})^3 \\
&= -4\frac{a2^{\beta} m_{\beta}}{n^{\beta}(n-1)^{\beta}} [a^3 E(\tau_{kl}^{3\beta}) + 3a^2 b E(\tau_{kl}^{2\beta+\gamma}) + 3ab^2 E(\tau_{kl}^{\beta+2\gamma}) + b^3 E(\tau_{kl}^{3\gamma})] \\
&= -4\left[\frac{a^4 2^{4\beta} m_{\beta} m_{3\beta}}{n^{4\beta}(n-1)^{4\beta}} + \frac{3a^2 b^2 2^{2\beta+2\gamma} m_{\beta} m_{\beta+2\gamma}}{n^{2\beta+2\gamma}(n-1)^{2\beta+2\gamma}}\right] \\
&= O\left(\frac{1}{n^{2\beta}}\right) + O\left(\frac{1}{n^{\beta+\gamma}}\right), \\
R_3 &= 6\left[\frac{a2^{\beta} m_{\beta}}{n^{\beta}(n-1)^{\beta}}\right]^2 E(a\tau_{kl}^{\beta} + b\tau_{kl}^{\gamma})^2 \\
&= 6\left[\frac{a2^{\beta} m_{\beta}}{n^{\beta}(n-1)^{\beta}}\right]^2 [a^2 E(\tau_{kl}^{2\beta}) + b^2 E(\tau_{kl}^{2\gamma}) + 2abE(\tau_{kl}^{\beta+\gamma})]
\end{aligned}$$

$$\begin{aligned}
&= 6\frac{a^4 2^{4\beta} {m_\beta}^2 m_{2\beta}}{n^{4\beta}(n-1)^{4\beta}} + 6\frac{a^2 b^2 2^{2\beta+2\gamma} {m_\beta}^2 m_{2\gamma}}{n^{2\beta+2\gamma}(n-1)^{2\beta+2\gamma}} \\
&= O\left(\frac{1}{n^{2\beta}}\right) + O\left(\frac{1}{n^{\beta+\gamma}}\right), \\
R_4 &= -3\left[\frac{a 2^\beta m_\beta}{n^\beta (n-1)^\beta}\right]^4 = -3\frac{a^4 2^{4\beta} {m_\beta}^4}{n^{4\beta}(n-1)^{4\beta}} \\
&= O\left(\frac{1}{n^{2\beta}}\right).
\end{aligned}$$

Then we can conclude that

$$\omega_4^* = O\left(\frac{1}{n^{2\beta}}\right) + O\left(\frac{1}{n^{\beta+\gamma}}\right) + O\left(\frac{1}{n^{2\gamma}}\right) = O\left(\frac{1}{n^{2\min(\beta,\gamma)}}\right). \tag{13}$$

Similar to the proof procedure of Theorem 2.1,

$$\sum_{k=2}^{p} E({Y_{nk}^*}^4) = \frac{4(p-2)}{p(p-1)} + \frac{2\omega_4^*}{p(p-1){\omega_2^*}^2}.$$

By (12) and (13), we can obtain that $\frac{\omega_4^*}{{\omega_2^*}^2}$ only relies on n and it converges to a finite constant when $n \to \infty$. Therefore, $\sum_{k=2}^{p} E(Y_{nk}^4) \to 0$ when $p \to \infty, n \to \infty$ or n is fixed. Thus we have proved the condition (9).

We similarly can prove that

$$E\left(\sum_{k=2}^{p} {Y_{nk}^*}^2 - 1\right) = 0,$$

$$E\left(\sum_{k=2}^{p} {Y_{nk}^*}^2 - 1\right)^2 \to 0,$$

which means $\sum_{k=2}^{p} {Y_{nk}^*}^2 \xrightarrow{P} 1$. Then the condition (10) is proved.

We can conclude that $at(\beta) + bt(\gamma)$ is asymptotically normal distributed, this together with the fact $Cov[t(\beta), t(\gamma)] = 0$ can draw the conclusion that $t(\beta)$ is asymptotically independent of $t(\gamma)$.

By the similar arguments of Theorem 2 in Wang et al. (2024), Theorem 3.1 in Shi et al. (2024) and Theorem 2.3 in He et al. (2021), we can prove that $t(\infty)$ is asymptotically independent of $t(\gamma), \gamma \in \Gamma = \{1, 2, 3, 4\}$. Thus the proof of Theorem 2.2 is complete. ■

Acknowledgments

The authors would like to thank two anonymous referees for many valuable comments and suggestions that have led to improvements in the paper.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by the National Natural Science Foundation of China [grant number 12471252]; and the Natural Science Foundation of Henan Province of China [grant number 242300421379].

References

Anderson, T. (2003), *An Introduction to Multivariate Statistical Analysis*, New York: John Wiley.

Bao, Z., Lin, L.C., Pan, G., and Zhou, W. (2013), 'Spectral Statistics of Large Dimensional Spearman's Rank Correlation Matrix and Its Application', *Annals of Statistics*, 43, 2588–2623.

