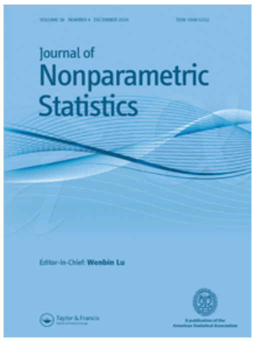

Journal of Nonparametric Statistics

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

An adaptive test based on Kendall's tau for independence in high dimensions

Xiangyu Shi, Yuanyuan Jiang, Jiang Du & Zhuqing Miao

To cite this article: Xiangyu Shi, Yuanyuan Jiang, Jiang Du & Zhuqing Miao (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

To link to this article: https://doi.org/10.1080/10485252.2023.2296521

Published online: 28 Dec 2023.

JOURNAL OF NONPARAMETRIC STATISTICS
2024, VOL. 36, NO. 4, 1064–1087
https://doi.org/10.1080/10485252.2023.2296521

Taylor & Francis
Taylor & Francis Group

Check for updates

An adaptive test based on Kendall's tau for independence in high dimensions

Xiangyu Shi[a], Yuanyuan Jiang[a], Jiang Du[a,b] and Zhuqing Miao[a]

[a]Faculty of Science, Beijing University of Technology, Beijing, People's Republic of China; [b]Beijing Institute of Scientific and Engineering Computing, Beijing, People's Republic of China

ABSTRACT

We consider testing the mutual independency for high-dimensional data. It is known that L_2-type statistics have lower power under sparse alternatives and L_∞-type statistics have lower power under dense alternatives in high dimensions. In this paper, we develop an adaptive test based on Kendall's tau to compromise both situations of the alternative, which can automatically be adapted to the underlying data. An adaptive test is very useful in practice as the sparsity or density for a data set is usually unknown. In addition, we establish the asymptotic joint distribution of L_2-type and L_∞-type statistics based on Kendall's tau under mild assumptions and the asymptotic null distribution of the proposed statistic. Simulation studies show that our adaptive test performs well in either dense or sparse cases. To illustrate the usefulness and effectiveness of the proposed test, real data sets are also analysed.

ARTICLE HISTORY
Received 5 May 2023
Accepted 11 December 2023

KEYWORDS
Asymptotic independence; high dimensions; complete independence; Kendall's tau; rank correlation

2010 MATHEMATICS SUBJECT CLASSIFICATIONS
62G10; 62G20; 62H15

1. Introduction

Testing the null hypothesis of mutual independence between the components of a random vector is an important topic in statistics and machine learning. In real data analysis, some useful modelling tools depend heavily on the mutual independence assumption, for example, independent component analysis, insurance portfolio (Dhaene, Denuit, Goovaerts, Kaas, and Vyncke 2002a, 2002b), blind identification techniques (Cardoso and Souloumiac 1993) and so on. Let $\boldsymbol{X} = (X_1, \ldots, X_d)^\top \in \mathbb{R}^d$ be a continuous random vector and let $\boldsymbol{x}_{\cdot,i} (i \in \{1, \ldots, n\})$ be the independent and identically distributed (i.i.d) observations of $\boldsymbol{X}$. We are interested in testing the independence of the components of $\boldsymbol{X}$, namely

$$H_0 : X_1, \ldots, X_d \quad \text{are mutually independent.}$$

The bivariate case, i.e. $d = 2$, has been studied by many scholars. Two classical rank-based tests for nonparametric independence are Spearman's rank correlation coefficient (Hotelling and Pabst 1936) and Kendall's rank correlation coefficient (Kendall 1938).

CONTACT Jiang Du dujiang84@163.com Faculty of Science, Beijing University of Technology, Beijing 100124, People's Republic of China; Beijing Institute of Scientific and Engineering Computing, Beijing 100124, People's Republic of China

Another type of nonparametric test depends on distance. Hoeffding (1948) and Blum, Kiefer, and Rosenblatt (1961) studied the integration of the squared distance between the joint cumulative distribution function and the product of marginal distribution functions. Rosenblatt (1975) presented a distance measure based on the density function and used kernel density estimates to construct a test statistic. In addition, Székely, Rizzo, and Bakirov (2007) and Székely and Rizzo (2009) developed the distance correlation coefficient based on the joint characteristic function and the marginal characteristic functions, and the statistic has good performance for non-monotonic and nonlinear data. Recently, Shen, Priebe, and Vogelstein (2020) extended distance correlation to Multiscale Graph Correlation, and there was almost no loss of power in the monotonic dependence.

With rapid development in data collection and storage techniques, it is popular nowadays to test the independence of high-dimensional data, in which the dimension is large relative to the sample size, i.e. $d > n$. The two types of constructed statistics commonly used are L_2-type (see Schott 2005; Mao 2018; Yao, Zhang, and Shao 2018) and L_∞-type (see Cai and Jiang 2011; Han, Chen, and Liu 2017).

When implementing testing the mutual independency for high-dimensional data, the practitioner could encounter with two common types of alternatives, dense and sparse cases. The dense alternatives mean that there are many small non-diagonal elements in the correlation matrix. However, the sparse alternatives mean that there are only a few large elements in the correlation matrix. There is a large body of work done on the two alternatives. In both cases, the researchers proposed the L_2-type and L_∞-type test statistics.

The L_2-type statistics are used to deal with dense alternative problems. Specifically, Schott (2005) first proposed a squared sum test statistic based on the Pearson correlation coefficient, which has good power under linear correlation and normal distribution. Mao (2017) proved that the squared sum test statistic based on Spearman's rho converges weakly to standard normal distribution under the null hypothesis. A statistic constructed by the pairwise distance covariance was introduced in Yao et al. (2018) to account for the nonlinear and non-monotonic dependence of the data. Leung and Drton (2018) considered three rank-based L_2-type test statistics without specific distribution or moment assumptions. Unfortunately, the L_2-type statistics may perform poorly under sparsity, see Yao et al. (2018).

On the contrary, the L_∞-type test statistics are usually powerful against sparse alternatives. For example, Cai and Jiang (2011) studied the statistics of the largest off-diagonal entries in the sample covariance matrix and applied the theoretical results to the construction of compressed sensing matrices. Zhou (2007) proposed the test statistic based on Spearman's rank correlation coefficient and proved that the asymptotic distribution was the Gumbel distribution under the null hypothesis. For nonlinear and non-monotone dependent data, a statistic is constructed from the maximum value of the degenerate U-statistic (see Drton, Han, and Shi 2020).

In real applications, the underlying data generation process is usually unknown, which could be either sparse, dense or in-between. However, both L_2-type and L_∞-type statistics are sensitive to the type of the alternative hypothesis. To solve this problem, Fan, Liao, and Yao (2015) proposed a test statistic named 'power enhancement component' based on Pearson correlation coefficient on the basis of the screening technique to enhance the power of quadratic statistics under sparse alternatives. He, Xu, Wu, and Pan (2021) presented an adaptive test combining p-value calculated from U-statistics of different orders, based on the asymptotic independence of L_2-type and L_∞-type statistics. Nevertheless, high moment assumptions are required for random vectors, which are a stringent condition for real data analysis. In addition, the asymptotic distributions of the above-mentioned test methods rely on the underlying data generation process. Therefore, it is in need of estimating the corresponding asymptotic variance. Unfortunately, as well known to all, it is difficult to obtain a good estimator of the asymptotic variance, especially for high dimensional data.

To get rid of the stringent high order moment condition and obtain a distribution free test method, we propose an adaptive test statistic based on Kendall's tau that combines L_2-type with L_∞-type test statistics in this paper. Under null hypothesis, we prove that L_2-type and L_∞-type test statistics based on Kendall's tau are asymptotically independent. Further, for d-dimensional continuous random vector, we obtain the asymptotic null distribution of the proposed test statistic. Compared with other existing works (Schott 2005; Cai and Jiang 2011; He et al. 2021), the proposed test works well when the second moment of the random vector $\boldsymbol{X}$ is infinite. In addition, the proposed test statistic is exactly distribution-free because the resulting test method is based on rank. Consequently, the null distribution of the proposed test statistic can be approximated via Monte Carlo simulation. We can obtain a simulation-based critical value table with sample size n and dimension d. This improves the efficiency of the calculation compared to the permutation test or bootstrap test methods. Importantly, the proposed test method can be automatically adapted to the underlying data, whether sparse or dense alternatives.

The rest of this paper is organised as follows. Section 2 displays the background materials about the L_2-type and L_∞-type statistics based on Kendall's tau. In Section 3, we give the adaptive test statistic and prove the asymptotic independence of the L_2-type and L_∞-type statistics. On this basis, we further give the asymptotic null distribution of the proposed statistic. Finite sample performance of the resulting test statistic is examined by simulation experiments in Section 4. Applications to studies of welding data and biochemical data are given in Section 5. Concluding remarks are given in Section 6. All technical details are deferred to the Appendixes.

2. Preliminary: L_2-type and L_∞-type statistics based on Kendall's tau

Consider a d-dimensional continuous random vector $\boldsymbol{X} = (X_1, \dots, X_d)^\top$ taking values in $\mathbb{R}^d$. Let $\{\boldsymbol{x}_{\cdot,i} = (x_{1i}, \dots, x_{di})^\top, i - 1, \dots, n\}$ be n i.i.d. observations of $\boldsymbol{X}$. For any two entries $k \neq \ell \in \{1, \dots, d\}$, let R_{ki} and $R_{\ell i}$ be the rank of x_{ki} and $x_{\ell i}$ $(i = 1, \dots, n)$. Recall the classic metric, Kendall's tau of X_k and X_ℓ is defined as

$$\tau_{k\ell} = \frac{2V_{k\ell}}{n(n-1)},$$

where $V_{k\ell} = \sum_{i=2}^{n} \sum_{j=1}^{i-1} \operatorname{sign}(R_{ki} - R_{kj})\operatorname{sign}(R_{\ell i} - R_{\ell j})$ and the sign function $\operatorname{sign}(\cdot)$ is defined as $\operatorname{sign}(x) = x/|x|$ with the convention $0/0 = 0$. As a nonparametric measure, Kendall's tau can be invoked to detect independence among two variables, see Sidak, Sen, and Hajek (1999).

To introduce the test statistic below, define

$$\omega_1 = \frac{2(2n+5)}{9n(n-1)}$$

and

$$\omega_2 = \frac{4d(d-1)(n-2)(100n^3+492n^2+731n+279)}{2025n^3(n-1)^3}.$$

Motivated by Schott (2005) and Mao (2018) presented an L_2-type statistic for testing complete independence with Kendall's tau. The statistic is

$$S_\tau = \omega_2^{-1/2}\left(\sum_{k=2}^{d}\sum_{\ell=1}^{k-1}\tau_{k\ell}^2 - \frac{d(d-1)}{2}\omega_1\right). \tag{1}$$

In Equation (1), ω_1 denotes the exact expectation of $\tau_{k\ell}^2 (k \neq \ell \in \{1,\dots,d\})$ under the null hypothesis and ω_2 denotes the exact variance of $\sum_{k=2}^{d}\sum_{\ell=1}^{k-1}\tau_{k\ell}^2$ under the null hypothesis.

