



**MONASH** University

**Australia**

Department of Econometrics and Business Statistics

<http://www.buseco.monash.edu.au/depts/ebs/pubs/wpapers/>

**Independence Test for High Dimensional Random Vectors**

**G. Pan, J. Gao, Y. Yang and M. Guo**

**January 2012**

**Working Paper 01/12**

# Independence Test for High Dimensional Random Vectors

G. Pan, J. Gao\*, Y. Yang and M. Guo

January 20, 2012

## Abstract

This paper proposes a new mutual independence test for a large number of high dimensional random vectors. The test statistic is based on the characteristic function of the empirical spectral distribution of the sample covariance matrix. The asymptotic distributions of the test statistic under the null and local alternative hypotheses are established as dimensionality and the sample size of the data are comparable. We apply this test to examine multiple MA(1) and AR(1) models, panel data models with some spatial cross-sectional structures. In addition, in a flexible applied fashion, the proposed test can capture some dependent but uncorrelated structures, for example, nonlinear MA(1) models, multiple ARCH(1) models and vandermonde matrices. Simulation results are provided for detecting these dependent structures. An empirical study of dependence between closed stock prices of several companies from New York Stock Exchange (NYSE) demonstrates that the feature of cross-sectional dependence is popular in stock markets.

**Keywords:** Independence test, cross-sectional dependence, empirical spectral distribution, characteristic function, Marcenko-Pastur Law.

**JEL Classifications:** C12, C21, C22.

## 1 Introduction

A prominent feature of data collection nowadays is that the number of variables is comparable with the sample size. This is the opposite of the classical situations where many observations are made on low-dimensional data. Specific examples of such high dimensional data include microarray expression, images, multiuser detection, climate data and financial data (see, for

---

\*Correspondence: Jiti Gao, Department of Econometrics and Business Statistics, Monash University, Caulfield East Victoria 3145, Australia. Email: jiti.gao@monash.edu.

example, Donoho (2000), Fan, Lv and Qi (2011), Johnstone (2001), Johnstone and Titterton (2009)). This type of data trend poses great challenges because traditional multivariate approaches do not necessarily work, which were established for the case of the sample size  $n$  tending to infinity and the dimension  $p$  remaining fixed (See Anderson (1984)). There have been a substantial set of research works dealing with high dimensional data (see, for example, Bai and Saranadasa (1996), Fan and Lv (2010), Huang, Horowitz and Ma (2008), Fan and Li (2001)).

Measuring mutual dependence is important in time series analysis and cross-sectional panel data analysis. While serial dependence can be characterized by the general spectral density function (see Hong (1998); Hong (1999)), mutual dependence is difficult to be described by a single criteria. This paper proposes a new statistic to test mutual dependence for a large number of high dimensional random vectors, including multiple time series and cross-sectional panel data.

Suppose that  $\{X_{ji}, j = 1, \dots, n; i = 1, \dots, p\}$  are real-valued random variables. For  $1 \leq i \leq p$ , let  $\mathbf{x}_i = (X_{1i}, \dots, X_{ni})^T$  denote the  $i$ -th time series and  $\mathbf{x}_1, \dots, \mathbf{x}_p$  be a panel of  $p$  time series, where  $n$  usually denotes the sample size in each of the time series data. In both theory and practice, it is not uncommon to assume that each of the time series  $(X_{1i}, X_{2i}, \dots, X_{ni})$  is statistically independent, but it may be unrealistic to assume that  $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_p$  are independent or even uncorrelated. This is because there is no natural ordering for cross-sectional indices. There are such cases in various disciplines. In economics and finance, for example, it is not unreasonable to expect that there is significant evidence of cross-sectional dependence in output innovations across  $p$  countries and regions in the world. In the field of climatology, there is also some evidence to show that climatic variables in different stations may be cross-sectionally dependent and the level of cross-sectional dependence may be determined by some kind of physical distance. Moreover, one would expect that climatic variables, such as temperature and rainfall variables, in a station in Australia have higher-level dependence with the same type of climatic variables in a station in New Zealand than those in the United States.

In such situations, it may be necessary to test whether  $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_p$  are independent before a statistical model is used to model such data. In the econometrics and statistics literature, several papers have basically considered testing for cross-sectional uncorrelatedness for the residuals involved in some specific regression models. Such studies include Pesaran (2004) for the parametric linear model case, Hsiao, Pesaran and Pick (2007) for the parametric nonlinear case, and Chen, Gao and Li (2009) for the nonparametric nonlinear case. Other related papers include

Su and Ullah (2009) for testing conditional uncorrelatedness through examining a covariance matrix in the case where  $p$  is fixed. As the main motivation of this paper, we will propose using an empirical spectral distribution function based test statistic for cross-sectional independence of  $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_p$ .

The aim is to test

$$\mathbf{H}_0 : \mathbf{x}_1, \dots, \mathbf{x}_p \text{ are independent; against } \mathbf{H}_1 : \mathbf{x}_1, \dots, \mathbf{x}_p \text{ are not independent,} \quad (1.1)$$

where  $\mathbf{x}_i = (X_{1i}, \dots, X_{ni})^T$  for  $i = 1, \dots, p$ .

In time series analysis, mutual independence test for multiple time series has long been of interest. Moreover, time series always display various kinds of dependence. For example, an autoregressive conditional heteroscedastic (ARCH(1)) model involves a martingale difference sequence (MDS); a nonlinear moving average (MA) model is not a MDS, but its autocorrelations are zero; a linear moving average (MA) model and an autoregressive (AR) model are both models with correlated structures. In this paper, we also employ the proposed statistic to test dependence for multiple time series.

Section 8.5 of Anderson (1984) also consider a similar problem but with fixed dimensions. His problem and approach are as follows. Let the  $pm$ -component vector  $\mathbf{x}$  be distributed according to  $N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ . Partition  $\mathbf{x}$  into  $m$  subvectors with  $p$  components respectively, that is,  $(\mathbf{h}_1^T, \dots, \mathbf{h}_p^T)^T$ . The vector of means  $\boldsymbol{\mu}$  and the covariance matrix  $\boldsymbol{\Sigma}$  can be partitioned respectively,  $(\boldsymbol{\mu}_1^T, \dots, \boldsymbol{\mu}_p^T)^T$ , and

$$\boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} & \cdots & \boldsymbol{\Sigma}_{1p} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} & \cdots & \boldsymbol{\Sigma}_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{\Sigma}_{p1} & \boldsymbol{\Sigma}_{p2} & \cdots & \boldsymbol{\Sigma}_{pp} \end{pmatrix}. \quad (1.2)$$

The null hypothesis to be tested is that the subvectors  $\mathbf{h}_1, \mathbf{h}_2, \dots, \mathbf{h}_p$  are mutually independently distributed. If  $\mathbf{h}_1, \mathbf{h}_2, \dots, \mathbf{h}_p$  are independent subvectors,

$$E(\mathbf{x}_i - \boldsymbol{\mu}_i)(\mathbf{x}_j - \boldsymbol{\mu}_j)' = \boldsymbol{\Sigma}_{ij} = \mathbf{0}, \quad \forall i \neq j. \quad (1.3)$$

Thus, the null hypothesis is equivalent to testing  $\mathbf{H}_0 : \boldsymbol{\Sigma}_{ij} = \mathbf{0}, \forall i \neq j$ . This can be stated

alternatively as the hypothesis that  $\Sigma$  is of the form

$$\Sigma_0 = \begin{pmatrix} \Sigma_{11} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \Sigma_{22} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \Sigma_{pp} \end{pmatrix}. \quad (1.4)$$

To tackle the problem above, draw  $n$  observations from the population  $\mathbf{x}$  to form a sample covariance matrix. The likelihood ratio criterion proposed is

$$\lambda = \frac{|\mathbf{Q}|^{\frac{1}{2}n}}{\prod_{i=1}^p |\mathbf{Q}_{ii}|^{\frac{1}{2}n}}, \quad (1.5)$$

where

$$\mathbf{Q} = \begin{pmatrix} \mathbf{Q}_{11} & \mathbf{Q}_{12} & \cdots & \mathbf{Q}_{1p} \\ \mathbf{Q}_{21} & \mathbf{Q}_{22} & \cdots & \mathbf{Q}_{2p} \\ \vdots & \vdots & & \vdots \\ \mathbf{Q}_{p1} & \mathbf{Q}_{p2} & \cdots & \mathbf{Q}_{pp} \end{pmatrix}, \quad (1.6)$$

in which  $\mathbf{Q}_{ik}$  is the sample covariance matrix of the random vector  $\mathbf{h}_i$  and  $\mathbf{h}_k$ .

Our approach essentially uses the characteristic function of the empirical spectral distribution of sample covariance matrices in large random matrix theory. When  $\mathbf{x}_1, \dots, \mathbf{x}_p$  are mutual independent, the limiting spectral distribution (LSD) of the corresponding sample covariance matrix is the Marcenko-Pastur (M-P) law (see, for example, Marcenko and Pastur (1967), Bai and Silverstein (2009)). From this point, any deviation of the LSD from M-P law is evidence of dependence. Indeed, Silverstein (1995) and Bai and Zhou (2008) report the LSD of the sample covariance matrix with correlations in columns and it is different from the M-P law. In dependent but uncorrelated cases, the LSD may also be the M-P law, e.g. ARCH(1) model. However, the proposed test can still capture the dependence by utilizing the nonzero correlation between high-order series. Unlike the Anderson's test, we need not re-draw observations from the set of vectors of  $\mathbf{x}_1, \dots, \mathbf{x}_p$  due to the high dimensionality.

The rest of the paper is organized as follows. Section 2 introduces the proposed test statistic and some related large dimensional random matrix results. We present the asymptotic distributions of the test statistic under both the null and local alternative hypotheses in Section 3. Moreover, its applications to some panel data models with spatial dependent structures are also

illustrated. Section 4 studies a general panel data model whose dependent structure is different from that investigated in Section 3 and develops the asymptotic distribution of the proposed statistic under the null hypothesis for this model. Several simulated results which demonstrate the effectiveness of the proposed test under the two dependent structures are shown in Section 5. Furthermore, this section illustrates simulation results for some dependent but uncorrelated structures, though we have not developed asymptotic theory for these kinds of models. An empirical analysis of daily closed stock prices from NYSE is provided in Section 6. We conclude with a discussion in Section 7. All mathematical proofs are relegated to Section 8.

## 2 Theory and Methodology

Before we establish the main theory and methodology, we first make the following assumptions:

**Assumption 1.** For each  $i = 1, \dots, p$ ,  $X_{1i}, \dots, X_{ni}$  are independent and identically distributed (i.i.d) random variables with mean zero, variance one and finite fourth moment. When  $X_{ji}$ 's are complex random variables we require  $EX_{ji}^2 = 0$ .

**Assumption 2.**  $p = p(n)$  with  $p/n \rightarrow c \in (0, \infty)$ .

We stack  $p$  time series one by one to form a data matrix  $\mathbf{X} = (\mathbf{x}_1, \dots, \mathbf{x}_p)$ . Moreover denote the sample covariance matrix by

$$\mathbf{A}_n = \frac{1}{n} \mathbf{X}^* \mathbf{X}$$

where  $\mathbf{X} = (\mathbf{x}_1, \dots, \mathbf{x}_p) = (\mathbf{y}_1, \dots, \mathbf{y}_n)^T$ ,  $\mathbf{y}_j^T$  denotes the  $j$ -th row of the matrix  $\mathbf{X}$  and  $\mathbf{X}^*$  is the Hermitian transform of the matrix  $\mathbf{X}$ . The empirical spectral distribution (ESD) of the sample covariance matrix  $\mathbf{A}_n$  is defined as

$$F^{\mathbf{A}_n}(x) = \frac{1}{p} \sum_{i=1}^p I(\lambda_i \leq x), \quad (2.1)$$

where  $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_p$  are eigenvalues of  $\mathbf{A}_n$ .

It is well-known that if  $\mathbf{x}_1, \dots, \mathbf{x}_p$  are independent and  $p/n \rightarrow c \in (0, \infty)$  then  $F^{\mathbf{A}_n}(x)$  converges with probability one to the Marcenko-Pastur Law  $F^c(x)$  (see Marcenko-Pastur (1967) and Bai and Silverstein (2009)) whose density has an explicit expression

$$f_c(x) = \begin{cases} \frac{1}{2\pi xc} \sqrt{(b-x)(x-a)}, & a \leq x \leq b; \\ 0, & \text{otherwise}; \end{cases} \quad (2.2)$$

and a point mass  $1 - 1/c$  at the origin if  $c > 1$ , where  $a = (1 - \sqrt{c})^2$  and  $b = (1 + \sqrt{c})^2$ .

When there is some correlation among  $\mathbf{x}_1, \dots, \mathbf{x}_p$ , denote by  $\mathbf{T}_p$  the covariance matrix of the first row,  $\mathbf{y}_1^T$ , of  $\mathbf{X}$ . Then, under Assumption 1, when  $F^{\mathbf{T}_p}(x) \xrightarrow{D} H(x)$ ,  $F^{\mathbf{A}_n}$  converges with probability one to a non random distribution function  $F^{c,H}$  whose Stieltjes transform satisfies (see Silverstein (1995))

$$m(z) = \int \frac{1}{x(1-c-czm(z))-z} dH(x), \quad (2.3)$$

where the Stieltjes transform  $m_G$  for any c.d.f  $G$  is defined as

$$m_G(z) = \int \frac{1}{\lambda - z} dG(\lambda), \quad \text{Im}z > 0 \quad (2.4)$$

and  $G$  can be recovered by the inversion formula

$$G\{[x_1, x_2]\} = \frac{1}{\pi} \lim_{\varepsilon \rightarrow 0^+} \int_{x_1}^{x_2} \text{Im}(m_G(x + i\varepsilon)) dx,$$

where  $x_1$  and  $x_2$  are continuity points of  $G$ .

Moreover, equation (2.3) takes a simpler form when  $F^{c,H}$  is replaced by

$$\underline{F}^{c,H} = (1-c)I_{[0,\infty)} + cF^{c,H}, \quad (2.5)$$

which is the limiting ESD of  $\underline{\mathbf{A}}_n = \frac{1}{n} \mathbf{X} \mathbf{X}^*$ . Its Stieltjes transform

$$\underline{m}(z) = -\frac{1-c}{z} + cm(z) \quad (2.6)$$

has an inverse

$$z = z(\underline{m}) = \frac{1}{\underline{m}} + c \int \frac{x}{1+x\underline{m}} dH(x). \quad (2.7)$$

The construction of our test statistic relies on the following observation: the limit of the ESD of the sample covariance matrix  $\mathbf{A}_n$  is the M-P law by (2.2) when  $\mathbf{x}_1, \dots, \mathbf{x}_p$  are independent and satisfy Assumption 1, while the limit of the ESD is determined from (2.3) when there is some correlation among  $\mathbf{x}_1, \dots, \mathbf{x}_p$  with the covariance matrix  $\mathbf{T}_p$  different from the identity matrix. Moreover, preliminary investigations indicate that when  $\mathbf{x}_1, \dots, \mathbf{x}_p$  are only uncorrelated (without any further assumptions), the limit of the ESD of  $\mathbf{A}_n$  is not the M-P law (see Ryan and Debban (2009)). So, these motivate us to use the ESD of  $\mathbf{A}_n$ ,  $F^{\mathbf{A}_n}(x)$ , as a test statistic. However, there is no central limit theorem for  $(F^{\mathbf{A}_n}(x) - F^{c,H}(x))$ , as argued by Bai and Silverstein (2004). Therefore, instead, we consider the characteristic function of  $F^{\mathbf{A}_n}(x)$ .

The characteristic function of  $F^{\mathbf{A}_n}(x)$  is

$$s_n(t) \triangleq \int e^{itx} dF^{\mathbf{A}_n}(x) = \frac{1}{p} \sum_{i=1}^p e^{it\lambda_i}, \quad (2.8)$$

where  $\lambda_i, i = 1, \dots, p$  are eigenvalues of the sample covariance matrix  $\mathbf{A}_n$ .

Our test statistic is then proposed as follows:

$$M_n = \int_{T_1}^{T_2} |s_n(t) - s(t)|^2 dU(t), \quad (2.9)$$

where  $s(t) := s(t, c_n)$  is the characteristic function of  $F^{c_n}(x)$ , obtained from the M-P law  $F^c(x)$  with  $c$  replaced by  $c_n = p/n$ , and  $U(t)$  is a distribution function with its support on a compact interval, say  $[T_1, T_2]$ .

To develop the asymptotic distribution of  $M_n$  under a local alternative, the following assumption is needed.

**Assumption 3.** Let  $\mathbf{T}_p$  be a  $p \times p$  random Hermitian nonnegative definite matrix with a bounded spectral norm. Let  $\mathbf{y}_j^T = \mathbf{w}_j^T \mathbf{T}_p^{1/2}$ , where  $\mathbf{T}_p^{1/2}$  is the  $p \times p$  Hermitian matrix that satisfies  $(\mathbf{T}_p^{1/2})^2 = \mathbf{T}_p$  and  $\mathbf{w}_j^T = (W_{j1}, \dots, W_{jp})$ ,  $j = 1, \dots, n$  are i.i.d random vectors, in which  $W_{ji}, j \leq n, i \leq p$  are i.i.d with mean zero, variance one and finite fourth moment.

The empirical spectral distribution  $F^{\mathbf{T}_p}$  of  $\mathbf{T}_p$  converges weakly to a distribution  $H$  on  $[0, \infty)$  as  $n \rightarrow \infty$ ; all the diagonal elements of the matrix  $\mathbf{T}_p$  are equal to 1.

