Statistica Sinica Preprint No: SS-2022-0382	
Title	Rank Based Tests for High Dimensional White Noise
Manuscript ID	SS-2022-0382
URL	http://www.stat.sinica.edu.tw/statistica/
DOI	10.5705/ss.202022.0382
Complete List of Authors	Dachuan Chen, Fengyi Song and Long Feng
Corresponding Authors	Long Feng
E-mails	flnankai@nankai.edu.cn
Notice: Accepted version subject to English editing.	

Statistica Sinica

Rank Based Tests for High Dimensional White Noise

Dachuan Chen, Fengyi Song and Long Feng

Nankai University

Abstract: The development of high-dimensional white noise test is important in both statistical theories and applications, where the dimension of the time series can be comparable to or exceed the length of the time series. This paper proposes several distribution-free tests using the rank based statistics for testing the high-dimensional white noise, which are robust to the heavy tails and do not quire the finite-order moment assumptions for the sample distributions. Three families of rank based tests are analyzed in this paper, including the simple linear rank statistics, non-degenerate U-statistics and degenerate U-statistics. The asymptotic null distributions and rate optimality are established for each family of these tests. Among these tests, the test based on degenerate U-statistics can also detect the non-linear and non-monotone relationships in the autocorrelations. Moreover, this is the first result on the asymptotic distributions of rank correlation statistics which allowing for the cross-sectional dependence in high dimensional data.

Key words and phrases: Key Words: White noise test, Serial correlation, High dimensionality, Simple linear rank statistics, Non-degenerate U-statistics, Degenerate U-statistics.

1. Introduction

The hypothesis test for white noise is a critical methodology in statistical inference and modeling. It is necessary in diagnosis checking for the linear regression and time series modeling. There has been a vast increase in the amount of high-dimensional data available in recent years, which has received increasing attention from statisticians. The existence of such high-dimensional data is widespread, including the areas of genomics, neuroscience, finance, economics and so on. This brings additional difficulties for the problem of diagnosis checking, which means that in the theoretical development of the test for high-dimensional white noise, the dimension of the time series can be comparable to or exceed the length of the time series.

For the white noise tests designed for univariate time series, many commonly used methodologies are well documented in Li (2004). The alternative hypothesis of these tests can be grouped into two different classes: (i) specified alternative in form of some explicit parametric model; (ii) completely unspecified alternative, which means that the departure from white noise can be arbitrary. It is well known that likelihood based tests are more powerful than the omnibus tests under the first class of the alternatives, see, e.g. Chang et al. (2017). Under the second class of alternatives, the Box-Pierce portmanteau test and its variations are most popular because of its ease of use in practice and motivate the white noise tests for multivariate time series such as Hosking (1980) and Li and McLeod (1981). Specially, the Ljung–Box test is a type of statistical test of whether any of a group of autocorrelations of a time series are different from zero. Instead of testing randomness at each distinct lag, it tests the "overall" randomness based on a number of lags. These tests enjoy the theoretical benefits of asymptotically distribution-free and χ^2-distributed properties under null hypothesis, see, e.g. Li (2004) and Lütkepohl (2005).

There are also some white noise tests constructed for the multivariate time series which assuming that the dimension of the times series is smaller than the length of the time series in asymptotics, see, e.g., Hosking (1980) and Li and McLeod (1981). However, the existing literature suggests that these tests suffer from the slow convergence to their asymptotic null distributions, see Li et al. (2019). This fact calls for the more efficient testing methodologies for multivariate time series, or even high-dimensional time series.

Several omnibus tests for high-dimensional white noise have been developed in recent years, see, e.g., Chang et al. (2017), Li et al. (2019), Tsay (2020) and Feng et al. (2022b). Among these existing theories, the tests proposed in Chang et al. (2017), Li et al. (2019) and Feng et al. (2022b) are distribution-dependent, while the test in Tsay (2020) is distribution-free. Chang et al. (2017) developed a max-type test for this purpose based on the maximum absolute auto-correlations and cross-correlations of the component series. Li et al. (2019) proposed a sum-type test for high dimensional white noise by summing up the squared singular values of the first several lagged sample auto-covariance matrices. In general, the max-type test can only work well under the sparse alternatives where only a few elements in the auto-correlations are nonzero. In contrast, the sum-type test can only work well under the dense alternatives. To test the high dimensional white noise, Feng et al. (2022b) show the asymptotic independence between the max-type test statistic and a new sum-type test statistic. Based on this theoretical result, this paper constructed the Fisher's combination test which is robust to both sparse and dense alternatives. As a distribution-free approach, Tsay (2020) developed the high-dimensional white noise test based on the Spearman's rank correlation and the theory of extreme values.

More accurately, in this paper we consider the following hypothesis testing problem. Let $\boldsymbol{\varepsilon}_t$ be a p-dimensional weakly stationary time series with mean zero. We want to test the following hypothesis:

$$H_0 : \{\boldsymbol{\varepsilon}_t\} \text{ is } \textit{white noise versus } H_1 : \{\boldsymbol{\varepsilon}_t\} \textit{ is not white noise} \tag{1.1}$$

In this paper, we said a time series $x_1, \cdots, x_T$ are white noise if they are all independent and identically distributed. So, under the null hypothesis, $\boldsymbol{\varepsilon}_{t+k}$ is independent of $\boldsymbol{\varepsilon}_t$ for all $k > 0$. Here the dimension of the time series p is comparable to or even larger than the sample size n.

In this paper, we develop the rank based tests for testing the high-dimensional white noise, which are distribution-free. The proposed tests are robust to the heavy tails and do not require the finite-order moment assumptions or any tail assumptions for the sample distribution. There are three families of rank based tests investigated in this paper, including the simple linear rank statistics, non-degenerate U-statistics and degenerate U-statistics, with the examples of Spearman's rho, Kendall's tau, Hoeffding's D, Blum-Kiefer-Rosenblatt's R and Bergsma-Dassios-Yanagimoto's τ^*. Among these tests, simple linear rank statistics and non-degenerate U-statistics can only work well with the linear or monotone relationships in autocorrelations. In contrast, the degenerate U-statistics can also work well with the non-linear and non-monotone relationships in autocorrelations. As the theoretical results of this paper, we have established the asymptotic null distribution, the power analysis and the rate optimality in terms of power for each family of the rank based test statistics.

Because this paper shows one possible application of the rank correlation statistics in high-dimensional data analysis, we here provide a brief literature review for the rank correlation statistics and point out the theoretical contribution of this paper. Han et al. (2017) proposed the rank based tests based on the simple linear rank statistics and non-degenerate U-statistics for testing the mutual independence among all elements in the high-dimensional random vectors. Drton et al. (2020) proposed the hypothesis test based on the degenerate U-statistics with the same purpose as Han et al. (2017). As mentioned earlier, Tsay (2020) applied the Spearman's rank correlation to the test of high-dimensional white noise. However, the asymptotic distributions of the rank correlation statistics in these three existing literature are all derived based on the assumption of cross-sectional independence in high-dimensional data. Therefore, as the theoretical contribution of our result, this is the first paper in existing literature which established the asymptotic distribution of the rank correlation statistics without assuming the cross-sectional independence.

The main contributions of this paper are summarized as follows.

1. We develop the rank based tests for testing the high dimensional white noise, which are distribution free. Our test are robust to the heavy tails and do not require the finite-order moment assumptions.

2. Besides the simple linear rank statistics and non-degenerate U-statistics, we also develop the tests for the degenerate U-statistics, which are

very useful to detect the non-linear and non-monotone relationships in autocorrelations. Limiting null distributions and the rate optimality in terms of power of these three families of tests are established in this paper.

3. In the existing literature concerning the asymptotic distribution of rank correlation statistics, this paper is the first one on this topic which allowing for the cross-sectional dependence in the high-dimensional data. In contrast, the other existing results are all based on the assumption of cross-sectional independence of the data, see, e.g., Han et al. (2017), Drton et al. (2020) and Tsay (2020).

This paper is organized as follows. Section 2 proposes the theoretical results about three families of distribution-free test statistics, including the simple linear rank statistics, non-degenerate U-statistics and degenerate U-statistics. The limiting null distributions of these tests are derived and their rate-optimality in terms of power is also analyzed. Section 3 shows the empirical sizes and the power comparison of the proposed test statistics based on Monte Carlo simulation. Section 4 concludes this paper and discusses several possible directions for the research in the future. All mathematical proofs of the theoretical results in this paper are collected in supplementary material. In the supplementary material, we also consider high dimensional white noise test based on Chatterjee's rank Correlation (Chatterjee , 2021) and L-statistics with the above three-type rank based correlations (Chang et al. , 2023).

2. Rank based tests

In this section, we state the theoretical results for three families of rank based methodologies for testing the high-dimensional white noise, including simple linear rank statistics, non-degenerate U-statistics and degenerate U-statistics.

2.1 Simple linear rank statistics

First, we restate the definition of relative ranks in Han et al. (2017). Consider the dependence between $\{(\varepsilon_{1,i}, \varepsilon_{k+1,j}), \cdots, (\varepsilon_{n-k,i}, \varepsilon_{n,j})\}$ for any two entries $i, j \in \{1, \ldots, p\}$. Let $Q^i_{n-k,t}(k)$ be the rank of $\varepsilon_{t,i}$ in $\{\varepsilon_{1,i}, \ldots, \varepsilon_{n-k,i}\}$ and let $\tilde{Q}^j_{n-k,t+k}(k)$ be the rank of $\varepsilon_{t+k,j}$ in $\{\varepsilon_{k+1,j}, \cdots, \varepsilon_{n,j}\}$. Let $R^{ij}_{n-k,t+k}(k)$ be the relative rank of $\varepsilon_{t+k,j}$ compared to $\varepsilon_{t,i}$; that is, $R^{ij}_{n-k,t+k}(k) \equiv \tilde{Q}^j_{n-k,t'+k}(k)$ subject to the constraint that $Q^i_{n-k,t'}(k) = t$ for $t = 1, \cdots, n-k$.

