Measuring and testing independence for multivariate time series by auto multivariate distance covariance

Xuejun Ma* Yao Dong† Jingli Wang‡

Abstract

We propose an auto multivariate distance covariance for time series, which extends the concept of joint high distance covariance. Base on the proposed auto multivariate distance covariance, we develop two new procedures for testing mutual dependence for multivariate time series, which can be taken as combinations of the auto multivariate distance covariance and Box and Pierce (1970) test and Li and McLeod (1981) test. Simulation results indicate that the proposed methods are very effective compared with other similar approaches. The proposed methods are illustrated by analyzing the relationships of real gross domestic product of the United Kingdom, Canada, and the United States.

Keywords: auto joint high distance covariance; multivariate time series; mutual independence.

MSC2010 subject classifications: Primary 62G10; secondary 62H20.

1 Introduction

In time series analysis, a fundamental problem is to explore serial dependence. The auto correlation function (ACF) plays a central role in the univariate time series, which can measure the linear relationship between variables at fixed lags. Based on ACF, many researchers developed a variety of correlation-based statistics to test the serial dependence such as Box and Pierce (1970) and Ljung and Box (1978). As is known, ACF can not detect the nonlinear relationship since it is established via the Pearson correlation which can only measure the linear relationship between variables. For multivariable time series, there are also many statistics to test the pairwise independence. Chitturi (1974) and Hosking (1980) have extended the univariate portmanteau test (Box and Pierce, 1970) to multivariate time series. Li and McLeod (1981) modified the multivariate portmanteau test to improve the accuracy in small sample. Based on the determinant of

*School of Mathematical Sciences, Soochow University, Suzhou, China. xuejunma@suda.edu.cn

†School of Mathematical Sciences, Soochow University, Suzhou, China. 20184207042@stu.suda.edu.cn

‡The corresponding author. School of Statistics and Data Sciences, Nankai University, Tianjin, China. jlwang@nankai.edu.cn

the standardized multivariate residual autocorrelations, Mahdi and McLeod (2012) proposed a generalized variance test which is also an extension of the univariate portmanteau test. However, these methods are used for testing the linear relationships, not nonlinear relationships.

To test the nonlinear dependence, Zhou (2012) proposed an auto-distance correlation function (ADCF) which is an extension of the distance correlation function (Szekely et al., 2007). It is a nonparametric method and tests the independence between two random vectors via the distance correlation between the joint characteristic function and the product of marginal characteristic functions. Fokianos and Pitsillou (2017) combined ADCF and the generalized spectral density approach of Hong (1999) to test pairwise dependence for both linear and nonlinear time series. Fokianos and Pitsillou (2018) developed the matrix multivariate auto-distance covariance function as a statistic with the help of the generalized spectral density approach (Hong, 1999). Recently, Chakraborty and Zhang (2019) proposed a high-order distance covariance to measure the Lancaster interaction dependence and applied it to causal inference. Obviously, above approaches are based on auto-correlation matrix or auto-distance covariance matrix, and the elements of these matrices are the covariance or correlation between two variables.

It is known that pairwise independent collections are not mutually independent, but any collection of mutually independent random variables are pairwise independent. There is a well-known example given by Stojanov (1987). For any time t, η_t and ξ_t are two independent variables uniformly distributed on the interval $(0, 1)$. Let $X_{t1} = \tan(\eta_t), X_{t2} = \tan(\xi_t), X_{t3} = -\tan(\eta_t+\xi_t)$, X_{t1} and X_{t2} are independent because of the independence of η_t and ξ_t. From the distribution of X_{t3}, we can get that X_{t3} and X_{t1} are independent, as are X_{t3} and X_{t2}. However, X_{t1}, X_{t2} and X_{t3} are not mutually independent because $X_{t1} + X_{t2} + X_{t3} = X_{t1}X_{t2}X_{t3}$. Therefore, for multivariable time series, not limited to pairwise independence, mutually independence should be considered.

This paper is concerned with serial independence, especially, in the setting of multivariable time series. We will introduce the concepts of auto high order distance covariance (AHDCov) and auto joint high order conditional distance covariance (AJHCDCov), which are extended the notions of high order distance covariance and joint distance covariance introduced by Chakraborty and Zhang (2019). AJHCDCov can test both linear and nonlinear dependence for multivariate time series at fixed lags. Furthermore, we combine AJHCDCov with Box and Pierce (1970) and Li and McLeod (1981) test, respectively, to test independence for multivariate time series.

The rest of this paper is organized as follows. In Section 2, we introduce two auto multivariate distance covariance: AHDCov and AJHCDCov. Asymptotic properties of the resulting estimators are also established in this section with the proofs sketched in the Appendix. Section 3 is devoted to two statistics for testing independence of multivariate time series. Sections 4 and 5 demonstrate the simulation studies and a real data analysis.

2 Method

2.1 Some Preliminaries

Let $\{X_t, t \in \mathbb{Z}\}$ be a p-dimensional stationary time series, where $X_t = (X_{t,1}, \ldots, X_{t,p})^\top$. $\phi_Z(u)$ be the characteristic function of $Z, u \in \mathbb{R}$, that is $\phi_Z(u) = E[\exp(\imath\langle u, Z\rangle)]$, where $\imath$ is the imaginary unit. Define $\omega_p(u) = (c_1|u|_p^{p+1})^{-1}$ with $c_p = \pi^{(1+p)/2}/\Gamma((1+p)/2)$, where $\Gamma(\cdot)$ is the gamma function, $|\cdot|_p$ is the Euclidean norm of $\mathbb{R}^p$. Let $d\omega = (c_1 \times \cdots \times c_p \times |u_1|^2 \times \cdots \times |u_p|^2)^{-1} dt_1 \times \cdots \times dt_p$. Denote by I_k^q the collection of k-tuples of indices from $\{1, \ldots, q\}$ such that all indices are different.

