ISSN 1440-771X

MONASH University

Australia

Department of Econometrics and Business Statistics

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

Cross-sectional Independence Test for a Class of Parametric Panel Data Models

Guangming Pan, Jiti Gao, Yanrong Yang and Meihui Guo

September 2015

Working Paper 17/15

Cross-sectional Independence Test for a Class of Parametric Panel Data Models

Guangming Pan[*], Jiti Gao, Yanrong Yang[†] and Meihui Guo[‡]

September 30, 2015

Abstract

This paper proposes a new statistic to conduct cross-sectional independence test for the residuals involved in a parametric panel data model. The proposed test statistic, which is called linear spectral statistic (LSS), is established based on the characteristic function of the empirical spectral distribution (ESD) of the sample correlation matrix of the residuals. The main advantage of the proposed test statistic is that it can capture nonlinear cross-sectional dependence. Asymptotic theory for a general class of linear spectral statistics is established, as the cross-sectional dimension N and time length T go to infinity proportionally. This type of statistics covers many classical statistics, including the bias-corrected Lagrange Multiplier (LM) test statistic and the likelihood ratio test statistic. Furthermore, the power under a local alternative hypothesis is analyzed and the asymptotic distribution of the proposed statistic under this local hypothesis is also established. Finite sample performance shows that the proposed test statistic works well numerically in each individual case and it can also distinguish some dependent but uncorrelated structures, for example, nonlinear MA(1) models and multiple ARCH(1) models.

Keywords: Characteristic function, cross–sectional independence, empirical spectral distribution, linear panel data models, Marcenko-Pastur Law.

JEL: C12, C21, C22.

[*]Associate Professor Guangming Pan is from School of Physical and Mathematical Sciences, Nanyang Technological University, Singapore;

[†]Professor Jiti Gao and Dr. Yanrong Yang are from Department of Econometrics and Business Statistics, Monash University, Australia;

[‡]Professor Meihui Guo is from Department of Applied Mathematics, National Sun Yat-sen University, Taiwan.

1 Introduction

Cross-sectional dependence has been widely studied in panel data analysis. It plays an important role in economic and financial models and creates great challenges to classical statistical inference. For example, the existence of cross-sectional dependence can lead to the loss of efficiency of the classical least-square estimation method. Before imposing any structure on models under study, it is necessary to test whether there is a type of cross-sectional dependence. The econometrics literature basically discusses about how to test for cross-sectional uncorrelatedness in panel data analysis. Under the case of fixed N and large T, Breusch and Pagan (1980) proposed Lagrange multiplier (LM) test statistic which is based on the average of correlation coefficients of the residuals. For large N and large T, Pesaran, Ullah and Yamagata (2008) developed a bias-adjusted LM test using finite sample approximations. Recently, Baltagi, Feng and Kao (2012) derived the asymptotic distribution of a scaled LM test statistic proposed in Pesaran (2004). However, both papers assume normally distributed error components. Pesaran (2004) provided a diagnostic test for parametric linear models based on the average of the sample correlations as N and T are comparable, which is called the CD test. Chen, Gao and Li (2012) extended the CD test to nonparametric nonlinear models. Other related studies include Su and Ullah (2009) for testing conditional uncorrelation through examining a covariance matrix in the case of N being fixed. Meanwhile, Schott (2005) also established an asymptotic distribution for a scaled LM test statistic for high dimensional normally distributed data. Bai and Silverstein (2004) analyzed this kind of statistics based on sample covariance matrices, and Bai, el. (2009) utilized it to develop an asymptotic theory for likelihood ratio (LR) statistics under high dimensional settings.