The first family includes tests based on simple linear rank statistics of the form

$$V_{ij}(k) \equiv (n-k)^{1/2} \sum_{t=1}^{n-k} c_{n-k,t}\, g\left\{R^{ij}_{n-k,t+k}(k)/(n-k+1)\right\} \quad (i,j \in \{1,\dots,p\})$$

where $\{c_{n-k,t}, t = 1, \dots, n-k\}$ form an array of constants called the regression constants and $g(\cdot)$ is a Lipschitz function called the score function. We assume $\sum_{t=1}^{n-k} c^2_{n-k,t} > 0$ to avoid triviality. It is immediately clear that Spearman's rho belongs to the family of simple linear rank statistics. To accommodate tests of high-dimensional white noise, we further pose the alignment assumption

$$c_{n-k,t} = (n-k)^{-1} f\{t/(n-k+1)\}$$

where $f(\cdot)$ is a Lipschitz function. Under this assumption, the simple linear rank statistic is a general measure of the agreement between the ranks of two sequences. The Spearman's rho belongs to the family of simple linear rank statistics with $g(x) = f(x) = x - \frac{1}{2}$.

Under H_0, the distribution of $V_{ij}(k)$ is irrelevant to the specific distribution of ε_t for all $i, j \in \{1, \dots, p\}$. Accordingly, the mean and variance of $V_{ij}(k)$ are calculable without knowing the true distribution. Let $E_{H_0}(\cdot)$ and $\mathrm{var}_{H_0}(\cdot)$ be the expectation and variance of a certain statistic under H_0. We have

$$E_{H_0}\left(V_{ij}(k)\right) = (n-k)^{1/2} \bar{g}_{n-k} \sum_{t=1}^{n-k} c_{n-k,t}, \tag{2.2}$$

$$\sigma_V^2 = \operatorname{var}_{H_0}(V_{ij}(k)) = \frac{n-k}{n-k-1} \sum_{t=1}^{n-k} [g\{i/(n-k+1)\} - \bar{g}_{n-k}]^2 \sum_{t=1}^{n-k} (c_{n-k,t} - \bar{c}_{n-k})^2 \tag{2.3}$$

where $\bar{g}_{n-k} \equiv (n-k)^{-1} \sum_{t=1}^{n-k} g\{t/(n-k+1)\}$ is the sample mean of $g\left\{R_{n-k,t}^{ij}/(n-k+1)\right\}(t = 1, \ldots, n-k)$ and $\bar{c}_{n-k} = (n-k)^{-1} \sum_{t=1}^{n-k} c_{n-k,t}$. Based on $\{V_{ij}(k), 1 \leqslant i, j \leqslant p, 1 \leq k \leq K\}$, we propose the following statistic for testing H_0 :

$$V_n \equiv \max_{1 \leq k \leq K} \max_{1 \leq i,j \leq p} |V_{ij}(k) - E_{H_0}(V_{ij}(k))|$$

Note that we can allow K to grow as n increase. Let $N = Kp^2$. We define the following assumption for any matrix $\mathbf{\Sigma}$:

(A1) Let $\mathbf{\Sigma} = (\sigma_{ij})_{1 \leq i,j \leq N}$. For some $\varrho \in (0,1)$, assume $|\sigma_{ij}| \leq \varrho$ for all $1 \leq i < j \leq N$ and $N \geq 2$. Suppose $\{\delta_N; N \geq 1\}$ and $\{\varsigma_N; N \geq 1\}$ are positive constants with $\delta_N = o(1/\log N)$ and $\varsigma = \varsigma_N \to 0$ as $N \to \infty$. For $1 \leq i \leq N$, define $B_{N,i} = \{1 \leq j \leq N; |\sigma_{ij}| \geq \delta_N\}$ and $C_N = \{1 \leq i \leq N; |B_{N,i}| \geq N^{\varsigma}\}$. We assume that $|C_N|/N \to 0$ as $N \to \infty$.

Here we define $\{\nu_1, \cdots, \nu_{Kp^2}\} = \{V_{ij}(k)/\sigma_V\}_{1 \leq i,j \leq p, 1 \leq k \leq K}$. Define $\sigma_{ij}^V = \operatorname{cor}(\nu_i, \nu_j)$ and $\mathbf{\Sigma}_V = (\sigma_{ij}^V)_{1 \leq i,j \leq N}$.

To derive the limiting null distribution of simple linear rank statistics, we need the following conditions.

(C1) The regression constants $\{c_{n-k,1}, \ldots, c_{n-k,n-k}\}$ satisfying

$$\max_{1\leqslant i\leqslant n-k} |c_{n-k,i} - \bar{c}_{n-k}|^2 \leqslant \frac{C_1^2}{n-k} \sum_{i=1}^{n-k} (c_{n-k,i} - \bar{c}_{n-k})^2,$$

$$\left| \sum_{i=1}^{n-k} (c_{n-k,i} - \bar{c}_{n-k})^3 \right|^2 \leqslant \frac{C_2^2}{n-k} \left\{ \sum_{i=1}^{n-k} (c_{n-k,i} - \bar{c}_{n-k})^2 \right\}^3$$

where $\bar{c}_{n-k} \equiv \sum_{i=1}^{n-k} c_{n-k,i}$ represents the sample mean of the regression constants and C_1 and C_2 are two constants.

(C2) The score function $g(\cdot)$ is differentiable with bounded Lipschitz constant.

(C3) The correlation matrix $\mathbf{\Sigma}_V$ satisfies Assumption (A1).

Remark: The assumption (A1) is the same as the condition (2.2) in Feng et al. (2022a), which demands the number of variables that are strongly-correlated with many other variables should not be too much. If the eigenvalues of $\mathbf{\Sigma}_V$ are all bounded, we have $\max_{1\le i\le N} \sum_{j=1}^N \sigma_{ij}^{V2} \le C$ for some constant $C > 0$. Then, let $\delta_N = (\log N)^{-2}$ for $N \ge e^e$, so for each $1 \leqslant i \leqslant N$, $\delta_N^2 \cdot |B_{N,i}| \leqslant \sum_{j=1}^N \sigma_{ij}^{V2} \leqslant C$. Hence, $|B_{N,i}| \leqslant C \cdot (\log N)^2 < N^\kappa$ where $\kappa = \kappa_N := 5(\log\log N)/\log N$ for large N . As a result, $|C_N| = 0$ and condition (C3) holds. Condition (C1) is commonly used to deviate the asymptotical normality of the simple linear rank statistics, see Hájek et al. (1999) and Kallenberg (1982). If f is a linear function, Condition (C1) will hold directly.

Next, we state the theoretical result about the limiting null distribution of simple linear rank statistics.

Theorem 1. *Suppose (C1)-(C3) hold. Then, under H_0, for any $y \in \mathbb{R}$, we have*

$$\left|\mathrm{P}\left(V_n^2/\sigma_V^2 - 2\log(Kp^2) + \log\log(Kp^2) \leqslant y\right) - \exp\left\{-\pi^{-1/2}\exp(-y/2)\right\}\right| = o(1)$$

where $\sigma_V^2 = \mathrm{var}_{H_0}(V_{ij}(k))$ *if* $N = o(n^\epsilon)$ *as* $n \to \infty$ *for some positive constant* ϵ.

We propose the following size-α test T_α^V of H_0:

$$T_\alpha^V \doteq I\left(V_n^2/\sigma_V^2 - 2\log(Kp^2) + \log\log(Kp^2) \geq q_\alpha\right), \qquad (2.4)$$

where $q_\alpha = -\log(\pi) - 2\log\log(1-\alpha)^{-1}$.

To specify the alternative hypothesis, we introduce a notation for a set of vectors which satisfying some specific condition. Define $N = Kp^2$. Let $\mathcal{U}(c)$ be a set of vectors indexed by a constant c:

$$\mathcal{U}(c) \equiv \left\{M = (m_l)_{1\leq l\leq N} \in \mathbb{R}^N \mid \max_{1\leq l\leq N} m_l \geqslant c(\log N)^{1/2}\right\}.$$

Based on the above definition, we know that $\mathcal{U}(c)$ is the set of vectors of which at least one element has magnitude greater than $c(\log N)^{1/2}$ for some large enough constant $c > 0$.

Next, we specify the sparse local alternative based on $\mathcal{U}(c)$. We define the random vector $\hat{V} = \left[\hat{V}_{ij}(k)\right] \in \mathbb{R}^N$ by

$$\hat{V}_{ij}(k) = \sigma_V^{-1}\left\{V_{ij}(k) - E_{H_0}\left(V_{ij}(k)\right)\right\}, \quad (1 \leqslant i, j \leqslant p; 1 \leqslant k \leqslant K)$$

where σ_V is defined in (2.3) and $\{V_{ij}(k), 1 \leqslant i, j \leqslant p, 1 \leq k \leq K\}$ are the simple linear rank statistics. Let the population version of $\hat{V}$ be $V \equiv E(\hat{V})$. We study the power of tests against the alternative

$$H_{\mathrm{a}}^V(c) \equiv \{F(\boldsymbol{\varepsilon}) : V\{F(\boldsymbol{\varepsilon})\} \in \mathcal{U}(c)\}$$

where $F(\boldsymbol{\varepsilon})$ is the joint distribution function of $\boldsymbol{\varepsilon}$ and we write $V\{F(\boldsymbol{\varepsilon})\}$ to emphasize that $V = E(\hat{V}) = \int \hat{V} \mathrm{d}F(\boldsymbol{\varepsilon})$ is a function of $F(\boldsymbol{\varepsilon})$.

The following theorem now describe the conditions under which the power of the test based on simple linear rank statistics converges to one as n and p going to infinity, under the sparse local alternative H_{a}^V.