Definition 2.1. For $j > 0$, the auto high order distance covariance (AHDCov) between X_t and X_{t+j} is defined as

$$\begin{aligned}\mathcal{D}(j) &= \mathrm{HDC}^2(X_{t,1}, \ldots, X_{t,p}, X_{t+j,1}, \ldots, X_{t+j,p}) \\ &= \int_{\mathbb{R}^{2p}} \Big| E\Big\{ \Big[\prod_{i=1}^{p} \Big(\phi_{X_{t,i}}(u_i) - \exp(\imath\langle u_i, X_{t,i}\rangle)\Big)\Big(\phi_{X_{t+j,i}}(u) - \exp(\imath\langle u_i, X_{t+j,i}\rangle)\Big)\Big]\Big\}\Big|^2 d\omega \quad (2.1)\end{aligned}$$

$\mathcal{D}(j)$ can measure both linear and nonlinear relationships among the variables $X_{t,1}, \ldots, X_{t,p}$, $X_{t+j,1}, \ldots, X_{t,p}$. When $p = 1$, $\mathcal{D}(j)$ is the autodistance covariance between X_t and X_{t+j} (Zhou, 2012; Fokianos and Pitsillou, 2017). As we known, when autodistance covariance equals to zero, X_t and X_{t+j} are independent. For $\mathcal{D}(j)$, if X_t and X_{t+j} are independent, $\mathcal{D}(j)$ is zero, but the converse doesn't hold. It motives us to study the auto joint multivariate distance covariance (AJMDC). For the sake of presentation, denote $Y_{t,1} = X_{t,1}, \ldots, Y_{t,p} = X_{t,p}, Y_{t,p+1} = X_{t+j,1}, \ldots, Y_{t,2p} = X_{t+j,p}$.

Definition 2.2. For $j > 0$, the auto joint high order distance covariance (AJHDCov) between X_t and X_{t+j} is defined as

$$\begin{aligned}\mathcal{J}(j) &= \mathrm{JHDC}^2(X_{t,1}, \ldots, X_{t,p}, X_{t+j,1}, \ldots, X_{t+j,p}) \\ &= \mathrm{JHDC}^2(Y_{t,1}, \ldots, Y_{t,p}, Y_{t,p+1}, \ldots, Y_{t,2p}) \\ &= C_2 \sum_{i_1, i_2 \in I_2^{2p}} \mathrm{HDC}^2(Y_{t,i_1}, Y_{t,i_2}) + C_3 \sum_{i_1, i_2, i_3 \in I_3^{2p}} \mathrm{HDC}^2(Y_{t,i_1}, Y_{t,i_2}, Y_{t,i_3}) \\ &\quad + \cdots + C_{2p}\mathrm{HDC}^2(Y_{t,1}, \ldots, Y_{t,2p}) \quad (2.2)\end{aligned}$$

for some specific constants $C_i > 0$ with $2 \le i \le 2p$. Besides, $\mathcal{J}(0) = \mathrm{JHDC}^2(X_{t,1}, \ldots, X_{t,p})$.

From Proposition 2.3 of Chakraborty and Zhang (2019), $\mathrm{JHDC}(Y_{t,1}, \ldots, Y_{t,p}, Y_{t,p+1}, \ldots, Y_{t,2p}) = 0$ if and only if $Y_{t,1}, \ldots, Y_{t,p}$, $Y_{t,p+1}, \ldots, Y_{2p}$ are mutually independent, which implies $\mathcal{J}(j)$ can test mutual independence for $X_{t,1}, \ldots, X_{t,p}, X_{t+j,1}, \ldots, X_{t+j,p}$.

Now, we provide a brief discussion on the relationship between $\mathcal{J}(j)$ and auto-distance covariance matrix $\mathcal{V}(j) = \{\mathcal{V}_{rm}(j)\}_{r,m=1}^p$ of Fokianos and Pitsillou (2018), where $\mathcal{V}_{rm}(j) = \mathrm{HDC}^2(X_{t,r}, X_{t+j,m})$. We have the following comments:

(1) $\mathcal{V}_{rm}(0)$ measures concurrent dependence between $X_{t,r}$ and $X_{t,m}$. For any $r, m = 1, \ldots, p$, $X_{t,r}$ and $X_{t,m}$ are independent if and only all elements are zero in $\mathcal{V}_{rm}(0)$ (i.e., $\mathcal{V}(j) = 0$).

If there is a nonzero element (say, $\mathcal{V}(j) \neq 0$), there exist r' and m' so that $X_{t,r'}$ and $X_{t,m'}$ are dependent. It implies that $\mathcal{V}(0)$ measures the pairwise dependence among $X_{t,1}, \dots, X_{t,p}$. However, when $\mathcal{J}(0) = 0$, $X_{t,1}, \dots, X_{t,p}$ are mutual independent.