Note that under Assumption 3,  $\mathbf{A}_n$  becomes  $\mathbf{T}_p^{1/2} \mathbf{W}^* \mathbf{W} \mathbf{T}_p^{1/2}$ , where  $\mathbf{W} = (\mathbf{w}_1, \dots, \mathbf{w}_n)^T$ . The assumption that all the diagonal elements of  $\mathbf{T}_p$  are equal to 1 is used to guarantee that  $EX_{ji}^2 = 1$ . Roughly speaking, this constraint is to ensure that the data are weakly stationary. However, we would remark that such a constraint may be removed if we could have known  $\mathbf{T}_p$  is not a diagonal matrix. Under Assumption 3, when  $\mathbf{T}_p = \mathbf{I}_p$ , the random vectors  $\mathbf{x}_1, \dots, \mathbf{x}_p$  are independent and when  $\mathbf{T}_p \neq \mathbf{I}_p$ , they are not independent. For convenience, we name this dependent structure as ‘linear dependent structure’.

To develop the asymptotic distribution of the test statistic, write

$$G_n(x) = p[F^{\mathbf{A}_n}(x) - F^{c_n}(x)]. \quad (2.10)$$

Then,  $p(s_n(t) - s(t))$  can be decomposed as sum of the random part and the non-random part as follows:

$$\begin{aligned} p(s_n(t) - s(t)) &= \int e^{itx} dG_n(x) \\ &= \int e^{itx} d(p[F^{\mathbf{A}_n}(x) - F^{c_n, H_n}(x)]) + \int e^{itx} d(p[F^{c_n, H_n}(x) - F^{c_n}(x)]), \end{aligned} \quad (2.11)$$

where  $F^{c_n, H_n}$  is obtained from  $F^{c, H}$  with  $c$  and  $H$  replaced by  $c_n = p/n$  and  $H_n = F^{\mathbf{T}_p}$ .



**Theorem 1.** *Suppose that Assumptions 1 and 2 are true.*

1) *Assume that Assumptions 3 and the following conditions hold:*

$$\frac{1}{n} \sum_{i=1}^n \mathbf{e}_i^* \mathbf{T}_p^{1/2} (\underline{\mathbf{m}}(z_1) \mathbf{T}_p + \mathbf{I})^{-1} \mathbf{T}_p^{1/2} \mathbf{e}_i \mathbf{e}_i^* \mathbf{T}_p^{1/2} (\underline{\mathbf{m}}(z_2) \mathbf{T}_p + \mathbf{I})^{-1} \mathbf{T}_p^{1/2} \mathbf{e}_i \rightarrow h_1(z_1, z_2) \quad (2.12)$$

and

$$\frac{1}{n} \sum_{i=1}^n \mathbf{e}_i^* \mathbf{T}_p^{1/2} (\underline{\mathbf{m}}(z) \mathbf{T}_p + \mathbf{I})^{-1} \mathbf{T}_p^{1/2} \mathbf{e}_i \mathbf{e}_i^* \mathbf{T}_p^{1/2} (\underline{\mathbf{m}}(z) \mathbf{T}_p + \mathbf{I})^{-2} \mathbf{T}_p^{1/2} \mathbf{e}_i \rightarrow h_2(z), \quad (2.13)$$

where  $\mathbf{e}_i^*$  is the  $n$ -dimensional row vector with the  $i$ -th element being 1 and others 0. Then the scaled proposed test statistic  $p^2 M_n$  converges in distribution to a random variable  $R$

$$R = \int_{T_1}^{T_2} (|V(t) + \delta_1(t)|^2 + |Z(t) + \delta_2(t)|^2) dU(t), \quad (2.14)$$

where  $(V(t), Z(t))$  is a Gaussian vector and  $\delta_j(t)$ ,  $j = 1, 2$  are defined as

$$\delta_1(t) = \lim_{n \rightarrow \infty} \int \cos(tx) dp(F^{c_n, H_n}(x) - F^{c_n}(x)), \quad (2.15)$$

$$\delta_2(t) = \lim_{n \rightarrow \infty} \int \sin(tx) dp(F^{c_n, H_n}(x) - F^{c_n}(x)). \quad (2.16)$$

The mean and variance of  $(V(t), Z(t))$  are specified as follows. If  $X_{11}$  is real, then

$$\begin{aligned} EV(t) &= \frac{1}{2\pi i} \oint_{\gamma} \cos(tz) \frac{c \int \underline{\mathbf{m}}^3(z) \tau^2 (1 + \tau \underline{\mathbf{m}}(z))^{-3} dH(\tau)}{(1 - c \int \underline{\mathbf{m}}^2(z) \tau^2 (1 + \tau \underline{\mathbf{m}}(z))^{-2} dH(\tau))^2} dz \\ &\quad - \frac{EX_{11}^4 - 3}{2\pi i} \oint_{\gamma} \cos(tjz) \frac{c \underline{\mathbf{m}}^3(z) h_2(z)}{1 - c \int \underline{\mathbf{m}}^2(z) \tau^2 (1 + \tau \underline{\mathbf{m}}(z))^{-2} dH(\tau)} dz. \end{aligned}$$

Replacing  $\cos(tz)$  in  $EV(t)$  by  $\sin(tjz)$  yields the expression of  $EZ(t)$ . Also

$$\begin{aligned} \text{Cov}(V(t_j), Z(t_h)) &= -\frac{1}{2\pi^2} \oint_{\gamma_1} \oint_{\gamma_2} \frac{\cos(t_j z_1) \sin(t_h z_2)}{(\underline{\mathbf{m}}(z_1) - \underline{\mathbf{m}}(z_2))^2} \frac{d}{dz_2} \underline{\mathbf{m}}(z_2) \frac{d}{dz_1} \underline{\mathbf{m}}(z_1) dz_1 dz_2 \\ &\quad - \frac{c(EX_{11}^4 - 3)}{4\pi^2} \oint_{\gamma_1} \oint_{\gamma_2} \cos(t_j z_1) \sin(t_h z_2) \frac{d^2}{dz_1 dz_2} [\underline{\mathbf{m}}(z_1) \underline{\mathbf{m}}(z_2) h_1(z_1, z_2)] dz_1 dz_2. \end{aligned}$$

If  $X_{11}$  is complex with  $EX_{11}^2 = 0$ , then

$$EV(t) = -\frac{E|X_{11}|^4 - 2}{2\pi i} \oint_{\gamma} \cos(tz) \frac{c \underline{\mathbf{m}}^3(z) h_2(z)}{1 - c \int \underline{\mathbf{m}}^2(z) \tau^2 dH(\tau) / (1 + \tau \underline{\mathbf{m}}(z))^2} dz,$$

and the covariance

$$\begin{aligned} \text{Cov}(V(t_j), Z(t_h)) &= -\frac{1}{4\pi^2} \oint_{\gamma_1} \oint_{\gamma_2} \frac{\cos(t_j z_1) \sin(t_h z_2)}{(\underline{\mathbf{m}}(z_1) - \underline{\mathbf{m}}(z_2))^2} \frac{d}{dz_2} \underline{\mathbf{m}}(z_2) \frac{d}{dz_1} \underline{\mathbf{m}}(z_1) dz_1 dz_2 \\ &\quad - \frac{c(E|X_{11}|^4 - 2)}{4\pi^2} \oint_{\gamma_1} \oint_{\gamma_2} \cos(t_j z_1) \sin(t_h z_2) \frac{d^2}{dz_1 dz_2} [\underline{\mathbf{m}}(z_1) \underline{\mathbf{m}}(z_2) h_1(z_1, z_2)] dz_1 dz_2. \end{aligned}$$

The covariances  $\text{Cov}(V_j, V_h)$  and  $\text{Cov}(Z_j, Z_h)$  are similar to  $\text{Cov}(V_j, Z_h)$  except replacing  $(\cos(t_j z), \sin(t_h z))$  by  $(\cos(t_j z), \cos(t_h z))$  and  $(\sin(t_j z), \sin(t_h z))$  respectively. The contours  $\gamma_1$  and  $\gamma_2$  above are all closed and are taken in the positive direction in the complex plane, each enclosing the support of  $F^{c,H}$ . Also  $\gamma_1$  and  $\gamma_2$  are disjoint.

2) Under the null hypothesis  $\mathbf{H}_0$  the scaled statistic  $p^2 M_n$  then converges in distribution to

$$R_0 = \int_{T_1}^{T_2} (|\tilde{V}(t)|^2 + |\tilde{Z}(t)|^2) dU(t), \quad (2.17)$$

where the distribution of  $(\tilde{V}(t), \tilde{Z}(t))$  can be obtained from  $(V(t), Z(t))$  with  $H(\tau)$  becoming the degenerate distribution at the point 1,  $\underline{m}(z)$  being the Stieltjes transform of the M-P law,  $h_1(z_1, z_2) = \frac{1}{(\underline{m}(z_1)+1)(\underline{m}(z_2)+1)}$  and  $h_2(z) = \frac{1}{(\underline{m}(z)+1)^3}$ .

**Remark 1.** Assumption 1 assumes that all the entries of  $\mathbf{x}_i$  are identically distributed. It is of practical interest to consider removing the identical distribution condition. Instead of assuming identically distributed entries for  $\mathbf{x}_i$ , we need only to impose the following additional assumptions: for any  $k = 1, \dots, p; j = 1, \dots, n$ ,  $EX_{jk} = 0$ ,  $EX_{jk}^2 = 1$ ,  $\sup_{j,k} EX_{jk}^4 < \infty$  and for any  $\eta > 0$ ,

$$\frac{1}{\eta^4 np} \sum_{j=1}^n \sum_{k=1}^p E(|X_{jk}|^4 I_{(|X_{jk}| \geq \eta \sqrt{n})}) \rightarrow 0. \quad (2.18)$$

A careful checking on the arguments of Theorem 1.1 of Bai and Silverstein (2004) and Theorem 1.4 of Pan and Zhou (2008) indicates that Lemma 6 still holds and hence Theorem 1 holds under (2.18). For the expressions of the mean and covariance of the asymptotic random vector, we substitute the fourth moment  $E|W_{11}|^4$  in Lemma 6 with the average of all the fourth moments of all the entries, i.e.  $\frac{\sum_{j=1, k=1}^{n,p} E|W_{jk}|^4}{np}$ .

Conditions (2.12) and (2.13) can be removed if  $E[W_{11}^4] = 3$  in the real-number case or  $E[W_{11}^2] = 2$  in the complex-number case (see, for example, Bai and Silverstein (2004)). The first part of the above theorem is concerned with asymptotic distributions of the test statistic under a local alternative hypothesis, i.e., Assumption 3. With respect to Assumption 3, we would make the following comments, which are useful in the subsequent application section.

If  $\mathbf{y}_j^T = \tilde{\mathbf{w}}_j^T \mathbf{C}$ , where  $\mathbf{C}$  is any  $q \times p$  nonrandom matrix and  $\tilde{\mathbf{w}}_j, j = 1, \dots, n$  are i.i.d.  $q \times 1$  random vectors with their respective entries being i.i.d random variables, then Theorem 1 is still applicable. This is because  $\frac{1}{n} \mathbf{X} \mathbf{X}^*$  in this case becomes  $\frac{1}{n} \tilde{\mathbf{W}} \mathbf{C} \mathbf{C}^* \tilde{\mathbf{W}}^*$  and the nonnegative definitive matrix  $\mathbf{C} \mathbf{C}^*$  can be decomposed into  $\mathbf{C} \mathbf{C}^* = \mathbf{T}_q^{1/2} \mathbf{T}_q^{1/2}$ , where  $\mathbf{T}_q^{1/2}$  is a  $q \times q$  Hermitian matrix and  $\tilde{\mathbf{W}} = (\tilde{\mathbf{w}}_1, \dots, \tilde{\mathbf{w}}_n)$ . Note that the eigenvalues of  $\frac{1}{n} \mathbf{X} \mathbf{X}^*$  differ from those of  $\frac{1}{n} \mathbf{T}_q^{1/2} \mathbf{W}^* \mathbf{W} \mathbf{T}_q^{1/2}$  by  $|p - q|$  zeros. Thus, we may instead resort to CLT of  $\frac{1}{n} \mathbf{T}_q^{1/2} \tilde{\mathbf{W}}^* \tilde{\mathbf{W}} \mathbf{T}_q^{1/2}$ .

### 3 Applications to Multiple MA(1), AR(1) and Spatial Cross-sectional dependent Structures

This subsection is to explore some applications of the proposed test.

**Example 3.1.** Consider a multiple moving average model of order 1(MA(1)) of the form:

$$\mathbf{v}_t = \mathbf{z}_t + \psi \mathbf{z}_{t-1}, t = 1, \dots, p, \quad (3.1)$$

where  $|\psi| < 1$ ;  $\mathbf{z}_t = (Z_{1t}, \dots, Z_{nt})^T$  is an  $n$ -dimensional random vector with i.i.d. elements, each of which has zero mean and unit variance; and  $\mathbf{v}_t = (V_{1t}, \dots, V_{nt})^T$ . Denote by  $\hat{\mathbf{V}}_j^T$  and  $\hat{\mathbf{Z}}_j^T$  respectively the  $j$ -th rows of  $\mathbf{V} = (V_{jt})_{n \times p}$  and  $\mathbf{Z} = (Z_{jt})_{n \times (p+1)}$ .

For each  $j = 1, \dots, n$ , the MA(1) model (3.1) can be written as

$$\hat{\mathbf{V}}_j^T = \hat{\mathbf{Z}}_j^T \mathbf{C}, \quad (3.2)$$

where

$$\mathbf{C} = \begin{pmatrix} \psi & 0 & 0 & \cdots & 0 & 0 \\ 1 & \psi & 0 & \cdots & 0 & 0 \\ 0 & 1 & \psi & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & \psi \\ 0 & 0 & 0 & \cdots & 0 & 1 \end{pmatrix}_{(p+1) \times p}. \quad (3.3)$$

From Assumption 3, the last paragraph of the preceding subsection and Theorem 1, our test is able to capture the dependence of  $\mathbf{v}_1, \dots, \mathbf{v}_p$  as  $n$  and  $p$  go to infinity in the same order.

**Example 3.2.** Consider a multiple autoregressive model of order 1(AR(1)) of the form:

$$\mathbf{v}_t = \phi \mathbf{v}_{t-1} + \mathbf{z}_t, t = 1, \dots, p, \quad \mathbf{v}_0 = \frac{1}{\sqrt{1 - \phi^2}} \mathbf{z}_0, \quad (3.4)$$

where  $|\phi| < 1$ ; for any  $t = 0, 1, \dots, p$ ,  $\mathbf{z}_t = (Z_{1t}, \dots, Z_{nt})^T$  is an  $n$ -dimensional random vector with i.i.d. elements, each of which has zero mean and unit variance; and  $\mathbf{v}_t = (V_{1t}, \dots, V_{nt})^T$ . Denote the  $j$ -th rows of  $\mathbf{V} = (V_{jt})_{n \times (p+1)}$  and  $\mathbf{Z} = (Z_{jt})_{n \times (p+1)}$  as  $\hat{\mathbf{V}}_j^T$  and  $\hat{\mathbf{Z}}_j^T$  respectively.

For each  $j = 1, \dots, n$ , the AR(1) model (3.4) can be written as

$$\hat{\mathbf{V}}_j^T = \hat{\mathbf{Z}}_j^T \mathbf{D}, \quad (3.5)$$

where

$$\mathbf{D} = \begin{pmatrix} 1 & -\phi & (-\phi)^2 & \cdots & (-\phi)^{p-2} & (-\phi)^{p-1}/\sqrt{1-\phi^2} \\ 0 & 1 & -\phi & \cdots & (-\phi)^{p-3} & (-\phi)^{p-2}/\sqrt{1-\phi^2} \\ 0 & 0 & 1 & \cdots & (-\phi)^{p-4} & (-\phi)^{p-3}/\sqrt{1-\phi^2} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & -\phi/\sqrt{1-\phi^2} \\ 0 & 0 & 0 & \cdots & 0 & 1/\sqrt{1-\phi^2} \end{pmatrix}. \quad (3.6)$$

By Theorem 1, we can also apply the proposed test  $M_n$  to this AR(1) model as well.

**Example 3.3.** We now consider a panel data case. Let  $\{v_{ji} : i = 1, \dots, p; j = 1, \dots, n\}$  be the error components in a panel data model. They may be cross-sectionally correlated. In panel data analysis, it is of interest to consider the cross-sectional independence hypothesis, i.e.

$$\mathbf{H}_{00} : Cov(v_{ji}, v_{jh}) = 0 \text{ for all } j = 1, \dots, n \text{ and all } i \neq h;$$

against

$$\mathbf{H}_{11} : Cov(v_{ji}, v_{jh}) \neq 0 \text{ for some } j \text{ and some } i \neq h. \quad (3.7)$$

Under the assumption that  $\{v_{ji} : i = 1, \dots, p; j = 1, \dots, n\}$  are normal distributed, this hypothesis is equivalent to the independence hypothesis that

$$\mathbf{H}_0 : \mathbf{v}_1, \dots, \mathbf{v}_p \text{ are independent; against } \mathbf{H}_1 : \mathbf{v}_1, \dots, \mathbf{v}_p \text{ are not independent,} \quad (3.8)$$

where  $\mathbf{v}_i = (v_{1i}, \dots, v_{ni})^T$ ,  $i = 1, \dots, p$ .