Theorem 2. *Assume Conditions (C1)-(C3) hold. And assume that* $\sigma_V^2 = A_1\{1 + o(1)\}$ *and* $\max\{|f(0)|, |g(0)|\} \leqslant A_2$ *for some positive constants* A_1 *and* A_2*. Further assume that* $f(\cdot)$ *and* $g(\cdot)$ *have bounded Lipschitz constants. Then, for some large scalar* B_1 *depending only on* A_1, A_2 *and the Lipschitz constants of* $f(\cdot)$ *and* $g(\cdot)$

$$\inf_{F(\boldsymbol{\varepsilon}) \in H_{\mathrm{a}}^V(B_1)} \operatorname{pr}\left(T_\alpha^V = 1\right) = 1 - o(1)$$

where the infimum is taken over all distributions $F(\boldsymbol{\varepsilon})$ *such that* $V\{F(\boldsymbol{\varepsilon})\} \in \mathcal{U}(B_1)$.

Define $r_{ij}(k)$ is the correlation between $\varepsilon_{t,i}$ and $\varepsilon_{t+k,j}$. To investigate the rate optimality of the test based on simple linear rank statistics, we need the following assumption for the distribution:

(A2) When $\boldsymbol{\varepsilon}$ is Gaussian, suppose that for large n and p, $cV_{ij}(k) \leqslant r_{ij}(k) \leqslant CV_{ij}(k)$ for $1 \le i, j \le p, 1 \le k \le K$ with probability tending to one, where c and C are two constants.

For each n, define $\mathcal{T}_\alpha$ to be the set of all measurable size-α tests. In other words, $\mathcal{T}_\alpha := \{T_\alpha : \mathrm{pr}(T_\alpha = 1|H_0) \le \alpha\}$.

Finally, the rate optimality result can be stated by the following theorem. Recall that T_α^V defined in (2.4) can correctly reject the null hypothesis provided that at least one element in V has magnitude greater than $c(\log N)^{1/2}$ for some constant c. In the following theorem, we show that the rate of the signal gap $(\log N)^{1/2}$ cannot be further relaxed.

Theorem 3. *Suppose that the simple linear rank statistics* $\{V_{ij}(k), 1 \leqslant i, j \leqslant p, 1 \le k \le K\}$ *satisfy all the conditions in Theorems 1 and 2. Suppose also that Assumption (A2) holds. Then, the corresponding size-*α *test* T_α^V *is rate-optimal. In other words, there exist two constants* $D_1 < D_2$ *such that:*

(i) $\sup_{F(\boldsymbol{\varepsilon})\in H_{\mathrm{a}}^{V}(D_2)} \mathrm{pr}\left(T_\alpha = 0\right) = o(1)$*;*

(ii) for any $\beta > 0$ *satisfying* $\alpha + \beta < 1$*, for large* n *and* p *we have*

$$\inf_{T_\alpha\in\mathcal{T}_\alpha} \sup_{F(\boldsymbol{\varepsilon})\in H_{\mathrm{a}}^{V}(D_1)} \mathrm{pr}\left(T_\alpha = 0\right) \geqslant 1-\alpha-\beta$$

The above theorem means that any measurable size-α test cannot distinguish between the null hypothesis and the sparse alternative when the coefficient c in $H_{\mathrm{a}}^{V}(c)$ is small enough.

As an example of simple linear rank statistic, we state the high-dimensional white noise test based on the Spearman's rho as follows.

Example 1 (Spearman's rho). Recall that $Q^i_{n-k,t}(k)$ and $\tilde{Q}^j_{n-k,t+k}(k)$ be the ranks of $\varepsilon_{t,i}$ and $\varepsilon_{t+k,j}$ among $\{\varepsilon_{1,i},\ldots,\varepsilon_{n-k,i}\}$ and $\{\varepsilon_{k+1,j},\cdots,\varepsilon_{n,j}\}$, respectively. Let $R^{ij}_{n-k,t+k}(k)$ be the relative rank of $\varepsilon_{t+k,j}$ compared to $\varepsilon_{t,i}$; that is, $R^{ij}_{n-k,t+k}(k) \equiv \tilde{Q}^j_{n-k,t'+k}(k)$ subject to the constraint that $Q^i_{n-k,t'}(k) = t$ for $t = 1,\cdots,n-k$. Spearman's rho is defined as

$$\begin{aligned}
\rho_{ij}(k) &= \frac{\sum_{t=1}^{n-k}\left(Q^i_{n-k,t}(k)-\bar{Q}^i_{n-k}(k)\right)\left(\tilde{Q}^j_{n-k,t+k}(k)-\bar{\tilde{Q}}^j_{n-k}(k)\right)}{\left\{\sum_{t=1}^{n-k}\left(Q^i_{n-k,t}(k)-\bar{Q}^i_{n-k}(k)\right)^2\sum_{t=1}^{n-k}\left(\tilde{Q}^j_{n-k,t+k}(k)-\bar{\tilde{Q}}^j_{n-k}(k)\right)^2\right\}^{1/2}}\\
&= \frac{12}{(n-k)\left((n-k)^2-1\right)}\sum_{t=1}^{n-k}\left(i-\frac{n-k+1}{2}\right)\left(R^{ij}_{n-k,t+k}(k)-\frac{n-k+1}{2}\right) \quad (i,j\in\{1,\ldots,p\})
\end{aligned}$$

where $\bar{Q}^i_{n-k}(k) = \tilde{Q}^j_{n-k}(k) \equiv (n-k+1)/2$. This is a simple linear rank statistic, and we have

$$E_{H_0}(\rho_{ij}(k)) = 0, \quad \mathrm{var}_{H_0}(\rho_{ij}(k)) = (n-k-1)^{-1} \quad (i, j \in \{1, \ldots, p\})$$

According to (2.4), the corresponding test statistic is

$$L_\rho = I\left\{ \max_{1 \le k \le K} \max_{1 \le i,j \le p} (n-k)\rho_{ij}(k)^2 - 2\log(Kp^2) + \log\log(Kp^2) \geqslant q_\alpha \right\}$$

where $q_\alpha \equiv -\log(\pi) - 2\log\log(1-\alpha)^{-1}$.

2.2 Non-degenerate U-statistics

The second family includes the tests based on non-degenerate U-statistics of the form (Han et al. 2017)

$$U_n = \max_{1 \le k \le K} \max_{1 \le i,j \le p} (n-k)^{1/2} \left| U_{ij}(k) - E_{H_0}(U_{ij}(k)) \right| \qquad (2.5)$$

where $A^m_{n-k} = (n-k)(n-k-1)\cdots(n-k-m+1)$,

$$U_{ij}(k) = \frac{1}{A^m_{n-k}} \sum_{1 \le t_1 \ne t_2, \cdots, \ne t_m \le n-k} h((\varepsilon_{t_1,i}, \varepsilon_{t_1+k,j})^\top, \cdots, (\varepsilon_{t_m,i}, \varepsilon_{t_m+k,j})^\top) \qquad (2.6)$$

Here $U_{ij}(k)$ depends only on $\{R^{ij}_{n-k,t}(k)\}^n_{t=k+1}$. For our purposes h may always be assumed to be bounded but not necessarily symmetric. The boundedness assumption is mild since correlation is the object of interest.

Further concepts concerning U-statistics are needed to state the assumption for the derivation of the limiting null distribution. For $m \in \mathbb{Z}^+$, we define $[m] = \{1, 2, \cdots, m\}$ and write $\mathcal{P}_m$ for the set of all $m!$ permutations of $[m]$. For any kernel $h(\cdot)$, any number $\ell \in [m]$, and any measure $\mathbb{P}_{\boldsymbol{Z}}$, we write

$$h_\ell\left(\boldsymbol{z}_1 \ldots, \boldsymbol{z}_\ell; \mathbb{P}_{\boldsymbol{Z}}\right) := \mathbb{E}h\left(\boldsymbol{z}_1 \ldots, \boldsymbol{z}_\ell, \boldsymbol{Z}_{\ell+1}, \ldots, \boldsymbol{Z}_m\right) \tag{2.7}$$

and

$$h^{(\ell)}\left(\boldsymbol{z}_1, \ldots, \boldsymbol{z}_\ell; \mathbb{P}_{\boldsymbol{Z}}\right) := h_\ell\left(\boldsymbol{z}_1, \ldots, \boldsymbol{z}_\ell; \mathbb{P}_{\boldsymbol{Z}}\right) - \mathbb{E}h - \sum_{k=1}^{\ell-1} \sum_{1 \le i_1 < \cdots < i_k \le \ell} h^{(k)}\left(\boldsymbol{z}_{i_1}, \ldots, \boldsymbol{z}_{i_k}; \mathbb{P}_{\boldsymbol{Z}}\right) \tag{2.8}$$

where $\boldsymbol{Z}_1, \ldots, \boldsymbol{Z}_m$ are m independent random vectors with distribution $\mathbb{P}_{\boldsymbol{Z}}$ and $\mathbb{E}h := \mathbb{E}h\left(\boldsymbol{Z}_1, \ldots, \boldsymbol{Z}_m\right)$. The kernel as well as the corresponding U-statistic is non-degenerate under $\mathbb{P}_{\boldsymbol{Z}}$ if the variance of $h_1(\cdot)$ is not zero.

Based on above definitions, we state the following conditions which are needed to derive the limiting null distribution.

(C4) The kernel function $h(\cdot)$ is bounded and non-degenerate.

(C5) The correlation matrix of $U_{ij}(k) - \boldsymbol{\Sigma}_U$ satisfies Assumption (A1).

The following theorem show the asymptotic distribution of the non-degenerate U-statistics under the null hypothesis.

Theorem 4. *Suppose (C4)-(C5) hold. Then under H_0, for any $y \in \mathbb{R}$ we have*

$$\left|\mathrm{P}\left(U_n^2/\sigma_U^2 - 2\log(Kp^2) + \log\log(Kp^2) \leqslant y\right) - \exp\left\{-\pi^{-1/2}\exp(-y/2)\right\}\right| = o(1)$$

where $\sigma_U^2 = (n-k)\mathrm{var}_{H_0}(U_{ij}(k))$ if $N = o(n^{\epsilon})$ as $n \to \infty$ for some positive constant ϵ.

We propose the following size-α test T_α^U of H_0:

$$T_\alpha^U \doteq I\left(U_n^2/\sigma_U^2 - 2\log(Kp^2) + \log\log(Kp^2) \geq q_\alpha\right) \tag{2.9}$$

where $q_\alpha = -\log(\pi) - 2\log\log(1-\alpha)^{-1}$.