Lemma 2.1 (Theorem 1 of Mao 2018**):** *Under H_0, S_τ converges in distribution to the standard normal distribution $\Phi(x)$ with $(n,d)\to\infty$ or $d\to\infty$ for fixed n.*

According to Lemma 2.1, we investigate the asymptotic null distribution of the test statistic S_τ. In particular, the next theorem characterises the convergence rate for S_τ.

Theorem 2.1: *Under the same condition of Lemma 2.1, it has*

$$\sup_x |\Pr(S_\tau \le x) - \Phi(x)| \le O(p^{-1/5}).$$

Han et al. (2017) examined distribution-free test statistics based on L_∞-type, which includes Kendall's tau as an important example. The test statistic they proposed based on Kendall's tau is

$$M_\tau = \omega_1^{-1}\max_{\ell<k}\tau_{k\ell}^2 - 4\ln d + \ln\ln d, \tag{2}$$

where $\mathrm{E}_{H_0}(\tau_{k\ell}) = 0$ and $\mathrm{Var}_{H_0}(\tau_{k\ell}) = \omega_1, k \neq \ell \in \{1,\dots,d\}$.

Lemma 2.2 (Theorem 1 of Han et al. 2017**):** *Under H_0, if $d\to\infty$ and $\ln d = o(n^{1/3})$ as $n\to\infty$, M_τ converges weakly to the Gumbel distribution,*

$$F(y) = \exp\{-(8\pi)^{-1/2}\exp(-y/2)\}.$$

Lemma 2.3 (Theorem A1 of Han et al. 2017**):** *Under the same condition of Lemma 2.2, it has*

$$\sup_y |\Pr(M_\tau \le y) - F(y)| = O\left(\frac{(\ln d)^{3/2}}{n^{1/2}} + \frac{1}{(\ln d)^{3/2}}\right).$$

Lemma 2.2 gives the asymptotic null distribution of the statistic M_τ. Lemma 2.3 characterises the convergence rate for the statistic M_τ. Obviously, M_τ suffers from size distortions due to the slow convergence to the extreme value distribution. The exact distribution of the proposed test statistic is free of the data generating process. Thus we can approximate the exact distribution of the statistic M_τ via simulation to accelerate the rate of convergence. Under H_0, $\tau_{k\ell}$ is independent and depends only on the ranks R_{ki} and $R_{\ell i}$ $(i \in \{1,\dots,n\})$. The L_∞-type test statistic has lower power for a similar problem under the dense alternatives.

3. The proposed test method

It is known that the L_2-type test statistics have lower power under sparse alternatives and the L_∞-type test statistics have lower power under dense alternatives. In practice, we usually have no knowledge of whether the data is sparsely dependent. To enable automatic adaptation to the underlying data, we introduce the adaptive test statistic,

$$C_\tau = \min\{1 - F(M_\tau), 1 - \Phi(S_\tau)\}. \tag{3}$$

The minimum combination is a popular method in which the idea is to approximate the maximum power by taking the minimum p-value (Pan, Kim, Zhang, Shen, and Wei 2014; Xu, Lin, Wei, and Pan 2016).

Since $\Phi(x)$ and $F(y)$ are continuous functions, the following results hold according to the continuous mapping theorem, i.e. Lemma A.5,

$$\begin{aligned} 1 - F(M_\tau) &\xrightarrow{d} U[0,1], \\ 1 - \Phi(S_\tau) &\xrightarrow{d} U[0,1], \end{aligned} \tag{4}$$

where $\xrightarrow{d}$ denotes convergence in distribution. Hence, the asymptotic distribution of C_τ can be obtained if we can prove that S_τ and M_τ are asymptotically independent. Before proving the asymptotic independence of S_τ and M_τ, we first give the following two lemmas: Lemmas 3.1 and 3.2. Lemma 3.1 gives an upper bound on the rth central moment of $\tau_{k\ell}^2$.

Lemma 3.1: *Let $r \geq 2$ be given, under H_0, then*

(1) $\mathrm{E}(|\tau_{k\ell}^2 - \mathrm{E}\tau_{k\ell}^2|^r) \leq Cn^{-r}$;
(2) $\mathrm{E}\{|\sum_{k=\ell+1}^{d}(\tau_{k\ell}^2 - \mathrm{E}\tau_{k\ell}^2)|^r\} \leq C(d-\ell)^{r/2}n^{-r}$;
(3) $\mathrm{E}\{|\sum_{\ell=1}^{k-1}(\tau_{k\ell}^2 - \mathrm{E}\tau_{k\ell}^2)|^r\} \leq C(k-1)^{r/2}n^{-r}$,

for all $1 \leq \ell < k \leq d$ and $d \geq 3$, where C is a constant depending on r only.

For notation simplicity, we set $\ell_d = \omega_1^{1/2}(4\ln d - \ln\ln d + y)^{1/2}$ and $\Lambda_d = \{(\ell, k); 1 \leq \ell < k \leq d\}$. Define $T_d(x) = \{S_\tau \leq x\}$ and $Z_I = \{|\tau_{k\ell}| > \ell_d\}$ for any $I = (\ell, k) \in \Lambda_d$. To make a clear presentation, we simply order the elements in Λ_d. For any $I_1 = (\ell_1, k_1) \in \Lambda_d$ and $I_2 = (\ell_2, k_2) \in \Lambda_d$, we say $I_1 < I_2$ if $\ell_1 < \ell_2$ or $\ell_1 = \ell_2$ but $k_1 < k_2$.

Lemma 3.2: *Under H_0, for a fixed positive integer $q \geq 1$, we have*

$$\lim_{d\to\infty} G(d,q) \to \frac{1}{q!}F^q(y) < \infty, \tag{5}$$

where $G(d,q) = \sum_{I_1<\cdots<I_q\in\Lambda_d} \Pr(Z_{I_1}\cdots Z_{I_q})$ and $F(y)$ represents the Gumbel distribution. And then, we have

$$\sum_{I_1<\cdots<I_q\in\Lambda_d} |\Pr(T_d Z_{I_1}\cdots Z_{I_q}) - \Pr(T_d)\Pr(Z_{I_1}\cdots Z_{I_q})| \to 0,$$

as $d \to \infty$.

Based on Lemmas 3.1 and 3.2, the following theorem shows that the statistics S_τ and M_τ have a distribution-free joint limit law under H_0. Note that, no assumptions about the moment of the random vector are needed in this theorem.

Theorem 3.1: *Under H_0, the statistics S_τ and M_τ are asymptotically independent as $n \to \infty$, i.e.*

$$\lim_{d\to\infty} \Pr\left(S_\tau \le x, M_\tau \le y\right) = \Phi(x) \cdot F(y),$$

when $\ln d = o(n^{1/3})$ *as* $n \to \infty$.

Remark 3.1: He et al. (2021) showed that the L_∞-type $\mathcal{U}(\infty)$ test statistic is asymptotically mutually independent with finite-order U-statistics with moment assumptions that

$$\lim_{d\to\infty} \min_{1\le j\le d} \mathrm{E}\left(X_j - \mu_j\right)^2 > 0, \quad \text{and} \quad \lim_{d\to\infty} \max_{1\le j\le d} \mathrm{E}\left(X_j - \mu_j\right)^8 < \infty.$$

The test statistic $\mathcal{U}(\infty)$ based on the Pearson correlation coefficient is as follows:

$$\mathcal{U}(\infty) = \max_{1\le j_1\ne j_2\le d} |\hat{\sigma}_{j_1,j_2}/\sqrt{\hat{\sigma}_{j_1,j_1}\hat{\sigma}_{j_2,j_2}}|,$$

where $(\hat{\sigma}_{j_1,j_2})_{d\times d} = \sum_{i=1}^n (\boldsymbol{x}_{\cdot,i} - \bar{\boldsymbol{x}})(\boldsymbol{x}_{\cdot,i} - \bar{\boldsymbol{x}})^\top/n$ and $\bar{\boldsymbol{x}} = \sum_{i=1}^n \boldsymbol{x}_{\cdot,i}/n$. Therefore, the resulting test statistic requires the second moment of the random vector $X = (X_1, \dots, X_d)^\top$. Moreover, the eighth moment condition is needed to establish the asymptotic joint distribution of different U-statistics. Although Kendall's tau is a U-statistic, we may not directly apply the conclusion in He et al. (2021) to derive Theorem 3.1 in our paper as the moment assumption on the random vector is strong. If one of the margin distributions has infinite second moment, the $\mathcal{U}(\infty)$ test method is consequently invalid. In fact, Theorem 3.1 works well when the variance of the random vector $\boldsymbol{X}$ is infinite. Moreover, it is insensitive to outliers and allows for a heavy-tailed distribution for the random vector.

The test statistic C_τ is constructed by the minimum value function, whose asymptotic distribution can be inscribed by the minimum value of two random variables. The following theorem establishes the asymptotic null distribution of C_τ.

Theorem 3.2: *Set $C_\tau = \min\{1 - F(M_\tau), 1 - \Phi(S_\tau)\}$. Under H_0, if $d \to \infty$ and $\ln d = o(n^{1/3})$ as $n \to \infty$, C_τ converges to $W = \min\{A, B\}$ in distribution as $n \to \infty$, where A and B are i.i.d. random variables with distribution $U[0, 1]$. The distribution function of W is given by $H(t) = 2t - t^2$ for $t \in [0, 1]$.*

According to Theorem 3.2, we will reject the null hypothesis if $C_\tau < 1 - \sqrt{1-\alpha}$, where $\alpha \in (0, 1)$ is the significance level.

4. Simulation studies

In this section, we use Monte Carlo simulations to investigate the finite sample performance of the proposed test statistic C_τ. We compare C_τ with the test statistics proposed by Mao (2018), Han et al. 2017, Schott (2005) and Cai and Jiang (2011) denoted as S_τ,

M_τ, S_r and M_r, respectively. S_r and M_r are the L_2-type and L_∞-type statistics based on the Pearson correlation coefficient, respectively. We also compare C_τ with the methods in Fan et al. (2015) and He et al. (2021), which are denoted as 'PE_r' and '$U_{\min}$', respectively. The null distributions of the test statistics based on rank are distribution free, thus the corresponding null distributions can be approximated by Monte Carlo simulations. In the simulation, 5000 Monte Carlo runs are used to obtain the critical values for each combination of (n, d). We use the notation MC_τ and TC_τ to denote the proposed test methods whose critical values are obtained by Monte Carlo simulations and the asymptotic null distributions, respectively. The same applies to S_τ and M_τ.

4.1. Finite sample performance under dense and sparse alternatives

Set the sample size $n = 50, 100$ and the dimension $d = 50, 100, 200, 400$. The significance level is taken as $\alpha = 0.05$ in each case, and all results are averaged over 2000 independent replications. We used the following seven models to generate the data $\mathcal{X} = (x_{ij})_{n\times d}$:

Model 1: $x_{ij} \overset{i.i.d}{\sim} N(0, 1)$.

Model 2: $x_{ij} \overset{i.i.d}{\sim} Cauchy(0, 1)$.

Model 3: $x_{ij} \overset{i.i.d}{\sim} t(4)$.

Model 4: $\boldsymbol{x}_{1,\cdot}, \ldots, \boldsymbol{x}_{n,\cdot}$ are independently generated from a normal distribution $N_d(0, \Sigma)$ with $\Sigma = (1 - \frac{\ln d}{2n})I_d + D_d$ and $D_d = (\frac{\ln d}{2n})_{d\times d}$.