Since the population mean and variance of the original data are usually unknown, sample covariance matrices cannot provide us with sufficient and correct information about the data. In order to address such issues, Gao, el. (2014) proposed using linear spectral statistics of sample correlation matrices. One of the main advantage of using sample correlation matrices over sample covariance matrices is that it does not require the first two population moments of the elements of the random vector under study to be known. In this paper, we further explore the idea of using the characteristic function of the empirical spectral distribution (ESD) of the sample correlation matrix of the data under study. We then propose a new test statistic for testing cross–sectional independence of the cross–sectional residuals involved in a class of parametric panel data models. The construction of the new test statistic is based on the fact that it is a sum of the characteristic function of each eigenvalue of the sample correlation matrix. In view of this, this statistic includes the high order moments of the residuals under investigation. Due to possible nonlinear dependence being reflected by the relationship among high order moments of the residuals, our proposed statistic is applicable to distinguish various dependent structures. In view of this point, we are able to test for cross–sectional independence rather than just cross–sectional uncorrelatedness, as has been discussed in the econometrics literature (see, for example, Pesaran (2004)).

In terms of the comparison with the work by Gao, el. (2014), we can stress the following points. First, this paper deals with the case where the cross–sectional residuals are unobservable. By contrast, Gao, el. (2014) considered a vector of observable random variables. Second, Gao, el. (2014) focused on the case where the observed random variables are all independent and identically distributed. By contrast, this paper allows that the cross–sectional residuals can be either independent or dependent. We then establish new asymptotic distributions for the proposed test statistic for such cases. The main difficulty involved in the establishment of the main results of this paper is that the estimated versions of the cross–sectional residuals are always highly dependent even when the cross–sectional residuals themselves are assumed to be independent in Sections 2 and 3. We should also point out that Section 4 then demonstrates both the effectiveness and the strength of the proposed test statistic for capturing some weak dependence structures. As a consequence, the proposed test is applicable to test for cross–sectional dependence among some commonly used econometric models, such as spatial moving average, dependent factor, nonlinear moving average and multiple ARCH models.

The rest of the paper is organized as follows. Section 2 introduces the proposed test statistic and some results related to large dimensional random matrix theory. Asymptotic theory is presented in Section 3, including the asymptotic distribution of the proposed test statistic under the null hypothesis and the power under a general class of local alternative hypotheses. Section 4 specifically studies a local alternative hypothesis, under which the asymptotic distribution of the new statistic is demonstrated. In Section 5, the finite sample performance illustrates the effectiveness of the proposed test statistic under different dependent structures, including some dependent but uncorrelated structures. Conclusions are in Section 6. All the mathematical proofs are given in Appendix A, and computation code functions are displayed in Appendix B.

2 The Model and test statistics

Consider a parametric linear panel data model of the form

$$y_{jt} = \alpha_j + \mathbf{x}_{jt}^{\tau}\boldsymbol{\beta} + u_{jt}, \quad j = 1, 2, \ldots, N; \ t = 1, 2, \ldots, T, \tag{2.1}$$

where j indexes the j-th cross-sectional unit and t indexes the t-th time series observation; y_{jt} is the dependent variable; $\mathbf{x}_{jt}$ denotes the p-dimension regressors with the slope parameter $\boldsymbol{\beta}$; α_j is the fixed effect with $\sum_{j=1}^{N} \alpha_j = 0$ for the identifiability of the model (2.1); and the error component u_{jt} is allowed to be cross-sectionally dependent but uncorrelated with $\mathbf{x}_{jt}$.

The aim of this paper is to conduct a cross-sectional independence test as follows.

$$\mathbb{H}_0 : \{u_{jt}\} \ \ is \ \ independent \ \ of \ \ \{u_{rt}\} \ \ for \ \ all \ \ j \neq r; \tag{2.2}$$

against

$$\mathbb{H}_a : \{u_{jt}\} \ and \ \{u_{rt}\} \ are \ dependent \ for \ some \ j \neq r. \tag{2.3}$$