To specify the sparse local alternative for the tests based on non-degenerate U-statistics, we first define the random vector $\hat{U} = \left[\hat{U}_{ij}(k)\right] \in \mathbb{R}^N$ by

$$\hat{U}_{ij}(k) = \sigma_U^{-1}(n-k)^{1/2}\left\{U_{ij}(k) - E_{H_0}\left(U_{ij}(k)\right)\right\}, \quad (1 \leqslant i,j \leqslant p; 1 \leqslant k \leqslant K)$$

where σ_U is defined in Theorem 4 and $\{U_{ij}(k), 1 \leqslant i,j \leqslant p, 1 \leq k \leq K\}$ are the non-degenerate U-statistics. Let the population version of $\hat{U}$ be $U \equiv E(\hat{U})$. We study the power of tests against the alternative

$$H_{\mathrm{a}}^U(c) \equiv \{F(\boldsymbol{\varepsilon}) : U\{F(\boldsymbol{\varepsilon})\} \in \mathcal{U}(c)\}$$

where $F(\boldsymbol{\varepsilon})$ is the joint distribution function of $\boldsymbol{\varepsilon}$ and we write $U\{F(\boldsymbol{\varepsilon})\}$ to emphasize that $U = E(\hat{U}) = \int \hat{U}\mathrm{d}F(\boldsymbol{\varepsilon})$ is a function of $F(\boldsymbol{\varepsilon})$.

The following theorem states the conditions which are required to establish the convergence of the power of T_α^U to one as n and p going to infinity under the sparse alternative.

Theorem 5. *Suppose that the kernel function $h(\cdot)$ in (2.6) is bounded with $|h(\cdot)| \leqslant A_3$ and*

$$m^2 \operatorname{var}_{H_0}\left[E_{H_0}\left\{h\left((X_{11}, X_{12})^\top, \ldots, (X_{m1}, X_{m2})^\top\right) \mid (X_{11}, X_{12})^\top\right\}\right] = A_4\{1+o(1)\}$$

for some positive constants A_3 and A_4. Then, for some large scalar B_2 depending only on A_3, A_4 and m,

$$\inf_{F(\boldsymbol{\varepsilon}) \in H_{\mathrm{a}}^U(B_2)} P\left(T_\alpha = 1\right) = 1 - o(1)$$

where the infimum is taken over all distributions $F(\boldsymbol{\varepsilon})$ such that $U\{F(\boldsymbol{\varepsilon})\} \in \mathcal{U}\left(B_2\right)$.

To study the rate optimality in terms of power for the tests based on non-degenerate U-statistics, we need the following assumption for the distribution:

(A3) When $\boldsymbol{\varepsilon}$ is Gaussian, suppose that for non-degenerate U-statistics $U_{ij}(k)$ and large n and p, $cU_{ij}(k) \leqslant r_{ij}(k) \leqslant CU_{ij}(k)$ for $1 \leq i, j \leq p, 1 \leq k \leq K$ with probability tending to one, where c and C are two constants.

The rate optimality result and related conditions for the tests based on the non-degenerate U-statistics can be shown as follows, which implies that the rate of the signal gap $(\log N)^{1/2}$ cannot be further relaxed.

Theorem 6. *Suppose that Non-degenerate U-statistics* $\{U_{ij}(k), 1 \leqslant i, j \leqslant p, 1 \leq k \leq K\}$ *satisfy all the conditions in Theorems 4 and 5. Suppose also that Assumption (A3) holds. Then, the corresponding size-*α *test* T_α^U *is rate-optimal. In other words, there exist two constants* $D_3 < D_4$ *such that:*

(i) $\sup_{F(\boldsymbol{\varepsilon}) \in H_{\mathrm{a}}^U(D_4)} \mathrm{pr}\left(T_\alpha = 0\right) = o(1)$*;*

(ii) for any $\beta > 0$ *satisfying* $\alpha + \beta < 1$*, for large* n *and* p *we have*

$$\inf_{T_\alpha \in \mathcal{T}_\alpha} \sup_{F(\boldsymbol{\varepsilon}) \in H_{\mathrm{a}}^U(D_3)} \mathrm{pr}\left(T_\alpha = 0\right) \geqslant 1 - \alpha - \beta.$$

As an example of non-degenerate U-statistics, we state the high-dimensional white noise test based on the Kendall's tau as follows.

Example 2 (Kendall's tau). Kendall's tau is defined, for $i, j \in \{1, \ldots, p\}$, by

$$\begin{aligned}
\tau_{ij}(k) &= \frac{2}{(n-k)(n-k-1)} \sum_{1 \leq l < l' \leq n-k} \mathrm{sign}\left(\varepsilon_{l',i} - \varepsilon_{l,i}\right) \mathrm{sign}\left(\varepsilon_{l'+k,j} - \varepsilon_{l+k,j}\right) \\
&= \frac{2}{(n-k)(n-k-1)} \sum_{1 \leq l < l' \leq n-k} \mathrm{sign}\left(R_{n-k,l'+k}^{ij}(k) - R_{n-k,l+k}^{ij}(k)\right)
\end{aligned}$$

where the sign function $\operatorname{sign}(\cdot)$ is defined as $\operatorname{sign}(x) = x/|x|$ with the convention $0/0 = 0$. This statistic is a function of the relative ranks $\left\{R^{ij}_{n-k,t+k}(k), t = 1, \ldots, n-k\right\}$ and is also a U-statistic with bounded kernel $h\left(x_{1,\{1,2\}}, x_{2,\{1,2\}}\right) \equiv \operatorname{sign}\left(x_{1,1} - x_{2,1}\right)\operatorname{sign}\left(x_{1,2} - x_{2,2}\right)$. Accordingly, Kendall's tau is a rank-type U-statistic. Moreover,

$$E_{H_0}\left(\tau_{ij}(k)\right) = 0, \quad \operatorname{var}_{H_0}\left(\tau_{ij}(k)\right) = \frac{2(2(n-k)+5)}{9(n-k)(n-k-1)} \quad (i, j \in \{1, \ldots, p\})$$

According to (8), the proposed test statistic based on Kendall's tau is

$$L_\tau = I\left\{\max_{1\leq k\leq K}\max_{1\leq i,j\leq p} \frac{9(n-k)(n-k-1)}{2(2(n-k)+5)}\tau_{ij}(k)^2 - 2\log(Kp^2) + \log\log(Kp^2) \geqslant q_\alpha\right\}$$

2.3 Degenerate U-statistics

The third family includes the tests based on degenerate U-statistics, which are very useful to detect the non-linear and non-monotone relationships in the autocorrelations. We use the term completely degenerate to indicate that the variances of $h_1(\cdot), \ldots, h_{m-1}(\cdot)$ are all zero. Finally, let $\mathbb{P}_0$ be the uniform distribution on $[0,1]$, and write $\mathbb{P}_0 \otimes \mathbb{P}_0$ for its product measure, the uniform distribution on $[0,1]^2$.

In order to derive the limiting null distribution and establish the theoretical results related to the power of the tests based on the degenerate U-statistics, we need the following assumption concerning the kernel function h.

(C6) The kernel h is rank-based, symmetric, and has the following three properties:

(i) h is bounded.

(ii) h is mean-zero and degenerate under independent continuous margins, i.e., $\mathbb{E}\left\{h_1\left(\boldsymbol{Z}_1; \mathbb{P}_0 \otimes \mathbb{P}_0\right)\right\}^2 = 0$ as $\boldsymbol{Z}_1 \sim \mathbb{P}_0 \otimes \mathbb{P}_0$

(iii) $h_2\left(\boldsymbol{z}_1, \boldsymbol{z}_2; \mathbb{P}_0 \otimes \mathbb{P}_0\right)$ has uniformly bounded eigenfunctions, that is, it admits the expansion

$$h_2\left(\boldsymbol{z}_1, \boldsymbol{z}_2; \mathbb{P}_0 \otimes \mathbb{P}_0\right) = \sum_{v=1}^{\infty} \lambda_v \phi_v\left(\boldsymbol{z}_1\right) \phi_v\left(\boldsymbol{z}_2\right)$$

where $\{\lambda_v\}$ and $\{\phi_v\}$ are the eigenvalues and eigenfunctions satisfying the integral equation

$$\mathbb{E} h_2\left(\boldsymbol{z}_1, \boldsymbol{Z}_2\right) \phi\left(\boldsymbol{Z}_2\right) = \lambda \phi\left(\boldsymbol{z}_1\right) \text{ for all } \boldsymbol{z}_1 \in \mathbb{R}^2$$

with $\boldsymbol{Z}_2 \sim \mathbb{P}_0 \otimes \mathbb{P}_0, \lambda_1 \geq \lambda_2 \geq \cdots \geq 0, \Lambda = \sum_{v=1}^{\infty} \lambda_v \in (0, \infty),$ and $\sup_v \|\phi_v\|_\infty < \infty.$

The first requirement about boundedness property can be easily verified for the rank correlations which are commonly used, for example, Spearman's rho, Kendall's tau and many others. The other two requirements are much more specific, but can be satisfied by some typical rank correlation measures as long as their consistency properties are known. Moreover, it is easy to see that the assumption $\Lambda > 0$ implies $\lambda_1 > 0$, so that $h_2(\cdot)$ is not a constant function.

We also need the following condition to derive the limiting null distribution for the degenerate U-statistics. We first make several definitions which will be used in the following condition. Define a quantity θ, which is any absolute constant such that

$$\theta < \sup \left\{ q \in [0, 1/3) : \sum_{v > [n^{(1-3q)/5}]} \lambda_v = O(n^{-q}) \right\}$$

if infinitely many eigenvalues λ_v are nonzero, and $\theta = 1/3$ otherwise. Define $\omega_{l,v} = (n - k_l)^{-1/2} \sum_{t=1}^{n-k_l} \phi_v(Z_{t,l})$ for $l = 1, \cdots, N$, $v = 1, \cdots, M$ where $M = [n^{(1-3\theta)/5}]$ and $Z_{t,l}$ is the corresponding Z of $U_{ij}(k)$ in Condition (C6). Let $b_{lv,rs} = \mathrm{cov}(\omega_{l,v}, \omega_{r,s})$ for $1 \le l, r \le N, 1 \le v, s \le M$. Let $\boldsymbol{\omega}_l = (\omega_{l,1}, \cdots, \omega_{l,M})$ and $\Xi_l = \boldsymbol{\Sigma}_l \boldsymbol{\Sigma}_l^\top$ where $\boldsymbol{\Sigma}_l \in \mathbb{R}^{M \times (N-1)M}$ is the covariance matrix between $\boldsymbol{\omega}_l$ with $\boldsymbol{\omega}_r, r \in \{1, \cdots, N\} \setminus \{l\}$.