(2) $\mathcal{V}_{rm}(j)$ measures the dependence between $X_{t,r}$ and $X_{t+j,m}$. Thus, when $\mathcal{V}_{rm}(j) = 0$, $X_{t,r}$ and $X_{t+j,m}$ has no j-lag relation. It implies $\mathcal{V}(j)$ measures the pairwise independence between $X_{t,r}$ and $X_{t+j,m}$ for $r, m = 1, \dots, p$. However, as comments (1), if $\mathcal{J}(j) = 0$, we can conclude that $X_{t,1}, \dots, X_{t,p}, X_{t+j,1}, \dots, X_{t+j,p}$ are mutual independent.

It should be noted that $\mathcal{V}(j)$ and $\mathcal{V}(0)$ are matrices, not numbers. In practice, matrices may lead to very complicated calculation. Fokianos and Pitsillou (2018) constructed a portmanteau statistic to transform these matrices to a scaled value. Delightedly, $\mathcal{J}(j)$ can not only measure mutual independence, but also be a value, which has a similar form as the autocovariance function of univariate time series.

For the sake of presentation, given $x, x' \in \mathbb{R}$, we define the conditional bivariate function

$$U_{t,i}(x, x') = E(|x - X'_{t,i}|) + E(|X_{t,i} - x'|) - |x - x'| - E(|X_{t,i} - X'_{t,i}|), \tag{2.3}$$

where $X'_{t,i}$ is an independent copy of $X_{t,i}$. From Chakraborty and Zhang (2019), $U_{t,i}(x, x')$ can be rewritten as

$$U_{t,i}(x, x') = \int_R \Big\{[\phi_{X_{t,i}}(u) - \exp(\imath\langle u, x\rangle)]\overline{[\phi_{X'_{t,i}}(u) - \exp(\imath\langle u, x'\rangle)]}\Big\}\omega(t)du. \tag{2.4}$$

By plugging (2.3) and (2.4) into (2.1) and (2.2), we have

$$\mathcal{D}(j) = E\Big[\prod_{i=1}^{p} U_{t,i}(X_{t,i}, X'_{t,i})U_{t+j,i}(X_{t+j,i}, X'_{t+j,i})\Big] \tag{2.5}$$

$$\mathcal{J}(j) = E\Big[\prod_{i=1}^{p}(U_{t,i}(X_{t,i}, X'_{t,i})U_{t+j,i}(X_{t+j,i}, X'_{t+j,i}) + c)\Big] - c^{2p} \tag{2.6}$$

Now, we consider the sample $\{(X^k_{t,1}, \dots, X^k_{t,p})\}^n_{k=1}$. Let $Y^k_{t,1} = X^k_{t,1}, \dots, Y^k_{t,p} = X^k_{t,p}$, $Y^{t,k}_{t,p+1} = X^{k+j}_{t,1}, \dots, Y^k_{t,2p} = X^{k+j}_{t,p}, k = 1, \dots, n - j$. The estimator of $\phi_{X_i|Z}(t_i)$ is defined by

$$\phi^n_{Y_{t,i}}(u_i) = \frac{1}{n-j}\sum_{k=1}^{n-j}\exp(\imath\langle u_i, Y^k_{t,i}\rangle).$$

Then the empirical estimator of $\mathcal{D}(j)$ can be presented as

$$\begin{aligned}\widehat{\mathcal{D}}(j) &= \mathrm{HCDCov}^2_n(Y_{t,1}, \dots, Y_{t,2p}) \\ &= \int_{R^{2p}} \Big|\frac{1}{n-j}\sum_{k=1}^{n-j}\Big\{\prod_{i=1}^{2p}\Big[\phi^n_{Y_{t,i}}(u_i) - \exp(\imath\langle u_i, Y^k_{t,i}\rangle)\Big]\Big\}\Big|^2 d\omega. \end{aligned}\tag{2.7}$$

Furthermore, we have

$$
\begin{aligned}
\widehat{\mathcal{J}}(j) = c^{2p-2} & \sum_{i_1,i_2 \in I_2^{2p}} \mathrm{HCDCov}_n^2(Y_{t,i_1}, Y_{t,i_2}) \\
& + c^{2p-3} \sum_{i_1,i_2,i_3 \in I_3^{2p}} \mathrm{HCDCov}_n^2(Y_{t,i_1}, Y_{t,i_2}, Y_{t,i_3}) \\
& + \cdots + \mathrm{HCDCov}_n^2(Y_{t,1}, \ldots, Y_{t,2p}). \qquad (2.8)
\end{aligned}
$$

The following theorem show the asymptotic properties of two estimators

Theorem 1. *Suppose that Conditions 1 and 2 hold (refer to Appendix), then*

$$
\begin{aligned}
\widehat{\mathcal{D}}(j) & \xrightarrow{a.s} \mathcal{D}(j), \\
\widehat{\mathcal{J}}(j) & \xrightarrow{a.s} \mathcal{J}(j),
\end{aligned}
$$

as $n \to \infty$, where $\xrightarrow{a.s}$ denotes the almost sure convergence.

However, it is not easy to calculate the estimators $\widehat{\mathcal{D}}(j)$ and $\widehat{\mathcal{J}}(j)$ because of the integrals. Following Chakraborty and Zhang (2019), we construct a bias-corrected estimator.