Modern panel data literature has mainly adopted two different approaches to model error cross-sectional dependence: the spatial approach and the factor-structure approach. For the spatial approach, there are three popular spatial models which are the Spatial Moving Average(SMA), Spatial Auto-Regressive(SAR) and Spatial Error Components(SEC) processes. They are defined as follows:

$$SMA : v_{ji} = \sum_{k=1}^p \omega_{ik} \varepsilon_{jk} + \varepsilon_{ji}, \quad (3.9)$$

$$SAR : v_{ji} = \sum_{k=1}^p \omega_{ik} v_{jk} + \varepsilon_{ji}, \quad (3.10)$$

$$SEC : v_{ji} = \sum_{k=1}^p \omega_{ik} \xi_{jk} + \varepsilon_{ji}, \quad (3.11)$$

where  $\omega_{ik}$  is the  $i$ -specific spatial weight attached to individual  $k$ ;  $\{\varepsilon_{ji} : i = 1, \dots, p; j = 1, \dots, n\}$  and  $\{\xi_{ji} : i = 1, \dots, p; j = 1, \dots, n\}$  are two sets with i.i.d. random components with zero mean and unit variance; moreover,  $\{\xi_{ji} : i = 1, \dots, p; j = 1, \dots, n\}$  are uncorrelated with  $\{\varepsilon_{ji}, i = 1, \dots, p; j = 1, \dots, n\}$ .

Denote the  $j$ -th row of  $\mathbf{V} = (v_{ji})_{n \times p}$ ,  $\boldsymbol{\varepsilon} = (\varepsilon_{ji})_{n \times p}$  and  $\boldsymbol{\xi} = (\xi_{ji})_{n \times p}$  by  $\hat{\mathbf{v}}_j^T$ ,  $\hat{\boldsymbol{\varepsilon}}_j^T$  and  $\hat{\boldsymbol{\xi}}_j^T$  respectively. Set  $\boldsymbol{\omega} = (\omega_{ik})_{p \times p}$ .

Model SMA (3.9) may be rewritten as  $\hat{\mathbf{v}}_j^T = \hat{\boldsymbol{\varepsilon}}_j^T(\boldsymbol{\omega}^T + \mathbf{I}_p), \forall j = 1, \dots, n$  and hence  $\mathbf{T}_p = (\boldsymbol{\omega} + \mathbf{I}_p)(\boldsymbol{\omega}^T + \mathbf{I}_p)$ . For model SAR (3.10), assume that  $\boldsymbol{\omega} - \mathbf{I}_p$  is invertible. We then write  $\hat{\mathbf{v}}_j^T = \hat{\boldsymbol{\varepsilon}}_j^T(\boldsymbol{\omega}^T - \mathbf{I}_p)^{-1}, \forall j = 1, \dots, n$ . Hence  $\mathbf{T}_p = (\boldsymbol{\omega} - \mathbf{I}_p)^{-1}(\boldsymbol{\omega}^T - \mathbf{I}_p)^{-1}$ . Therefore the test statistic  $M_n$  can be used to identify whether  $\mathbf{v}_1, \dots, \mathbf{v}_p$  of models (3.9) and (3.10) are independent. Hence it can capture the cross-sectional dependence for the SMA model and SAR model

As for the SEC model, whether the statistic  $M_n$  can detect the dependence of the SEC model relies on the properties of the sample covariance matrix in the form of

$$\mathbf{B}_n = \frac{1}{n}(\boldsymbol{\omega}\boldsymbol{\xi} + \boldsymbol{\varepsilon})(\boldsymbol{\omega}\boldsymbol{\xi} + \boldsymbol{\varepsilon})^T, \quad (3.12)$$

where  $\boldsymbol{\xi} = (\boldsymbol{\xi}_1, \dots, \boldsymbol{\xi}_p)^T$  and  $\boldsymbol{\varepsilon} = (\boldsymbol{\varepsilon}_1, \dots, \boldsymbol{\varepsilon}_p)^T$ .

Under the null hypothesis  $\mathbf{H}_0$ , Dozier and Silversten (2007) provides the LSD of the matrix  $\mathbf{B}_n$  whose Stieljes transform is

$$\hat{m}(z) = \int \frac{d\hat{H}(x)}{\frac{x}{1+c\hat{m}(z)} - (1+c\hat{m}(z))z + 1 - c}, \quad (3.13)$$

where  $\hat{H}(x)$  is the limit of  $F^{\frac{1}{n}}\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}^T$ .

With this result, we know that the LSD of the matrix  $\mathbf{B}_n$  is not the M-P law. In view of this, the proposed test  $M_n$  should capture the dependence of the SEC model (3.11) in theory. However, the asymptotic distribution of the proposed test statistic  $M_n$  for this case will need to be developed in future work.

## 4 A general panel data model

Note that the proposed test is based on the idea that the limits of ESDs under the null and local alternative hypotheses are different. Yet, it may be the case that there exists some dependence among the set of vectors of  $\mathbf{v}_1, \dots, \mathbf{v}_p$  but the limit of the ESD associated with such vectors is the M-P law. Then a natural question is whether the statistic  $M_n$  works in this case. We below investigate a panel data model as an example.

Consider a panel data model of the form

$$v_{ij} = \varepsilon_{ij} + \frac{1}{\sqrt{p}}u_i, \quad i = 1, \dots, p; \quad j = 1, \dots, n, \quad (4.1)$$

where  $\{\varepsilon_{ij}, i = 1, \dots, p; j = 1, \dots, n\}$  is a sequence of i.i.d. real random variables with  $E\varepsilon_{11} = 0$  and  $E\varepsilon_{11}^2 = 1$ , and  $\{u_i, i = 1, \dots, p\}$  are real random variables, and independent of  $\{\varepsilon_{ij}, i = 1, \dots, p; j = 1, \dots, n\}$ .

For any  $i = 1, \dots, p$ , set

$$\mathbf{v}_i = (v_{i1}, \dots, v_{in})^T. \quad (4.2)$$

The aim of this section is to test the null hypothesis specified in (3.8) for model (4.1).

Model (4.1) can be written as

$$\mathbf{v} = \boldsymbol{\varepsilon} + \mathbf{u}\mathbf{e}^T, \quad (4.3)$$

where  $\mathbf{v} = (\mathbf{v}_1, \dots, \mathbf{v}_p)^T$ ,  $\mathbf{u} = (\frac{1}{\sqrt{p}}u_1, \dots, \frac{1}{\sqrt{p}}u_p)^T$  and  $\mathbf{e}$  is  $p \times 1$  vector with all elements being one.

Consider the sample covariance matrix

$$\mathbf{S}_n = \frac{1}{n}\mathbf{v}\mathbf{v}^T = \frac{1}{n}(\boldsymbol{\varepsilon} + \mathbf{u}\mathbf{e}^T)(\boldsymbol{\varepsilon} + \mathbf{u}\mathbf{e}^T)^T. \quad (4.4)$$

By Lemma 5 in the Appendix and the fact that  $\text{rank}(\mathbf{u}\mathbf{e}^T) \leq 1$ , it can be concluded that the limit of the ESD of the matrix  $\mathbf{S}$  is the same as that of the matrix  $\frac{1}{n}\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}^T$ , i.e. the M-P law. Even so, we still would like to use the proposed statistic  $M_n$  to test the null hypothesis of mutual independence. However, this model does not satisfy Assumption 1 because the elements of each vector  $\mathbf{v}_i$  are not independent and they include the common random factor  $u_i$ , Theorem 1 thus can not be directly applied to this model. Therefore, we need to develop a new asymptotic theory for the proposed statistic  $M_n$  for this model.

**Theorem 2.** *For model (4.1), in addition to Assumption 1 and Assumption 2, we assume that*

$$E\|\mathbf{u}\|^4 < \infty, \quad \frac{1}{p^2}E \left[ \sum_{i \neq j}^p (u_i^2 - \bar{u})(u_j^2 - \bar{u}) \right] \rightarrow 0, \quad \text{as } n \rightarrow \infty, \quad (4.5)$$

where  $\bar{u}$  is a positive constant number.

Then, the proposed test statistic  $p^2M_n$  converges in distribution to the random variable  $R_2$  given by

$$R_2 = \int_{t_1}^{t_2} (|W(t)|^2 + |Q(t)|^2) dU(t), \quad (4.6)$$

where  $(W(t), Q(t))$  is a Gaussian vector whose mean and covariance are specified as follows:

$$\begin{aligned}
EW(t) &= -\frac{c}{2\pi i} \oint_{\gamma} \cos(t_j z) \frac{\underline{m}^3(z)(1 + \underline{m}(z))}{((1 + \underline{m}(z))^2 - c\underline{m}^2(z))^2} dz - \frac{c(EY_{11}^4 - 3)}{2\pi i} \oint_{\gamma} \cos(t_j z) \frac{\underline{m}^3(z)}{(1 + \underline{m}(z))^2 - c\underline{m}^2(z)} dz \\
&\quad + \frac{c}{2\pi i} \oint_{\gamma} \cos(t_j z) \frac{\underline{m}(z)}{z((1 + \underline{m}(z))^2 - c\underline{m}^2(z))} dz \\
&\quad - \frac{1}{2\pi i} \oint_{\gamma} \cos(t_j z) \frac{\frac{c\underline{m}(z)}{z[(1 + \underline{m}(z))^2 - c\underline{m}^2(z)]} + \bar{u} \int \frac{1}{(\lambda - z)^2} dF^{MP}(\lambda)}{\bar{u}z\underline{m}^2(z) - 1} dz; \tag{4.7}
\end{aligned}$$

and

$$\begin{aligned}
Cov(W(t_j), Q(t_h)) &= -\frac{1}{2\pi^2} \oint_{\gamma_1} \oint_{\gamma_2} \frac{\cos(t_j z_1) \sin(t_h z_2)}{(\underline{m}(z_1) - \underline{m}(z_2))^2} \frac{d}{dz_1} \underline{m}(z_1) \frac{d}{dz_2} \underline{m}(z_2) dz_1 dz_2 \\
&\quad - \frac{c(EY_{11}^4 - 3)}{4\pi^2} \oint_{\gamma_1} \oint_{\gamma_2} \cos(t_j z_1) \cos(t_h z_2) \frac{d^2}{dz_1 dz_2} \left[ \frac{\underline{m}(z_1) \underline{m}(z_2)}{(1 + \underline{m}(z_1))(1 + \underline{m}(z_2))} \right] dz_1 dz_2. \tag{4.8}
\end{aligned}$$

Replacing  $\cos(t_j z)$  in  $EW(t_j)$  by  $\sin(t_j z)$  yields the expression of  $EQ(t_j)$ . The covariances  $Cov(W(t_j), W(t_h))$  and  $Cov(Q(t_j), Q(t_h))$  are similar except replacing  $\sin(t_h z)$  and  $\cos(t_j z)$  by  $\cos(t_h z)$  and  $\sin(t_j z)$  respectively. The contours in (4.7) and (4.8) both enclose the interval  $[(1 - \sqrt{c})^2 + 2c\bar{u}, (1 + \sqrt{c})^2 + 2c\bar{u}]$ . Moreover, the contours  $\gamma_1$  and  $\gamma_2$  are disjoint.

**Remark 2.** When  $u_1, \dots, u_p$  are independent and hence  $\mathbf{v}_1, \dots, \mathbf{v}_p$  are independent, condition (4.5) is true.

In view of Theorem 2, we see that the proposed test statistic  $M_n$  still works mainly due to the involvement of the last term on the right hand of (2.11).

## 5 Small sample simulation studies

This section provides some simulated examples to show the finite sample performance of the proposed test. In addition, we also compare the performance with that of a likelihood ratio test proposed by Anderson (1984). Simulations are used to find the empirical sizes and powers of the proposed independence test and hence evaluate the performance of the test. To show the efficiency of our test, the two kinds of dependent structures investigated in section 3 and section 5 are detected, such as multiple MA(1) and AR(1) model, SMA and the general panel data model.

### 5.1 Empirical sizes and empirical powers

First we introduce the method of calculating empirical sizes and empirical powers. Let  $z_{\frac{1}{2}\alpha}$  and  $z_{1-\frac{1}{2}\alpha}$  be the  $100(\frac{1}{2}\alpha)\%$  and  $100(1 - \frac{1}{2}\alpha)\%$  quantiles of the asymptotic null distribution of the

test statistic  $M_n$  respectively. With  $K$  replications of the data set simulated under the null hypothesis, we calculate the empirical size as

$$\hat{\alpha} = \frac{\{\# \text{ of } M_n^H \geq z_{1-\frac{1}{2}\alpha} \text{ or } M_n^H \leq z_{\frac{1}{2}\alpha}\}}{K}, \quad (5.1)$$

where  $M_n^H$  represents the values of the test statistic  $M_n$  based on the data simulated under the null hypothesis.

In our simulation, we choose  $K = 1000$  as the number of repeated simulations. The significance level is  $\alpha = 0.05$ . Since the asymptotic null distribution of the test statistic is not a classical distribution, we need to estimate the quantiles  $z_{\frac{1}{2}\alpha}$  and  $z_{1-\frac{1}{2}\alpha}$ . Naturally, we do as follows: generate  $K$  replications of the asymptotic distributed random variable and then select the  $(K\frac{1}{2}\alpha)$ -th smallest value  $\hat{z}_{\frac{1}{2}\alpha}$  and  $(K\frac{1}{2}\alpha)$ -th largest value  $\hat{z}_{1-\frac{1}{2}\alpha}$  as the estimated  $100(\frac{1}{2}\alpha)\%$  and  $100(1 - \frac{1}{2}\alpha)\%$  quantiles of the asymptotic distributed random variable.

With the estimated critical points  $\hat{z}_{\frac{1}{2}\alpha}$  and  $\hat{z}_{1-\frac{1}{2}\alpha}$  under the null hypothesis, the empirical power is calculated as

$$\hat{\beta} = \frac{\{\# \text{ of } M_n^A \geq \hat{z}_{1-\frac{1}{2}\alpha} \text{ or } M_n^A \leq \hat{z}_{\frac{1}{2}\alpha}\}}{K}, \quad (5.2)$$

where  $M_n^A$  represents the values of the test statistic  $M_n$  based on the data simulated under the alternative hypothesis.

## 5.2 Comparisons with the classical likelihood ratio test

For the proposed independence test, we generate the number  $n$  of  $p$ -dimensional independent and identical distributed random vectors  $\{\mathbf{y}_j\}_{j=1}^n$ , each with the mean vector  $\mathbf{0}_p$  and the covariance matrix  $\Sigma$ . Under the null hypothesis,  $\{\mathbf{y}_j\}_{j=1}^n$  are generated in two scenarios:

1. Each  $\mathbf{w}_j$  is a  $p$ -dimensional normal random vector with the mean vector  $\mathbf{0}_p$  and the covariance matrix  $\Sigma = \mathbf{I}_p$ ;  $\forall j = 1, \dots, n$ ,  $\mathbf{y}_j = \mathbf{T}_p \mathbf{w}_j$  with  $\mathbf{T}_p = \mathbf{I}_p$ ;
2. Each  $\mathbf{w}_j$  consists of i.i.d. random variables with standardized Gamma(4,2) distribution, so they have zero means and unit variances;  $\forall j = 1, \dots, n$ ,  $\mathbf{y}_j = \mathbf{T}_p \mathbf{w}_j$  with  $\mathbf{T}_p = \mathbf{I}_p$ .

Under the alternative hypothesis, we consider the case:  $\mathbf{T}_p^{1/2} = (\sqrt{0.95}\mathbf{I}_p, \sqrt{0.05}\mathbf{1}_p)$ , where  $\mathbf{1}_p$  is a  $p$ -dimensional vector with 1 as entries. In this case, the population covariance matrix of  $\mathbf{y}_j$  is  $\Sigma = 0.95\mathbf{I}_p + 0.05\mathbf{1}_p\mathbf{1}_p'$ , which is called the compound symmetric covariance matrix.

For normal distributed data, the fourth moment of each element is  $E|X_{11}|^4 = 3$ ; for standardized Gamma(4,2) distributional data,  $E|X_{11}|^4 = 4.5$ . This can be calculated by a formula



as follows: the  $k$ -th moment of  $X_{11}$  which is  $Gamma(\alpha, \lambda)$  is

$$E|X_{11}|^k = \frac{(\alpha + k - 1) \cdots (\alpha + 1)\alpha}{\lambda^k}. \quad (5.3)$$

Anderson (1984) provides a likelihood ratio criterion(LRT) to test independence for a fixed number of fixed dimensional normal distributed random vectors. We compare it with the proposed test.

Under the null hypothesis, the distribution of  $L$  is the distribution of  $L_2 L_3 \cdots L_p$ , where  $L_2, \dots, L_p$  are independently distributed with  $L_k$  having the distribution of  $U_{m, (k-1)m, n-1-(k-1)m}$ . Furthermore, for any  $k = 2, \dots, p$ , as  $n \rightarrow \infty$ ,  $-(n - \frac{3}{2} - \frac{km}{2}) \log U_{m, (k-1)m, n-1-(k-1)m}$  has  $\chi^2$  distribution with  $(k-1)m^2$  degrees of freedom(See section 8.5 of Anderson (1984)).

From the construction of the LRT test, we can see that the LRT utilizes additional  $n$  observations of the random vectors  $\mathbf{x}_1, \dots, \mathbf{x}_p$  under investigation, while the proposed test does not need this information. However, we can choose  $m = 1$  and apply LRT to independence test of the random vectors  $\tilde{\mathbf{x}}_1, \dots, \tilde{\mathbf{x}}_p$ , where for any  $i = 1, \dots, p$ , the elements of the vector  $\tilde{\mathbf{x}}_i$  consist of its  $n$  observations. Hence the LRT test can test independence for the number of  $p$  random vectors with dimension  $n$  by choosing  $m = 1$ .