(C7) There exists a constant $\delta \in (0, 1)$ satisfying $\lambda_{max}(\Xi_l) \le \delta$ for all $1 \le l \le N$. Suppose $\{\delta_N; N \ge 1\}$ and $\{\varsigma_N; N \ge 1\}$ are positive constants with $\delta_N = o(1/\log N)$ and $\varsigma = \varsigma_N \to 0$ as $N \to \infty$. Let $\Xi_{ij} = \mathrm{cov}(\boldsymbol{\omega}_i, \boldsymbol{\omega}_j)$. For $1 \le i \le N$, define $B_{N,i} = \left\{1 \le j \le N \mid \lambda_{max}(\Xi_{ij}\Xi_{ij}^\top) \ge \delta_N^{2+2c}\right\}$ for some constant $c > 0$ and $C_N = \{1 \le i \le N; |B_{N,i}| \ge N^{\varsigma}\}$. We assume that $|C_N|/N \to 0$ as $N \to \infty$.

In the following theorem, we show the limiting null distribution and related conditions for the degenerate U-statistics.

Theorem 7. *Under conditions (C6)-(C7). Then for any absolute constant $y \in \mathbb{R}$ that*

$$\mathbb{P}\left\{\max_{1\le k\le K}\max_{1\le i,j\le p}\frac{n-k-1}{\lambda_1\binom{m}{2}}U_{ij}(k)-2\log(Kp^2)-(\mu_1-2)\log\log(Kp^2)+\frac{\Lambda}{\lambda_1}\le y\right\}$$

$$=\exp\left\{-\frac{\kappa}{\Gamma\left(\mu_1/2\right)}\exp\left(-\frac{y}{2}\right)\right\}+o(1)$$

for $\log N = o(n^{\theta})$ as $n\to\infty$. Here μ_1 is the multiplicity of the largest eigenvalue λ_1 in the sequence $\{\lambda_1,\lambda_2,\ldots\}$, $\kappa := \prod_{v=\mu_1+1}^{\infty}(1-\lambda_v/\lambda_1)^{-1/2}$ and $\Gamma(z) := \int_0^{\infty} x^{z-1}e^{-x}dx$ is the gamma function.

We propose the following size-α test T_α^D for degenerate U-statistics:

$$T_\alpha^D = I\left(\max_{1\le k\le K}\max_{1\le i,j\le p}\frac{n-k-1}{\lambda_1\binom{m}{2}}U_{ij}(k)-2\log(Kp^2)-(\mu_1-2)\log\log(Kp^2)+\frac{\Lambda}{\lambda_1}\ge \tilde{q}_\alpha\right) \tag{2.10}$$

where $\tilde{q}_\alpha$ is the $1-\alpha$ quantile of the Gumbel distribution function $\exp\left\{-\kappa/\Gamma\left(\mu_1/2\right)\exp\left(-y/2\right)\right\}$, i.e.,

$$\tilde{q}_\alpha = -\log\left(\frac{\Gamma^2\left(\mu_1/2\right)}{\kappa^2}\right)-2\log\log(1-\alpha)^{-1}.$$

It is easy to show that $\mathbb{P}_{H_0}\left(T_\alpha^D = 1\right) = \alpha + o(1)$

We study the power of the proposed test based on the degenerate U-statistics from now on. It is necessary to introduce a new distribution family which is also useful to specify the alternative. Recall the definition of $h^{(1)}(\cdot)$ in (2.8) . For any kernel function $h(\cdot)$ and constants $\gamma > 0$ and $n, p \in \mathbb{Z}^+$, define a general p-dimensional (not necessarily continuous) distribution family as follows:

$$\mathcal{D}(\gamma, np; h) := \left\{F(\boldsymbol{X}) : \boldsymbol{X} \in \mathbb{R}^{np}, \mathrm{Var}_{ijk}\left\{h^{(1)}\left(\cdot; \mathbb{P}_{ijk}\right)\right\} \leq \gamma \mathbb{E}_{ijk} h \text{ for all } 1 \leq i, j \leq p, 1 \leq k \leq K\right\}$$

where $F(\boldsymbol{X})$ is the distribution (law) of $\boldsymbol{X}$, and $\mathbb{P}_{ijk}, \mathbb{E}_{ijk}(\cdot)$, and $\mathrm{Var}_{ijk}(\cdot)$ stand for the probability measure, expectation, and variance operated on the bivariate distribution of $\left(\varepsilon_{ti}, \varepsilon_{t+k,j}\right)^\top$, respectively. The family $\mathcal{D}(\gamma, np; h)$ intrinsically characterizes the slope of the non-negative function $\mathrm{Var}_{ijk}\left\{h^{(1)}\left(\cdot; \mathbb{P}_{ijk}\right)\right\}$ with regard to the dependence between ε_{ti} and $\varepsilon_{t+k,j}$, characterized by the non-negative correlation measure $\mathbb{E}_{ijk}h$. Under the null hypothesis, we have

$$\mathrm{Var}_{ijk}\left\{h^{(1)}\left(\cdot; \mathbb{P}_{ijk}\right)\right\} = \mathbb{E}_{ijk}h = 0$$

provided that Condition (C6) holds for $h(\cdot)$. Therefore, as the dependence between ε_{ti} and $\varepsilon_{t+k,j}$ increasing, it can be expected that the variance $\mathrm{Var}_{ijk}\left\{h^{(1)}\left(\cdot; \mathbb{P}_{ijk}\right)\right\}$ will depart away from zero with the same or a slower rate compared to $\mathbb{E}_{ijk}h$.

In the following theorem, we show that the power of the proposed test T_α^D converges to one as n and p increasing to infinity under a newly specified sparse alternative.

Theorem 8. *Given any $\gamma > 0$ and a kernel $h(\cdot)$ satisfying Condition (C6), there exists some sufficiently large B_3 depending on γ such that*

$$\inf_{F(\boldsymbol{\varepsilon})\in\mathcal{D}(\gamma,np;h)\cap H_{\text{a}}^U(B_3)} \mathbb{P}\left(T_\alpha^D = 1\right) = 1 - o(1)$$

The establishment of the rate optimality of the tests based on degenerate U-statistics requires the following assumption for the distribution:

(A4) When $\boldsymbol{\varepsilon}$ is Gaussian, suppose that for degenerate U-statistics $U_{ij}(k)$ and large n and p, $cU_{ij}(k) \leqslant r_{ij}(k) \leqslant CU_{ij}(k)$ for $1 \le i, j \le p, 1 \le k \le K$ with probability tending to one, where c and C are two constants.

Under the new type of sparse local alternative, we could show the rate optimality in terms of power for the proposed test in the following theorem.

Theorem 9. *Suppose that Degenerate U-statistics $\{U_{ij}(k), 1 \leqslant i, j \leqslant p, 1 \le k \le K\}$ satisfy all the conditions in Theorems 7 and 8. Suppose also that Assumption (A4) holds. Then, the corresponding size-α test T_α^D is rate-optimal. In other words, there exist two constants $D_5 < D_6$ such that:*

(i) $\sup_{F(\boldsymbol{\varepsilon})\in\mathcal{D}(\gamma,np;h)\cap H_{\text{a}}^U(D_6)} \text{pr}\left(T_\alpha = 0\right) = o(1)$;

(ii) for any $\beta > 0$ satisfying $\alpha + \beta < 1$, for large n and p we have

$$\inf_{T_\alpha \in \mathcal{T}_\alpha} \sup_{F(\boldsymbol{\varepsilon}) \in \mathcal{D}(\gamma, np; h) \cap H_{\mathrm{a}}^U(D_5)} \mathrm{pr}(T_\alpha = 0) \geqslant 1 - \alpha - \beta$$

Three examples belonging to the family of degenerate U-statistics are provided to test the high dimensional white noise as follows.

Example 3 (Hoeffding's D). The Hoeffding's D statistic is a rank-based U-statistic of order 5, which is based on the symmetric kernel

$$h_D(z_1, \ldots, z_5) := \frac{1}{16} \sum\nolimits_{(i_1, \ldots, i_5) \in \mathcal{P}_5}$$
$$[\{I(z_{i_1,1} \le z_{i_5,1}) - I(z_{i_2,1} \le z_{i_5,1})\}\{I(z_{i_3,1} \le z_{i_5,1}) - I(z_{i_4,1} \le z_{i_5,1})\}]$$
$$[\{I(z_{i_1,2} \le z_{i_5,2}) - I(z_{i_2,2} \le z_{i_5,2})\}\{I(z_{i_3,2} \le z_{i_5,2}) - I(z_{i_4,2} \le z_{i_5,2})\}].$$

Thus, the Hoeffding's D correlation measure is given by $\mathbb{E}h_D$. Based on Weihs et al. (2018, Proposition 7) or Nandy et al. (2016, Theorem 4.4), under the measure $\mathbb{P}_0 \otimes \mathbb{P}_0$, the eigenvalues and corresponding eigenfunctions of $h_{D,2}(\cdot)$ are:

$$\lambda_{i,j;D} = 3/\left(\pi^4 i^2 j^2\right) > 0, \quad i, j \in \mathbb{Z}^+$$

and

$$\phi_{i,j;D}\left\{(z_{1,1}, z_{1,2})^\top\right\} = 2\cos(\pi i z_{1,1})\cos(\pi j z_{1,2}), \quad i, j \in \mathbb{Z}^+,$$

where $\Lambda_D := \sum_{i,j} \lambda_{i,j;D} = 1/12$ and $\sup_{i,j} \|\phi_{i,j;D}\|_\infty \le 2$. Therefore, by considering the results in Hoeffding (1948), the kernel $h_D(\cdot)$ satisfies the three properties in Condition (C6). Based on the result in Hoeffding (1948, p. 547), the correlation measure $\mathbb{E}h_D$ is non-negative for arbitrary pair of random variables. Moreover, as shown by Hoeffding (1948) and Yanagimoto (1970), for a pair of random variables which is absolutely continuous in $\mathbb{R}^2$, the sufficient and necessary condition for their independence is that $\mathbb{E}h_D = 0$. However, this result does not hold when the data is discrete or is continuous but not absolute continuous, e.g. a counter example is given in Remark 1 of Yanagimoto (1970).