For $1 \le k, l \le n-j$,

$$
\begin{aligned}
\widehat{U}_i(k,l) = & \frac{1}{n-j-2} \sum_{v=1}^{n-j} |Y_{t,i}^k - Y_{t,i}^v| + \frac{1}{n-j-2} \sum_{s=1}^{n-j} |Y_{t,i}^s - Y_{t,i}^l| \\
& - |Y_{t,i}^k - Y_{t,i}^l| - \frac{1}{(n-j-1)(n-j-2)} \sum_{s,v=1}^{n-j} |Y_{t,i}^s - Y_{t,i}^v| \qquad (2.9)
\end{aligned}
$$

when $k \neq l$, and $\widehat{U}_i(k,l) = 0$ when $k = l$. Hence $\sum_{v \neq k} \widehat{U}_i(k,v) = \sum_{u \neq l} \widehat{U}_i(u,l) = 0$, which lead to the double-centered property $E[\widehat{U}_i(Y_{t,i}, Y_{t,i}^{'})|Y_{t,i}] = E[\widehat{U}_i(Y_{t,i}, Y_{t,i}^{'})|Y_{t,i}^{'}] = 0$ for its population counterpart. Let $DCov\widehat{(Y_{t,i}, Y_{t,j})} = \sum_{k \neq l} \widehat{U}_i(k,l)\widehat{U}_j(k,l)/(n(n-3))$ and write $D\widehat{Cov(Y_{t,i})} = DCov\widehat{(Y_{t,i}, Y_{t,i})}$. Then the bias-corrected estimator is defined as

$$
\widehat{\mathcal{J}}(j) = \frac{1}{(n-j)^2} \sum_{k,l=1}^{n-j} \prod_{i=1}^{2p} \Big[\widehat{U}_i(k,l) + c\Big] - c^{2p} \qquad (2.10)
$$

and

$$
\widehat{\mathcal{J}}_S(j) = \frac{1}{(n-j)^2} \sum_{k,l=1}^{n-j} \prod_{i=1}^{2p} \Big[\frac{\widehat{U}_i(k,l)}{D\widehat{Cov(Y_{t,i})}} + c\Big] - c^{2p} \qquad (2.11)
$$

$\widehat{\mathcal{J}}_S(j)$ is designed to eliminate the dimensional effects. c reflects the relative importance of main effect and the higher order effects. We have not present the specific criterion to select c since it is very challenging. We set $c = 0.5$ in the simulations and the real data analysis.

3 Main Results

In the section, we construct test statistics for detecting the dependence for multivariate time series, which is to test the null hypothesis:

$$H_0 : \{X_t\} \text{ is an independent and identically distributed sequence}$$
$$\text{versus}$$
$$H_1 : \text{negation of } H_0$$

From the above, AJHCDCov can measure the mutual independence at fixed lags. Based on Box and Pierce (1970) and Li and McLeod (1981) and AJHCDCov, we propose the following statistics for testing independence for multivariate time series:

$$T_{DBP,n} = n \sum_{j=1}^{m} \widehat{\mathcal{J}}(j), \tag{3.12}$$

$$T^S_{DBP,n} = n \sum_{j=1}^{m} \widehat{\mathcal{J}}_S(j) \tag{3.13}$$

and

$$T_{DLB,n} = n(n+2) \sum_{j=1}^{m} \frac{1}{n-j} \widehat{\mathcal{J}}(j), \tag{3.14}$$

$$T^S_{DLB,n} = n(n+2) \sum_{j=1}^{m} \frac{1}{n-j} \widehat{\mathcal{J}}_S(j) \tag{3.15}$$

where m is the lag.

Since $T_{DBP,n}$, $T^S_{DBP,n}$, $T_{DLB,n}$ and $T^S_{DLB,n}$ are asymmetric statistics, it is impossible to derive their asymptotic distributions. In this paper we use the bootstrap method to estimate the p-values for these statistics. For simplification, let T_n represent the asymmetric statistics $T_{DBP,n}$, $T^S_{DBP,n}$, $T_{DLB,n}$ and $T^S_{DLB,n}$. Let W_n^* be a bootstrap time series sample. Based on W_n^*, we can obtain $T_n(W_n^*)$. In the same way, we can get $\{T_n^1(W_n^*), \ldots, T_n^B(W_n^*)\}$, where B is the replication time. Then the empirical p-value is as follows

$$\hat{p} = \frac{1}{B+1} \sum_{b=1}^{B} I_{[T_n(W_n), +\infty)}(T_n^b(W_n^*)),$$

where

$$I_A(x) = \begin{cases} 1 & \text{if } x \in A, \\ 0 & \text{if } x \notin A. \end{cases}$$

4 Simulations

In this section, we evaluate the finite sample performance of the proposed method, compared with some existing alternatives in the literature, HOSK (Hosking, 1980), LIM (Li and McLeod, 1981), MAH (Mahdi and McLeod, 2012) and MAD (Fokianos and Pitsillou, 2018). The preceding three methods are implemented in the R package `portes`. mADCFtest function of the R package `dCovTS` can be used for the last method. As Fokianos and Pitsillou (2018), we choose the sample size $n = 500$. The nominal level is 0.05. The lag m is $[3n^{\lambda}]$ with $\lambda = 0.1, 0.2$ and 0.3 where $[a]$ denotes the smallest integer not less than a. The bootstrap replication is B=199.

Example 1. $(X_{t1}, X_{t2}, X_{t3})^{\top}$ follows the standard multivariate normal distribution.

Example 2. η_t and ξ_t are i.i.d. from the uniform distribution on the interval $(0, 1)$. Let $X_{t1} = \tan(\eta_t), X_{t2} = \tan(\xi_t), X_{t3} = -\tan(\eta_t + \xi_t)$. According to the section 1, X_{t1}, X_{t2} and X_{t3} are pairwise independent, but not mutually independent.