Table 1 and Table 2 show the empirical sizes and empirical powers of our proposed test and the LRT test for normal distributed random vectors respectively. From Table 1 and Table 2, we can see that the LRT test does not work when  $p$  and  $n$  are both large while the proposed test possesses good performance when  $p$  and  $n$  go to infinity in the same order. The LRT test is only applicable to the case that  $p$  is fixed and  $n$  tends to infinity. From Table 2, it can be seen that the LRT fails when  $p$  is comparable with  $n$ . When the difference between  $p$  and  $n$  are large, the sizes and the powers of the proposed test become worse. This is because our test is proposed under the case that  $p$  and  $n$  are in the same order when they approximate to infinity. The proposed test also works well for gamma random vectors while the LRT test is not applicable to gamma case, since, in theory, LRT test is provided for normal random vectors. Table 3 provides the empirical sizes and empirical powers of the proposed test for the gamma case. In our simulation, we choose  $p, n = 5, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100$  for the proposed test and the LRT test. The significant level  $\alpha$  is chosen as 0.05. In each case, we run  $K = 1000$  repeated simulations. Reports for empirical powers show that the proposed test can check independence for both normal and gamma vectors well. Moreover, the empirical powers converge to 1 as  $n, p \rightarrow \infty$ .

### 5.3 Multiple MA(1), AR(1) and SMA model

Consider multiple MA(1) model

$$\mathbf{v}_t = \mathbf{z}_t + \psi \mathbf{z}_{t-1}, t = 1, \dots, p. \quad (5.4)$$

We choose  $\psi = 0.5$  and the simulation results in Table 4 show that the proposed test performs well for this model.

Consider multiple AR(1)

$$\mathbf{v}_t = \phi \mathbf{v}_{t-1} + \mathbf{z}_t, t = 1, \dots, p, \mathbf{v}_0 = \frac{1}{\sqrt{1 - \phi^2}} \mathbf{z}_0. \quad (5.5)$$

Let  $\phi = 0.5$ . The empirical powers for this model are provided in Table 5. As  $n$  and  $p$  increase in the same order, the empirical power tends to 1.

As for the Spatial Moving Average(SMA) model, i.e.

$$v_{ji} = \sum_{k=1}^p \omega_{ik} \varepsilon_{jk} + \varepsilon_{ji}, \quad (5.6)$$

generate  $\varepsilon_{jk} \stackrel{i.i.d.}{\sim} normal(1, 1), \forall j = 1, \dots, n; k = 1, \dots, p$ . Apply the proposed statistic  $M_n$  for the sample matrix  $\frac{1}{n} \mathbf{V}^* \mathbf{V}$  and the empirical powers are illustrated in Table 6, which show that  $M_n$  performs well for capturing the cross-sectional dependence for SMA model.

### 5.4 The general panel data model

We examine the finite sample performance of the proposed test for the general panel data model (4.1), i.e.

$$v_{ij} = \varepsilon_{ij} + \frac{1}{\sqrt{p}} u_i, i = 1, \dots, p; j = 1, \dots, n, \quad (5.7)$$

where  $\{\varepsilon_{ij}, i = 1, \dots, p; j = 1, \dots, n\}$  is a sequence of i.i.d. random variables and  $E\varepsilon_{11} = 0, E\varepsilon_{11}^2 = 1; \{u_i, i = 1, \dots, p\}$  are independent of  $\{\varepsilon_{ij}, i = 1, \dots, p; j = 1, \dots, n\}$ .

Under the null hypothesis, we generate  $u_i \stackrel{i.i.d.}{\sim} normal(1, 1), i = 1, \dots, p$  and under the alternative hypothesis, we experiment with  $\mathbf{u} = (\frac{1}{\sqrt{p}} u_1, \frac{1}{\sqrt{p}} u_2, \dots, \frac{1}{\sqrt{p}} u_p) \sim \frac{1}{\sqrt{p}} N(\mathbf{1}_p, \Sigma)$ , where  $\Sigma = \mathbf{T} \mathbf{T}^T$  and  $\mathbf{T}$  is a  $p \times p$  matrix whose elements are generated  $t_{ik} \stackrel{i.i.d.}{\sim} U(0, 1), i, k = 1, \dots, p$ .

The simulation results including empirical sizes and powers in Table 7 show that the proposed test can capture the dependence for the general panel data model (4.1).

## 5.5 Some time series and Vandermonde matrix

Dependent structures of a set of random vectors are always described by non-zero correlations among them, such as the linear dependent structure developed in Section 3. However, there are some data which are not independent but uncorrelated. We consider three such examples and test their dependence by the proposed test statistic although we have not developed central limit theorems for this type of local alternatives.

### 5.5.1 Nonlinear MA model

Consider nonlinear MA models of the form

$$R_{tj} = Z_{t-1,j}Z_{t-2,j}(Z_{t-2,j} + Z_{tj} + 1), \quad t = 1, \dots, p; \quad j = 1, \dots, n; \quad (5.8)$$

where  $\mathbf{z}_t = (Z_{t1}, \dots, Z_{tn})$  is an  $n$ -dimensional random vector with i.i.d. elements, each of which has zero mean and unit variance; and  $\mathbf{r}_t = (R_{t1}, \dots, R_{tn})$ . For any  $j = 1, \dots, n$ , the correlation matrix of  $(R_{j1}, R_{j2}, \dots, R_{jp})$  is a diagonal matrix. This model is provided by Kuan and Lee (2004) which tests the martingale difference hypothesis. Our proposed independence test can be applied to this nonlinear MA model, and the powers in Table 8 show that this test performs well for this model.

This result also implies that the limit of the empirical spectral distribution of the nonlinear MA model (5.8) is not the M-P law since the proposed test statistic is established on the characteristic function of the M-P law.

### 5.5.2 Multiple ARCH(1) model

Consider the multiple autoregressive conditional heteroscedastic(ARCH(1)) model:

$$W_{tj} = Z_{tj} \sqrt{\alpha_0 + \alpha_1 W_{t-1,j}^2}, \quad t = 1, \dots, p; \quad j = 1, \dots, n; \quad (5.9)$$

where  $\mathbf{z}_t = (Z_{t1}, \dots, Z_{tn})$  is an  $n$ -dimensional random vector with i.i.d. elements, each of which has zero mean and unit variance; and  $\boldsymbol{\omega}_t = (W_{t1}, \dots, W_{tn})$ .

For each  $j = 1, \dots, n$ , ARCH(1) model  $(W_{1j}, W_{2j}, \dots, W_{pj})$  is a martingale difference sequence. ARCH(1) model has many applications in financial analysis. There exists no theoretical results stating that the LSD of the sample covariance matrix for ARCH(1) model is M-P Law, but from Figure 1, we can see that the LSD of the sample covariance matrix for ARCH(1) model is indeed M-P Law. A rigorous study is under investigation. For the ARCH(1) model, the proposed test can not capture the dependence of  $(\boldsymbol{\omega}_1, \boldsymbol{\omega}_2, \dots, \boldsymbol{\omega}_p)$  directly, but we can test the

dependence of  $(\omega_1^2, \omega_2^2, \dots, \omega_p^2)$ . Since this test can tell us that  $(\omega_1^2, \omega_2^2, \dots, \omega_p^2)$  are independent either, naturally it can be concluded that  $(\omega_1, \omega_2, \dots, \omega_p)$  are also not independent. Here we take  $\alpha_0 = 0.9$  and  $\alpha_1 = 0.1$ . Table 9 shows the powers of our test for testing dependence of ARCH(1) model.

### 5.5.3 Vandermonde matrix

Consider the  $n \times p$  vandermonde matrix  $\mathbf{V}$  of the form

$$\mathbf{V} = \frac{1}{\sqrt{n}} \begin{pmatrix} 1 & 1 & \cdots & 1 \\ e^{-i\omega_1} & e^{-i\omega_2} & \cdots & e^{-i\omega_p} \\ \vdots & \vdots & \ddots & \vdots \\ e^{-i(n-1)\omega_1} & e^{-i(n-1)\omega_2} & \cdots & e^{-i(n-1)\omega_p} \end{pmatrix} \quad (5.10)$$

where  $\omega_i, i = 1, \dots, p$  are called phased distributions and are assumed i.i.d on  $[0, 2\pi)$ . Then the entries of  $\mathbf{V}$  lie on the unit circle. Obviously, all the entries of the rows of  $\mathbf{V}$  are not independent while the columns are independent. Denote the sample covariance matrix of  $\mathbf{V}$  by  $\mathbf{D} = \mathbf{V}^H \mathbf{V}$ .

Vandermonde matrices play an important role in signal processing and wireless applications such as direction of arrival estimation, pre-coding or sparse sampling theory, etc.. Ryan and Debban (2009) established that as both  $n, p$  go to  $\infty$  with their ratio being a positive constant, the limiting spectral distribution of  $\mathbf{D} = \mathbf{V}^H \mathbf{V}$  is not the M-P law. This result reminds us that the proposed test should be applied to capture the dependence structure of the rows of the matrix  $\mathbf{V}$ . It is easy to see that, for any  $k = 1, \dots, n - 1$  and  $j = 1, \dots, p$ ,  $E(e^{-ik\omega_j})^2 = 0$  and  $E|e^{-ik\omega_j}|^4 = 1$ . The empirical powers in table 10 show that the proposed test works well in detecting dependence of Vandermonde matrices.

## 6 Empirical application for financial data

As an application of the proposed independence test, we test the cross-sectional dependence of daily closed stock prices between several companies from New York Stock Exchange(NYSE) during the period 1990.1.1 – 2002.1.1. The data set is obtained from Wharton Research Data Services(WRDS) database. It consists of daily closed stock prices of 632 companies at 5001 days.

We randomly choose  $p$  companies from the total 632 companies in the data set. For each company  $i = 1, 2, \dots, p$ , there exist 5001 daily closed stock prices and denote it as

$\mathbf{s}_i = (s_{i1}, s_{i2}, \dots, s_{i5001})$ . From the price vector  $\mathbf{s}_i$ , we construct an time series of closed stock prices with length  $n$ :  $\mathbf{x}_i = (s_{i1}, s_{i,1+50}, s_{i,1+2 \times 50}, \dots, s_{i,1+50n})$ . Of course,  $p \leq 632$  and  $n \leq 100$ . The construction of  $\mathbf{x}_i, i = 1, \dots, p$  is based on the idea that the prices tend to being independent as the length of the time between them is large. Hence for each company  $i$ , the elements of stock price time series  $\mathbf{x}_i$  are independent.

The proposed test  $M_n$  is applied to testing dependence of  $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_p$ . For each  $(p, n)$ , we randomly choose  $p$  companies from the database, construct the corresponding vectors  $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_p$ , and then calculate the P-value by applying the proposed test. Repeat this procedure  $num = 100$  times and plot the P-value graph to see whether the cross-sectional 'dependence' feature is popular in NYSE.

From Fig. 2, we can see that, as the number of companies  $p$  increases, more experiments are rejected in terms of the P-values below 0.05. When  $p = 30, n = 35$ , all the 100 experiments are rejected. This phenomenon is reasonable since as  $p$  increases, the opportunity that we choose dependent companies increases as the number of total companies is invariant. It shows that cross-sectional dependence exists and is popular in NYSE. This suggests that the assumption that cross-sectional independence in such empirical studies may not be appropriate.

As comparison, we select 40 companies from the transportation section of NYSE and investigate the dependence of daily stock prices of these companies from the same industry section—transportation section. We also perform the same experiments as above for those total 40 companies. The P-value graphs of  $n = 10, p = 5, n = 15, p = 10$  and  $n = 20, p = 15$  are provided in Fig. 3. Compared with the corresponding P-value graphs  $n = 10, p = 5, n = 15, p = 10$  and  $n = 20, p = 15$  in Fig. 2, the daily closed stock prices of companies from the same section are more likely to be dependent.

## 7 Conclusions

This paper provides an approach for testing independence among a large number of high dimensional random vectors based on the characteristic function of the empirical spectral distribution of the sample covariance matrix of the random vectors. This test can capture various kinds of dependent structures, e.g. MA(1), AR(1) model, nonlinear MA(1) model, ARCH(1) model and the general panel data model established in the simulation section. The conventional method(LRT proposed by Anderson (1984)) utilizes the correlated relationship between random vectors to capture their dependence. This idea is only efficient for normal distributed data. It may be an inappropriate tool for non-Gaussian distributed data, such as martingale difference

sequences (e.g. ARCH(1) model), nonlinear MA(1) model, the Vandermonde matrix, etc., which possess dependent but uncorrelated structures. The proposed test is not restricted to normal distributed data. Moreover, it is designed for a large number of high dimensional random vectors. An empirical application to closed stock prices of several companies from New York Stock Exchange highlights this approach.

## 8 Appendix

### 8.1 Some useful lemmas

**Lemma 1** (Theorem 8.1 of Billingsley (1999)). *Let  $P_n$  and  $P$  be probability measures on  $(C, \varphi)$ . If the finite dimensional distributions of  $P_n$  converge weakly to those of  $P$ , and if  $\{P_n\}$  is tight, then  $P_n \Rightarrow P$ .*

**Lemma 2** (Theorem 12.3 of Billingsley (1999)). *The sequence  $\{X_n\}$  is tight if it satisfies these two conditions*

(I) *The sequence  $\{X_n(0)\}$  is tight.*

(II) *There exists constants  $\gamma \geq 0$ ,  $\alpha > 1$ , and a nondecreasing, continuous function  $F$  on  $[0, 1]$  such that*

$$E\{|X_n(t_2) - X_n(t_1)|^\gamma\} \leq |F(t_2) - F(t_1)|^\alpha \quad (8.1)$$

*holds for all  $t_1, t_2$ , and  $n$ .*

**Lemma 3** (Continuous Theorem). *Let  $X_n$  and  $X$  be random elements defined on a metric space  $S$ . Suppose  $g : S \rightarrow S'$  has a set of discontinuous points  $D_g$  such that  $P(X \in D_g) = 0$ . Then*

$$X_n \xrightarrow{d} X \Rightarrow g(X_n) \xrightarrow{d} g(X). \quad (8.2)$$

**Lemma 4** (Complex mean value theorem (see Lemma 2.4 of Guo and Nicolus (2006))). *Let  $\Omega$  be an open convex set in  $\mathbb{C}$ . If  $f : \Omega \rightarrow \mathbb{C}$  is an analytic function and  $a, b$  are distinct points in  $\Omega$ , then there exist points  $u, v$  on  $L(a, b)$  such that*

$$\operatorname{Re}\left(\frac{f(a) - f(b)}{a - b}\right) = \operatorname{Re}(f'(u)), \quad \operatorname{Im}\left(\frac{f(a) - f(b)}{a - b}\right) = \operatorname{Im}(f'(v)), \quad (8.3)$$

*where  $\operatorname{Re}(z)$  and  $\operatorname{Im}(z)$  are the real and imaginary parts of  $z$  respectively; and  $L(a, b) \triangleq \{a + t(b - a) : t \in (0, 1)\}$ .*

**Lemma 5** (Theorem A.44 of Bai and Silverstein (2009)). *Let  $\mathbf{A}$  and  $\mathbf{B}$  be two  $p \times n$  complex matrices. Then,*

$$\|F^{\mathbf{A}\mathbf{A}^*} - F^{\mathbf{B}\mathbf{B}^*}\| \leq \frac{1}{p} \operatorname{rank}(\mathbf{A} - \mathbf{B}). \quad (8.4)$$

## 8.2 Proofs of main theorems

Bai and Silverstein (2004) established the remarkable central limit theorem for functional of eigenvalues of  $\mathbf{A}_n$  under the additional assumption that  $E|X_{11}|^4 = 3$  while Pan and Zhou (2008) provided a supplement to this theorem by eliminating the condition to some extent. From Theorem 4 of Pan and Zhou (2008) and (2.11) we can directly obtain the following lemma.