Bao, Z., Pan, G., and Zhou, W. (2012), 'Tracy-Widom Law for the Extreme Eigenvalues of Sample Correlation Matrices', *Electronic Journal of Probability*, 17, 1–32.

Bilodeau, M., and Nangue, A.G. (2017), 'Tests of Mutual Or Serial Independence of Random Vectors with Applications', *Journal of Machine Learning Research*, 18(74), 1–40.

Birke, M., and Dette, H. (2005), 'A Note on Testing the Covariance Matrix for Large Dimension', *Statistics and Probability Letters*, 74(3), 281–289.

Blum, J.R., Kiefer, J., and Rosenblatt, M. (1961), 'Distribution Free Tests of Independence Based on the Sample Distribution Function', *The Annals of Mathematical Statistics*, 32(2), 485–498.

Brown, T.C., and Eagleson, G.K. (1984), 'A Useful Property of Some Symmetric Statistics', *The American Statistician*, 38(1), 63–65.

Cai, T., and Jiang, T. (2011), 'Limiting Laws of Coherence of Random Matrices with Applications to Testing Covariance Structure and Construction of Compressed Sensing Matrices', *Annals of Statistics*, 39(3), 1496–1525.

Cai, T., and Jiang, T. (2012), 'Phase Transition in Limiting Distributions of Coherence of High-dimensional Random Matrices', *Journal of Multivariate Analysis*, 107(3), 24–39.

Chakraborty, S., and Zhang, X. (2019), 'Distance Metrics for Measuring Joint Dependence with Application to Causal Inference', *Journal of the American Statistical Association*, 114(528), 1638–1650.

Chen, S., Zhang, L., and Zhong, P. (2010), 'Tests for High-dimensional Covariance Matrices', *Journal of the American Statistical Association*, 105(490), 810–819.

Drton, M., Han, F., and Shi, H. (2020), 'High-dimensional Consistent Independence Testing with Maxima of Rank Correlations', *The Annals of Statistics*, 48(6), 3206–3227.

Fan, J., Liao, Y., and Yao, J. (2015), 'Power Enhancement in High-dimensional Cross-sectional Tests', *Econometrica*, 83(4), 1497–1541.

Hájek, J., Šidák, Z., and Sen, P.K. (1999), *Theory of Rank Tests*, New York: Academic Press.

Han, F., Chen, S., and Liu, H. (2017), 'Distribution-free Tests of Independence in High Dimensions', *Biometrika*, 104(4), 813–828.

He, D., Liu, H., Xu, K., and Cao, M. (2021), 'Generalized Schott Type Tests for Complete Independence in High Dimensions', *Journal of Multivariate Analysis*, 183, 104731.

He, Y., Xu, G., Wu, C., and Pan, W. (2021), 'Asymptotically Independent U-statistics in High-dimensional Testing', *The Annals of Statistics*, 49(1), 154–181.

Hoeffding, W. (1948), 'A Class of Statistics with Asymptotically Normal Distribution', *The Annals of Mathematical Statistics*, 19(3), 293–325.

Jiang, T. (2004), 'The Asymptotic Distributions of the Largest Entries of Sample Correlation Matrices', *Annals of Probability*, 14, 865–880.

Jiang, D., Bai, Z., and Zheng, S. (2013), 'Testing the Independence of Sets of Large-dimensional Variables', *Science China Mathematics*, 56, 135–147.

Jiang, T., and Qi, Y. (2015), 'Limiting Distributions of Likelihood Ratio Tests for High-dimensional Normal Distributions', *Scandinavian Journal of Statistics*, 42(4), 988–1009.

Jiang, T., and Yang, F. (2013), 'Central Limit Theorems for Classical Likelihood Ratio Tests for High-dimensional Normal Distributions', *Annals of Statistics*, 41(4), 2029–2074.

Johnstone, I.M. (2001), 'On the Distribution of the Largest Eigenvalue in Principal Components Analysis', *The Annals of Statistics*, 29(2), 295–327.