Model 5: $(x_{i1}, \ldots, x_{id})$ is generated by $(x_{i1}, \ldots, x_{id})^\top = (y_{i0}, \ldots, y_{i(d-1)})^\top + (\frac{2\ln d}{n})(y_{i1}, \ldots, y_{id})^\top$ where $y_{ij} \overset{i.i.d}{\sim} Cauchy(0, 1)$ across j and i.

Model 6: $\boldsymbol{x}_{1,\cdot}, \ldots, \boldsymbol{x}_{n,\cdot}$ are independently generated from a normal distribution $N_d(0, \Sigma)$ with $\Sigma = (\sigma_{ij})_{d\times d}, \sigma_{ii} = 1, i = 1, 2, \ldots .d, \sigma_{12} = \sigma_{21} = \frac{1}{8}\ln d$ and $\sigma_{ij} = 0$ for others.

Model 7: $\boldsymbol{x}_{1,\cdot}, \ldots, \boldsymbol{x}_{n,\cdot}$ are independently generated. We first generate $Z_1, \ldots, Z_d$ from standard Cauchy distribution, and then construct $\boldsymbol{X} = (X_1, \ldots, X_d)^\top$ by setting $X_1 = Z_1 + (\frac{3\ln d}{n})Z_2, X_2 = Z_2 + (\frac{3\ln d}{n})Z_1$ and $X_j = Z_j$ for $3 \le j \le d$.

We investigate the empirical sizes of the tests in *Models* 1–3. To study the power properties of the different test statistics, we consider two sets of dense alternatives (*Models* 4–5) and sparse alternatives (*Models* 6–7).

Tables 1–3 describe the empirical size and power of different test statistics. Table 1 shows the empirical sizes for the cases in which the generated data are constant-tailed, thick-tailed and expectation non-existent. The empirical sizes for TS_τ, MS_τ and MC_τ are close to the nominal sizes. The empirical size of TC_τ is slightly smaller than the nominal size for n or p due to the fact that the asymptotic critical value is used. The empirical size of TM_τ is relatively conservative, which is mainly due to its slower convergence to the Gumbel distribution. Further, all sizes of MM_τ with simulation-based critical values are now close to the nominal size. Test S_r has good performance in *Model 1* and cannot test independence in *Model 2*. M_r has a lower size under *Model 1*, primarily owing to the slower rate of convergence. M_r is unable to test independence under *Model 3*, and the empirical size is even 1 under *Model 2*. The empirical size of test PE_r in *Model 1* approaches the nominal size

Table 1. Empirical sizes of tests.

n	50				100			
d	50	100	200	400	50	100	200	400
Model 1								
S_r	0.042	0.055	0.048	0.053	0.047	0.044	0.047	0.049
TS_τ	0.044	0.053	0.049	0.049	0.050	0.043	0.053	0.053
MS_τ	0.046	0.057	0.052	0.051	0.056	0.045	0.055	0.055
M_r	0.013	0.007	0.001	0.001	0.021	0.020	0.013	0.009
TM_τ	0.029	0.028	0.018	0.013	0.029	0.027	0.027	0.033
MM_τ	0.044	0.063	0.052	0.051	0.041	0.047	0.044	0.052
TC_τ	0.037	0.037	0.031	0.029	0.040	0.036	0.037	0.044
MC_τ	0.042	0.056	0.047	0.040	0.049	0.048	0.056	0.053
PE_r	0.168	0.135	0.080	0.073	0.068	0.058	0.053	0.051
$U_{\min}$	0.060	0.073	0.065	0.072	0.062	0.060	0.061	0.055
Model 2								
S_r	0.418	0.439	0.432	0.440	0.578	0.568	0.577	0.574
TS_τ	0.040	0.057	0.054	0.044	0.047	0.053	0.049	0.045
MS_τ	0.043	0.057	0.056	0.047	0.051	0.054	0.053	0.045
M_r	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
TM_τ	0.024	0.020	0.016	0.014	0.032	0.036	0.037	0.028
MM_τ	0.041	0.056	0.052	0.040	0.054	0.054	0.058	0.051
TC_τ	0.038	0.040	0.038	0.033	0.044	0.043	0.049	0.035
MC_τ	0.045	0.055	0.052	0.045	0.052	0.055	0.072	0.043
PE_r	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
$U_{\min}$	NA	NA	NA	NA	NA	NA	NA	NA
Model 3								
S_r	0.056	0.057	0.061	0.059	0.051	0.062	0.051	0.059
TS_τ	0.048	0.045	0.047	0.056	0.047	0.049	0.046	0.048
MS_τ	0.049	0.049	0.050	0.057	0.052	0.052	0.049	0.049
M_r	0.091	0.156	0.225	0.360	0.141	0.264	0.493	0.765
TM_τ	0.033	0.020	0.016	0.017	0.034	0.026	0.030	0.030
MM_τ	0.052	0.045	0.053	0.042	0.055	0.043	0.051	0.052
TC_τ	0.044	0.031	0.033	0.034	0.041	0.041	0.042	0.031
MC_τ	0.047	0.046	0.052	0.048	0.053	0.049	0.063	0.041
PE_r	0.387	0.448	0.564	0.731	0.198	0.284	0.427	0.677
$U_{\min}$	NA	0.057	NA	NA	0.046	0.056	0.053	NA

with the increase of n and d. In *Model 2* and *Model 3*, PE_r cannot test the independence of the sample. This fact is attributed to the assumption that the Pearson correlation coefficient has finite second moment conditions. The $U_{\min}$ method overestimates the experience size under the normal distribution and is invalid under the Cauchy distribution. Moreover, it is sometimes invalid under the t-distribution, mainly because the sixth order of the U-statistic is beyond the scope of the calculation.

The empirical powers of the 10 test statistics under the dense alternative are shown in Table 2. Under *Model 4*, the powers of S_r, PE_r, TS_τ, MS_τ, TC_τ and MC_τ increase to 1 as increases. While the test methods based on L_2-type statistics are more powerful than the correspondingly L_∞-type statistics, since the L_2-type statistic is suitable for dense alternative. Under *Model 5*, TS_τ, MS_τ, TC_τ and MC_τ have higher power than M_τ using distribution-based and simulation-based thresholds and the empirical power increases as n increases. In addition, the empirical power of MM_τ is higher than that of TM_τ, because the statistic M_τ converges weakly to the extreme value distribution. Although the empirical power of S_r increases as d increases and the empirical powers of M_r and PE_r are 1, S_r, M_r and PE_r are not apply to the underlying data generated by the Cauchy distribution, mainly

Table 2. Empirical powers of tests in dense cases.

n	50				100			
d	50	100	200	400	50	100	200	400
Model 4								
S_r	0.434	0.918	0.999	1.000	0.178	0.651	0.993	1.000
TS_τ	0.375	0.876	0.998	1.000	0.158	0.574	0.986	1.000
MS_τ	0.362	0.873	0.998	1.000	0.155	0.577	0.986	1.000
M_r	0.015	0.018	0.008	0.003	0.036	0.026	0.021	0.024
TM_τ	0.036	0.044	0.040	0.041	0.040	0.044	0.046	0.053
MM_τ	0.071	0.099	0.120	0.113	0.063	0.069	0.079	0.094
TC_τ	0.380	0.878	0.999	1.000	0.168	0.582	0.988	1.000
MC_τ	0.393	0.894	0.999	1.000	0.192	0.621	0.989	1.000
PE_r	0.510	0.925	0.999	1.000	0.207	0.654	0.994	1.000
$U_{\min}$	0.999	1.000	1.000	1.000	0.993	1.000	1.000	1.000
Model 5								
S_r	0.891	0.927	0.956	0.971	0.866	0.914	0.947	0.972
TS_τ	0.856	0.952	0.990	1.000	0.752	0.887	0.976	0.995
MS_τ	0.853	0.951	0.990	0.999	0.748	0.888	0.977	0.995
M_r	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
TM_τ	0.426	0.514	0.594	0.722	0.308	0.407	0.530	0.684
MM_τ	0.545	0.691	0.823	0.907	0.377	0.492	0.639	0.801
TC_τ	0.888	0.967	0.994	1.000	0.794	0.908	0.987	0.997
MC_τ	0.896	0.973	0.997	1.000	0.819	0.923	0.992	0.998
PE_r	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
$U_{\min}$	NA	NA	NA	NA	NA	NA	NA	NA

Table 3. Empirical powers of tests in sparse cases.

n	50				100			
d	50	100	200	400	50	100	200	400
Model 6								
S_r	0.053	0.060	0.054	0.055	0.082	0.069	0.056	0.054
TS_τ	0.050	0.058	0.051	0.050	0.077	0.062	0.059	0.054
MS_τ	0.048	0.059	0.054	0.048	0.077	0.064	0.061	0.055
M_r	0.201	0.307	0.504	0.757	0.845	0.963	0.999	1.000
TM_τ	0.210	0.329	0.533	0.793	0.786	0.936	0.996	1.000
MM_τ	0.260	0.425	0.645	0.861	0.809	0.944	0.997	1.000
TC_τ	0.182	0.281	0.473	0.746	0.734	0.918	0.993	1.000
MC_τ	0.194	0.323	0.531	0.767	0.754	0.926	0.995	1.000
PE_r	0.492	0.605	0.760	0.926	0.860	0.956	0.997	1.000
$U_{\min}$	0.233	0.289	0.371	0.347	0.763	0.874	0.946	0.976
Model 7								
S_r	0.433	0.435	0.431	0.436	0.577	0.568	0.578	0.575
TS_τ	0.086	0.077	0.057	0.052	0.075	0.057	0.057	0.043
MS_τ	0.081	0.076	0.056	0.049	0.073	0.059	0.062	0.045
M_r	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
TM_τ	0.806	0.869	0.924	0.951	0.646	0.727	0.783	0.836
MM_τ	0.834	0.904	0.952	0.967	0.687	0.760	0.820	0.870
TC_τ	0.755	0.833	0.895	0.933	0.592	0.682	0.755	0.798
MC_τ	0.769	0.853	0.918	0.941	0.615	0.704	0.778	0.812
PE_r	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
$U_{\min}$	NA	NA	NA	NA	NA	NA	NA	NA

due to the fact that the test methods based on the Pearson correlation coefficient are invalid in this case. The $U_{\min}$ method has high empirical power under normal distribution, but it is invalid under Cauchy distribution for the reason that $U_{\min}$ test statistic requires finite second moment of the random vector.

Table 3 reports the empirical power under the sparse alternative. In *Model* 6, the L_2-type test statistics is less powerful, and the empirical power of M_r, PE_r, TM_τ, MM_τ, TC_τ and MC_τ increases as the sample size n or dimension d increases. In *Model 7*, TM_τ, MM_τ, TC_τ and MC_τ have higher empirical power than S_r, TS_τ and MS_τ. The empirical powers of M_r and PE_r are 1 mainly due to the property of the Pearson correlation coefficient. The test method PE_r has high empirical power when the data are generated from a normal distribution, but invalids when the data are generated from a Cauchy distribution. Therefore, this method is not universal. The $U_{\min}$ method has lower empirical power than L_∞-type test, TC_τ test, MC_τ test and PE_r test under normal distribution, and it is invalid under Cauchy distribution. Furthermore, we find that the powers of tests TS_τ, TM_τ and TC_τ based on asymptotic null distributions are only slightly different from the ones using Monte Carlo simulated approximate null distribution.