**Lemma 6.** *Under Assumptions 2 and 3, we have, for any positive integer  $k$ ,*

$$\left( \int \cos(t_1 x) dG_n(x), \dots, \int \cos(t_k x) dG_n(x), \int \sin(t_1 x) dG_n(x), \dots, \int \sin(t_k x) dG_n(x) \right) \quad (8.5)$$

*converges in distribution to Gaussian vectors  $(V(t_1) + \delta_1(t_1), \dots, V(t_k) + \delta_1(t_k), Z_1 + \delta_2(t_1), \dots, Z(t_k) + \delta_2(t_k))$ , where  $\delta_1(t)$ ,  $\delta_2(t)$  are, respectively, defined as*

$$\delta_1(t) = \lim_{n \rightarrow \infty} \int \cos(tx) dp(F^{c_n, H_n}(x) - F^{c_n}(x)), \quad (8.6)$$

$$\delta_2(t) = \lim_{n \rightarrow \infty} \int \sin(tx) dp(F^{c_n, H_n}(x) - F^{c_n}(x)). \quad (8.7)$$

*The means and covariances of  $V(t_j)$  and  $Z(t_j)$  are specified in Theorem 1:*

**Remark 3.** *When  $\mathbf{T}_n = \mathbf{I}$ , the mean and variance of the asymptotic Gaussian distribution for power functions  $f(x) = x^r$ ,  $\forall r \in \mathbb{Z}^+$  is calculated in Pan and Zhou (2008) and Bai and Silverstein (2004). Hence the corresponding means and covariances for  $f_1(x) = \sin tx$  and  $f_2(x) = \cos tx$  can be derived by Taylor series of  $\sin tx$  and  $\cos tx$ .*

*Proof of Theorem 1.* Let  $t$  belong to a closed interval  $I = [T_1, T_2]$ . To finish Theorem 1, in view of Lemma 1 and Lemma 6, it suffices to prove the tightness of  $\{(\phi_n(t), \psi_n(t)) : t \in I\}$ . Thus it suffices to prove the tightness of  $p(s_n(t) - s(t))$ . Repeating the same truncation and centralization steps as those in Bai and Silverstein (2004), we may assume that

$$|X_{ij}| < \delta_n \sqrt{n}, \quad EX_{ij} = 0, \quad E|X_{ij}|^2 = 1, \quad E|X_{ij}|^4 < \infty. \quad (8.8)$$

Set  $M_n(z) = n[m_{F^{\mathbf{A}_n}}(z) - m_{F^{c_n, H_n}}(z)]$ . By the Cauchy theorem

$$f(x) = -\frac{1}{2\pi i} \oint \frac{f(z)}{z - x} dz, \quad (8.9)$$

we have, with probability one, for all  $n$  large,

$$\int e^{itz} dp(F^{\mathbf{A}_n}(x) - F^{c_n, H_n}(x)) = -\frac{1}{2\pi i} \oint_{\mathcal{C}} e^{itz} M_n(z) dz. \quad (8.10)$$

The contour  $\mathcal{C}$  involved in the above integral is specified as follows. Let

$$\mathcal{C}_u = \{x + iv_0 : x \in [x_l, x_r]\}, \quad (8.11)$$

where  $v_0 > 0$ ,  $x_r$  is any number greater than  $\limsup \lambda_{max}(\mathbf{T}_n)(1 + \sqrt{c})^2$ ,  $x_l$  is any negative number if  $c \geq 1$  and otherwise choose  $x_l \in (0, \limsup \lambda_{min}(\mathbf{T}_n)(1 - \sqrt{c})^2)$ . Then the contour  $\mathcal{C}$  is defined by the union of  $\mathcal{C}_+$  and its symmetric part  $\mathcal{C}_-$  with respect to the  $x$ -axis, where

$$\mathcal{C}_+ = \{x_l + iv : v \in [0, v_0]\} \cup \mathcal{C}_u \cup \{x_r + iv : v \in [0, v_0]\}. \quad (8.12)$$

From Lemma 1.1 and the argument regarding equivalence in probability of  $M_n(z)$  and its truncation version given in Page 563 in Bai and Silverstein (2004) and Lemma 3 we have

$$\oint_{\mathcal{C}} |M_n(z)||dz| \xrightarrow[n]{D} \oint_{\mathcal{C}} |M(z)||dz|, \quad (8.13)$$

where  $M(z)$  is a Gaussian process, the limit of  $M_n(z)$ .

We conclude from Lemma 4 that for any  $\delta > 0$

$$\begin{aligned} & \sup_{|t_1 - t_2| < \delta, t_1, t_2 \in I} \left| \oint_{\mathcal{C}} (e^{it_1 z} - e^{it_2 z}) M_n(z) dz \right| \\ & \leq \sup_{|t_1 - t_2| < \delta, t_1, t_2 \in I} \left| \oint_{\mathcal{C}} \sqrt{(Re(ize^{it_3 z}))^2 + (Im(ize^{it_4 z}))^2} \delta |M_n(z)||dz| \right| \\ & \leq K\delta \left| \oint_{\mathcal{C}} |M_n(z)||dz| \right| \xrightarrow{D} K\delta \left| \oint_{\mathcal{C}} |M(z)||dz| \right|, \quad \text{as } n \rightarrow \infty, \end{aligned} \quad (8.14)$$

where  $t_3$  and  $t_4$  lies in the interval  $[T_1, T_2]$ , the last inequality uses (8.13) and the fact that  $Re(ize^{it_3 z})$ ,  $Im(ize^{it_4 z})$  are bounded on the contour  $\mathcal{C}$  and  $K$  (and in the sequel) is a constant number which may be different from line to line.

By (8.14), we have for any  $\varepsilon > 0$ ,

$$P\left(\sup_{|t_1 - t_2| < \delta, t_1, t_2 \in [0, 1]} \left| \oint_{\mathcal{C}} (e^{it_1 z} - e^{it_2 z}) M_n(z) dz \right| \geq \varepsilon\right) \leq P\left(K\delta \left| \oint_{\mathcal{C}} |M_n(z)||dz| \right| \geq \varepsilon\right) \quad (8.15)$$

and

$$\lim_{\delta \rightarrow 0} \limsup_{n \rightarrow \infty} P\left(K\delta \left| \oint_{\mathcal{C}} |M_n(z)||dz| \right| \geq \varepsilon\right) = \lim_{\delta \rightarrow 0} P\left(K\delta \left| \oint_{\mathcal{C}} |M(z)||dz| \right| \geq \varepsilon\right) = 0. \quad (8.16)$$

Hence (8.15) and (8.16) imply that

$$\lim_{\delta \rightarrow 0} \limsup_{n \rightarrow \infty} P\left(\sup_{|t_1 - t_2| < \delta, t_1, t_2 \in I} \left| \oint_{\mathcal{C}} (e^{it_1 z} - e^{it_2 z}) M_n(z) dz \right| \geq \varepsilon\right) = 0. \quad (8.17)$$

By Theorem 7.3 of Billingsley (1999),  $\int e^{itx} dp(F^{\mathbf{A}_n}(x) - F^{c_n, H_n}(x))$  is tight. Moreover from the assumption we see that  $\int e^{itx} dp(F^{c_n, H_n}(x) - F^{c_n}(x))$  is tight by Lemma 4.

□



To prove Theorem 2, we first need to establish Lemma 7 below. To this end, write

$$H_n(x) = p[F^{\mathbf{S}_n}(x) - F^{c_n}(x)], \quad (8.18)$$

where  $\mathbf{S}_n$  is defined in (4.4).

**Lemma 7.** *Under the assumptions of Theorem 2, we have for any positive integer  $k$ ,*

$$\left( \int \cos(t_1 x) dH_n(x), \dots, \int \cos(t_k x) dH_n(x), \int \sin(t_1 x) dH_n(x), \dots, \int \sin(t_k x) dH_n(x) \right) \quad (8.19)$$

*converges in distribution to a Gaussian vector  $(W(t_1), \dots, W(t_k), Q(t_1), \dots, Q(t_k))$  whose mean and covariance function are given in Theorem 2.*

*Proof of Lemma 7.* Repeating the same truncation and centralization steps as those in Bai and Silverstein (2004), we may assume that

$$|\varepsilon_{ij}| < \delta_n \sqrt{n}, \quad E\varepsilon_{ij} = 0, \quad E|\varepsilon_{ij}|^2 = 1, \quad E|\varepsilon_{ij}|^4 < \infty. \quad (8.20)$$

For the panel data model (4.1), let

$$\mathbf{v}_j = (v_{1j}, \dots, v_{pj})^T, \quad \boldsymbol{\varepsilon}_j = (\varepsilon_{1j}, \dots, \varepsilon_{pj})^T, \quad \mathbf{u} = \left( \frac{1}{\sqrt{p}}u_1, \dots, \frac{1}{\sqrt{p}}u_p \right)^T, \quad j = 1, \dots, n. \quad (8.21)$$

The model can be written in the vector form as

$$\mathbf{v}_j = \boldsymbol{\varepsilon}_j + \mathbf{u}, \quad j = 1, \dots, n. \quad (8.22)$$

We then define the sample covariance matrix by  $\mathbf{S}_n = \frac{1}{n} \sum_{j=1}^n \mathbf{v}_j \mathbf{v}_j^T$ . Moreover write

$$\bar{\mathbf{v}} = \frac{1}{n} \sum_{j=1}^n \mathbf{v}_j, \quad \bar{\boldsymbol{\varepsilon}} = \frac{1}{n} \sum_{j=1}^n \boldsymbol{\varepsilon}_j, \quad (8.23)$$

and

$$\mathbf{D}_n = \frac{1}{n} \sum_{j=1}^n \boldsymbol{\varepsilon}_j \boldsymbol{\varepsilon}_j^T, \quad \mathcal{S}_n = \frac{1}{n} \sum_{j=1}^n (\mathbf{v}_j - \bar{\mathbf{v}})(\mathbf{v}_j - \bar{\mathbf{v}})^T, \quad \mathcal{D}_n = \frac{1}{n} \sum_{j=1}^n (\boldsymbol{\varepsilon}_j - \bar{\boldsymbol{\varepsilon}})(\boldsymbol{\varepsilon}_j - \bar{\boldsymbol{\varepsilon}})^T. \quad (8.24)$$

Note that  $\mathcal{S}_n = \mathcal{D}_n$ . The sample covariance matrix  $\mathbf{S}_n$  can be then expressed as

$$\mathbf{S}_n = \mathcal{S}_n + \bar{\mathbf{v}} \bar{\mathbf{v}}^T = \mathcal{D}_n + \bar{\mathbf{v}} \bar{\mathbf{v}}^T. \quad (8.25)$$

By (4.5) and the Burkholder inequality we have

$$E|\mathbf{u}^T \bar{\boldsymbol{\varepsilon}}|^4 \leq \frac{1}{n^4} E \left| \sum_{j=1}^n \mathbf{u}^T \boldsymbol{\varepsilon}_j \right|^4 \leq \frac{K}{n^4} E \left| \sum_{j=1}^n \mathbf{u}^T \mathbf{u} \right|^2 + \frac{K}{n^4} \sum_{j=1}^n E |\mathbf{u}^T \boldsymbol{\varepsilon}_j|^4 = O\left(\frac{1}{n^2}\right),$$

which, together with Borel-Cantelli's Lemma, implies that

$$\mathbf{u}^T \bar{\boldsymbol{\varepsilon}} \xrightarrow{a.s.} 0.$$

Also, condition (4.5) ensures that

$$\mathbf{u}^T \mathbf{u} \rightarrow \bar{u}. \quad (8.26)$$

Therefore by (4.5) and Theorem 2 of Pan and Zhou (2011), we have, as  $n \rightarrow \infty$ ,

$$\lambda_{max}(\bar{\mathbf{v}}\bar{\mathbf{v}}^T) = \bar{\mathbf{v}}^T \bar{\mathbf{v}} = \bar{\boldsymbol{\varepsilon}}^T \bar{\boldsymbol{\varepsilon}} + \mathbf{u}^T \mathbf{u} + 2\mathbf{u}^T \bar{\boldsymbol{\varepsilon}} \xrightarrow{a.s.} c + \bar{u}, \quad as \ n \rightarrow \infty. \quad (8.27)$$

Furthermore, Jiang (2004) proved that

$$\lambda_{max}(\mathcal{D}_n) \rightarrow (1 + \sqrt{c})^2, \quad a.s. \ as \ n \rightarrow \infty \quad (8.28)$$

and Xiao and Zhou (2010) proved that, when  $c \leq 1$

$$\lambda_{min}(\mathcal{D}_n) \rightarrow (1 - \sqrt{c})^2, \quad a.s. \ as \ n \rightarrow \infty. \quad (8.29)$$

By (8.27) (8.28) and (8.29), the maximal and minimal eigenvalues of  $\mathbf{S}_n$  satisfy with probability one

$$\limsup_{n \rightarrow \infty} \lambda_{max}(\mathbf{S}_n) \leq c + \bar{u} + (1 + \sqrt{c})^2, \quad (8.30)$$

and

$$\liminf_{n \rightarrow \infty} \lambda_{min}(\mathbf{S}_n) \geq (1 - \sqrt{c})^2. \quad (8.31)$$

As in the proof of Theorem 2, we obtain from Cauchy's formula, with probability one, for  $n$  large,

$$\begin{aligned} p \int f(x) d[F^{\mathbf{S}_n}(x) - F^{c_n}(x)] &= \frac{p}{2\pi i} \int \oint_{\gamma} \frac{f(z)}{z-x} dz d[F^{\mathbf{S}_n}(x) - F^{c_n}(x)] \\ &= \frac{p}{2\pi i} \oint_{\gamma} f(z) dz \int \frac{1}{z-x} d[F^{\mathbf{S}_n}(x) - F^{c_n}(x)] \\ &= -\frac{1}{2\pi i} \oint_{\gamma} f(z) (tr(\mathbf{S}_n - z\mathbf{I}_p)^{-1} - pm_{c_n}(z)) dz, \end{aligned} \quad (8.32)$$

where  $m_{c_n}(z)$  is obtained from  $m(z)$  with  $c$  replaced by  $c_n$ . The contour  $\gamma$  is specified as follows: Let  $v_0 > 0$  be arbitrary and set  $\gamma_{\mu} = \{\mu + iv_0, \mu \in [\mu_{\ell}, \mu_r]\}$ , where  $\mu_r > c + \bar{u} + (1 + \sqrt{c})^2$  and  $0 < \mu_{\ell} < I_{(0,1)}(c)(1 - \sqrt{c})^2$  or  $\mu_{\ell}$  is any negative number if  $c \geq 1$ . Then define

$$\gamma^+ = \{\mu_{\ell} + iv : v \in [0, v_0]\} \cup \gamma_{\mu} \cup \{\mu_r + iv : v \in [0, v_0]\} \quad (8.33)$$

and let  $\gamma^-$  be the symmetric part of  $\gamma^+$  about the real axis. Then set  $\gamma = \gamma^+ \cup \gamma^-$ .

Set

$$\begin{aligned}\mathcal{S}_n^{-1}(z) &= (\mathcal{S}_n - z\mathbf{I}_p)^{-1}, \quad \mathbf{S}_n^{-1}(z) = (\mathbf{S}_n - z\mathbf{I}_p)^{-1}, \\ \mathcal{D}_n^{-1}(z) &= (\mathcal{D}_n - z\mathbf{I}_p)^{-1}, \quad \mathbf{D}_n^{-1}(z) = (\mathbf{D}_n - z\mathbf{I}_p)^{-1}.\end{aligned}\tag{8.34}$$

Then we have

$$\text{tr}\mathbf{S}_n^{-1}(z) - pm_{c_n}(z) = (\text{tr}\mathcal{D}_n^{-1}(z) - pm_{c_n}(z)) - \frac{\bar{\mathbf{v}}^T \mathcal{D}_n^{-2}(z) \bar{\mathbf{v}}}{1 + \bar{\mathbf{v}}^T \mathcal{D}_n^{-1}(z) \bar{\mathbf{v}}},\tag{8.35}$$

where we have used the identity

$$(\mathbf{C} + \mathbf{r}\mathbf{r}^T)^{-1} = \mathbf{C}^{-1} - \frac{\mathbf{C}^{-1}\mathbf{r}\mathbf{r}^T\mathbf{C}^{-1}}{1 + \mathbf{r}^T\mathbf{C}^{-1}\mathbf{r}},\tag{8.36}$$

where  $\mathbf{C}$  and  $(\mathbf{C} + \mathbf{r}\mathbf{r}^T)$  are both invertible; and  $\mathbf{r} \in \mathbb{R}^p$ . The first term on the right hand of (8.35) was investigated in Pan (2011). In what follows we consider the second term on the right hand of (8.35).

One may verify that

$$(\mathbf{C} + q\mathbf{r}\mathbf{v}^T)^{-1} = \frac{\mathbf{C}^{-1}}{1 + q\mathbf{v}^T\mathbf{C}^{-1}\mathbf{r}},\tag{8.37}$$

where  $\mathbf{C}$  and  $(\mathbf{C} + q\mathbf{r}\mathbf{v}^T)$  are both invertible,  $q$  is a scalar and  $\mathbf{r}, \mathbf{v} \in \mathbb{R}^p$ . This, together with (8.22) and (8.36), yields

$$\begin{aligned}\bar{\mathbf{v}}^T \mathcal{D}_n^{-1}(z) \bar{\mathbf{v}} &= \bar{\boldsymbol{\varepsilon}}^T \mathcal{D}_n^{-1}(z) \bar{\boldsymbol{\varepsilon}} + 2\mathbf{u}^T \mathcal{D}_n^{-1}(z) \bar{\boldsymbol{\varepsilon}} + \mathbf{u}^T \mathcal{D}_n^{-1}(z) \mathbf{u} \\ &= \frac{\bar{\boldsymbol{\varepsilon}}^T \mathbf{D}_n^{-1}(z) \bar{\boldsymbol{\varepsilon}}}{1 - \bar{\boldsymbol{\varepsilon}}^T \mathbf{D}_n^{-1}(z) \bar{\boldsymbol{\varepsilon}}} + 2\frac{\mathbf{u}^T \mathbf{D}_n^{-1}(z) \bar{\boldsymbol{\varepsilon}}}{1 - \bar{\boldsymbol{\varepsilon}}^T \mathbf{D}_n^{-1}(z) \bar{\boldsymbol{\varepsilon}}} + \mathbf{u}^T \mathcal{D}_n^{-1}(z) \mathbf{u}\end{aligned}\tag{8.38}$$

and

$$\begin{aligned}\bar{\mathbf{v}}^T \mathcal{D}_n^{-2}(z) \bar{\mathbf{v}} &= \bar{\boldsymbol{\varepsilon}}^T \mathcal{D}_n^{-2}(z) \bar{\boldsymbol{\varepsilon}} + 2\mathbf{u}^T \mathcal{D}_n^{-2}(z) \bar{\boldsymbol{\varepsilon}} + \mathbf{u}^T \mathcal{D}_n^{-2}(z) \mathbf{u} \\ &= \frac{\bar{\boldsymbol{\varepsilon}}^T \mathbf{D}_n^{-2}(z) \bar{\boldsymbol{\varepsilon}}}{(1 - \bar{\boldsymbol{\varepsilon}}^T \mathbf{D}_n^{-1}(z) \bar{\boldsymbol{\varepsilon}})^2} + \frac{2\mathbf{u}^T \mathbf{D}_n^{-2}(z) \bar{\boldsymbol{\varepsilon}}}{1 - \bar{\boldsymbol{\varepsilon}}^T \mathbf{D}_n^{-1}(z) \bar{\boldsymbol{\varepsilon}}} + \frac{2\mathbf{u}^T \mathbf{D}_n^{-1}(z) \bar{\boldsymbol{\varepsilon}} \bar{\boldsymbol{\varepsilon}}^T \mathbf{D}_n^{-2}(z) \bar{\boldsymbol{\varepsilon}}}{(1 - \bar{\boldsymbol{\varepsilon}}^T \mathbf{D}_n^{-1}(z) \bar{\boldsymbol{\varepsilon}})^2} + \mathbf{u}^T \mathcal{D}_n^{-2}(z) \mathbf{u}.\end{aligned}\tag{8.39}$$