Before proposing a new test statistic, we write the model (2.1) into a centralized form that is suitable for deriving our test statistic. Minus the average $\bar{y}_{j\cdot} := \frac{1}{T}\sum_{t=1}^{T} y_{jt}$ on both sides of (2.1), we have

$$\widetilde{y}_{jt} = \widetilde{\mathbf{x}}_{jt}^{\tau}\boldsymbol{\beta} + \widetilde{u}_{jt}, \quad j = 1, 2, \ldots, N; \; t = 1, 2, \ldots, T,$$

where $\widetilde{y}_{jt} = y_{jt} - \bar{y}_{j\cdot}$, $\widetilde{\mathbf{x}}_{jt} = \mathbf{x}_{jt} - \bar{\mathbf{x}}_{j\cdot}$ and $\widetilde{u}_{jt} = u_{jt} - \bar{u}_{j\cdot}$ with $\bar{\mathbf{x}}_{j\cdot} = \frac{1}{T}\sum_{t=1}^{T}\mathbf{x}_{jt}$ and $\bar{u}_{j\cdot} = \frac{1}{T}\sum_{t=1}^{T} u_{jt}$.

The vector form of model (2.1) is

$$\widetilde{\mathbf{Y}}_j = \widetilde{\mathbf{X}}_j\boldsymbol{\beta} + \widetilde{\mathbf{u}}_j, \quad j = 1, 2, \ldots, N, \tag{2.4}$$

where $\widetilde{\mathbf{Y}}_j = (\widetilde{y}_{j1}, \widetilde{y}_{j2}, \ldots, \widetilde{y}_{jT})^{\tau}$, $\widetilde{\mathbf{X}}_j = (\widetilde{\mathbf{x}}_{j1}, \widetilde{\mathbf{x}}_{j2}, \ldots, \widetilde{\mathbf{x}}_{jT})^{\tau}$ and $\widetilde{\mathbf{u}}_j = (\widetilde{u}_{j1}, \widetilde{u}_{j2}, \ldots, \widetilde{u}_{jT})^{\tau}$.

Under the null hypothesis, the least squares estimator of $\boldsymbol{\beta}$ is

$$\hat{\boldsymbol{\beta}} = (\sum_{j=1}^{N}\widetilde{\mathbf{X}}_j^{\tau}\widetilde{\mathbf{X}}_j)^{-1}\sum_{j=1}^{N}\widetilde{\mathbf{X}}_j^{\tau}\widetilde{\mathbf{Y}}_j.$$

Then the estimator for $\mathbf{u}_j$, $j = 1, 2, \ldots, N$ is

$$\hat{\mathbf{u}}_j = \widetilde{\mathbf{Y}}_j - \widetilde{\mathbf{X}}_j(\sum_{r=1}^{N}\widetilde{\mathbf{X}}_r^{\tau}\widetilde{\mathbf{X}}_r)^{-1}(\sum_{r=1}^{N}\widetilde{\mathbf{X}}_r^{\tau}\widetilde{\mathbf{Y}}_r). \tag{2.5}$$

We are now ready to introduce linear spectral statistics for cross-sectional independence test (2.2). Consider the sample correlation matrix

$$\hat{\mathbf{R}}_N = (\hat{\rho}_{rj})_{N\times N}, \quad \textit{with} \quad \hat{\rho}_{rj} = \frac{\hat{\mathbf{u}}_r^{\tau}\hat{\mathbf{u}}_j}{||\hat{\mathbf{u}}_r|| \cdot ||\hat{\mathbf{u}}_j||}. \tag{2.6}$$

Let us study a class of statistics related to eigenvalues of the matrix $\hat{\mathbf{R}}_N$. First, the empirical spectral distribution (ESD) of the sample correlation matrix $\hat{\mathbf{R}}_N$ is defined as

$$F^{\hat{\mathbf{R}}_N}(x) = \frac{1}{N}\sum_{j=1}^{N} I(\lambda_j \le x),$$

where $\lambda_1 \le \lambda_2 \le \ldots \le \lambda_N$ are the eigenvalues of $\hat{\mathbf{R}}_N$ and $I(\cdot)$ is an indicator function.