In general, the proposed test statistic C_τ has good empirical power under either sparse or dense alternatives, no matter whether the critical values are obtained by Monte Carlo simulations or the asymptotic null distributions. Therefore, the test statistic C_τ is adaptive to the underlying data and is a suitable alternative test in practice.

4.2. Power performance under various strengths of dependence

In this section, we evaluate the power performance of the test statistic under different dependence strengths. The empirical powers are calculated based on 1000 replications and the nominal significance level is set to be 5%. Here, we assume that the sample size n and dimension d take values from $\{50, 100\}$ and $\{50, 100, 200, 400\}$, respectively. We set the following two models:

Model 8(a): $\boldsymbol{x}_{1,\cdot}, \ldots, \boldsymbol{x}_{n,\cdot}$ are independently generated from a normal distribution $N_d(0, \Sigma)$ with $\Sigma = (1-\rho)I_d + \rho e_d e_d^T$, $\rho \in \{0.02, 0.04, 0.06, 0.08\}$.
Model 8(b): $\boldsymbol{x}_{1,\cdot}, \ldots, \boldsymbol{x}_{n,\cdot}$ are independently generated from a normal distribution $N_d(0, \Sigma)$ with $\Sigma = (\sigma_{ij})_{d\times d}$, $\sigma_{ii} = 1(i \in \{1, \ldots, d\})$, $\sigma_{12} = \sigma_{21} = \rho$, $\rho \in \{0.6, 0.7, 0.8, 0.9\}$ and $\sigma_{ij} = 0$ for others.

Tables 4 and 5 show that under different dependence intensities, TS_τ and MS_τ perform best in dense cases with empirical powers, but poorly in sparse cases. In contrast, the proposed L_∞-type tests TM_τ and MM_τ perform best in sparse cases but poorly in dense cases. As can be seen from Tables 4 and 5, the empirical power performance of our proposed adaptive test TC_τ is always very close to optimal in all tests, regardless of whether the local alternative is sparse or not. This shows a very attractive property of C_τ, which compromises testing for random variables in both sparse and dense cases.

4.3. Power performance of non-central or non-negative distribution

In this section, we consider random variables generated by non-central and non-negative distributions. Other settings are the same as Section 4.2.

Model 9(a)*:* $(x_{i1}, \ldots, x_{id})$ is generated by $(x_{i1}, \ldots, x_{id})^\top = (y_{i0}, \ldots, y_{i(d-1)})^\top + (\frac{2\ln d}{n})(y_{i1}, \ldots, y_{id})^\top$ where $y_{ij} \overset{i.i.d}{\sim} t(4, 1)$ across j and i.

Table 4. Empirical powers under various strengths of dependence in dense cases.

n	50				100			
d	50	100	200	400	50	100	200	400
$\rho = 0.02$								
TS_τ	0.078	0.185	0.400	0.782	0.168	0.413	0.822	0.993
MS_τ	0.072	0.179	0.399	0.776	0.168	0.420	0.828	0.993
TM_τ	0.033	0.024	0.014	0.018	0.045	0.054	0.045	0.029
MM_τ	0.056	0.062	0.055	0.057	0.073	0.077	0.074	0.057
TC_τ	0.093	0.191	0.404	0.783	0.179	0.430	0.827	0.993
MC_τ	0.101	0.225	0.450	0.815	0.201	0.461	0.856	0.994
$\rho = 0.04$								
TS_τ	0.387	0.764	0.972	0.998	0.809	0.990	1.000	1.000
MS_τ	0.375	0.759	0.972	0.998	0.803	0.990	1.000	1.000
TM_τ	0.043	0.044	0.022	0.029	0.089	0.082	0.087	0.101
MM_τ	0.080	0.108	0.081	0.074	0.126	0.130	0.141	0.160
TC_τ	0.395	0.766	0.974	0.998	0.814	0.991	1.000	1.000
MC_τ	0.415	0.796	0.981	0.998	0.826	0.991	1.000	1.000
$\rho = 0.06$								
TS_τ	0.781	0.981	0.998	1.000	0.993	1.000	1.000	1.000
MS_τ	0.772	0.980	0.998	1.000	0.993	1.000	1.000	1.000
TM_τ	0.061	0.065	0.055	0.048	0.142	0.183	0.172	0.177
MM_τ	0.110	0.128	0.149	0.132	0.211	0.257	0.270	0.279
TC_τ	0.786	0.981	0.998	1.000	0.993	1.000	1.000	1.000
MC_τ	0.799	0.983	0.998	1.000	0.995	1.000	1.000	1.000
$\rho = 0.08$								
TS_τ	0.958	0.998	1.000	1.000	1.000	1.000	1.000	1.000
MS_τ	0.957	0.998	1.000	1.000	1.000	1.000	1.000	1.000
TM_τ	0.125	0.106	0.091	0.074	0.283	0.299	0.356	0.376
MM_τ	0.182	0.201	0.245	0.176	0.379	0.402	0.486	0.527
TC_τ	0.961	0.998	1.000	1.000	1.000	1.000	1.000	1.000
MC_τ	0.964	0.998	1.000	1.000	1.000	1.000	1.000	1.000

Model 9(b): $\boldsymbol{x}_{1,\cdot}, \dots, \boldsymbol{x}_{n,\cdot}$ are independently generated. We first generate $Z_1, \dots, Z_d$ from $t(4,1)$ distribution with degrees of freedom 4 and non-centrality parameter 1, and then construct $\boldsymbol{X} = (X_1, \dots, X_d)^\top$ by setting $X_1 = Z_1 + (\frac{3\ln d}{n})Z_2$, $X_2 = Z_2 + (\frac{3\ln d}{n})Z_1$ and $X_j = Z_j$ for $3 \le j \le d$.
Model 10(a): $(x_{i1}, \dots, x_{id})$ is generated by $(x_{i1}, \dots, x_{id})^\top = (y_{i0}, \dots, y_{i(d-1)})^\top + (\frac{2\ln d}{n})(y_{i1}, \dots, y_{id})^\top$ where $y_{ij} \overset{i.i.d}{\sim} \chi^2_5$ across j and i.
Model 10(b): $\boldsymbol{x}_{1,\cdot}, \dots, \boldsymbol{x}_{n,\cdot}$ are independently generated. We first generate $Z_1, \dots, Z_d$ from χ^2_5 distribution, and then construct $\boldsymbol{X} = (X_1, \dots, X_d)^\top$ by setting $X_1 = Z_1 + (\frac{3\ln d}{n})Z_2$, $X_2 = Z_2 + (\frac{3\ln d}{n})Z_1$ and $X_j = Z_j$ for $3 \le j \le d$.

When the data in *Model* 9 are generated from multivariate non-centralised $t(4,1)$ distribution, Table 6 includes a comparison of the test statistics. The data in *Model 10* are generated from the non-negative chi-square distribution, and the results are also included in Table 6. In *Models* 9(a) and 10(a), it can be seen that L_2-type test statistics, MC_τ, and TC_τ have the ability to test the independence of random variables, while L_∞-type test statistics do not. In contrast, in *Models* 9(b) *and* 10(b), L_∞-type test statistics, TC_τ, and MC_τ have the ability to test the independence of random variables, while L_2-type statistics do not. In addition, as the dimensionality increases, the empirical power of the test statistics has an increasing tending in their corresponding applicable and valid cases.

Table 5. Empirical powers under various strengths of dependence in sparse cases.

n	50				100			
d	50	100	200	400	50	100	200	400
$\rho = 0.6$								
TS_τ	0.056	0.064	0.054	0.042	0.111	0.078	0.051	0.062
MS_τ	0.057	0.062	0.055	0.044	0.108	0.079	0.055	0.059
TM_τ	0.571	0.408	0.271	0.174	0.990	0.973	0.952	0.891
MM_τ	0.636	0.511	0.399	0.274	0.993	0.979	0.957	0.911
TC_τ	0.512	0.363	0.238	0.155	0.984	0.962	0.926	0.866
MC_τ	0.534	0.399	0.287	0.179	0.986	0.965	0.942	0.875
$\rho = 0.7$								
TS_τ	0.085	0.070	0.055	0.045	0.204	0.095	0.055	0.058
MS_τ	0.077	0.070	0.056	0.045	0.203	0.097	0.055	0.062
TM_τ	0.876	0.828	0.698	0.561	1.000	1.000	0.999	0.997
MM_τ	0.902	0.875	0.806	0.651	1.000	1.000	0.999	0.998
TC_τ	0.902	0.875	0.806	0.651	1.000	1.000	0.999	0.998
MC_τ	0.860	0.803	0.690	0.535	1.000	1.000	0.999	0.997
$\rho = 0.8$								
TS_τ	0.129	0.087	0.060	0.045	0.356	0.122	0.062	0.059
MS_τ	0.117	0.080	0.061	0.044	0.354	0.126	0.066	0.061
TM_τ	0.992	0.988	0.973	0.951	1.000	1.000	1.000	1.000
MM_τ	0.995	0.996	0.988	0.969	1.000	1.000	1.000	1.000
TC_τ	0.987	0.983	0.960	0.929	1.000	1.000	1.000	1.000
MC_τ	0.987	0.987	0.974	0.942	1.000	1.000	1.000	1.000
$\rho = 0.9$								
TS_τ	0.197	0.097	0.062	0.050	0.621	0.201	0.082	0.065
MS_τ	0.187	0.096	0.064	0.046	0.616	0.204	0.089	0.065
TM_τ	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
MM_τ	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
TC_τ	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
MC_τ	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

5. Application

5.1. Welding data

The installation of the drive shaft of the automobile requires circular welding of the crown nesting. The input of the automatic welding machine must be controlled within a certain operating range to ensure the quality of welding. To control this process, monitoring engineers observe four key variables: voltage, electric current, feed speed and air flow. Tables 5–9 of Applied Multivariate Statistical Analysis Sixth Edition (Johnson and Wichern 2002) gave data on these variables over a 5-s period. Furthermore, for continuous observations of each variable, there is no clear serial correlation. The corresponding data matrix has a total of $n = 40$ rows and $d = 4$ columns.