Example 3. We consider a vector autoregressive model

$$\begin{pmatrix} x_{1t} \\ x_{2t} \\ x_{3t} \end{pmatrix} = \begin{bmatrix} -0.50 & 0.50 & -0.10 \\ 0.60 & 0.20 & 0.30 \\ -0.40 & 0.30 & -0.40 \end{bmatrix} \begin{pmatrix} x_{1,t-1} \\ x_{2,t-1} \\ x_{3,t-1} \end{pmatrix} + \begin{bmatrix} -0.10 & 0.05 & -0.20 \\ -0.20 & 0.20 & 0.30 \\ 0.10 & 0.50 & -0.10 \end{bmatrix} \begin{pmatrix} x_{1,t-2} \\ x_{2,t-2} \\ x_{3,t-2} \end{pmatrix} + \begin{pmatrix} \varepsilon_{t1} \\ \varepsilon_{t2} \\ \varepsilon_{t3} \end{pmatrix}$$

$\varepsilon_{t1}, \varepsilon_{t2}$ and ε_{t3} are i.i.d. from the normal distribution $N(0, 0.5^2)$.

The simulation results for Example 1,2 and 3 are presented in Table 1. From Table 1, we can get the following comments.

(1) $T^S_{DBP,n}$ and $T^S_{DLB,n}$ are simliar, and outperform than $T_{DBP,n}$, and $T_{DLB,n}$ respectively, especially in Example 2, since they scale the data via distance variance.

(2) For Example 1, HOSK, LIM, MAH and MAD have slightly inflated rejection probabilities under the null. Our proposed methods are close to 0.05, which implied they provide better performance in comparison with the other four competitors.

(3) For Example 2, HOSK, LIM, MAH and MAD work not well, especially MAD, since they aim to test the pairwise independence instead of the mutually independence.

(4) Example 3 is pairwise dependence so the performances of all methods are good.

5 Real data

In this section, we consider the quarterly growth rates, in percentages, of real gross domestic product (GDP) of the United Kingdom, Canada, and the United States from the second quarter of 1980 to the second quarter of 2011, say x_{t1}, x_{t2} and x_{t3}. The data were seasonally adjusted and downloaded from the database of Federal Reserve Bank at St. Louis. The GDP were in millions of local currency, and the growth rate denotes the differenced series of log GDP. Tsay (2013) built

Table 1: Empirical size or power of simulation studies.

	Example 1			Example 2			Example 3		
Lag	6	11	20	6	11	20	6	11	20
HOSK	0.075	0.085	0.060	0.190	0.190	0.185	1.000	1.000	1.000
LIM	0.075	0.080	0.060	0.190	0.190	0.190	1.000	1.000	1.000
MAH	0.080	0.105	0.145	0.185	0.210	0.235	1.000	1.000	1.000
MAD	0.055	0.075	0.065	0.000	0.000	0.000	1.000	1.000	1.000
$T_{DBP,n}$	0.050	0.040	0.045	0.630	0.620	0.620	1.000	1.000	1.000
$T^S_{DBP,n}$	0.050	0.040	0.060	0.945	0.940	0.910	1.000	1.000	1.000
$T_{DLB,n}$	0.050	0.040	0.045	0.630	0.620	0.620	1.000	1.000	1.000
$T^S_{DLB,n}$	0.050	0.040	0.060	0.945	0.940	0.910	1.000	1.000	1.000

vector autoregression VAR(2) model following the expression

$$\begin{pmatrix} x_{1t} \\ x_{2t} \\ x_{3t} \end{pmatrix} = \begin{bmatrix} 0.16 \\ 0.00 \\ 0.28 \end{bmatrix} + \begin{bmatrix} 0.47 & 0.21 & 0.00 \\ 0.33 & 0.27 & 0.50 \\ 0.47 & 0.23 & 0.23 \end{bmatrix} \begin{pmatrix} x_{1,t-1} \\ x_{2,t-1} \\ x_{3,t-1} \end{pmatrix} + \begin{bmatrix} 0.00 & 0.00 & 0.00 \\ -0.20 & 0.00 & 0.00 \\ -0.30 & 0.00 & 0.00 \end{bmatrix} \begin{pmatrix} x_{1,t-2} \\ x_{2,t-2} \\ x_{3,t-2} \end{pmatrix} + \begin{pmatrix} \hat{\varepsilon}_{t1} \\ \hat{\varepsilon}_{t2} \\ \hat{\varepsilon}_{t3} \end{pmatrix}$$

Now we test the mutually independence of the residuals by above methods in the literature. Table 2 summarizes the results. HOSK, LIM, MAH and MAD are similar. However our proposed methods reject the null, which implies that there are strong serial correlations in the residuals of the VAR(2) model. It is consistent with the result of Tsay (2013), which suggests to try VAR(4) model.

Table 2: Results of real data.

Lag	HOSK	LIM	MAH	MAD	$T^S_{DBP,n}$	$T^S_{DLB,n}$
1	0.994	0.994	0.994	0.995	0.080	0.080
2	0.823	0.824	0.932	0.270	0.000	0.000
3	0.706	0.706	0.838	0.200	0.005	0.005
4	0.391	0.400	0.584	0.175	0.005	0.005
5	0.615	0.615	0.466	0.160	0.005	0.005
6	0.805	0.797	0.455	0.165	0.035	0.035

Appendix

The following technical conditional conditions are imposed in the section:

Condition 1. X_{ti} is a strictly stationary α-mixing process with mixing coefficients $\alpha(j), \quad j \geq 1, i = 1, \ldots, p$.