It is proved in Section 2.5 and (4.3) of Pan (2011) that as  $n \rightarrow \infty$ ,

$$\sup_{z \in \gamma} \left| \bar{\boldsymbol{\varepsilon}}^T \mathbf{D}_n^{-2}(z) \bar{\boldsymbol{\varepsilon}} - \frac{c\bar{m}^2(z)}{(1 + \bar{m}(z))^2 - c\bar{m}^2(z)} \right| \xrightarrow{i.p.} 0;\tag{8.40}$$

$$\sup_{z \in \gamma} \left| \bar{\boldsymbol{\varepsilon}}^T \mathbf{D}_n^{-1}(z) \bar{\boldsymbol{\varepsilon}} - (1 + z\bar{m}(z)) \right| \xrightarrow{i.p.} 0;\tag{8.41}$$

and

$$\sup_{z \in \gamma} \left| \mathbf{u}^T \mathbf{D}_n^{-1}(z) \bar{\boldsymbol{\varepsilon}} \right| \xrightarrow{i.p.} 0,\tag{8.42}$$

(where we also use an argument similar to (2.28) of Pan (2011)). By Lemma 2 of Bai, Miao and Pan (2007), (3.4) and (4.3) in Pan (2011)), and (4.5) we have, as  $n \rightarrow \infty$ ,

$$\sup_{z \in \gamma} \left| \mathbf{u}^T \mathcal{D}_n^{-1}(z) \mathbf{u} - \bar{u} \underline{m}(z) \right| \xrightarrow{i.p.} 0. \quad (8.43)$$

The next aim is to prove that

$$\sup_{z \in \gamma} \left| \mathbf{u}^T \mathcal{D}_n^{-2}(z) \mathbf{u} - \bar{u} \int \frac{1}{(\lambda - z)^2} dF^{MP}(\lambda) \right| \xrightarrow{i.p.} 0 \quad (8.44)$$

and that

$$\sup_{z \in \gamma} \left| \mathbf{u}^T \mathbf{D}_n^{-2}(z) \bar{\boldsymbol{\varepsilon}} \right| \xrightarrow{i.p.} 0. \quad (8.45)$$

Consider (8.44) first. By the formula (8.36), we have an expansion

$$\mathbf{u}^T \mathcal{D}_n^{-2}(z) \mathbf{u} = \mathbf{u}^T \mathbf{D}_n^{-2}(z) \mathbf{u} + \frac{2\mathbf{u}^T \mathbf{D}_n^{-2}(z) \bar{\boldsymbol{\varepsilon}} \bar{\boldsymbol{\varepsilon}}^T \mathbf{D}_n^{-1}(z) \mathbf{u}}{1 - \bar{\boldsymbol{\varepsilon}}^T \mathbf{D}_n^{-1}(z) \bar{\boldsymbol{\varepsilon}}} + \frac{\mathbf{u}^T \mathbf{D}_n^{-1}(z) \bar{\boldsymbol{\varepsilon}} \bar{\boldsymbol{\varepsilon}}^T \mathbf{D}_n^{-2}(z) \bar{\boldsymbol{\varepsilon}} \bar{\boldsymbol{\varepsilon}}^T \mathbf{D}_n^{-1}(z) \mathbf{u}}{(1 - \bar{\boldsymbol{\varepsilon}}^T \mathbf{D}_n^{-1}(z) \bar{\boldsymbol{\varepsilon}})^2}.$$

For any given  $z \in \gamma$ , we conclude from Theorem 1 of Pan (2011) and Helly-Bray's theorem that

$$\mathbf{u}^T \mathcal{D}_n^{-2}(z) \mathbf{u} - \bar{u} \int \frac{1}{(\lambda - z)^2} dF^{MP}(\lambda) \xrightarrow{i.p.} 0 \quad \text{as } n \rightarrow \infty. \quad (8.46)$$

By the expansion of  $\mathbf{u}^T \mathcal{D}_n^{-2}(z) \mathbf{u}$  and (8.40)-(8.42), to prove (8.44), it suffices to prove the tightness of  $\left\{ K_n^{(1)}(z) = \mathbf{u}^T \mathbf{D}_n^{-2}(z) \mathbf{u} - \bar{u} \int \frac{1}{(\lambda - z)^2} dF^{MP}(\lambda), z \in \gamma \right\}$  and  $\left\{ \mathbf{u}^T \mathbf{D}_n^{-2}(z) \bar{\boldsymbol{\varepsilon}}, z \in \gamma \right\}$ .

To this end, as in Bai and Silverstein (2004), below introduce the truncated version of  $\mathbf{u}^T \mathbf{D}_n^{-2}(z) \mathbf{u}$ . Define  $\gamma_r = \{\mu_r + iv : v \in [n^{-1}\rho_n, v_0]\}$ ,

$$\gamma_\ell = \begin{cases} \{\mu_\ell + iv : v \in [n^{-1}\rho_n, v_0]\}, & \mu_\ell > 0, \\ \{\mu_\ell + iv : v \in [0, v_0]\}, & \mu_\ell < 0, \end{cases} \quad (8.47)$$

where

$$\rho_n \downarrow 0, \quad \rho_n \geq n^{-\theta}, \quad \text{for some } \theta \in (0, 1). \quad (8.48)$$

Let  $\gamma_n^+ = \gamma_\ell \cup \gamma_\mu \cup \gamma_r$  and  $\gamma_n^-$  denote the symmetric part of  $\gamma_n^+$  with respect to the real axis. We then define the truncated process  $\widehat{\mathbf{u}^T \mathbf{D}_n^{-2}(z) \mathbf{u}}$  of the process  $\mathbf{u}^T \mathbf{D}_n^{-2}(z) \mathbf{u}$  for  $z = \alpha + iv$  by

$$\widehat{\mathbf{u}^T \mathbf{D}_n^{-2}(z) \mathbf{u}} = \begin{cases} \mathbf{u}^T \mathbf{D}_n^{-2}(z) \mathbf{u} & z \in \gamma_n = \gamma_n^+ \cup \gamma_n^-, \\ \frac{nv + \rho_n}{2\rho_n} \mathbf{u}^T \mathbf{D}_n^{-2}(z_{r_1}) \mathbf{u} + \frac{\rho_n - nv}{2\rho_n} \mathbf{u}^T \mathbf{D}_n^{-2}(z_{r_2}) \mathbf{u} & \mu = \mu_r, v \in [-n^{-1}\rho_n, n^{-1}\rho_n], \\ \frac{nv + \rho_n}{2\rho_n} \mathbf{u}^T \mathbf{D}_n^{-2}(z_{\ell_1}) \mathbf{u} + \frac{\rho_n - nv}{2\rho_n} \mathbf{u}^T \mathbf{D}_n^{-2}(z_{\ell_2}) \mathbf{u} & \mu = \mu_\ell > 0, v \in [-n^{-1}\rho_n, n^{-1}\rho_n], \end{cases} \quad (8.49)$$

where  $z_{r_1} = \mu_r + in^{-1}\rho_n$ ,  $z_{r_2} = \mu_r - in^{-1}\rho_n$ ,  $z_{\ell_1} = \mu_\ell + in^{-1}\rho_n$  and  $z_{\ell_2} = \mu_\ell - in^{-1}\rho_n$ . We then have

$$\sup_{z \in \gamma} \left| \mathbf{u}^T \widehat{\mathbf{D}_n^{-2}}(z) \mathbf{u} - \mathbf{u}^T \mathbf{D}_n^{-2}(z) \mathbf{u} \right| \leq K \rho_n \|\mathbf{u}\|^2 \left( \frac{1}{|\lambda_{\max}(\mathbf{D}_n) - \mu_r|} + \frac{1}{|\lambda_{\min}(\mathbf{D}_n) - \mu_l|} \right) \xrightarrow{i.p.} 0. \quad (8.50)$$

It is proved in Section 3 of Bai and Silverstein (2004) that, for any positive integer  $k$  and  $z \in \gamma_n^+ \cup \gamma_n^-$ ,

$$\max(E \|\mathbf{D}_n^{-1}(z)\|^k) \leq K. \quad (8.51)$$

It follows from independence between  $\mathbf{u}$  and  $\varepsilon_j, j = 1, \dots, n$  that

$$\begin{aligned} & E \left| \mathbf{u}^T \mathbf{D}_n^{-2}(z) \mathbf{u} - \bar{u} \int \frac{1}{(\lambda - z)^2} dF^{MP}(\lambda) \right| \\ & \leq E \left| \mathbf{u}^T \mathbf{D}_n^{-2}(z) \mathbf{u} \right| + \left| \bar{u} \int \frac{1}{(\lambda - z)^2} dF^{MP}(\lambda) \right| \\ & \leq E \|\mathbf{u}^T\|^2 E \|\mathbf{D}_n^{-2}(z)\|^2 + K \leq K, \end{aligned} \quad (8.52)$$

which ensures Condition (1) of Lemma 2. Similarly, we can derive  $E |\mathbf{u}^T \mathbf{D}_n^{-2}(z) \bar{\varepsilon}|^2 \leq K$ .

Next, we prove condition (2) of Lemma 2, i.e.

$$\sup_{n, z_1, z_2 \in \gamma_n^+ \cup \gamma_n^-} \frac{E |K_n^{(i)}(z_1) - K_n^{(i)}(z_2)|^2}{|z_1 - z_2|^2} < \infty, \quad i = 1, 2. \quad (8.53)$$

Note that

$$\mathbf{A}^{-1} - \mathbf{B}^{-1} = \mathbf{A}^{-1}(\mathbf{B} - \mathbf{A})\mathbf{B}^{-1}, \quad (8.54)$$

where  $\mathbf{A}$  and  $\mathbf{B}$  are any two nonsingular matrices. We then conclude that

$$\mathbf{D}_n^{-2}(z_1) - \mathbf{D}_n^{-2}(z_2) = (z_1 - z_2) \mathbf{D}_n^{-2}(z_1) \mathbf{D}_n^{-1}(z_2) + (z_1 - z_2) \mathbf{D}_n^{-1}(z_1) \mathbf{D}_n^{-2}(z_2). \quad (8.55)$$

Then

$$\begin{aligned} \frac{K_n^{(1)}(z_1) - K_n^{(1)}(z_2)}{z_1 - z_2} &= \mathbf{u}^T \mathbf{D}_n^{-2}(z_1) \mathbf{D}_n^{-1}(z_2) \mathbf{u} + \mathbf{u}^T \mathbf{D}_n^{-1}(z_1) \mathbf{D}_n^{-2}(z_2) \mathbf{u} \\ &\quad - \bar{u} \int \frac{(\lambda - z_1) + (\lambda - z_2)}{(\lambda - z_1)^2 (\lambda - z_2)^2} dF^{MP}(\lambda). \end{aligned} \quad (8.56)$$

As in (8.52), we can obtain

$$E |\mathbf{u}^T \mathbf{D}_n^{-2}(z_1) \mathbf{D}_n^{-1}(z_2) \mathbf{u}|^2 \leq K, \quad E |\mathbf{u}^T \mathbf{D}_n^{-1}(z_1) \mathbf{D}_n^{-2}(z_2) \mathbf{u}|^2 \leq K. \quad (8.57)$$

Since  $f(\lambda) = \frac{(\lambda - z_1) + (\lambda - z_2)}{(\lambda - z_1)^2 (\lambda - z_2)^2}$  is a continuous function when  $z_1, z_2 \in \gamma$ , the integral  $\int f(\lambda) dF^{MP}(\lambda)$  is bounded. This, together with (8.57), implies

$$\sup_{n, z_1, z_2 \in \gamma} \frac{E |K_n^{(1)}(z_1) - K_n^{(1)}(z_2)|^2}{|z_1 - z_2|^2} < \infty. \quad (8.58)$$

By (3.17) in Pan and Zhou (2011) and an argument similar to (8.55)-(8.58) we may verify that  $\mathbf{u}^T \mathbf{D}_n^{-2}(z) \bar{\boldsymbol{\varepsilon}}$  is tight for  $z \in \gamma$ . Summarizing the above we obtain (8.44).

Consider (8.45) now. From the last paragraph we see that it is enough to consider the pointwise convergence of  $\mathbf{u}^T \mathbf{D}_n^{-2}(z) \bar{\boldsymbol{\varepsilon}}$ . As in (8.49) and (8.50) we may define the truncated process  $\widehat{\mathbf{u}^T \mathbf{D}_n^{-2}(z) \bar{\boldsymbol{\varepsilon}}}$  of the process  $\mathbf{u}^T \mathbf{D}_n^{-2}(z) \bar{\boldsymbol{\varepsilon}}$  and then prove that their difference tends to zero in probability. As in (4.3) of Pan and Zhou (2011) one may prove that for given  $z \in \gamma_n^+$   $\mathbf{u}^T \mathbf{D}_n^{-2}(z) \bar{\boldsymbol{\varepsilon}} \xrightarrow{i.p.} 0$ .

From (8.38) to (8.45) we have

$$\sup_{z \in \gamma} \left| \bar{\mathbf{v}}^T \mathcal{D}_n^{-1}(z) \bar{\mathbf{v}} - \left( \frac{1 + z \underline{m}(z)}{-z \underline{m}(z)} + \bar{u} \underline{m}(z) \right) \right| \xrightarrow{i.p.} 0, \quad (8.59)$$

and

$$\sup_{z \in \gamma} \left| \bar{\mathbf{v}}^T \mathcal{D}_n^{-2}(z) \bar{\mathbf{v}} - \left( \frac{c}{z^2 [(1 + \underline{m}(z))^2 - c \underline{m}^2(z)]} + \bar{u} \int \frac{1}{(\lambda - z)^2} dF^{MP}(\lambda) \right) \right| \xrightarrow{i.p.} 0. \quad (8.60)$$

We then conclude from Slutsky's theorem that

$$\sup_{z \in \gamma} \left| \frac{\bar{\mathbf{v}}^T \mathcal{D}_n^{-2}(z) \bar{\mathbf{v}}}{1 + \bar{\mathbf{v}}^T \mathcal{D}_n^{-1}(z) \bar{\mathbf{v}}} - \frac{\frac{cm(z)}{z[(1+m(z))^2 - cm^2(z)]} + \bar{u} \int \frac{1}{(\lambda - z)^2} dF^{MP}(\lambda)}{\bar{u} z \underline{m}^2(z) - 1} \right| \xrightarrow{i.p.} 0. \quad (8.61)$$

The arguments of Theorem 1 of Pan (2011) show that the truncation version of  $(tr \mathcal{D}_n^{-1}(z) - pm_{c_n}(z))$  converges in distribution to a two-dimensional Gaussian process and that the difference between  $(tr \mathcal{D}_n^{-1}(z) - pm_{c_n}(z))$  and its truncation version goes to zero in probability (see Page 563 of Bai and Silverstein (2004) and (2.28) of Pan (2011)). Theorem 3 then follows from (8.61), (8.32) and (8.35), Slutsky's theorem and Lemma 3 (one may refer to Page 563 of Bai and Silverstein (2004)).  $\square$

*Proof of Theorem 2.* As in the proof of Theorem 1, in view of Lemma 7 it suffices to prove the tightness of  $\{p(s_n(t) - s(t)) : t \in I\}$ . As before, write

$$\begin{aligned} p(s_n(t) - s(t)) &= p \int e^{itx} d[F^{\mathbf{S}_n}(x) - F^{c_n}(x)] \\ &= -\frac{1}{2\pi i} \oint_{\gamma} e^{itz} (tr(\mathbf{S}_n - z \mathbf{I}_p)^{-1} - pm_{c_n}(z)) dz, \end{aligned} \quad (8.62)$$

where the contour  $\gamma$  is specified in Lemma 7.