Define $\{\boldsymbol{X}_{tijk}\} = \{(\varepsilon_{t,i}, \varepsilon_{t+k,j})^\top\}_{1\le t\le n-k}$. According to (2.10), the corresponding test is

$$\widehat{D}_{ij}(k) := \begin{pmatrix} n-k \\ 5 \end{pmatrix}^{-1} \sum_{t_1<\cdots<t_5} h_D\left(\boldsymbol{X}_{t_1,ijk}, \ldots, \boldsymbol{X}_{t_5,ijk}\right) \tag{2.11}$$

and

$$L_D := I\left\{\max_{1\le k\le K}\max_{1\le i,j\le p} \frac{\pi^4(n-k-1)}{30}\widehat{D}_{ij}(k) - 2\log(Kp^2) + \log\log(Kp^2) + \frac{\pi^4}{36} > Q_{D,\alpha}\right\} \tag{2.12}$$

where $Q_{D,\alpha} := \log\{\kappa_D^2/\pi\} - 2\log\log(1-\alpha)^{-1}$ and

$$\kappa_D := \left\{2\prod_{n=2}^{\infty}\frac{\pi/n}{\sin(\pi/n)}\right\}^{1/2} \approx 2.467.$$

Example 4 (Blum-Kiefer-Rosenblatt's R). The Blum-Kiefer-Rosenblatt's R statistic (Blum et al. (1961)) is a rank-based U-statistic of order 6, which is based on the symmetric kernel:

$$h_R(z_1,\ldots,z_6) := \frac{1}{32}\sum_{(i_1,\ldots,i_6)\in\mathcal{P}_6}$$
$$[\{I(z_{i_1,1}\le z_{i_5,1}) - I(z_{i_2,1}\le z_{i_5,1})\}\{I(z_{i_3,1}\le z_{i_5,1}) - I(z_{i_4,1}\le z_{i_5,1})\}]$$
$$[\{1(z_{i_1,2}\le z_{i_6,2}) - I(z_{i_2,2}\le z_{i_6,2})\}\{I(z_{i_3,2}\le z_{i_6,2}) - 1(z_{i_4,2}\le z_{i_6,2})\}].$$

The three properties in Condition (C6) can be easily verified based on the fact that $h_{R,2} = 2h_{D,2}$. Similarly, the correlation measure $\mathbb{E}h_R$ is non-negative for arbitrary pair of random variables. $\mathbb{E}h_R = 0$ if and only if the pair of random variables are independent (without requiring the continuity properties), see, e.g. page 490 of Blum et al. (1961).

According to (2.10), the corresponding test is

$$\widehat{R}_{ij}(k) := \begin{pmatrix} n-k \\ 6 \end{pmatrix}^{-1} \sum_{t_1<\cdots<t_6} h_R(\boldsymbol{X}_{t_1,ijk},\ldots,\boldsymbol{X}_{t_6,ijk})$$

and

$$L_R := I\left\{\max_{1\le k\le K}\max_{1\le i,j\le p}\frac{\pi^4(n-k-1)}{90}\widehat{R}_{ij}(k) - 2\log(Kp^2) + \log\log(Kp^2) + \frac{\pi^4}{36} > Q_{R,\alpha}\right\}$$

where $Q_{R,\alpha} := Q_{D,\alpha}$

Example 5 (Bergsma-Dassios-Yanagimoto's τ^*). Bergsma and Dassios (2014) introduced a rank correlation statistic as a U-statistic of order 4 with the symmetric kernel

$$h_{\tau^*}(z_1,\ldots,z_4)$$

$$
\begin{aligned}
:= & \frac{1}{16} \sum_{(i_1,\ldots,i_4)\in\mathcal{P}_4} \{1\left(z_{i_1,1}, z_{i_3,1} < z_{i_2,1}, z_{i_4,1}\right) + I\left(z_{i_2,1}, z_{i_4,1} < z_{i_1,1}, z_{i_3,1}\right) \\
& -I\left(z_{i_1,1}, z_{i_4,1} < z_{i_2,1}, z_{i_3,1}\right) - I\left(z_{i_2,1}, z_{i_3,1} < z_{i_1,1}, z_{i_4,1}\right)\} \\
& \{I\left(z_{i_1,2}, z_{i_3,2} < z_{i_2,2}, z_{i_4,2}\right) + I\left(z_{i_2,2}, z_{i_4,2} < z_{i_1,2}, z_{i_3,2}\right) \\
& -I\left(z_{i_1,2}, z_{i_4,2} < z_{i_2,2}, z_{i_3,2}\right) - I\left(z_{i_2,2}, z_{i_3,2} < z_{i_1,2}, z_{i_4,2}\right)\},
\end{aligned}
$$

where $I\left(y_1, y_2 < y_3, y_4\right) := I\left(y_1 < y_3\right) I\left(y_1 < y_4\right) I\left(y_2 < y_3\right) I\left(y_2 < y_4\right)$. Based on the fact that $h_{\tau^*,2} = 3h_{D,2}$, all properties in Condition (C6) can be verified for $h_{\tau^*}(\cdot)$. As shown by Theorem 1 in Bergsma and Dassios (2014), for a pair of random variables whose distribution is discrete, absolutely continuous, or a mixture of both, the correlation measure $\mathbb{E}h_{\tau^*}$ is non-negative and $\mathbb{E}h_{\tau^*} = 0$ if and only if the pair is independent.

According to (2.10), it yields the test

$$
\widehat{\tau}^*_{ij}(k) := \begin{pmatrix} n-k \\ 4 \end{pmatrix}^{-1} \sum_{t_1<\cdots<t_4} h_{\tau^*}\left(\boldsymbol{X}_{t_1,ijk}, \ldots, \boldsymbol{X}_{t_4,ijk}\right)
$$

and

$$
L_{\tau^*} := I\left\{\max_{1\le k\le K} \max_{1\le i,j\le p} \frac{\pi^4(n-k-1)}{54}\widehat{\tau}^*_{ij}(k) - 2\log(Kp^2) + \log\log(Kp^2) + \frac{\pi^4}{36} > Q_{\tau^*,\alpha}\right\}
$$

where $Q_{\tau^*,\alpha} := Q_{D,\alpha}$

3. Simulation

In this section, we evaluate the empirical sizes and powers of several test statistics based on Monte Carlo simulation. We mainly compare the performance of the following test statistics:

- L_r: the max-type test statistic provided by Chang et al. (2017);
- S_r: the sum-type test statistic provided by Li et al. (2019);
- L_ρ: the Spearman's rho statistic defined in Example 1;
- L_τ: the Kendall's tau statistic defined in Example 2;
- L_D: the Hoeffding's D statistic defined in Example 3;
- L_R: the Blum-Kiefer-Rosenblatt's R statistic defined in Example 4;
- L_{τ^*}: the Bergsma-Dassios-Yanagimoto's τ^* statistic defined Example 5.

3.1 Empirical sizes

Let $\boldsymbol{\varepsilon}_t = \mathbf{A}\boldsymbol{z}_t$. We consider the following four distribution for $\boldsymbol{z}_t$: (a) $\boldsymbol{z}_t \sim N(\mathbf{0}, \mathbf{I}_p)$; (b) $\boldsymbol{z}_t = \boldsymbol{w}_t^{1/3}$ with $\boldsymbol{w}_t \sim N(\mathbf{0}, \mathbf{I}_p)$; (c) $\boldsymbol{z}_t = \boldsymbol{w}_t^3$ with $\boldsymbol{w}_t \sim N(\mathbf{0}, \mathbf{I}_p)$; (d) $\boldsymbol{z}_t = (z_{t1}, \cdots, z_{tp})^\top$ with $z_{ti} \overset{i.i.d}{\sim} t(3)/\sqrt{3}$. For the Models (i)-(iv), we consider $\mathbf{A} = \mathbf{\Sigma}^{1/2}$ with $\mathbf{\Sigma} = (0.5^{|i-j|})_{1\le i,j\le p}$ and $\boldsymbol{z}_t$ follows the settings of (a)-(d). For the Models (v)-(viii), we set $\mathbf{A} = (a_{ij})_{1\le i,j\le p}$ with $a_{ij} \overset{i.i.d}{\sim} U(-1,1)$ and $\boldsymbol{z}_t$ follows the settings of (a)-(d).

Here we use parametric bootstrap method to determine the critical value of L_r; The empirical sizes of the seven test statistics listed above are reported in Tables 1 for $K = 2$. The simulation results with $K = 4, 6$ are in the Supplementary Material. In each table, the results are summarized for Models (i)-(viii) with different combinations of n and p, i.e., $n = 100, 200$ and $p = 30, 60, 120, 240$. It is easy to see that the degenerate U-statistics L_D, L_R and L_{τ^*} can control the sizes very well in most of the cases. However, the empirical sizes of the sum-type statistic S_r, the simple linear rank statistic L_ρ and non-degenerate U-statistic L_τ are a little smaller than the nominal level. The parametric bootstrap method proposed by Chang et al. (2017) can control the empirical sizes of L_r in most cases. In the supplemental material, we also proposed a permutation procedure to control the empirical sizes of the above rank-based test statistics. Additional simulation studies show the good performance of the permutation procedure.