To analyse whether the four variables are mutual independent, we can intuitively implement the likelihood ratio test due to the fact that 'big n, small p' for this dataset. Unfortunately, it is difficult for us to determine the joint distribution of the random vector. In addition, from Figure 1, we can conclude that the joint distribution is far away from multivariate normal distribution. More specifically, the distributions of air flow and feed speed are highly skewed and heavy tailed, respectively. Therefore, compared with existing test methods, the proposed method is more suitable this dataset. According to the calculation, we obtain that the p-values of both the test statistics S_τ and C_τ are close to 0. Therefore, we take the logarithm of the p-value of all statistics for comparison with $\log\alpha = -2.996$ ($\alpha = 0.05$). Consequently, the log-p-values are -9.464 for S_τ, -3.198 for M_τ and -9.464

Table 6. Empirical powers of non-central or non-negative distribution.

n	50				100			
d	50	100	200	400	50	100	200	400
Model 9(a)								
TS_τ	0.258	0.364	0.585	0.729	0.122	0.173	0.232	0.354
MS_τ	0.238	0.354	0.583	0.716	0.122	0.175	0.233	0.344
TM_τ	0.061	0.053	0.062	0.047	0.047	0.056	0.045	0.045
MM_τ	0.094	0.107	0.158	0.135	0.073	0.080	0.075	0.074
TC_τ	0.275	0.377	0.598	0.731	0.139	0.193	0.256	0.369
MC_τ	0.293	0.433	0.648	0.783	0.159	0.226	0.309	0.399
Model 9(b)								
TS_τ	0.058	0.042	0.043	0.054	0.053	0.051	0.059	0.052
MS_τ	0.057	0.046	0.044	0.057	0.057	0.052	0.061	0.055
TM_τ	0.181	0.272	0.378	0.479	0.083	0.093	0.085	0.094
MM_τ	0.235	0.348	0.499	0.593	0.108	0.122	0.123	0.132
TC_τ	0.163	0.212	0.328	0.433	0.091	0.080	0.080	0.092
MC_τ	0.175	0.246	0.384	0.466	0.102	0.099	0.107	0.110
Model 10(a)								
TS_τ	0.242	0.380	0.511	0.708	0.113	0.166	0.220	0.344
MS_τ	0.230	0.364	0.505	0.690	0.112	0.172	0.224	0.327
TM_τ	0.056	0.064	0.041	0.053	0.054	0.046	0.053	0.048
MM_τ	0.091	0.118	0.123	0.123	0.081	0.077	0.084	0.087
TC_τ	0.259	0.395	0.519	0.714	0.130	0.186	0.243	0.360
MC_τ	0.274	0.445	0.597	0.761	0.154	0.220	0.310	0.385
Model 10(b)								
TS_τ	0.057	0.058	0.043	0.072	0.055	0.060	0.043	0.057
MS_τ	0.057	0.060	0.044	0.070	0.062	0.061	0.048	0.057
TM_τ	0.224	0.282	0.338	0.434	0.071	0.073	0.106	0.110
MM_τ	0.274	0.367	0.458	0.566	0.096	0.096	0.142	0.144
TC_τ	0.179	0.246	0.277	0.376	0.083	0.081	0.100	0.110
MC_τ	0.193	0.279	0.348	0.416	0.094	0.091	0.128	0.120

for C_τ. In contrast, the Schott (2005) test results in a higher log-p-value of -0.1445. The Schott method is the only test statistic that does not reject the null hypothesis. The reason for this difference is that the Schott (2005) test method is based on the normal distribution specification. This implies that in the welding data, the samples of these four variables, voltage, electric current, feed speed and air flow, are generated in non-normal distributions. To better illustrate this guess, we plot the frequency histograms of the four variables in Figure 1. From Figure 1, (a) and (b) are close to the histogram of the normal distribution. While (c) and (d) are far from the histogram of the normal distribution.

5.2. Biochemical data

In this section, we apply the proposed test statistic C_τ to biochemical data sets to illustrate the usefulness of the resulting test in practice. Beerstecher et al. (1950) collected biochemical data that initially explored various characteristics of alcoholic and non-alcoholic individuals. These data are made up of 62 measurements taken on each of 12 people, 8 of whom served as controls and the remaining 4 as alcoholics. Although the results from this small sample size may not be definitive, experts assure us that they can at least provide strong information. We will restrict attention to a subset of the 62 variables, a set of 8 serum measures. For each group, the control group and the alcohol group, we will test

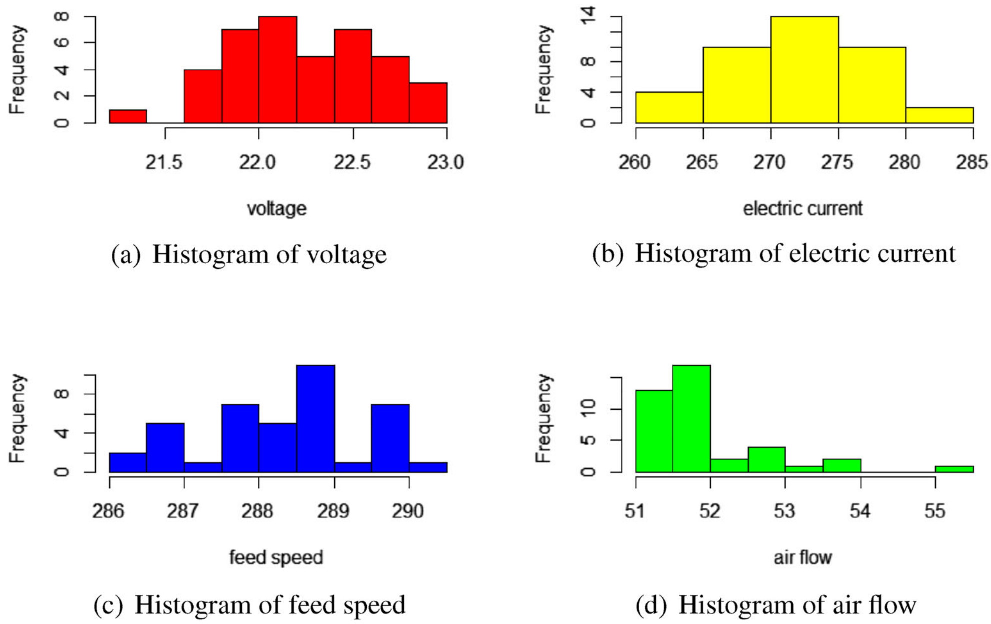

Figure 1. Histogram of four variables in welding data: (a) Histogram of voltage, (b) Histogram of electric current, (c) Histogram of feed speed and (d) Histogram of air flow.

completely independent hypotheses. The control group used $d = 8$ and $n = 8$, while the alcohol group used $n = 4$.

We discover that $C_\tau = 0.008$ for the control group corresponds to a p-value of 0.016. Therefore, there is significant evidence that these eight variables are related in some way. We obtain $C_\tau = 0.042$ with a p-value of 0.081 for the alcohol group. As a result, for the alcohol group, the evidence of a relationship between variables was not nearly as robust.

6. Conclusion

In this paper, we propose the adaptive test statistic C_τ based on Kendall's tau for the high dimensional independence test, which is constructed by combining the L_2-type statistic S_τ with the L_∞-type statistic M_τ. We first develop the asymptotic independence of the L_2-type and L_∞-type statistics based on Kendall's tau under mild conditions so that it can be conveniently implemented in practice. Based on this result, the asymptotic theory of the proposed statistic under the null hypothesis is obtained by the continuous mapping theorem. We propose a simple distribution-free independence test that rejects null when the adaptive statistic is less than $1 - \sqrt{1-\alpha}$. Simulation studies and the real data analysis show that the new test statistic can be adapted to the underlying data to test complete independence.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (Nos. 11971045 and 12271014), the Young Talent program of Beijing Municipal Education Commission (No.

CIT&TCD201904021), the Science and Technology Project of Beijing Municipal Education Commission (No. KM202210005012), and Beijing Postdoctoral Research Foundation, China.

Disclosure statement

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

References

Beerstecher, J.E., Sutton, H.E., Berry, H.K., Brown, W.D., Reed, J., Rich, G.B., Berry, L.J., and Williams, R.J. (1950), 'Biochemical Individuality. V. Explorations with Respect to the Metabolic Patterns of Compulsive Drinkers', *Archives of Biochemistry*, 29, 27–40.

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.

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

Callaert, H., and Veraverbeke, N. (1981), 'The Order of the Normal Approximation for A Studentized U-statistic', *The Annals of Statistics*, 9(1), 194–200.

Cardoso, J.F., and Souloumiac, A. (1993), 'Blind Beamforming for Non-Gaussian Signals', *IEE Proceedings F*, 140(6), 362–370.

Chow, Y.S., and Teicher, H. (1997), *Probability Theory: Independence, Interchangeability, Martingales*, Springer.

Davidson, J. (2021), *Stochastic Limit Theory: An Introduction for Econometricians*, OUP Oxford.

Dhaene, J., Denuit, M., Goovaerts, M.J., Kaas, R., and Vyncke, D. (2002a), 'The Concept of Comonotonicity in Actuarial Science and Finance: Applications', *Insurance: Mathematics and Economics*, 31(2), 133–161.

Dhaene, J., Denuit, M., Goovaerts, M.J., Kaas, R., and Vyncke, D. (2002b), 'The Concept of Comonotonicity in Actuarial Science and Finance: Theory', *Insurance: Mathematics and Economics*, 31(1), 3–33.

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.

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

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 Non-parametric Test of Independence', *The Annals of Mathematical Statistics*, 19(4), 546–557.

Hotelling, H., and Pabst, M.R. (1936), 'Rank Correlation and Tests of Significance Involving No Assumption of Normality', *The Annals of Mathematical Statistics*, 7(1), 29–43.

Johnson, R.A., and Wichern, D.W. (2002), *Applied Multivariate Statistical Analysis*, Prentice Hall India Learning Private Limited.

Kendall, M.G. (1938), 'A New Measure of Rank Correlation', *Biometrika*, 30(1–2), 81–93.

Leung, D., and Drton, M. (2018), 'Testing Independence in High Dimensions with Sums of Rank Correlations', *The Annals of Statistics*, 46(1), 280–307.

Mao, G. (2017), '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 & Data Analysis*, 117, 128–137.

McLeish, D.L. (1974), 'Dependent Central Limit Theorems and Invariance Principles', *The Annals of Probability*, 2(4), 620–628.

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.
Rosenblatt, M. (1975), 'A Quadratic Measure of Deviation of Two-dimensional Density Estimates and A Test of Independence', *The Annals of Statistics*, 3(1), 1–14.
Schott, J.R. (2005), 'Testing for Complete Independence in High Dimensions', *Biometrika*, 92(4), 951–956.
Shen, C., Priebe, C.E., and Vogelstein, J.T. (2020), 'From Distance Correlation to Multiscale Graph Correlation', *Journal of the American Statistical Association*, 115(529), 280–291.
Sidak, Z., Sen, P.K., and Hajek, J. (1999), *Theory of Rank Tests*, Elsevier.
Székely, G.J., and Rizzo, M.L. (2009), 'Brownian Distance Covariance', *The Annals of Applied Statistics*, 3(4), 1236–1265.
Székely, G.J., Rizzo, M.L., and Bakirov, N.K. (2007), 'Measuring and Testing Dependence by Correlation of Distances', *The Annals of Statistics*, 35(6), 2769–2794.
Van der Vaart, A.W. (1998), *Asymptotic Statistics*, Cambridge Series in Statistical and Probabilistic Mathematics (Vol. 3), UK: Cambridge University Press.
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.
Zhou, W. (2007), 'Asymptotic Distribution of the Largest Off-diagonal Entry of Correlation Matrices', *Transactions of the American Mathematical Society*, 359(11), 5345–5363.

Appendix 1.

In this section, a few essential technical lemmas are collected without proof. Throughout this paper, C represents a generic constant, whose values may vary at different locations.

Lemma A.1 (Lemma 2 of Callaert and Veraverbeke 1981**):** *Let*

$$U_n = \binom{n}{2}^{-1} \sum_{i<j} h\left(X_i, X_j\right)$$

be a U-statistic with $\mathrm{E}h(X_1, X_2) = \varpi$ *and* $\nu_r = \mathrm{E}|h(X_1, X_2) - \varpi|^r < \infty$ *for* $r \geqslant 1$. *Then* $\mathrm{E}|U_N - \varpi|^r \leqslant C\nu_r n^{-s}$ *where* C *is a constant,* $s = r - 1$ *for* $1 \leqslant r \leqslant 2$, *and* $s = r/2$ *for* $r \geqslant 2$.