Before we establish the proposed test statistic and the main asymptotic theory, we introduce two assumptions:

Assumption 1. *For each $j = 1, 2\ldots, N$, $\{u_{j1}, \cdots, u_{jT}\}$ are independent and identical distributed (i.i.d) random variables with $\mathbb{E}u_{jt} = 0$ and $\mathbb{E}u_{jt}^4 < \infty$. $\{\mathbf{x}_{jt} : j = 1, 2, \ldots, N; t = 1, 2, \ldots, T\}$ are i.i.d and each $\mathbf{x}_{jt} : p \times 1$ has i.i.d components with zero mean and finite fourth moments. Moreover, $\{u_{jt} : j = 1, 2, \ldots, N; t = 1, 2, \ldots, T\}$ and $\{\mathbf{x}_{jt} : j = 1, 2, \ldots, N; t = 1, 2, \ldots, T\}$ are independent.*

Assumption 2. *$N = N(T)$ with $\frac{N}{T} \to c \in (0, \infty)$ as $T \to \infty$.*

The strategy of analyzing the ESD of $\hat{\mathbf{R}}_N$ is divided into two steps. The first step is to investigate the eigenvalues of the matrix $\mathbf{R}_N = (\rho_{rj})_{N\times N}$ with ρ_{rj} being $\hat{\rho}_{rj}$ by replacing $\hat{\mathbf{u}}_r$ with $\widetilde{\mathbf{u}}_r$, while the second step compares the eigenvalues of $\hat{\mathbf{R}}_N$ with those of $\mathbf{R}_N$.

If $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_N$ are independent, $F^{\mathbf{R}_N}(x)$ converges with probability one to the Marcenko-Pastur (simply called M-P) law $F_c(x)$ with $c = \lim_{T\to\infty} N/T$ (see Jiang (2004)), whose density has an explicit expression of the form

$$f_c(x) = \begin{cases} \frac{1}{2\pi x c}\sqrt{(b-x)(x-a)}, & a \le x \le b; \\ 0, & otherwise; \end{cases} \tag{2.7}$$

and a point mass $1-1/c$ at the origin if $c > 1$, where $a = (1-\sqrt{c})^2$ and $b = (1+\sqrt{c})^2$. In the following section, we will prove that $F^{\hat{\mathbf{R}}_N}(x)$ has the same limit as $F^{\mathbf{R}_N}(x)$.

Based on the difference between the empirical spectral distribution $F^{\hat{\mathbf{R}}_N}(x)$ and M-P law $F_{c_N}(x)$ (which is $F_c(x)$ with c replaced by $c_N = N/T$), our test statistic is proposed to distinguish $\mathbb{H}_0$ from $\mathbb{H}_a$. Next, we study a new class of statistics called linear spectral statistics (LSS). LSS for the sample correlation matrix $\hat{\mathbf{R}}_N$ is of the form

$$\frac{1}{N}\sum_{j=1}^{N} f(\lambda_j) = \int f(x)dF^{\hat{\mathbf{R}}_N}(x),$$

where $f(\cdot)$ is an analytic function on $[0,\infty)$.

Consider a modified linear spectral statistic of the form:

$$T_N(f) = \int f(x)dG_N(x), \tag{2.8}$$

where $G_N(x) = N[F^{\hat{\mathbf{R}}_N}(x) - F_{c_N}(x)]$.

The linear spectral statistic $T_N(f)$ is a general statistic in the sense that it covers some classical statistics as special cases.

1. Schott's Statistic (Schott (2005)):

$$f_1(x) = x^2 - x: \quad T_N(f_1) = tr(\hat{\mathbf{R}}_N^2) - N - N\cdot\int (x^2+x)dF_{c_N}(x). \tag{2.9}$$

2. The Likelihood Ratio Test Statistic (Morrison (2005)):

$$f_2(x) = \log(x): \quad T_N(f_2) = \sum_{j=1}^{N}\log(\lambda_j) - N\cdot\int \log(x)dF_{c_N}(x), \tag{2.10}$$

   where $\lambda_i : i = 1, 2, \ldots, N$ are eigenvalues of $\hat{\mathbf{R}}_N$.