3.2 Power comparison

We consider the following eight examples as the data generation procedure in order to investigate the powers of different test statistics. Let

Table 1: Sizes of tests with $K = 2$ under Model (i)-(viii).

n	p	i							ii						
		L_r	L_τ	L_ρ	L_{τ^*}	L_D	L_R	S_r	L_r	L_τ	L_ρ	L_{τ^*}	L_D	L_R	S_r
100	30	0.043	0.013	0.014	0.054	0.067	0.05	0.02	0.045	0.016	0.015	0.046	0.065	0.04	0.027
100	60	0.036	0.01	0.008	0.052	0.073	0.041	0.018	0.039	0.016	0.011	0.047	0.076	0.031	0.015
100	120	0.034	0.007	0.006	0.029	0.073	0.023	0.002	0.042	0.009	0.007	0.033	0.068	0.026	0.004
100	240	0.044	0.009	0.009	0.034	0.081	0.027	0	0.043	0.009	0.005	0.04	0.082	0.028	0
200	30	0.035	0.017	0.016	0.045	0.05	0.04	0.034	0.056	0.016	0.015	0.04	0.056	0.035	0.042
200	60	0.038	0.022	0.023	0.045	0.054	0.046	0.025	0.042	0.014	0.016	0.034	0.044	0.03	0.027
200	120	0.039	0.012	0.008	0.032	0.056	0.025	0.015	0.046	0.015	0.018	0.04	0.044	0.034	0.008
200	240	0.043	0.008	0.01	0.037	0.06	0.032	0.005	0.047	0.018	0.014	0.051	0.067	0.046	0
		iii							iv						
100	30	0.022	0.009	0.007	0.03	0.052	0.026	0.085	0.038	0.011	0.012	0.044	0.061	0.034	0.038
100	60	0.024	0.006	0.004	0.034	0.05	0.027	0.089	0.042	0.008	0.005	0.028	0.057	0.022	0.05
100	120	0.032	0.01	0.011	0.044	0.087	0.029	0.055	0.061	0.011	0.009	0.055	0.097	0.042	0.057
100	240	0.036	0.033	0.01	0.035	0.086	0.03	0.034	0.054	0.009	0.008	0.04	0.088	0.028	0.035
200	30	0.051	0.014	0.015	0.041	0.056	0.04	0.08	0.037	0.018	0.019	0.044	0.055	0.046	0.071
200	60	0.035	0.015	0.014	0.04	0.054	0.035	0.08	0.043	0.019	0.018	0.048	0.059	0.045	0.075
200	120	0.036	0.016	0.011	0.046	0.062	0.039	0.086	0.047	0.016	0.017	0.039	0.066	0.039	0.056
200	240	0.042	0.021	0.016	0.041	0.062	0.035	0.034	0.036	0.014	0.012	0.045	0.068	0.041	0.062
		v							vi						
100	30	0.044	0.018	0.017	0.049	0.067	0.041	0.031	0.051	0.017	0.012	0.049	0.067	0.043	0.033
100	60	0.038	0.015	0.013	0.05	0.082	0.044	0.015	0.043	0.011	0.012	0.042	0.066	0.039	0.016
100	120	0.035	0.008	0.006	0.034	0.073	0.026	0.001	0.036	0.011	0.007	0.035	0.076	0.025	0.004
100	240	0.044	0.012	0.009	0.045	0.097	0.036	0	0.029	0.009	0.003	0.045	0.091	0.027	0
200	30	0.057	0.013	0.012	0.031	0.046	0.031	0.055	0.061	0.018	0.022	0.054	0.063	0.044	0.035
200	60	0.037	0.009	0.008	0.038	0.053	0.034	0.038	0.043	0.014	0.013	0.043	0.06	0.043	0.029
200	120	0.044	0.018	0.017	0.049	0.075	0.043	0.021	0.039	0.016	0.012	0.046	0.061	0.04	0.016
200	240	0.037	0.008	0.009	0.032	0.065	0.034	0.008	0.052	0.014	0.016	0.052	0.068	0.048	0.003
		vii							viii						
100	30	0.041	0.01	0.009	0.037	0.056	0.031	0.09	0.053	0.013	0.015	0.055	0.072	0.05	0.047
100	60	0.038	0.011	0.009	0.045	0.07	0.037	0.063	0.041	0.010	0.006	0.041	0.075	0.035	0.057
100	120	0.037	0.017	0.008	0.046	0.087	0.036	0.055	0.045	0.008	0.008	0.035	0.075	0.026	0.04
100	240	0.053	0.009	0.005	0.034	0.085	0.022	0.023	0.038	0.009	0.005	0.042	0.094	0.028	0.037
200	30	0.061	0.024	0.024	0.053	0.063	0.053	0.087	0.048	0.011	0.013	0.032	0.04	0.03	0.067
200	60	0.036	0.02	0.017	0.049	0.066	0.046	0.075	0.045	0.015	0.016	0.046	0.058	0.046	0.058
200	120	0.048	0.021	0.02	0.05	0.077	0.042	0.064	0.039	0.019	0.016	0.055	0.081	0.047	0.063
200	240	0.049	0.01	0.011	0.039	0.056	0.036	0.041	0.054	0.02	0.023	0.041	0.063	0.036	0.056

$\boldsymbol{z}_t \sim N(\mathbf{0}, \mathbf{I}_p)$. In the following, with slight abuse of notation, we write $f(v) = (f(v_1), \ldots, f(v_p))^\top$ for any univariate function $f : \mathbb{R} \to \mathbb{R}$ and $v = (v_1, \ldots, v_p)^\top \in \mathbb{R}^p$. That is, (I) $\boldsymbol{\varepsilon}_t = \mathbf{A}\boldsymbol{\varepsilon}_{t-1} + \boldsymbol{z}_t$; (II) $\boldsymbol{\varepsilon}_t = \sin(\frac{2\pi}{3}\mathbf{A}\boldsymbol{\varepsilon}_{t-1}) + \boldsymbol{z}_t$; (III) $\boldsymbol{\varepsilon}_t = \sin(\frac{\pi}{3}(\mathbf{A}\boldsymbol{\varepsilon}_{t-1})^{1/3}) + \boldsymbol{z}_t$; (IV) $\boldsymbol{\varepsilon}_t = (\mathbf{A}\boldsymbol{\varepsilon}_{t-1})^{1/3} + \boldsymbol{z}_t$; (V) $\boldsymbol{\varepsilon}_t = \boldsymbol{z}_t + \mathbf{A}\boldsymbol{z}_{t-1}$; (VI) $\boldsymbol{\varepsilon}_t = \boldsymbol{z}_t + \sin(\frac{2\pi}{3}\mathbf{A}\boldsymbol{z}_{t-1})$; (VII) $\boldsymbol{\varepsilon}_t = \boldsymbol{z}_t + \sin(\frac{\pi}{3}(\mathbf{A}\boldsymbol{z}_{t-1})^{1/3})$; (VIII) $\boldsymbol{\varepsilon}_t = \boldsymbol{z}_t + (\mathbf{A}\boldsymbol{z}_{t-1})^{1/3}$.

We consider $\mathbf{A} = (a_{ij})_{1\le i,j\le p}$ with $a_{ij} \sim U(-\rho, \rho)$ if $1 \le i, j \le k_0$ and $a_{ij} = 0$ otherwise. Models (I) and (V) can be classified as the linear relationship in autocorrelations, while Models (IV) and (VIII) can be classified as the monotone relationship in autocorrelations. The rest of the models are the non-linear and non-monotone relationships in autocorrelation.

Figures 1 report the power curves with different ρ for $K = 2$. The power curves with $K = 4, 6$ are in the Supplementary Material. For these three figures, we set $k_0 = 2$, $n = 100$ and $p = 30$. It is clear that the parameter ρ controls the level of the autocorrelation. Therefore, as ρ increasing, the power curves show the upward trend as well for most of the models. Moreover, the power curves of the degenerate U-statistics L_D, L_R and L_{τ^*} are higher than that of other max-type test statistics, i.e., L_r, L_ρ and L_τ. The sum-type test statistic S_r has the lowest power curve in most cases. It is not surprising because k_0 was set as 2 here and sum-type test cannot work well under the sparse alternatives.

Figure 2 shows the power curves with different k_0. For fixed p and ρ, the parameter k_0 is used to control the sparsity of the autocorrelations. The higher value of k_0 yields the lower level of the sparsity in the autocorrelations. As expected, the power curves of the max-type test statistics, i.e., L_r, L_ρ, L_τ, L_D, L_R and L_{τ^*} have the downward trend when k_0 increasing in most of the models. Moreover, among the six max-type test statistics, the power curves of the degenerate U-statistics L_D, L_R and L_{τ^*} are relatively higher than that of the other three test statistics. In contrast, the power curve of the sum-type test statistic S_r has the upward trend as k_0 increasing in most models.

Figure 3 shows the power curves with different p. For fixed ρ and k_0, as the parameter p increasing, the signal strength tends to decrease. Therefore, it is not surprising that all power curves of the seven test statistics show the downward trend as p increasing. The power curves of degenerate U-statistics L_D, L_R and L_{τ^*} are the highest in the seven test statistics. In contrast, the power curve of the sum-type test statistic S_r is the lowest among the seven statistics.

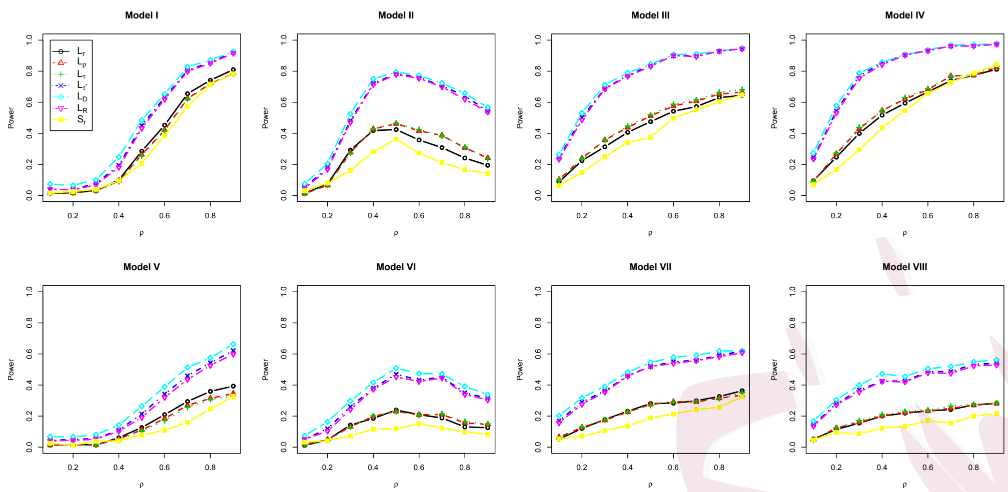

Figure 1: Power curves of different methods with different ρ and $k_0 = 2, n = 100, p = 30, K = 2$.