Lemma A.2 (Lemma C4 of Han et al. 2017**):** *Let* $d > 2$. *Under the null hypothesis* H_0, $\{\tau_{k\ell}, 1 \leq \ell < k \leq d\}$ *are identically and pairwise independently distributed.*

Lemma A.3 (Loève's c_r inequality, Davidson 2021**):** *For* $r > 0$,

$$\mathrm{E}\left|\sum_{i=1}^{m} X_i\right|^r \leq c_r \sum_{i=1}^{m} \mathrm{E}\left|X_i\right|^r$$

where $c_r = 1$ *when* $r \leq 1$ *and* $c_r = m^{r-1}$ *when* $r \geq 1$.

Lemma A.4 (Marcinkiewicz–Zygmund inequality, Chow and Teicher 1997**):** *Let* $m \geq 1$ *and* $\{\zeta_i; 1 \leq i \leq m\}$ *be independent random variables with* $\mathrm{E}\zeta_i = 0$ *for each* i *and* $\sup_{1\leq i\leq m} \mathrm{E}(|\zeta_i|^r) < \infty$ *for some* $r \geq 2$. *Then there exists a constant* $K_r > 0$ *depending on* r *only such that*

$$\begin{aligned}\mathrm{E}\left(|\zeta_1 + \cdots + \zeta_m|^r\right) &\leq K_r \cdot \mathrm{E}\left\{\left(\zeta_1^2 + \cdots + \zeta_m^2\right)^{r/2}\right\} \\ &\leq K_r \cdot m^{(r/2)-1}\left(\mathrm{E}\left|\zeta_1\right|^r + \cdots + \mathrm{E}\left|\zeta_m\right|^r\right).\end{aligned}$$

$Y = (Y_1, \ldots, Y_k)$ denotes a random vector in $\mathbb{R}^k$. A sequence of random vectors Y_n converges in distribution to a random vector Y, which is denoted by $Y_n \xrightarrow{d} Y$. Convergence in probability is denoted by $Y_n \xrightarrow{P} Y$. Convergence in almost surely is denoted by $Y_n \xrightarrow{\text{as}} Y$.

Lemma A.5 (Continuous mapping, Van der Vaart 1998**):** *Let $h : \mathbb{R}^k \mapsto \mathbb{R}^m$ be continuous at every point of a set A such that $P(Y \in A) = 1$.*

(i) *If $Y_n \xrightarrow{d} Y$, then $h(Y_n) \xrightarrow{d} h(Y)$;*
(ii) *If $Y_n \xrightarrow{P} Y$, then $h(X_n) \xrightarrow{P} h(Y)$;*
(iii) *If $Y_n \xrightarrow{\text{as}} Y$, then $h(Y_n) \xrightarrow{\text{as}} h(Y)$.*

Appendix 2.

We prove the main results in this section.

Proof of Theorem 2.1: Define $\omega_4 = \mathrm{E}[(\tau_{k\ell}^2 - \mathrm{E}\tau_{k\ell}^2)^4]$ and $Z_{nk} = \omega_2^{-1/2} \sum_{\ell=1}^{k-1} \gamma_{k\ell}$, where $\gamma_{k\ell} = \tau_{k\ell}^2 - \mathrm{E}\tau_{k\ell}^2$. Let $\{Z_{nk}, \mathcal{F}_{nk}, 2 \le k \le d, n \ge 2\}$ be the σ-field generated by $\{\boldsymbol{R}_1, \ldots, \boldsymbol{R}_k\}$. Set $\mathrm{E}_k = \mathrm{E}(\cdot \mid \mathcal{F}_{nk})$. By using the martingale method of McLeish (1974) to show the Berry–Esseen bound for the general case, we know

$$\sup_x |\Pr(S_\tau \le x) - \Phi(x)| \le O(\vartheta^{1/5}),$$

where $\vartheta = \mathrm{E}\{\sum_{k=2}^d \mathrm{E}_{k-1}(Z_{nk}^2) - 1\}^2 + \sum_{k=2}^d \mathrm{E}(Z_{nk}^4)$.

Now it only remains to bound by the following two terms:

$$\mathrm{E}\left\{\sum_{k=2}^d \mathrm{E}_{k-1}(Z_{nk}^2) - 1\right\}^2 \quad \text{and} \quad \sum_{k=2}^d \mathrm{E}(Z_{nk}^4). \tag{A1}$$

First, we restrict the first term of (A1). It is not difficult to find that

$$\left\{\sum_{k=2}^d \mathrm{E}_{k-1}(Z_{nk}^2) - 1\right\}^2 = \left\{\sum_{k=2}^d \mathrm{E}_{k-1}(Z_{nk}^2)\right\}^2 - 2\sum_{k=2}^d \mathrm{E}_{k-1}(Z_{nk}^2) + 1.$$

According to Lemma 1 of Mao (2018), we may obtain the conditional expectation

$$\sum_{k=2}^d \mathrm{E}_{k-1}(Z_{nk}^2) = \sum_{k=2}^d \mathrm{E}(Z_{nk}^2 \mid \boldsymbol{R}_\ell) = \sum_{k=2}^d \mathrm{E}(Z_{nk}^2), \quad \ell \in \{1, \ldots, k-1\}.$$

By proof of Lemma 2 of Mao (2018), one has

$$\left\{\sum_{k=2}^d \mathrm{E}_{k-1}(Z_{nk}^2) - 1\right\}^2 = 0.$$

Therefore,

$$E\left\{\sum_{k=2}^d \mathrm{E}_{k-1}(Z_{nk}^2) - 1\right\}^2 = 0. \tag{A2}$$

Next, we restrict the second term of (A1). According to proof of Lemma 2 of Mao (2018), we have

$$\sum_{k=2}^d \mathrm{E}\left(Z_{nk}^4\right) = \frac{4(d-2)}{d(d-1)} + \frac{2\omega_4}{d(d-1)\omega_2^2}. \tag{A3}$$

By (A2) and (A3), we obtain that

$$\sup_x |\Pr(S_\tau \le x) - \Phi(x)| \le O(d^{-1/5}).$$

Thus this completes the proof. ■

Proof of Lemma 3.1: (1) Define $h(R_{k,\{1,2\}}, R_{\ell,\{1,2\}}) = \mathrm{sign}(R_{k1} - R_{k2})\mathrm{sign}(R_{\ell 1} - R_{\ell 2})$, $1 \le \ell < k \le d$. By simple computation, $Eh(R_{k,\{1,2\}}, R_{\ell,\{1,2\}}) = 0$. Next,

$$\begin{aligned}
\nu_r &= \mathrm{E}\{h(R_{k,\{1,2\}}, R_{\ell,\{1,2\}}) - \mathrm{E}h(R_{k,\{1,2\}}, R_{\ell,\{1,2\}})\}^{2r} \\
&= \mathrm{E}\{h(R_{k,\{1,2\}}, R_{\ell,\{1,2\}})\}^{2r} \\
&= \sum_{k=1}^{n}\sum_{\ell=1}^{n}\sum_{i=1}^{n}\sum_{j=1}^{n}\{\mathrm{sign}(k-\ell)\mathrm{sign}(i-j)\}^{2r}\Pr(R_{k1}=k)\Pr(R_{k2}=\ell) \\
&\quad \cdot \Pr(R_{\ell 1}=i)\Pr(R_{\ell 2}=j) \\
&= \frac{1}{n^4}\sum_{k=1}^{n}\sum_{\ell=1}^{n}\sum_{i=1}^{n}\sum_{j=1}^{n}\{\mathrm{sign}(k-\ell)\mathrm{sign}(i-j)\}^{2r} \\
&\le 1.
\end{aligned}$$

Combining Lemma A.1 with Jensen's inequality, we have

$$\begin{aligned}
&\mathrm{E}(|\tau_{k\ell}^2 - \mathrm{E}\tau_{k\ell}^2|^r) \\
&\quad \le \mathrm{E}\left\{\left(\tau_{k\ell}^2 + \mathrm{E}\tau_{k\ell}^2\right)^r\right\} \\
&\quad \le \mathrm{E}\left[\left\{2\max(\tau_{k\ell}^2, \mathrm{E}\tau_{k\ell}^2)\right\}^r\right] \\
&\quad \le 2^r\mathrm{E}\left[\max\left\{\tau_{k\ell}^{2r}, (\mathrm{E}\tau_{k\ell}^2)^r\right\}\right] \\
&\quad \le 2^r\left\{\mathrm{E}(\tau_{k\ell}^{2r}) + (\mathrm{E}\tau_{k\ell}^2)^r\right\} \\
&\quad \le 2^{r+1}\mathrm{E}(\tau_{k\ell}^{2r}) \\
&\quad \le C\nu_r n^{-r} \\
&\quad \le Cn^{-r}.
\end{aligned}$$

(2) Invoking Lemma A.2 and the Marcinkiewicz–Zygmund inequality, i.e. Lemma A.4, one has

$$\begin{aligned}
&\mathrm{E}\left\{\left|\sum_{k=\ell+1}^{d}(\tau_{k\ell}^2 - \mathrm{E}\tau_{k\ell}^2)\right|^r\right\} \\
&\quad \le C(d-\ell)^{(r/2)-1}\sum_{k=\ell+1}^{d}\mathrm{E}|\tau_{k\ell}^2 - \mathrm{E}\tau_{k\ell}^2|^r \\
&\quad \le C(d-\ell)^{(r/2)-1}(d-\ell)\mathrm{E}|\tau_{k\ell}^2 - \mathrm{E}\tau_{k\ell}^2|^r \\
&\quad = C(d-\ell)^{r/2}n^{-r}.
\end{aligned}$$

(3) Similar to (2), we have

$$\mathrm{E}\left\{\left|\sum_{\ell=1}^{k-1}(\tau_{k\ell}^2 - \mathrm{E}\tau_{k\ell}^2)\right|^r\right\}$$

$$\leq C(k-1)^{(r/2)-1}\sum_{\ell=1}^{k-1}\mathrm{E}|\tau_{k\ell}^2-\mathrm{E}\tau_{k\ell}^2|^r$$
$$\leq C(k-1)^{(r/2)-1}(k-1)\mathrm{E}|\tau_{k\ell}^2-\mathrm{E}\tau_{k\ell}^2|^r$$
$$= C(k-1)^{r/2}n^{-r}.$$

This completes the proof of Lemma 3.1. ■

Proof of Lemma 3.2: With Equation (C9) in Han et al. (2017), it is natural to obtain $\frac{d^2}{2}\Pr(|\tau_{21}| > \ell_d) \to F(y)$ as $d \to \infty$. By the independence of $\tau_{k\ell}$ from Lemma A.2, we have

$$G(d,q) = \sum_{I_1<\cdots<I_q\in\Lambda_d}\Pr(Z_{I_1}\cdots Z_{I_q})$$
$$= \sum_{I_1<\cdots<I_q\in\Lambda_d}\prod_{t=1}^{q}\Pr(Z_{I_t})$$
$$\lesssim \binom{d}{q}\{d^{-2}F(y)\}^q$$
$$\lesssim \frac{1}{q!}F^q(y),$$

where $\lesssim$ denotes smaller than up to a constant. Hence,

$$\lim_{d\to\infty} G(d,q) \to \frac{1}{q!}F^q(y) < \infty. \tag{A4}$$

This completes the proof of (5).