Condition 2. $EX_{ti} < \infty, i = 1, \ldots, p$

The above conditions are general, see Fokianos and Pitsillou (2017)

Lemma 1. *For $z_1, \ldots, z_q$ with $n \geq 2$, then for any $r > 1$*

$$\Big|\sum_{i=1}^{n} z_i\Big|^r \leq n^{r-1} \sum_{i=1}^{n} |z_i|^r$$

The lemma can be found in Chakraborty and Zhang (2019).

Proof of Theorem 1 Define

$$\xi_j^n(u_1, \ldots, u_{2p}) = \frac{1}{n-j} \sum_{k=1}^{n-j} \Big\{ \prod_{i=1}^{2p} \Big[\phi_{Y_{t,i}}^{n-j}(u_i) - \exp(\imath \langle u_i, Y_{t,i}^k \rangle) \Big] \Big\}.$$

Then $\widehat{\mathcal{D}}(j)$ can be rewritten as

$$\widehat{\mathcal{D}}(j) = \int_{\mathbb{R}^{2p}} \Big| \xi_j^n(u_1, \ldots, u_{2p}) \Big|^2 d\omega.$$

Let $h_i^k = \exp(\imath \langle u_i, Y_{t,i}^k \rangle) - \phi_{Y_{t,i}}(u_i)$. Then

$$\xi_j^n(u_1, \ldots, u_{2p}) = \frac{1}{n-j} \sum_{k=1}^{n-j} \prod_{i=1}^{p} \Big(\frac{1}{n-j} \sum_{s=1}^{n-j} h_i^s - h_i^k \Big).$$

Since $|h_i^k| \leq 2$ for $k = 1, \ldots, n-j$, we have

$$|\xi_j^n(u_1, \ldots, u_{2p})| \leq 2^{2p}. \tag{A.1}$$

For any given $\delta > 0$, define the region

$$R(\delta) = \{(u_1, \ldots, u_{2p}) : \delta \leq |u_i| \leq 1/\delta, i = 1, \ldots, 2p\}.$$

and

$$\begin{aligned}
\widehat{\mathcal{D}}(j) &= \int_{R(\delta)} |\xi_j^n(u_1, \ldots, u_{2p})|^2 d\omega + \int_{R^c(\delta)} |\xi_j^n(u_1, \ldots, u_{2p})|^2 d\omega \\
&= R_{j,\delta}^{(n1)} + R_{j,\delta}^{(n2)}
\end{aligned}$$

where A^c is the complement of A.

From (A.1), we can get

$$\int_{R(\delta)} |\xi_j^n(u_1, \ldots, u_{2p})|^2 d\omega \leq \int_{R(\delta)} 4^{2p} d\omega < \infty.$$

For each δ, according to the Lebesgue dominated convergence theorem for α-mixing random variables, it follows that almost surely

$$\lim_{n\to\infty} R_{j,\delta}^{(n1)} = R_{j,\delta}^{(1)} = \int_{R(\delta)} |\xi_j(u_1,\dots,u_{2p})|^2 d\omega.$$

Clearly, $R_{j,\delta}^{(1)}$ converges to $\mathcal{D}(j)$ as δ tends to zero. So, it remains to prove that almost surely

$$\limsup_{\delta\to 0}\limsup_{n\to\infty} R_{j,\delta}^{(n2)} = 0. \tag{A.2}$$

Now, we write $R^c(\delta) = \cup_{i=1}^{2p}(A_i^1 \cup A_i^2)$, where $A_i^1 = \{|u_i| < \delta\}$ and $A_i^2 = \{|u_i| > 1/\delta\}$ for $i = 1,\dots,2p$. Then we have

$$R_{j,\delta}^{(n2)} = \int_{R^c(\delta)} |\xi_j^n(u_1,\dots,u_{2p})|^2 d\omega \le \sum_{\substack{i=1,\dots,2p\\ s=1,2}} \int_{A_i^s} |\xi_j^n(u_1,\dots,u_{2p})|^2 d\omega.$$

Some algebra yields that

$$\begin{aligned}
&\xi_j^n(u_1,\dots,u_{2p})\\
&= \sum_{B\subset\{1,\cdots,2p\}} \frac{1}{n-j}\sum_{k=1}^{n-j}\Big[\prod_{i\in B^c}\Big(\frac{1}{n-j}\sum_{s=1}^{n-j} h_i^s\Big)\cdot\prod_{i\in B}(-h_i^k)\Big]\\
&= \sum_{B\subset\{1,\cdots,2p\}} \prod_{i\in B^c}\Big(\frac{1}{n-j}\sum_{s=1}^{n-j} h_i^s\Big)\Big[\frac{1}{n-j}\sum_{k=1}^{n-j}\prod_{i\in B}(-h_i^k)\Big],
\end{aligned}$$