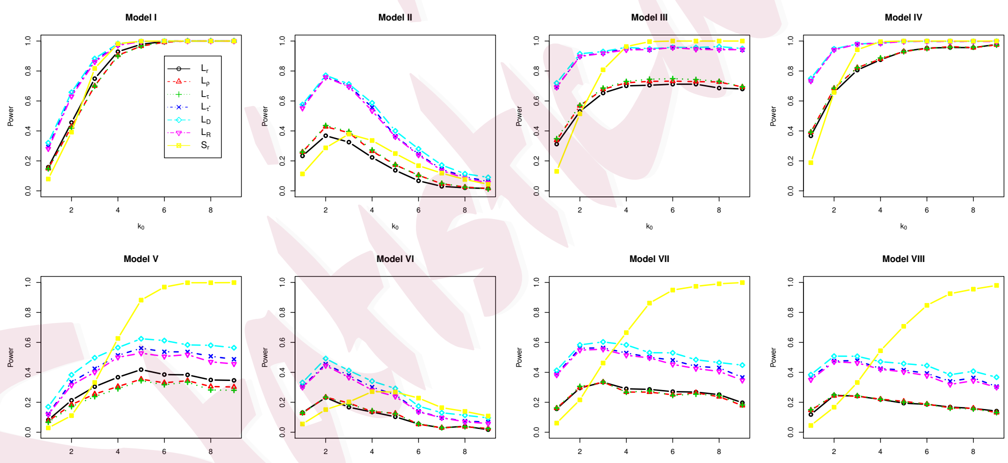

Figure 2: Power curves of different methods with different k_0 and $\rho = 0.6, n = 100, p = 30, K = 2$.

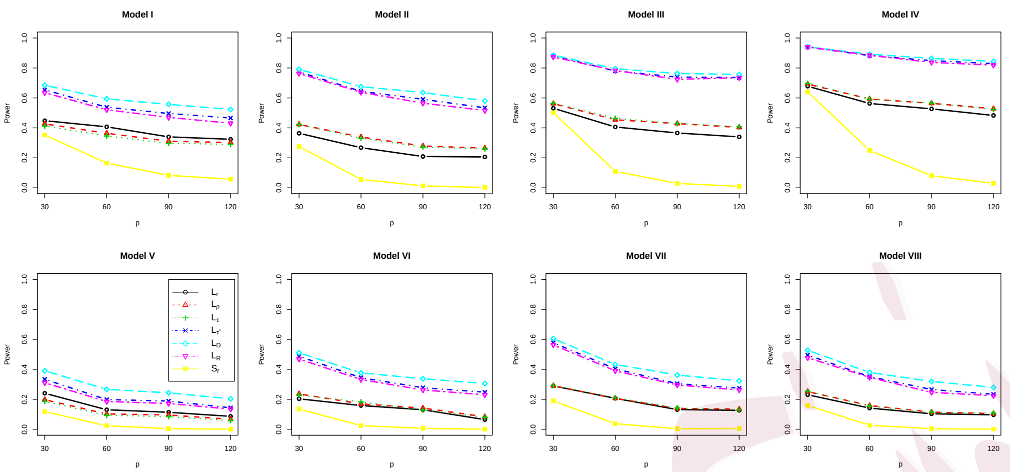

Figure 3: Power curves of different methods with different p and $\rho = 0.6, n = 100, k_0 = 2, K = 2$.

4. Conclusion

To test the high-dimensional white noise, we develop the max-type tests based on three families of rank based statistics, including the simple linear rank statistics, non-degenerate U-statistics and degenerate U-statistics. The proposed tests are distribution free and in particular, the degenerate U-statistics can be used to detect the non-linear and non-monotone relationships in autocorrelations. Finally, as the theoretical contribution of this paper, we have relaxed the cross-sectional independence assumption in existing literature when deriving the asymptotic distributions for the rank correlation statistics. From the simulation studies, we found that the power of degenerate U-statistics L_D, L_R, L_{τ^*} have the best performance. So we suggest the degenerate U-statistics proposed in subsection 2.3 in practice.

For the future directions related to the high-dimensional white noise test, it is also important to develop the theory for the sum-type tests based on the rank based statistics. The asymptotic independence between the max-type test and sum-type test based on the rank based statistics is also necessary to be established because of its usefulness in constructing some combination test which can be robust to both sparse and dense alternatives.

Acknowledgement

The research of Dachuan Chen is supported by the National Natural Science Foundation of China (Grants 12101335 and 12271271), the Natural Science Foundation of Tianjin (Grant 21JCQNJC00020), the Fundamental Research Funds for the Central Universities, Nankai University (Grants 63211088, 63221050, and 63231013) and Wukong Investment Research Funds. Long Feng was partially supported by Shenzhen Wukong Investment Company, the Fundamental Research Funds for the Central Universities under Grant No. ZB22000105 and 63233075, the China National Key R&D Program (Grant Nos. 2019YFC1908502, 2022YFA1003703, 2022YFA1003802, 2022YFA1003803) and the National Natural Science Foundation of China

Grants (Nos. 12271271, 11925106, 12231011, 11931001 and 11971247). Fengyi Song and Long Feng are co-corresponding authors and equally contributed to this paper.

References

M. A. Arcones and E. Giné. Limit theorems for u-processes. *The Annals of Probability*, pages 1494–1542, 1993.

W. Bergsma and A. Dassios. A consistent test of independence based on a sign covariance related to kendall tau. *Bernoulli*, 20(2):1006–1028, 2014.

J. R. Blum, J. Kiefer, and M. Rosenblatt. *Distribution free tests of independence based on the sample distribution function.* Sandia Corporation, 1961.

J. Chang, Q. Jiang and X. Shao Testing the martingale difference hypothesis in high dimension, *Journal of Econometrics*, 235(2): 972–1000,2023.

J. Chang, Q. Yao, and W. Zhou. Testing for high-dimensional white noise using maximum cross-correlations. *Biometrika*, 104(1):111–127, 2017.

S. Chatterjee, A new coefficent of correlation. *Journal of the American Statistical Association*, 116(536), 2009-2022, 2021.

M. Drton, F. Han, and H. Shi. High-dimensional consistent independence testing with maxima of rank correlations. *The Annals of Statistics*, 48(6):3206–3227, 2020.

L. Feng, T. Jiang, P. Li, and B. Liu. Asymptotic independence of the sum and maximum of dependent random variables with applications to high-dimensional tests. *arXiv 2205.01638*, 2022a.

L. Feng, B. Liu, and Y. Ma. Testing for high-dimensional white noise. *arXiv 2211.02964*, 2022b.

F. Han, S. Chen, and H. Liu. Distribution-free tests of independence in high dimensions. *Biometrika*, 104(4):813–828, 2017.

J. Hájek, Z. Sidak and P. K. Sen. *Theory of Rank Tests*. New York: Academic Press, 2nd ed, 1999.

W. Hoeffding. A non-parametric test of independence. *The annals of mathematical statistics*, pages 546–557, 1948.

J. R. Hosking. The multivariate portmanteau statistic. *Journal of the American Statistical Association*, 75(371):602–608, 1980.

W. C. M. Kallenberg. Cramér type large deviations for simple linear rank statistics. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete* 60, 403–9, 1982.

W. Li and A. McLeod. Distribution of the residual autocorrelations in multivariate arma time series models. *Journal of the Royal Statistical Society: Series B (Methodological)*, 43(2):

231–239, 1981.

W. K. Li. *Diagnostic checks in time series.* Chapman and Hall/CRC, 2004.

Z. Li, C. Lam, J. Yao, and Q. Yao. On testing for high-dimensional white noise. *The Annals of Statistics*, 47(6):3382–3412, 2019.

H. Lütkepohl. *New introduction to multiple time series analysis.* Springer Science & Business Media, 2005.

T. Malevich and B. Abdalimov. Large deviation probabilities for u-statistics. *Theory of Probability & Its Applications*, 24(1):215–219, 1979.

P. Nandy, L. Weihs, and M. Drton. Large-sample theory for the bergsma-dassios sign covariance. *Electronic Journal of Statistics*, 10(2):2287–2311, 2016.

R. S. Tsay. Testing serial correlations in high-dimensional time series via extreme value theory. *Journal of Econometrics*, 216(1):106–117, 2020.

L. Weihs, M. Drton, and N. Meinshausen. Symmetric rank covariances: a generalized framework for nonparametric measures of dependence. *Biometrika*, 105(3):547–562, 2018.

T. Yanagimoto. On measures of association and a related problem. *Annals of the Institute of Statistical Mathematics*, 22(1):57–63, 1970.

A. Y. Zaitsev. On the gaussian approximation of convolutions under multidimensional analogues of sn bernstein's inequality conditions. *Probability theory and related fields*, 74(4):535–566, 1987.

V. M. Zolotarev. Concerning a certain probability problem. *Theory of Probability & Its Applications*, 6(2):201–204, 1962.

School of Statistics and Data Science, KLMDASR, LEBPS, and LPMC, Nankai University

E-mail: dchen@nankai.edu.cn

School of Statistics and Data Science, KLMDASR, LEBPS, and LPMC, Nankai University

E-mail: sauntbai@163.com

School of Statistics and Data Science, KLMDASR, LEBPS, and LPMC, Nankai University

E-mail: flnankai@nankai.edu.cn