From the formula (8.36), we have

$$tr \mathcal{D}_n^{-1}(z) = tr \mathbf{D}_n^{-1}(z) + \frac{\bar{\boldsymbol{\varepsilon}}^T \mathbf{D}_n^{-2}(z) \bar{\boldsymbol{\varepsilon}}}{1 - \bar{\boldsymbol{\varepsilon}}^T \mathbf{D}_n^{-1}(z) \bar{\boldsymbol{\varepsilon}}}. \quad (8.63)$$

This, together with (8.35), yields

$$\operatorname{tr}\mathbf{S}_n^{-1}(z) - pm_{c_n}(z) = \operatorname{tr}\mathbf{D}_n^{-1}(z) - pm_{c_n}(z) + \frac{\bar{\boldsymbol{\varepsilon}}^T \mathbf{D}_n^{-2}(z) \bar{\boldsymbol{\varepsilon}}}{1 - \bar{\boldsymbol{\varepsilon}}^T \mathbf{D}_n^{-1}(z) \bar{\boldsymbol{\varepsilon}}} - \frac{\bar{\mathbf{v}}^T \mathcal{D}_n^{-2}(z) \bar{\mathbf{v}}}{1 + \bar{\mathbf{v}}^T \mathcal{D}_n^{-1}(z) \bar{\mathbf{v}}}. \quad (8.64)$$

By (8.32) and noting that  $M_n(z) = \operatorname{tr}\mathbf{D}_n^{-1}(z) - pm_{c_n}(z)$ , it is sufficient to prove the tightness of the following three terms:

$$g_{n1}(t) = -\frac{1}{2\pi i} \oint_{\gamma} e^{itz} M_n(z) dz, \quad (8.65)$$

$$g_{n2}(t) = -\frac{1}{2\pi i} \oint_{\gamma} e^{itz} \frac{\bar{\boldsymbol{\varepsilon}}^T \mathbf{D}_n^{-2}(z) \bar{\boldsymbol{\varepsilon}}}{1 - \bar{\boldsymbol{\varepsilon}}^T \mathbf{D}_n^{-1}(z) \bar{\boldsymbol{\varepsilon}}} dz, \quad (8.66)$$

$$g_{n3}(t) = -\frac{1}{2\pi i} \oint_{\gamma} e^{itz} \frac{\bar{\mathbf{v}}^T \mathcal{D}_n^{-2}(z) \bar{\mathbf{v}}}{1 + \bar{\mathbf{v}}^T \mathcal{D}_n^{-1}(z) \bar{\mathbf{v}}} dz, \quad (8.67)$$

The tightness of  $\{g_{n1}(t) : t \in I = [T_1, T_2]\}$  has been proved in Theorem 1. Next, via the same method adopted by Theorem 1, we prove the tightness of  $\{g_{ni}(t) : t \in I = [T_1, T_2]\}$ ,  $i = 2, 3$  as follows.

By (8.40), (8.41) and Slutsky's theorem, we have

$$\sup_{z \in \gamma} \left| \frac{\bar{\boldsymbol{\varepsilon}}^T \mathbf{D}_n^{-2}(z) \bar{\boldsymbol{\varepsilon}}}{1 - \bar{\boldsymbol{\varepsilon}}^T \mathbf{D}_n^{-1}(z) \bar{\boldsymbol{\varepsilon}}} + \frac{cm(z)}{z((1 + \underline{m}(z))^2 - c\underline{m}^2(z))} \right| \xrightarrow{i.p.} 0. \quad (8.68)$$

We conclude from (8.68), (8.61) and Lemma 3 that, as  $n \rightarrow \infty$ ,

$$\oint_{\gamma} \left| \frac{\bar{\boldsymbol{\varepsilon}}^T \mathbf{D}_n^{-2}(z) \bar{\boldsymbol{\varepsilon}}}{1 - \bar{\boldsymbol{\varepsilon}}^T \mathbf{D}_n^{-1}(z) \bar{\boldsymbol{\varepsilon}}} \right| |dz| \xrightarrow{a.s.} \oint_{\gamma} \left| \frac{cm(z)}{z((1 + \underline{m}(z))^2 - c\underline{m}^2(z))} \right| |dz| \quad (8.69)$$

and

$$\oint_{\gamma} \left| \frac{\bar{\mathbf{v}}^T \mathcal{D}_n^{-2}(z) \bar{\mathbf{v}}}{1 + \bar{\mathbf{v}}^T \mathcal{D}_n^{-1}(z) \bar{\mathbf{v}}} \right| |dz| \xrightarrow{a.s.} \oint_{\gamma} \left| \frac{\frac{cm(z)}{z[(1 + \underline{m}(z))^2 - c\underline{m}^2(z)]} + \bar{u} \int \frac{1}{(\lambda - z)^2} dF^{MP}(\lambda)}{\bar{u} z \underline{m}^2(z) - 1} \right| |dz|. \quad (8.70)$$

By (8.69), (8.70) and the same proof as (8.14) to (8.17), the tightness of  $\{g_{ni}(t) : t \in I = [T_1, T_2]\}$ ,  $i = 2, 3$  can be derived. □

## References

T. W. Anderson, An introduction to multivariate statistical analysis, *New York: Wiley, 2th edition*(1984).

- Z. D. Bai, B. Q. Miao, G. M. Pan, On asymptotics of eigenvectors of large sample covariance matrix, *Ann. Probab.*, **35(4)**(2007), 1532-1572.
- Z. D. Bai, H. Saranadasa, Effect of high dimension comparison of significance tests for a high dimensional two sample problem, *Statist. Sinica*, **6**(1996), 311-329.
- Z. D. Bai, J.W. Silverstein, CLT for linear spectral statistics of large-dimensional sample covariance matrices, *Ann. Probab.*, **32(1A)**(2004), 553-605.
- Z. D. Bai, J.W. Silverstein, Spectral analysis of large dimensional random matrices, *Springer*, **second edition**(2009).
- Z. D. Bai, W. Zhou, Large sample covariance matrices without independence structures in columns, *Statist. Sinica.*, **18**(2008), 425-442.
- P. Billingsley, Convergence of Probability Measures, Wiley, New York **first edition** (1968).
- P. Billingsley, Convergence of probability measures, *Wiley-Interscience Publication*, **second edition**(1999).
- T. W. Epps, A test for normality based on the empirical characteristic function, *Biometrika*, **70(3)**(1983), 723-726.
- J. Chen, J. Gao, D. Li, A new diagnostic test for cross-sectional uncorrelatedness in nonparametric panel data models, forthcoming in *Econometric Theory*, (2011).
- D. L. Donoho, High-dimensional data analysis: The curses and blessings of dimensionality, *American Math. Society Lecture, Math. Challenges of the 21st Century*, (2000).
- R. B. Dozier, J.W. Silverstein, On the empirical distribution of eigenvalues of large dimensional information-plus-noise type matrices, *J. Multivariate Anal.*, **98(4)**(2007), 678-694.
- J. Fan, J. Lv, A selective overview of variable selection in high dimensional feature space, *Statist. Sinica*, **20**(2010), 101-148.
- J. Fan, R. Li, Variable selection via nonconcave penalized likelihood and its oracle properties, *J. Am. Statist. Ass.*, **96(456)**(2001), 1348-1360.
- J. Fan, J. Lv, L. Qi, Sparse high dimensional models in economics, *Available at SSRN: <http://ssrn.com/abstract=1659322>*, (2011).
- C. H. Guo, N.J. Higham, A Schur-Newton method for the matrix  $p$ -th root and its inverse, *SIAM Journal on Matrix Analysis and Applications*, **28(3)**(2006).
- Y. M. Hong, Testing for pairwise serial independence via the empirical distribution function, *J. R. Statist. Soc. B*, **60(2)**(1998), 429-453.



- Y. M. Hong, Hypothesis testing in time series via the empirical characteristic function: a generalized spectral density approach, *J. Am. Statist. Ass.*, **94(448)**(1999), 1201-1220.
- C. Hsiao, Analysis of panel data, *Cambridge University Press*, **2nd edition**(2003).
- T. F. Jiang, The limiting distributions of eigenvalues of sample correlation matrices, *Sankhya.*, **66(1)**(2004), 35-48.
- I. Johnstone, on the distribution of the largest principal component, *Ann. Statist.*, **29**(2001), 295-327.
- I. Johnstone, D.M. Titterton, Statistical challenges of high-dimensional data, *Phil. Trans. R. Soc. A*, **367**(2009), 4237-4253.
- C. Hsiao, M. H. Pesaran, A. Pick, Diagnostic tests of cross section independence for nonlinear panel data models, *IZA discussion paper*, **No.27**(2007).
- C. M. Kuan and W.M. Lee, A new test for the martingale difference hypothesis, *Studies in Nonlinear Dynamics and Econometrics*, **8(4)**(2004), 1-26.
- V. A. Marcenko, L.A. Pastur, Distribution for some sets of random matrices, *Math. USSR-Sb.*, **1**(1967), 45783.
- J. Huang, J.L. Horowitz, S. Ma, Asymptotic properties of bridge estimators in sparse high-dimensional regression models, *Ann. Statist.*, **36(2)**(2008), 587-613.
- G. M. Pan, W. Zhou, Central limit theorem for signal-to-interference ratio of reduced rank linear receiver, *Ann. Appl. Probab.*, **18(3)**(2008), 1232-1270.
- G. M. Pan, W. Zhou, Central limit theorem for Hotelling's  $T^2$  statistic under large dimension, *Ann. Appl. Probab.*, **21(5)**(2011), 1860-1910.
- G. M. Pan, Comparison between two types of large sample covariance matrices, *submitted*, (2011).
- M. H. Pesaran, General diagnostic tests for cross section dependence in panels, *Cambridge Working Paper in Economics*, **No.0435**(2004).
- Ø. Ryan, M. Debbah, Asymptotic behaviour of random vandermonde matrices with entries on the unit circle, *IEEE Trans. on Information Theory*, **55(7)**(2009), 3115-3148.
- J. W. Silverstein, Strong convergence of the empirical distribution of eigenvalues of large dimensional random matrices, *J. Multivariate Anal.* **55**(1995), 331-339.
- L. Su and A. Ullah, Testing conditional uncorrelatedness. *Journal of Business and Economic Statistics* **27** (2009), 18-29.
- H. Xiao, W. Zhou, On the limit of the smallest eigenvalue of some sample covariance matrix, *J. Theor. Probab.*, **23**(2010), 1-20.

Figure 1: Empirical Spectral Distribution of ARCH(1) and I.I.D Normal Vector

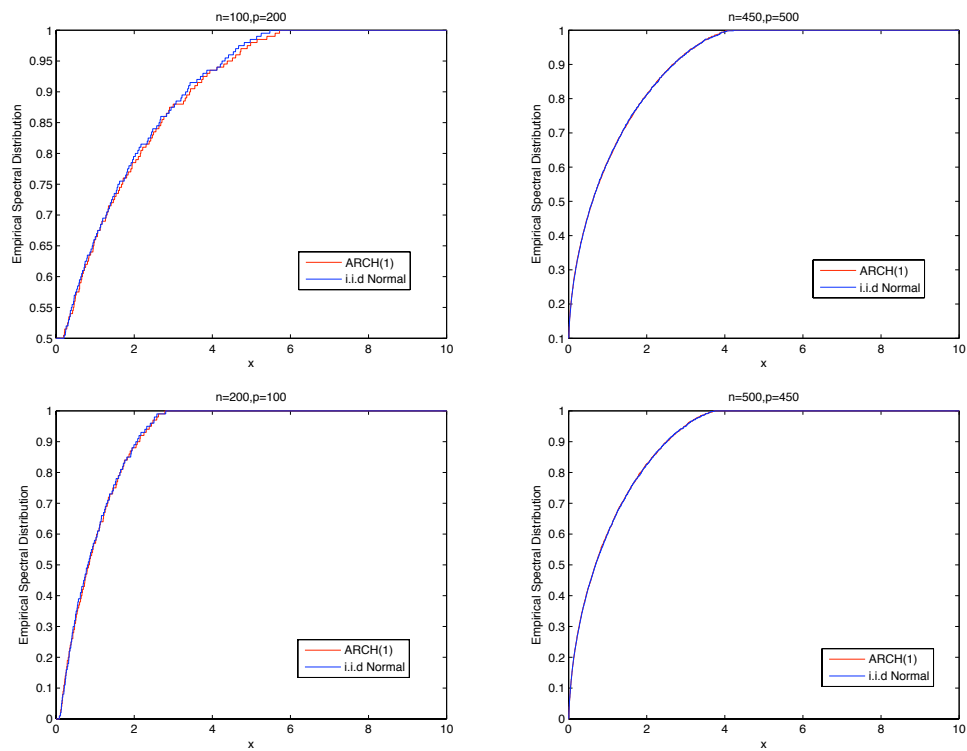
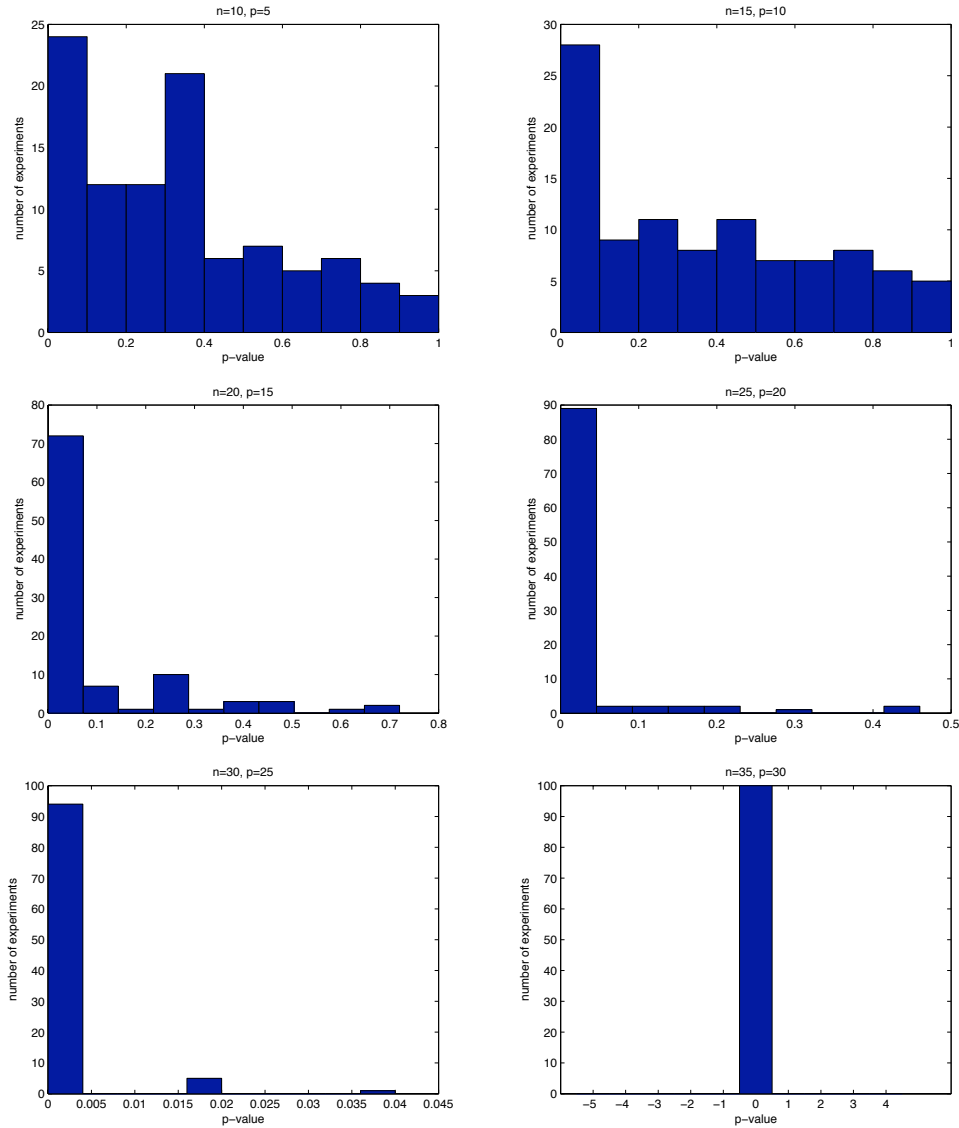
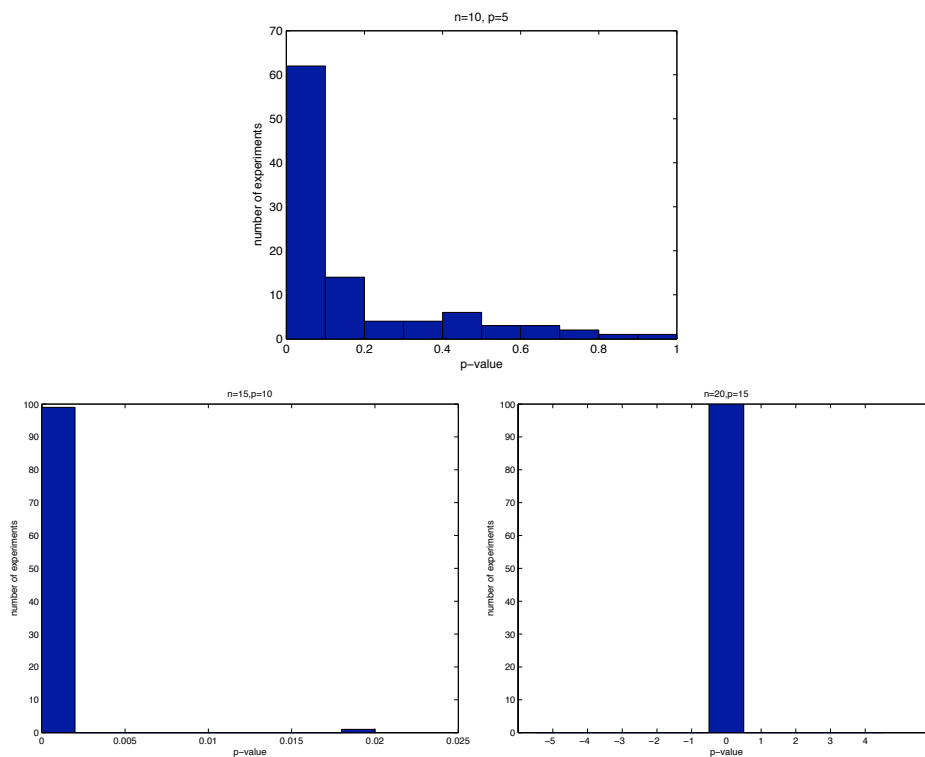


Figure 2: P-values of the proposed test for daily closed stock prices from NYSE



\*These are P-values for  $p$  companies from NYSE, each of which has  $n$  closed stock prices during the period 1990.1.1 – 2002.1.1. The number of repeated experiments are 100. All the closed stock prices are from WRDS database.