Next, for $I_1 < \cdots < I_q \in \Lambda_d$, write $I_u = (\ell_u, k_u)$ for $u \in \{1,\dots,q\}$. Let $\Lambda_{d,q} = \{(\ell_u,k); \ell_u < k \leq d, 1 \leq u \leq q\}\bigcup\{(\ell,k_u); 1 \leq \ell < k_u, 1 \leq u \leq q\}$ for $q \geq 1$. Since $|\Lambda_{d,q}| = \sum_{u=1}^{q}(d-\ell_u+k_u-2)$ and $\ell_u < k_u$ for each u, these results imply that

$$q(d-1) \leq |\Lambda_{d,q}| \leq \sum_{u=1}^{q}(d+k_u) \leq 2qd.$$

Define

$$s_\tau = \sum_{k=2}^{d}\sum_{\ell=1}^{k-1}\tau_{k\ell}^2,$$

for $d \geq 3$ and

$$s_{\tau,q} = \sum_{(\ell,k)\in\Lambda_{d,q}}\tau_{k\ell}^2,$$

for $d \geq q \geq 1$. Note that $Z_{I_1}\cdots Z_{I_q}$ is an event generated by random vectors $\{\tau_{k\ell}; (\ell,k) \in \Lambda_{d,q}\}$. An interesting observation is that $s_\tau - s_{\tau,q}$ is independent of $Z_{I_1}\cdots Z_{I_q}$. Thus

$$s_{\tau,q} = \sum_{u=1}^{q}\sum_{k=\ell_u+1}^{d}\tau_{k\ell_u}^2 + \sum_{u=1}^{q}\sum_{\ell=1}^{k_u-1}\tau_{k_u\ell}^2 - \sum_{t=1}^{q}\sum_{u=1}^{q}\tau_{k_u\ell_t}^2$$
$$\equiv B_{d,1} + B_{d,2} - B_{d,3}.$$

For any integer $r \geq 2$, using the c_r inequality, i.e. Lemma A.3, and (2) in Lemma 3.1, we obtain that

$$\mathrm{E}(|B_{d,1}-\mathrm{E}B_{d,1}|^r) = \mathrm{E}\left\{\left|\sum_{u=1}^{q}\sum_{k=\ell_u+1}^{d}(\tau_{k\ell_u}^2-\mathrm{E}\tau_{k\ell_u}^2)\right|^r\right\}$$

$$\le q^{r-1}\sum_{u=1}^{q}\mathrm{E}\left\{\left|\sum_{k=\ell_u+1}^{d}(\tau_{k\ell_u}^2-\mathrm{E}\tau_{k\ell_u}^2)\right|^r\right\}$$
$$\le q^{r-1}qC(d-\ell)^{r/2}n^{-r}$$
$$\le Cq^r d^{r/2}n^{-r}.$$

Similarly, using the c_r inequality and (3) in Lemma 3.1, we can show that

$$\mathrm{E}(|B_{d,2}-\mathrm{E}B_{d,2}|^r)=\mathrm{E}\left\{\left|\sum_{u=1}^{q}\sum_{\ell=1}^{k_u-1}(\tau_{k_u\ell}^2-\mathrm{E}\tau_{k_u\ell}^2)\right|^r\right\}$$
$$\le q^{r-1}\sum_{u=1}^{q}\mathrm{E}\left\{\left|\sum_{\ell=1}^{k_u-1}(\tau_{k_u\ell}^2-\mathrm{E}\tau_{k_u\ell}^2)\right|^r\right\}$$
$$\le q^{r-1}qC(k-1)^{r/2}n^{-r}$$
$$\le Cq^r d^{r/2}n^{-r}.$$

Lastly, combining (1) of Lemma 3.1 with the c_r inequality, we have

$$\mathrm{E}(|B_{d,3}-\mathrm{E}B_{d,3}|^r)=\mathrm{E}\left\{\left|\sum_{t=1}^{q}\sum_{u=1}^{q}(\tau_{k_u\ell_t}^2-\mathrm{E}\tau_{k_u\ell_t}^2)\right|^r\right\}$$
$$\le q^{2(r-1)}\sum_{t=1}^{q}\sum_{u=1}^{q}\mathrm{E}(|\tau_{k_u\ell_t}^2-\mathrm{E}\tau_{k_u\ell_t}^2|^r)$$
$$\le Cq^{2r}n^{-r}.$$

Therefore, with the above discussions, we obtain that

$$\begin{aligned}&\mathrm{E}|s_{\tau,q}-\mathrm{E}s_{\tau,q}|^r\\&\quad=\mathrm{E}|(B_{d,1}+B_{d,2}-B_{d,3})-\mathrm{E}(B_{d,1}+B_{d,2}-B_{d,3})|^r\\&\quad\le\mathrm{E}(|B_{d,1}-\mathrm{E}B_{d,1}|^r)+\mathrm{E}(|B_{d,2}-\mathrm{E}B_{d,2}|^r)+\mathrm{E}(|B_{d,3}-\mathrm{E}B_{d,3}|^r)\\&\quad\le Cq^r d^{r/2}n^{-r}+Cq^r d^{r/2}n^{-r}+Cq^{2r}n^{-r}\\&\quad\le C(q^r n^{-r}d^{r/2}+q^{2r}n^{-r}).\end{aligned}$$

For fixed $\epsilon\in(0,1)$, by the Markov inequality, one has

$$\begin{aligned}\Pr\left(\omega_2^{-1/2}|s_{\tau,q}-\mathrm{E}s_{\tau,q}|\ge\epsilon\right)&\le\frac{\mathrm{E}[|s_{\tau,q}-\mathrm{E}s_{\tau,q}|^r]}{(\omega_2^{1/2}\epsilon)^r}\\&\le\frac{C(q^r n^{-r}d^{r/2}+q^{2r}n^{-r})n^r}{d^r\epsilon^r}\\&=C\frac{q^r d^{r/2}+q^{2r}}{d^r\epsilon^r}\\&\le C'\frac{q^r}{d^{r/2}},\end{aligned}\tag{A5}$$

for all $d\ge q^2$, where C' is a constant depending on ϵ but free of dimension d, sample size n or indices $\{I_1,\ldots,I_q\}$. Fix $I_1<\cdots<I_q\in\Lambda_d$. By (A5) and the definition of $T_d(x)$,

$$\Pr(T_d(x)Z_{I_1}\cdots Z_{I_q})$$

$$
\begin{aligned}
&\le \Pr\left(T_d(x)Z_{I_1}\cdots Z_{I_q}, \omega_2^{-1/2}|s_{\tau,q} - \mathrm{E}s_{\tau,q}| < \epsilon\right) + C'\frac{q^r}{d^{r/2}}\\
&\le \Pr\left(\omega_2^{-1/2}\{(s_\tau - s_{\tau,q}) - \mathrm{E}(s_\tau - s_{\tau,q})\} \le x+\epsilon, Z_{I_1}\cdots Z_{I_q}\right) + C'\frac{q^r}{d^{r/2}}\\
&\le \Pr\left(\omega_2^{-1/2}\{(s_\tau - s_{\tau,q}) - \mathrm{E}(s_\tau - s_{\tau,q})\} \le x+\epsilon\right)\Pr(Z_{I_1}\cdots Z_{I_q}) + C'\frac{q^r}{d^{r/2}},
\end{aligned}
$$

where the last inequality holds owing to the fact that $s_\tau - s_{\tau,q}$ and $Z_{I_1}\cdots Z_{I_q}$ are independent. Furthermore, one has

$$
\begin{aligned}
&\Pr\left(\omega_2^{-1/2}\{(s_\tau - s_{\tau,q}) - \mathrm{E}(s_\tau - s_{\tau,q})\} \le x+\epsilon\right)\\
&\quad\le \Pr\left(\omega_2^{-1/2}\{(s_\tau - s_{\tau,q}) - \mathrm{E}(s_\tau - s_{\tau,q})\} \le x+\epsilon, \omega_2^{-1/2}|s_{\tau,q} - \mathrm{E}s_{\tau,q}| < \epsilon\right) + C'\frac{q^r}{d^{r/2}}\\
&\quad\le \Pr\left(\omega_2^{-1/2}(s_\tau - \mathrm{E}s_\tau) \le x+2\epsilon\right) + C'\frac{q^r}{d^{r/2}}\\
&\quad\le \Pr\left(T_d(x+2\epsilon)\right) + C'\frac{q^r}{d^{r/2}}.
\end{aligned}
$$

By the above two inequalities, we obtain that

$$
\Pr(T_d(x)Z_{I_1}\cdots Z_{I_q}) \le \Pr\left(T_d(x+2\epsilon)\right)\Pr(Z_{I_1}\cdots Z_{I_q}) + 2C'\frac{q^r}{d^{r/2}}. \tag{A6}
$$

Similarly,

$$
\begin{aligned}
&\Pr\left(\omega_2^{-1/2}\{(s_\tau - s_{\tau,q}) - \mathrm{E}(s_\tau - s_{\tau,q})\} \le x-\epsilon, Z_{I_1}\cdots Z_{I_q}\right)\\
&\quad\le \Pr\left(\omega_2^{-1/2}\{(s_\tau - s_{\tau,q}) - \mathrm{E}(s_\tau - s_{\tau,q})\} \le x-\epsilon, Z_{I_1}\cdots Z_{I_q}, \omega_2^{-1/2}|s_{\tau,q} - \mathrm{E}s_{\tau,q}| < \epsilon\right) + C'\frac{q^r}{d^{r/2}}\\
&\quad\le \Pr\left(T_d(x)Z_{I_1}\cdots Z_{I_q}\right) + C'\frac{q^r}{d^{r/2}}.
\end{aligned}
$$

Due to independence, the above inequality can be written as

$$
\Pr(T_d(x)Z_{I_1}\cdots Z_{I_q}) \ge \Pr\left(\omega_2^{-1/2}\{(s_\tau - s_{\tau,q}) - \mathrm{E}(s_\tau - s_{\tau,q})\} \le x-\epsilon\right)\Pr(Z_{I_1}\cdots Z_{I_q}) - C'\frac{q^r}{d^{r/2}}.
$$

Moreover,

$$
\begin{aligned}
&\Pr\left(\omega_2^{-1/2}(s_\tau - \mathrm{E}s_\tau) \le x-2\epsilon\right)\\
&\quad\le \Pr\left(\omega_2^{-1/2}(s_\tau - \mathrm{E}s_\tau) \le x-2\epsilon, \omega_2^{-1/2}(s_{\tau,q} - \mathrm{E}s_{\tau,q}) < \epsilon\right) + C'\frac{q^r}{d^{r/2}}\\
&\quad\le \Pr\left(\omega_2^{-1/2}\{(s_\tau - s_{\tau,q}) - \mathrm{E}(s_\tau - s_{\tau,q})\} \le x-\epsilon\right) + C'\frac{q^r}{d^{r/2}}.
\end{aligned}
$$