Kendall, M.G. (1938), 'A New Measure of Rank Correlation', *Biometrika*, 30, 81–93.
Ledoit, O., and Wolf, M. (2002), 'Some Hypothesis Tests for the Covariance Matrix when the Dimension is Large Compared to the Sample Size', *The Annals of Statistics*, 30(4), 1081–1102.
Leung, D., and Drton, M. (2017), 'Testing Independence in High Dimensions with Sums of Rank Correlations', *Annals of Statistics*, 46(1), 280–307.
Li, D., Liu, W., and Rosalsky, A. (2010), 'Necessary and Sufficient Conditions for the Asymptotic Distribution of the Largest Entry of a Sample Correlation Matrix', *Probability Theory and Related Fields*, 148(1), 5–35.
Li, D., Qi, Y., and Rosalsky, A. (2012), 'On Jiang's Asymptotic Distribution of the Largest Entry of a Sample Correlation Matrix', *Journal of Multivariate Analysis*, 111, 256–270.
Liu, W., Lin, Z., and Shao, Q. (2008), 'The Asymptotic Distribution and Berry-Esseen Bound of a New Test for Independence in High Dimension with An Application to Stochastic Optimization', *The Annals of Applied Probability*, 18(6), 2337–2366.
Lucaks, E. (1955), 'Applications of Faàdi Bruno's Formula in Mathematical Statistics', *American Mathematical Monthly*, 62(5), 340–348.
Mao, G. (2014), 'A New Test of Independence for High-dimensional Data', *Statistics & Probability Letters*, 93, 14–18.
Mao, G. (2016), 'Robust Test for Independence in High Dimensions', *Communications in Statistics-Theory and Methods*, 46(20), 10036–10050.
Mao, G. (2018), 'Testing Independence in High Dimensions Using Kendall's Tau', *Computational Statistics and Data Analysis*, 117, 128–137.
Mao, G. (2020), 'On High-dimensional Tests for Mutual Independence Based on Pearson's Correlation Coefficient', *Communications in Statistics-Theory and Methods*, 49(14), 3572–3584.
McLeish, D.L. (1974), 'Dependent Central Limit Theorems and Invariance Principles', *The Annals of Probability*, 2(4), 620–628.
Mosteller, F., and Fisher, R.A. (1948), 'Questions and Answers', *The American Statistician*, 2(5), 30–31.
Nguyen, D., and Eisenstein, J. (2017), 'A Kernel Independence Test for Geographical Language Variation', *Computational Linguistics*, 43, 567–592.
Pan, W., Kim, J., Zhang, Y., Shen, X., and Wei, P. (2014), 'A Powerful and Adaptive Association Test for Rare Variants', *Genetics*, 197(4), 1081–1095.
Pearson, K. (1920), 'Notes on the History of Correlation', *Biometrika*, 13(1), 25–45.
Pfister, N., Bühlmann, P., Schölkopf, B., and Peters, J. (2021), 'Kernel-based Tests for Joint Independence', *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 80, 5–31.
Pillai, N.S., and Yin, J. (2012), 'Edge Universality of Correlation Matrices', *The Annals of Statistics*, 40(3), 1737–1763.
Schott, J.R. (2005), 'Testing for Complete Independence in High Dimensions', *Biometrika*, 92(4), 951–956.
Sheng, W., and Yin, X. (2016), 'Sufficient Dimension Reduction Via Distance Covariance', *Journal of Computational and Graphical Statistics*, 25, 91–104.
Shi, X., Jiang, Y., Du, J., and Miao, Z. (2024), 'An Adaptive Test Based on Kendall's Tau for Independence in High Dimensions', *Journal of Nonparametric Statistics*, 36(4), 1064–1087. DOI: 10.1080/10485252.2023.2296521.
Shi, X., Xu, M., and Du, J. (2023), 'Max-sum Test Based on Spearman's Footrule for High-dimensional Independence Tests', *Computational Statistics and Data Analysis*, 185, 107768.
Shi, X., Zhang, W., Du, J., and Kwessi, E. (2024), 'Testing Independence Based on Spearman's Footrule in High Dimensions', *Communications in Statistics-Theory and Methods*. DOI: 10.1080/03610926.2024.2369313
Silva, P.F., Marcal, A.R., and Silva, R.M. (2013), 'Evaluation of Features for Leaf Discrimination', *Springer Lecture Notes in Computer Science*, 7950, 197–204.
Silverstone, H. (1950), 'A Note on the Cumulants of Kendall's S-distribution', *Biometrika*, 37(3-4), 231–235.
Spearman, C. (1904), 'The Proof and Measurement of Association between Two Things', *The American Journal of Psychology*, 15(1), 72–101.

Srivastava, M.S. (2005), 'Some Tests concerning the Covariance Matrix in High Dimensional Data', *Journal of the Japan Statistical Society*, 35(2), 251–272.

Tsanas, A., Little, M.A., Fox, C., and Ramig, L.O. (2014), 'Objective Automatic Assessment of Rehabilitative Speech Treatment in Parkinson's Disease', *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 22(1), 181–190.

Wang, H., Liu, B., Feng, L., and Ma, Y. (2024), 'Rank-based Max-sum Tests for Mutual Independence of High-dimensional Random Vectors', *Journal of Econometrics*, 238(1), 105578.

Xu, G., Lin, L., Wei, P., and Pan, W. (2016), 'An Adaptive Two-sample Test for High-dimensional Means', *Biometrika*, 103(3), 609–624.

Yao, S., Zhang, X., and Shao, X. (2018), 'Testing Mutual Independence in High Dimension Via Distance Covariance', *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 80(3), 455–480.

Yu, K., Li, Q., Bergen, A.W., Pfeiffer, R.M., Rosenberg, P.S., Caporaso, N., Kraft, P., and Chatterjee, N. (2009), 'Pathway Analysis by Adaptive Combination of P-values', *Genetic Epidemiology*, 33(8), 700–709.

Zhou, W. (2007), 'Asymptotic Distribution of the Largest off-diagonal Entry of Correlation Matrices', *Transactions of the American Mathematical Society*, 359(11), 5345–5363.