The construction of our proposed test statistic mainly comes from the following observation: under the null hypothesis, the limit of the ESD of the sample correlation matrix $\hat{\mathbf{R}}_N$ is the M-P law defined in (2.7) when $\mathbf{u}_1, \cdots, \mathbf{u}_N$ satisfy Assumptions 1 and 2. Moreover, numerical investigations indicate that when $\mathbf{u}_1, \cdots, \mathbf{u}_N$ are only uncorrelated instead of independent, the limit of the ESD of $\hat{\mathbf{R}}_N$ is not the M-P law (see Ryan and Debbah (2009)). From this point, any deviation of the limit of the ESD from the M-P law is evidence of dependence. Hence these motivate us to use the ESD of $\hat{\mathbf{R}}_N$, $F^{\hat{\mathbf{R}}_N}(x)$, as a test statistic. However, there is no central limit theorem available for $(F^{\hat{\mathbf{R}}_N}(x) - F_{c_N}(x))$, as argued by Bai and Silverstein (2004). Therefore, instead, we consider the difference between the respective characteristic functions of $F^{\hat{\mathbf{R}}_N}(x)$ and $F_{c_N}(x)$.

The characteristic function of $F^{\hat{\mathbf{R}}_N}(x)$ is

$$\hat{s}_N(\ell) \triangleq \int e^{i\ell x} dF^{\hat{\mathbf{R}}_N}(x) = \frac{1}{N}\sum_{j=1}^{N} e^{i\ell\lambda_j}, \tag{2.11}$$

where $\lambda_j, j = 1, 2, \ldots, N$ are the eigenvalues of the sample correlation matrix $\hat{\mathbf{R}}_N$.

Our test statistic is then proposed as follows:

$$S_N = \int |\hat{s}_N(\ell) - s_{c_N}(\ell)|^2 dU(\ell), \tag{2.12}$$

where $s_{c_N}(\ell)$ is the characteristic function of $F_{c_N}(x)$, obtained from the M-P law $F_c(x)$ with c being replaced by $c_N = N/T$, and $U(\ell)$ is a weight function with its support on a compact interval, say $[L_1, L_2]$.

An important concept related to the spectral analysis of large dimensional random matrix theory is the Stieltjes transform. For any cumulative distribution function (CDF) G, its Stieltjes transform is defined as

$$m_G(z) = \int \frac{1}{\lambda - z} dG(\lambda), \quad \Im(z) \neq 0.$$

Linear spectral statistics and the Stieltjes transform of any CDF G have the relation

$$\int f(x) dG(x) = -\frac{1}{2\pi i}\oint_{\mathcal{C}} f(z) m_G(z) dz,$$

where f is analytic on an open set containing the support of G; $\mathcal{C}$ is a contour which is closed and is taken in the positive direction in the complex plane enclosing the support of G.

3 Asymptotic Theory

In this section, we will establish a new CLT for a general class of linear spectral statistics and then apply the CLT to the proposed test statistic S_N.

Before stating the main results, we specify some notation. Let $\underline{\hat{\mathbf{R}}}_N = \sum_{j=1}^{N} \frac{\hat{\mathbf{u}}_j \hat{\mathbf{u}}_j^\tau}{||\hat{\mathbf{u}}_j||^2}$. The Stieltjes transforms of ESD and LSD for $\underline{\hat{\mathbf{R}}}_N$ are denoted by $\underline{m}_N(z)$ and $\underline{m}_c(z)$, and the corresponding transforms for $\hat{\mathbf{R}}_N$ are denoted by $m_N(z)$ and $m_c(z)$, respectively. Moreover, $\underline{m}_{c_N}(z)$ and $m_{c_N}(z)$ are the respective $\underline{m}_c(z)$ and $m_c(z)$ with c replaced by c_N. For ease of notation, we denote $m_c(z)$ and $\underline{m}_c(z)$ by $m(z)$ and $\underline{m}(z)$, respectively.