Now observe from the two inequalities above, we have

$$
\begin{aligned}
&\Pr(T_d(x)Z_{I_1}\cdots Z_{I_q})\\
&\quad\ge \Pr\left(\omega_2^{-1/2}(s_\tau - \mathrm{E}s_\tau) \le x-2\epsilon\right)\Pr(Z_{I_1}\cdots Z_{I_q}) - 2C'\frac{q^r}{d^{r/2}}\\
&\quad= \Pr\left(T_d(x-2\epsilon)\right)\Pr(Z_{I_1}\cdots Z_{I_q}) - 2C'\frac{q^r}{d^{r/2}}.
\end{aligned}
$$

Combining this inequality with (A6), we can obtain that

$$|\Pr(T_d(x)Z_{I_1}\cdots Z_{I_q}) - \Pr(T_d(x))\Pr(Z_{I_1}\cdots Z_{I_q})| \le \Delta_{d,\epsilon}P(Z_{I_1}\cdots Z_{I_q}) + 4C'\frac{q^r}{d^{r/2}},$$

where

$$\Delta_{d,\epsilon} := |\Pr(T_d(x)) - \Pr(T_d(x+2\epsilon))| + |\Pr(T_d(x)) - \Pr(T_d(x-2\epsilon))|.$$

Hence,

$$\begin{aligned}
\xi(d,q) &:= \sum_{I_1<\cdots<I_q\in\Lambda_d}\{\Pr(T_d(x)Z_{I_1}\cdots Z_{I_q}) - \Pr(T_d(x))\Pr(Z_{I_1}\cdots Z_{I_q})\}\\
&\le \sum_{I_1<\cdots<I_q\in\Lambda_d}\left[\Delta_{d,\epsilon}\Pr(Z_{I_1}\cdots Z_{I_q}) + 4C'\frac{q^r}{d^{r/2}}\right]\\
&\le \Delta_{d,\epsilon}G(d,q) + (4C')\binom{\frac{1}{2}d(d-1)}{q}\frac{q^r}{d^{r/2}}.
\end{aligned}$$

Choosing $r = 6q$, and using the straightforward fact $\binom{a}{b} \le b^a$ for any integers $1 \le a \le b$, we obtain that

$$\binom{\frac{1}{2}d(d-1)}{q}\frac{q^r}{d^{r/2}} \le d^{2q}\frac{q^r}{d^{r/2}} \le \frac{q^r}{d^q}.$$

By Lemma 2.1, we have $\Pr(T_d(x)) \xrightarrow{d} \Phi(x)$ as $d \to \infty$. Hence,

$$\Delta_{d,\epsilon} \to |\Phi(x+2\epsilon) - \Phi(x)| + |\Phi(x-2\epsilon) - \Phi(x)|.$$

From (A4), we know $\lim_{d\to\infty} G(d,q) \le \frac{C}{q!}$, where C is a universal constant. Hence, for any $\epsilon > 0$, we conclude

$$\begin{aligned}
\limsup_{d\to\infty}\xi(d,q) &\le \frac{C}{q!}\limsup_{d\to\infty}\Delta_{d,\epsilon}\\
&= \frac{C}{q!}\{|\Phi(x+2\epsilon) - \Phi(x)| + |\Phi(x-2\epsilon) - \Phi(x)|\}\\
&= \frac{C}{q!}\{\phi(x_1)2\epsilon + \phi(x_2)2\epsilon\}\\
&\le \frac{C}{q!}\phi(0)\epsilon,
\end{aligned}$$

where $x_1 \in (x, x+2\epsilon)$, $x_2 \in (x-2\epsilon, x)$, and $\phi(\cdot)$ denotes the probability density function of the standard normal distribution. The second equation holds by Lagrange mean value theorem. By sending $\epsilon \to 0$, we prove $\limsup_{d\to\infty}\xi(d,q) \to 0$. ■

Proof of Theorem 3.1: To show asymptotic independence, it suffices to prove that

$$\lim_{d\to\infty}\Pr\left(S_\tau \le x, \max_{1\le\ell<k\le d}|\tau_{k\ell}| > \ell_d\right) = \Phi(x)\{1 - F(y)\}, \tag{A7}$$

for any $x \in \mathbb{R}$ and $y \in \mathbb{R}$. Review notations Λ_d, T_d and Z_I for any $I = (\ell, k) \in \Lambda_d$. Write

$$\Pr\left(S_\tau \le x, \max_{1\le\ell<k\le d}|\tau_{k\ell}| > \ell_d\right) = \Pr\left(\bigcup_{I\in\Lambda_d} T_dZ_I\right), \tag{A8}$$

where the notation T_dZ_I indicates $T_d \cap Z_I$. From the inclusion–exclusion principle,

$$\Pr\left(\bigcup_{I\in\Lambda_d} T_dZ_I\right) \le \sum_{I_1\in\Lambda_d}\Pr(T_dZ_{I_1}) - \sum_{I_1<I_2\in\Lambda_d}\Pr(T_dZ_{I_1}Z_{I_2})$$

$$+\cdots+\sum_{I_1<\cdots<I_{2a+1}\in\Lambda_d}\Pr(T_dZ_{I_1}\cdots Z_{I_{2a+1}}) \tag{A9}$$

and

$$\Pr\left(\bigcup_{I\in\Lambda_d}T_dZ_I\right)\geq\sum_{I_1\in\Lambda_d}\Pr(T_dZ_{I_1})-\sum_{I_1<I_2\in\Lambda_d}\Pr(T_dZ_{I_1}Z_{I_2})$$
$$+\cdots-\sum_{I_1<\cdots<I_{2a}\in\Lambda_d}\Pr(T_dZ_{I_1}\cdots Z_{I_{2a}}), \tag{A10}$$

for any integer $a\geq 1$. Recall the definition

$$G(d,q)=\sum_{I_1<\cdots<I_q\in\Lambda_d}\Pr(T_dZ_{I_1}\cdots Z_{I_q}),$$

for $q\geq 1$ in Lemma 3.2, which further obtains

$$\lim_{q\to\infty}\limsup_{d\to\infty}G(d,q)=0. \tag{A11}$$

Let

$$\xi(d,q)=\sum_{I_1<\cdots<I_q\in\Lambda_d}\left\{\Pr\left(T_dZ_{I_1}\cdots Z_{I_q}\right)-\Pr\left(T_d\right)\Pr\left(Z_{I_1}\cdots Z_{I_q}\right)\right\},$$

for $q\geq 1$. By Lemma 3.2, we have

$$\lim_{d\to\infty}\xi(d,q)=0,\quad\text{for some } q\geq 1. \tag{A12}$$

The assertion (A9) implies that

$$\Pr\left(\bigcup_{I\in\Lambda_d}T_dZ_I\right)\leq\Pr\left(T_d\right)\left\{\sum_{I_1\in\Lambda_d}\Pr\left(Z_{I_1}\right)-\sum_{I_1<I_2\in\Lambda_d}\Pr\left(Z_{I_1}Z_{I_2}\right)\right.$$
$$\left.+\cdots-\sum_{I_1<\cdots<I_{2a}\in\Lambda_d}\Pr\left(Z_{I_1}\cdots Z_{I_{2a}}\right)\right\}+\left\{\sum_{q=1}^{2a}\xi(d,q)\right\}+G(d,2a+1)$$
$$\leq\Pr\left(T_d\right)\Pr\left(\bigcup_{I\in\Lambda_d}Z_I\right)+\left\{\sum_{q=1}^{2a}\xi(d,q)\right\}+G(d,2a+1), \tag{A13}$$

where the last inequality uses again the inclusion–exclusion formula, that is,

$$\Pr\left(\bigcup_{I\in\Lambda_d}Z_I\right)\geq\sum_{I_1\in\Lambda_d}\Pr(Z_{I_1})-\sum_{I_1<I_2\in\Lambda_d}\Pr(Z_{I_1}Z_{I_2})+\cdots-\sum_{I_1<\cdots<I_{2a}\in\Lambda_d}\Pr(Z_{I_1}\cdots Z_{I_{2a}}),$$

for all $a\geq 1$. By the definition of l_d and Lemma 2.2,

$$\Pr\left(\bigcup_{I\in\Lambda_d}Z_I\right)=\Pr\left(\max_{1\leq\ell<k\leq d}|\tau_{k\ell}|>\ell_d\right)=\Pr\left(M_\tau>y\right)\xrightarrow{d}1-F(y),$$

as $n\to\infty$. By Lemma 2.1, $\Pr(T_d)\xrightarrow{d}\Phi(x)$ as $n\to\infty$. From (A8), (A12) and (A13), by first fixing a and letting $d\to\infty$ we obtain that

$$\limsup_{d\to\infty}\Pr\left(S_\tau\leq x,\max_{1\leq\ell<k\leq d}|\tau_{k\ell}|>\ell_d\right)\leq\Phi(x)\{1-F(y)\}+\lim_{d\to\infty}G(d,2a+1).$$

Now, sending $a \to \infty$ and using (A11), we further have

$$\limsup_{n\to\infty} \Pr\left(S_\tau \le x, \max_{1\le \ell<k\le d} |\tau_{k\ell}| > \ell_d\right) \le \Phi(x)\{1-F(y)\}. \tag{A14}$$

In a similar way, we next prove the lower bound. Using the same argument to (A10), we see that the counterpart of (A13) becomes

$$\begin{aligned}\Pr\left(\bigcup_{I\in\Lambda_d} T_d Z_I\right) &\ge \Pr\left(T_d\right)\left\{\sum_{I_1\in\Lambda_d}\Pr\left(Z_{I_1}\right) - \sum_{I_1<I_2\in\Lambda_d}\Pr\left(Z_{I_1}Z_{I_2}\right)\right.\\ &\quad \left. + \cdots + \sum_{I_1<\cdots<I_{2a-1}\in\Lambda_d}\Pr\left(Z_{I_1}\cdots Z_{I_{2a-1}}\right)\right\} + \left\{\sum_{q=1}^{2a-1}\xi(d,q)\right\} - G(d,2a)\\ &\ge \Pr\left(T_d\right)\Pr\left(\bigcup_{I\in\Lambda_d} Z_I\right) + \left\{\sum_{q=1}^{2a-1}\xi(d,q)\right\} - G(d,2a),\end{aligned}$$

where in the last step we use the inclusion–exclusion principle such that

$$\Pr\left(\bigcup_{I\in\Lambda_d} Z_I\right) \le \sum_{I_1\in\Lambda_d}\Pr(Z_{I_1}) - \sum_{I_1<I_2\in\Lambda_d}\Pr(Z_{I_1}Z_{I_2}) + \cdots + \sum_{I_1<\cdots<I_{2a-1}\in\Lambda_d}\Pr(Z_{I_1}\cdots Z_{I_{2a-1}}),$$

for all $a \ge 1$. Check (A8) and repeat the previous process to yield

$$\liminf_{n\to\infty} \Pr\left(S_\tau \le x, \max_{1\le \ell<k\le d} |\tau_{k\ell}| > \ell_d\right) \ge \Phi(x)\{1-F(y)\}, \tag{A15}$$

with $n \to \infty$ and then $a \to \infty$. Equation (A7) can be obtained by (A14) and (A15). ■

Proof of Theorem 3.2: By (4), $1 - F(M_\tau) \xrightarrow{d} U[0,1]$ and $1 - \Phi(S_\tau) \xrightarrow{d} U[0,1]$. Additionally, M_τ and S_τ are asymptotically independent as $n \to \infty$. According to the continuous mapping theorem (Theorem 2.3 of Van der Vaart 1998), it is straightforward to prove Theorem 3.2, so details are omitted. ■