Figure 3: P-values of the proposed test for daily closed stock prices from transportation sections of NYSE



\*These are P-values for  $p$  companies from transportation sections of NYSE, each of which has  $n$  closed stock prices during the period 1990.1.1 – 2002.1.1. The number of repeated experiments are 100. All the closed stock prices are from WRDS database.

Table 1: Empirical sizes and powers of the proposed test at significant level 0.05 for standardized normal distributed random vectors.

n	p										
	5	10	20	30	40	50	60	70	80	90	100
Empirical sizes											
5	0.069	0.064	0.027	0.039	0.034	0.052	0.035	0.033	0.037	0.043	0.047
10	0.049	0.059	0.048	0.046	0.055	0.044	0.043	0.048	0.048	0.038	0.046
20	0.045	0.042	0.047	0.051	0.052	0.052	0.051	0.041	0.042	0.046	0.041
30	0.034	0.035	0.052	0.053	0.068	0.063	0.050	0.049	0.047	0.040	0.046
40	0.055	0.050	0.052	0.061	0.056	0.038	0.042	0.064	0.054	0.044	0.066
50	0.038	0.041	0.048	0.044	0.050	0.052	0.050	0.058	0.046	0.053	0.041
60	0.064	0.061	0.043	0.049	0.047	0.049	0.050	0.049	0.059	0.046	0.048
70	0.047	0.042	0.051	0.039	0.052	0.055	0.058	0.044	0.037	0.038	0.049
80	0.028	0.032	0.033	0.041	0.052	0.054	0.047	0.042	0.047	0.046	0.040
90	0.042	0.037	0.032	0.048	0.054	0.044	0.053	0.041	0.037	0.051	0.042
100	0.037	0.046	0.036	0.035	0.045	0.029	0.045	0.056	0.048	0.047	0.056
Empirical powers											
5	0.044	0.053	0.072	0.070	0.065	0.048	0.053	0.063	0.057	0.053	0.080
10	0.054	0.065	0.046	0.051	0.052	0.049	0.049	0.037	0.049	0.060	0.062
20	0.056	0.066	0.091	0.059	0.062	0.051	0.045	0.075	0.078	0.078	0.063
30	0.051	0.062	0.054	0.068	0.051	0.081	0.092	0.087	0.115	0.128	0.148
40	0.041	0.036	0.052	0.055	0.126	0.183	0.144	0.208	0.209	0.276	0.254
50	0.068	0.051	0.054	0.066	0.151	0.262	0.329	0.420	0.474	0.432	0.574
60	0.046	0.027	0.055	0.098	0.154	0.310	0.455	0.592	0.596	0.733	0.715
70	0.050	0.044	0.018	0.110	0.211	0.379	0.565	0.747	0.793	0.846	0.841
80	0.053	0.046	0.036	0.102	0.233	0.439	0.687	0.892	0.906	0.944	0.977
90	0.028	0.040	0.054	0.122	0.226	0.484	0.752	0.878	0.963	0.980	0.999

\*The powers are under the alternative hypothesis that the population covariance matrix is  $\Sigma = 0.05I_p + 0.95\mathbf{1}_p\mathbf{1}'_p$ .

Table 2: Empirical sizes and powers of the LRT at significant level 0.05 for standardized normal distributed random vectors.

n	p										
	5	10	20	30	40	50	60	70	80	90	100
Empirical sizes											
5	0.005	0.475	0.495	0	0	0	0	0	0	0	0
10	0.030	0.115	0.504	0.454	0	0	0	0	0	0	0
20	0.047	0.030	0.522	1	0.500	0	0	0	0	0	0
30	0.040	0.032	0.076	0.873	1	1	0	0	0	0	0
40	0.044	0.034	0.057	0.148	0.982	1	1	1	0	0	0
50	0.057	0.034	0.044	0.062	0.299	1	1	1	1	1	0
60	0.049	0.053	0.040	0.065	0.101	0.578	1	1	1	1	1
70	0.038	0.060	0.049	0.054	0.108	0.229	0.784	1	1	1	1
80	0.055	0.040	0.043	0.072	0.079	0.147	0.442	0.950	1	1	1
90	0.054	0.061	0.047	0.033	0.048	0.129	0.233	0.691	0.991	1	1
100	0.055	0.048	0.066	0.069	0.076	0.085	0.185	0.402	0.844	0.997	1
Empirical powers											
5	0.005	0.504	0.505	0	0	0	0	0	0	0	0
10	0.033	0.125	0.506	0.485	0	0	0	0	0	0	0
20	0.051	0.049	0.587	1	0.496	0	0	0	0	0	0
30	0.056	0.062	0.121	0.912	1	1	0	0	0	0	0
40	0.061	0.077	0.139	0.311	0.990	1	1	1	0	0	0
50	0.092	0.090	0.150	0.206	0.576	1	1	1	1	1	0
60	0.088	0.142	0.220	0.266	0.420	0.849	1	1	1	1	1
70	0.102	0.170	0.222	0.320	0.463	0.691	0.984	1	1	1	1
80	0.133	0.169	0.246	0.392	0.478	0.666	0.883	0.994	1	1	1
90	0.122	0.190	0.331	0.404	0.526	0.666	0.819	0.983	1	1	1
100	0.153	0.232	0.411	0.550	0.658	0.713	0.842	0.952	0.995	1	1

\*The powers are under the alternative hypothesis that the population covariance matrix is  $\Sigma = 0.05I_p + 0.95\mathbf{1}_p\mathbf{1}'_p$ .

Table 3: Empirical sizes and powers of the proposed test at 0.05 significance level for standardized gamma distributed random vectors.

n	P										
	5	10	20	30	40	50	60	70	80	90	100
Empirical sizes											
5	0.089	0.078	0.068	0.059	0.060	0.063	0.061	0.047	0.050	0.061	0.044
10	0.066	0.075	0.068	0.054	0.046	0.051	0.047	0.048	0.050	0.042	0.052
20	0.057	0.058	0.072	0.062	0.053	0.054	0.053	0.052	0.049	0.051	0.048
30	0.069	0.069	0.069	0.076	0.058	0.056	0.054	0.047	0.067	0.066	0.054
40	0.061	0.049	0.058	0.040	0.065	0.048	0.063	0.065	0.068	0.047	0.067
50	0.053	0.054	0.055	0.057	0.059	0.048	0.067	0.066	0.043	0.059	0.053
60	0.059	0.052	0.057	0.060	0.052	0.067	0.058	0.064	0.064	0.061	0.069
70	0.044	0.050	0.064	0.055	0.071	0.054	0.067	0.064	0.051	0.077	0.048
80	0.045	0.050	0.061	0.043	0.071	0.055	0.071	0.053	0.056	0.070	0.060
90	0.041	0.067	0.034	0.056	0.049	0.054	0.050	0.060	0.047	0.060	0.058
100	0.070	0.045	0.059	0.055	0.047	0.062	0.069	0.057	0.056	0.060	0.061
Empirical powers											
5	0.334	0.575	0.853	0.944	0.983	0.989	0.998	0.997	0.999	0.998	0.998
10	0.513	0.838	0.979	0.999	1	1	1	1	1	1	1
20	0.721	0.970	0.999	1	1	1	1	1	1	1	1
30	0.834	0.998	1	1	1	1	1	1	1	1	1
40	0.914	1	1	1	1	1	1	1	1	1	1
50	0.952	1	1	1	1	1	1	1	1	1	1
60	0.991	1	1	1	1	1	1	1	1	1	1
70	0.992	1	1	1	1	1	1	1	1	1	1
80	0.998	1	1	1	1	1	1	1	1	1	1
90	0.999	1	1	1	1	1	1	1	1	1	1
100	1	1	1	1	1	1	1	1	1	1	1

\*The powers are under the alternative hypothesis that the population covariance matrix is  $\Sigma = 0.05I_p + 0.95\mathbf{1}_p\mathbf{1}'_p$ .

Table 4: Empirical powers of the proposed test at 0.05 significance level for MA(1) model.

n	P										
	5	10	20	30	40	50	60	70	80	90	100
5	0.096	0.097	0.097	0.111	0.210	0.214	0.198	0.221	0.201	0.223	0.199
10	0.082	0.090	0.173	0.227	0.374	0.715	0.722	0.809	0.903	0.913	0.934
20	0.099	0.165	0.400	0.597	0.683	0.822	0.951	0.994	1	1	1
30	0.067	0.121	0.611	0.733	0.803	0.986	1	1	1	1	1
40	0.091	0.321	0.653	0.938	0.968	1	1	1	1	1	1
50	0.139	0.416	0.910	0.956	0.998	1	1	1	1	1	1
60	0.117	0.412	0.918	0.994	1	1	1	1	1	1	1
70	0.146	0.535	0.983	1	1	1	1	1	1	1	1
80	0.138	0.620	0.990	1	1	1	1	1	1	1	1
90	0.239	0.805	0.999	1	1	1	1	1	1	1	1
100	0.132	0.761	1	1	1	1	1	1	1	1	1

Table 5: Empirical powers of the proposed test at 0.05 significance level for AR(1) model.

n	$p$										
	5	10	20	30	40	50	60	70	80	90	100
5	0.130	0.094	0.264	0.253	0.181	0.169	0.191	0.134	0.144	0.129	0.113
10	0.167	0.289	0.482	0.640	0.668	0.609	0.563	0.524	0.427	0.384	0.321
20	0.230	0.544	0.878	0.271	0.954	0.994	0.999	0.997	0.996	0.984	0.974
30	0.205	0.602	0.993	0.999	0.671	0.936	1	1	1	1	1
40	0.344	0.916	0.998	1	1	0.982	1	1	1	1	1
50	0.541	0.984	1	1	1	1	1	1	1	1	1
60	0.571	0.986	1	1	1	1	1	1	1	1	1
70	0.672	0.997	1	1	1	1	1	1	1	1	1
80	0.673	0.999	1	1	1	1	1	1	1	1	1
90	0.882	1	1	1	1	1	1	1	1	1	1
100	0.801	1	1	1	1	1	1	1	1	1	1

Table 6: Empirical powers of the proposed test at 0.05 significance level for SMA(1) model.

n	$p$										
	5	10	20	30	40	50	60	70	80	90	100
5	0.330	0.431	0.782	0.999	1	1	1	1	1	1	1
10	0.647	1	1	1	1	1	1	1	1	1	1
20	0.967	1	1	1	1	1	1	1	1	1	1
30	0.962	1	1	1	1	1	1	1	1	1	1
40	0.998	1	1	1	1	1	1	1	1	1	1
50	1	1	1	1	1	1	1	1	1	1	1
60	1	1	1	1	1	1	1	1	1	1	1
70	1	1	1	1	1	1	1	1	1	1	1
80	1	1	1	1	1	1	1	1	1	1	1
90	1	1	1	1	1	1	1	1	1	1	1
100	1	1	1	1	1	1	1	1	1	1	1



Table 7: Empirical sizes and powers of the proposed test at 0.05 significance level for the general panel data model.

n	$p$										
	5	10	20	30	40	50	60	70	80	90	100
Empirical sizes											
5	0.038	0.045	0.047	0.057	0.058	0.067	0.069	0.069	0.072	0.076	0.074
10	0.041	0.042	0.046	0.049	0.056	0.050	0.065	0.047	0.068	0.063	0.069
20	0.035	0.042	0.049	0.056	0.052	0.046	0.063	0.043	0.069	0.055	0.057
30	0.043	0.049	0.043	0.059	0.048	0.068	0.059	0.057	0.040	0.055	0.047
40	0.048	0.052	0.043	0.060	0.057	0.046	0.049	0.054	0.046	0.058	0.061
50	0.057	0.048	0.046	0.058	0.052	0.055	0.048	0.049	0.050	0.040	0.041
60	0.058	0.056	0.055	0.048	0.047	0.045	0.053	0.066	0.058	0.049	0.050
70	0.062	0.060	0.059	0.056	0.049	0.057	0.049	0.068	0.052	0.036	0.043
80	0.071	0.063	0.067	0.047	0.048	0.058	0.056	0.044	0.059	0.057	0.055
90	0.065	0.068	0.065	0.048	0.053	0.048	0.056	0.048	0.048	0.066	0.060
100	0.037	0.046	0.036	0.035	0.045	0.043	0.045	0.056	0.048	0.047	0.055
Empirical powers											
5	0.150	0.238	0.345	0.417	0.484	0.529	0.549	0.615	0.611	0.668	0.692
10	0.125	0.247	0.452	0.526	0.568	0.633	0.669	0.737	0.765	0.751	0.800
20	0.206	0.343	0.493	0.615	0.673	0.752	0.788	0.813	0.860	0.864	0.876
30	0.111	0.404	0.535	0.684	0.757	0.756	0.855	0.875	0.882	0.909	0.953
40	0.308	0.393	0.605	0.698	0.786	0.820	0.878	0.898	0.944	0.959	0.953
50	0.207	0.450	0.603	0.718	0.815	0.889	0.923	0.938	0.966	0.973	0.980
60	0.268	0.430	0.594	0.780	0.826	0.918	0.913	0.926	0.974	0.976	0.984
70	0.144	0.434	0.649	0.798	0.888	0.883	0.944	0.968	0.971	0.982	0.996
80	0.171	0.454	0.678	0.796	0.872	0.921	0.938	0.967	0.989	0.992	0.995
90	0.204	0.431	0.683	0.834	0.874	0.916	0.963	0.985	0.985	0.994	0.994
100	0.291	0.398	0.687	0.836	0.884	0.931	0.973	0.987	0.992	0.994	1

Table 8: Empirical powers of the proposed test at 0.05 significance level for nonlinear MA model.

n	$p$										
	5	10	20	30	40	50	60	70	80	90	100
5	0.033	0.023	0.004	0.008	0.005	0.007	0.008	0.007	0.009	0.014	0.070
10	0.618	0.543	0.374	0.220	0.108	0.037	0.011	0	0	0	0
20	0.804	0.744	0.703	0.614	0.581	0.511	0.447	0.340	0.306	0.257	0.216
30	0.854	0.815	0.777	0.779	0.780	0.740	0.662	0.662	0.597	0.579	0.555
40	0.878	0.842	0.856	0.856	0.845	0.825	0.779	0.770	0.772	0.702	0.698
50	0.884	0.899	0.868	0.884	0.864	0.888	0.860	0.875	0.869	0.828	0.820
60	0.892	0.912	0.882	0.904	0.920	0.923	0.933	0.900	0.892	0.892	0.882
70	0.896	0.934	0.906	0.927	0.934	0.921	0.952	0.925	0.943	0.917	0.926
80	0.936	0.886	0.925	0.921	0.952	0.950	0.943	0.943	0.958	0.953	0.936
90	0.922	0.925	0.939	0.935	0.958	0.959	0.955	0.982	0.962	0.957	0.954
100	0.926	0.935	0.920	0.937	0.952	0.964	0.970	0.975	0.978	0.970	0.965

Table 9: Empirical powers of the proposed test at 0.05 significance level for ARCH(1) model.

n	<i>P</i>										
	5	10	20	30	40	50	60	70	80	90	100
5	0.148	0.093	0.073	0.076	0.058	0.099	0.113	0.088	0.106	0.112	0.141
10	0.454	0.399	0.315	0.215	0.128	0.121	0.112	0.121	0.093	0.095	0.092
20	0.401	0.428	0.508	0.464	0.444	0.468	0.436	0.338	0.356	0.337	0.302
30	0.272	0.364	0.616	0.712	0.753	0.759	0.793	0.743	0.638	0.639	0.598
40	0.232	0.357	0.572	0.823	0.790	0.874	0.915	0.885	0.876	0.855	0.860
50	0.222	0.339	0.622	0.757	0.891	0.969	0.957	0.975	0.967	0.957	0.975
60	0.216	0.448	0.617	0.862	0.901	0.976	0.983	0.987	0.992	0.997	0.997
70	0.200	0.339	0.592	0.864	0.931	0.982	0.993	0.998	0.997	0.998	0.996
80	0.194	0.376	0.566	0.824	0.950	0.967	0.997	1	1	0.999	1
90	0.184	0.456	0.721	0.839	0.960	0.995	0.999	1	0.999	1	1
100	0.128	0.306	0.802	0.859	0.934	0.992	1	1	1	1	1

Table 10: Empirical powers of the proposed test at 0.05 significance level for Vandermonde Matrix.

n	<i>P</i>										
	10	20	30	40	50	60	70	80	90	100	120
10	0.180	0.197	0.192	0.169	0.202	0.211	0.190	0.185	0.171	0.177	0.240
20	0.309	0.332	0.356	0.327	0.291	0.303	0.301	0.295	0.321	0.243	0.478
30	0.324	0.433	0.473	0.413	0.461	0.408	0.445	0.395	0.368	0.397	0.606
40	0.458	0.512	0.527	0.546	0.533	0.490	0.518	0.498	0.450	0.457	0.655
50	0.593	0.437	0.540	0.571	0.614	0.569	0.577	0.566	0.565	0.537	0.764
60	0.504	0.538	0.551	0.567	0.616	0.662	0.588	0.581	0.572	0.607	0.744
70	0.548	0.526	0.560	0.627	0.668	0.641	0.694	0.707	0.641	0.678	0.741
80	0.550	0.545	0.580	0.633	0.712	0.719	0.693	0.768	0.729	0.749	0.805
90	0.589	0.544	0.596	0.667	0.695	0.712	0.743	0.754	0.738	0.728	0.807
100	0.464	0.549	0.610	0.645	0.704	0.757	0.772	0.751	0.752	0.808	0.928
120	0.633	0.660	0.736	0.737	0.759	0.855	0.854	0.909	0.960	0.999	1