Applied the inequality $|\sum_{j=1}^n r_j|^2 \le n\sum_{j=1}^n r_j^2$, we can get

$$\begin{aligned}
|\xi_j^n(u_1,\dots,u_{2p})|^2 &\le 4^p \sum_{B\subset\{1,2,\cdots,q\}} \Big[\prod_{i\in B^c}\Big(\frac{1}{n-j}\sum_{s=1}^{n-j} h_i^s\Big)\Big]^2\cdot\Big[\frac{1}{n-j}\sum_{k=1}^{n-j}\prod_{i\in B}(-h_i^k)\Big]^2\\
&=: 2^q \sum_{B\subset\{1,2,\cdots,q\}} \Delta_B. \quad \text{(say)}
\end{aligned}$$

Now suppose $B = \{\gamma_1,\dots,\gamma_d\}$ and $B^c = \{\gamma_{d+1},\dots,\gamma_{2q}\}$ where $d = 1,\dots,2p-1$.

$$\begin{aligned}
\Delta_B &= \Big[\frac{1}{n-j}\sum_{k=1}^{n-j}\prod_{i\in B}(-h_i^k)\Big]^2\cdot\Big[\prod_{i\in B^c}\Big(\frac{1}{n-j}\sum_{s=1}^{n-j} h_i^s\Big)\Big]^2\\
&\le \frac{\sum_{k=1}^{n-j}|h_{\gamma_1}^k|^2\dots|h_{\gamma_d}^k|^2}{n-j}\prod_{i\in B^c}\frac{\sum_{s=1}^{n-j}|h_i^s|^2}{n-j}.
\end{aligned} \tag{A.3}$$

From the prooof of Theorem 2 of Szekely et al. (2007), one can obtain that for $i = 1,\dots,2p$

$$\int_R \frac{|h_i^k|^2}{c_i|u_i|^2}du_i \le 2(|Y_{t,i}^k| + E(|Y_{t,i}|)), \tag{A.4}$$

$$\int_{|t_i|<\delta} \frac{|h_i^k|^2}{c_i|u_i|^2}du_i \le 2E_{Y_{t,i}}(|Y_{t,i}^k - Y_{t,i}|)G(|Y_{t,i}^k - Y_{t,i}|\delta), \tag{A.5}$$

$$\int_{|t_i|>1/\delta} \frac{|h_i^k|^2}{c_i|u_i|^2}du_i \le 4\delta, \tag{A.6}$$

where $E_{Y_{t,i}}$ means the expectation with respect to $Y_{t,i}$, and for $z = (z_1, \ldots, z_d) \in \mathbb{R}^d$

$$G(y) = \int_{|z|_d < y} \frac{1 - \cos z_1}{|z|_d^{1+d}} dz$$

which is bounded by c_d and $\lim_{y \to 0} G(y) = 0$. .

Now, we consider integration of Δ_B in $R^c(\delta)$. Given i, we discuss this integration in two cases.

Case 1: $j \in B$. Without loss of generality, we suppose $\gamma_1 = j$. Firstly, we consider the integral on A_j^1. With A.4-A.6, we have

$$\begin{aligned}
&\int_{A_j^1} \Delta_B d\omega \\
&\leq \int_{A_{\gamma_1}^1} \Big(\frac{1}{n-j} \sum_{k=1}^{n-j} |h_{\gamma_1}^k|^2 \ldots |h_{\gamma_d}^k|^2 \Big) \prod_{i \in B^c} \frac{\sum_{k=1}^{n-j} |h_i^k|^2}{n-j} d\omega \\
&= \frac{1}{n-j} \sum_{k=1}^{n-j} \int_{A_{\gamma_1}^1} |h_{\gamma_1}^k|^2 d\omega_{\gamma_1} \int_R |h_{\gamma_2}^k|^2 d\omega_{\gamma_2} \cdots \int_R |h_{\gamma_d}^k|^2 d\omega_{\gamma_d} \times \prod_{i \in B^c} \frac{\sum_{k=1}^{n-j} \int_R |h_i^k|^2 d\omega_i}{n-j} \\
&\leq \frac{1}{n-j} \sum_{k=1}^{n-j} \Big\{ 2E_{Y_{t,\gamma_1}} (|Y_{t,\gamma_1}^k - Y_{t,\gamma_1}) G(|Y_{t,\gamma_1}^k - Y_{t,\gamma_1}|\delta) \times 2\Big[|Y_{t,\gamma_2}^k| + E(|Y_{t,\gamma_2}|)\Big] \\
&\quad \times \cdots \times 2\Big[|Y_{t,\gamma_d}^k| + E(|Y_{t,\gamma_d}|)\Big]\Big\} \times \prod_{i \in B^c} \frac{\sum_{k=1}^{n-j} 2(|Y_{t,\gamma_i}^k| + E(|Y_{t,\gamma_i}|)}{n-j}
\end{aligned} \tag{A.7}$$

So

$$\limsup_{n \to \infty} \int_{A_j^1} \Delta_B d\omega \leq 4^{2p-1} E_{Y_{t,\gamma_1}} (|Y_{t,\gamma_1} - Y'_{t,\gamma_1}|) G(|Y_{t,\gamma_1} - Y'_{t,\gamma_1}|\delta)) \prod_{i \neq 1} E(|Y_{t,\gamma_i}|)$$

Hence, almost surely,

$$\limsup_{\delta \to 0} \limsup_{n \to \infty} \int_{A_j^1} \Delta_B d\omega = 0 \tag{A.8}$$

Now we consider the integral on A_j^2.

$$\begin{aligned}
&\int_{A_j^2} \Delta_B d\omega \\
&\leq \int_{A_{\gamma_1}^2} \frac{\sum_{k=1}^{n-j} |h_{\gamma_1}^k|^2 \ldots |h_{\gamma_d}^k|^2}{n-j} \prod_{i \in B^c} \frac{\sum_{k=1}^{n} |h_i^k|^2}{n-j} d\omega \\
&= \frac{1}{n-j} \sum_{k=1}^{n-j} \Big[\int_{A_{\gamma_1}^2} |h_{\gamma_1}^k|^2 d\omega_{\gamma_1} \int_R |h_{\gamma_2}^k|^2 d\omega_{\gamma_2} \cdots \int_R |h_k^{\gamma_d}|^2 d\omega_{\gamma_d}\Big] \times \prod_{i \in B^c} \frac{\sum_{k=1}^{n-j} \int_R |h_i^k|^2 d\omega_i}{n-j} \\
&= \frac{1}{n-j} \Big[4\delta \times 2(|Y_{t,\gamma_2}^k| + E(|Y_{t,\gamma_2}|) \times \cdots \times 2(|Y_{t,\gamma_d}^k| + E(|Y_{t,\gamma_d}|\Big] \\
&\quad \times \prod_{i \in B^c} \frac{\sum_{k=1}^{n-j} 2(|Y_{t,\gamma_i}^k| + E(|Y_{t,\gamma_i}|)}{n-j}
\end{aligned}$$

Similarly,

$$\limsup_{\delta\to 0}\limsup_{n\to\infty}\int_{A_j^2}\Delta_B d\omega = 0 \tag{A.9}$$

Case 2: $j \notin B$. Following the similar steps as case 1, we can prove

$$\limsup_{\delta\to 0}\limsup_{n\to\infty}\int_{A_j^2}\Delta_B d\omega = 0, \tag{A.10}$$

$$\limsup_{\delta\to 0}\limsup_{n\to\infty}\int_{A_j^1}\Delta_B d\omega = 0. \tag{A.11}$$

Note that for $B = \emptyset$ or $B = \{1, 2, \cdots, 2p\}$,

$$\limsup_{\delta\to 0}\limsup_{n\to\infty}\int_{A_j^2}\prod_{i=1}^{2p}\Big(\frac{1}{n-j}\sum_{s=1}^{n-j}h_i^s\Big)^2 d\omega = 0, \tag{A.12}$$

$$\limsup_{\delta\to 0}\limsup_{n\to\infty}\int_{A_j^1}\prod_{i=1}^{2p}\Big(\frac{1}{n-j}\sum_{s=1}^{n-j}h_i^s\Big)^2 d\omega = 0, \tag{A.13}$$

$$\limsup_{\delta\to 0}\limsup_{n\to\infty}\int_{A_j^2}\Big(\frac{1}{n-j}\sum_{k=1}^{n-j}h_1^k\dots h_{2p}^k\Big)^2 d\omega = 0, \tag{A.14}$$

$$\limsup_{\delta\to 0}\limsup_{n\to\infty}\int_{A_j^1}\Big(\frac{1}{n-j}\sum_{k=1}^{n-j}h_1^k\dots h_{2p}^k\Big)^2 d\omega = 0. \tag{A.15}$$

From (A.8)-(A.15), we can prove (A.2). So we conclude that

$$\widehat{\mathcal{D}}(j) \xrightarrow{a.s} \mathcal{D}(j).$$

With the definition of $\widehat{\mathcal{J}}(j)$,

$$\widehat{\mathcal{J}}(j) \xrightarrow{a.s} \mathcal{J}(j).$$

References

Box, G. and Pierce, D. (1970). Distribution of Residual Autocorrelation in Autoregressive- Integrated Moving Average Time Series Models. Journal of the American Statistical Association, 65, 1509–1526.

Chakraborty, S. and Zhang, X. (2019). Distance metrics for measuring joint dependence with application to causal inference. Journal of the American Statistical Association, 114(528): 1638-1650.

Chitturi, Ratnam V. (1974). Distribution of Residual Autocorrelations in Multiple Autoregressive Schemes. Journal of the American Statistical Association, 69(348), 928-934.

Fokianos, K. and Pitsillou, M. (2017). Consistent testing for pairwise dependence in time series. Technometrics, 59, 262-270.

Fokianos, K. and Pitsillou, M. (2018). Testing independence for multivariate time series via the auto-distance correlation matrix. Biometrika, 105(2), 337-352.

Hong, Y. (1999). Hypothesis testing in time series via the empirical characteristic function: a generalized spectral density approach. Journal of the American Statistical Association, 94(448), 1201–1220.

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

Ljung, G. and Box, G. (1978). On a Measure of Lack of Fit in Time Series Models. Biometrika, 65, 297–303.

Li, W. and McLeod, A. (1981). Distribution of the residual autocorrelations in multivariate ARMA time series models. Journal of the Royal Statistical Society: Series B, 43(2), 231-239.

Mahdi, E. and McLeod, A. (2012). Improved multivariate portmanteau test. Journal of Time Series Analysis, 33(2), 211-222.

Stojanov, M. (1987). Counterexamples in probability (second edition), Wiley.

Szekely, G., Rizzo, M. and Bakirov, N.(2007). Measuring and testing dependence by correlation of distances. The Annals of Statistics, 35(6), 2769-2794.

Tsay S. (2013) Multivariate Time Series Analysis: With R and Financial Applications. Wiley.

Zhou, Z. (2012). Measuring nonlinear dependence in time-series, a distance correlation approach. Journal of Time Series Analysis, 33(3), 438-457.