In Theorem 1 below, and then Theorems 2–4 in Sections 3 and 4, we will establish some new asymptotic properties. Their proofs are given in Appendix A of a supplementary document. Theorem 1 provides the CLT for linear spectral statistics based on the sample correlation matrix $\hat{\mathbf{R}}_N$.

Theorem 1. *In addition to Assumptions 1 and 2, let $f_1, f_2, \ldots, f_k$ be functions on $\mathbb{R}$ analytic on an open interval containing*

$$\big[(1-\sqrt{c})^2, (1+\sqrt{c})^2\big].$$

Moreover, let $\kappa = \frac{\mathbb{E}(u_{11}^4)}{\mathbb{E}(u_{11}^2)}$. Then the random vector

$$\Big(\int f_1(x)dG_N(x), \ldots, \int f_k(x)dG_N(x)\Big)$$

converges weakly to a Gaussian vector $(U_{f_1}, \ldots, U_{f_k})$, with means

$$\begin{aligned}
E_r\left[U_{f_j}\right] &= \frac{\kappa-1}{2\pi i}\oint_{\mathcal{C}} f_j(z)\frac{c\underline{m}(z)\big(z(1+\underline{m}(z))+1-c\big)}{\big((z(1+\underline{m}(z))-c\big)^2-c\big)\big(z(1+\underline{m}(z))-c\big)}dz \\
&-\frac{\kappa-3}{2\pi i}\oint_{\mathcal{C}} f_j(z)\frac{cz\underline{m}(z)m^2(z)\big(1+\underline{m}(z)\big)\big(z(1+\underline{m}(z))+1-c\big)}{\big((z(1+\underline{m}(z))-c)^2-c\big)\big(1+cm(z)\big)}dz \\
&-\frac{1}{2\pi i}\oint_{\mathcal{C}} f_j(z)\frac{c\underline{m}^{(2)}(z)\Big(z(1+\underline{m}(z))+1-c\Big)}{\underline{m}(z)\big(z+z\underline{m}(z)-c\big)\Big(\big(z(1+\underline{m}(z))-c\big)^2-c\Big)}dz \\
&+\frac{1}{2\pi i}\oint_{\mathcal{C}} f_j(z)\frac{c\Big(1+z\underline{m}(z)-zm(z)\underline{m}(z)-z^2m(z)\underline{m}^2(z)\Big)\Big(1+\underline{m}(z)\Big)\Big(z(1+\underline{m}(z))+1-c\Big)}{z(1+cm(z))\big(z(1+\underline{m}(z))-c)^2-c\big)}dz \\
&+\frac{1}{2\pi i}\oint_{\mathcal{C}} f_j(z)\Big(\frac{cm(z)}{z}-czm(z)\underline{m}^{(2)}(z)\Big)dz \qquad (3.1)
\end{aligned}$$

and covariance function

$$\begin{aligned}
Cov(U_{f_j}, U_{f_r}) &= -\frac{1}{2\pi^2}\oint_{\mathcal{C}_1}\oint_{\mathcal{C}_2} f_j(z_1)f_r(z_2)\frac{cm^{(2)}(z_1)m^{(2)}(z_2)}{\left(1+c\left(m\left(z_1\right)+m\left(z_2\right)\right)+c\left(c-1\right)m\left(z_1\right)m\left(z_2\right)\right)^2}dz_1dz_2 \\
&+\frac{\kappa-1}{4\pi^2}\oint_{\mathcal{C}_1}\oint_{\mathcal{C}_2} f_j(z_1)f_r(z_2)\frac{c\underline{m}^{(2)}(z_1)\underline{m}^{(2)}(z_2)}{(1+\underline{m}(z_1))^2(1+\underline{m}(z_2))^2}dz_1dz_2 \\
&-\frac{\kappa-3}{4\pi^2}\oint_{\mathcal{C}_1}\oint_{\mathcal{C}_2} f_j(z_1)f_r(z_2)V(c, m(z_1), m(z_2))dz_1dz_2,
\end{aligned} \qquad (3.2)$$

where

$$m^{(2)}(z) = \int \frac{1}{(\lambda-z)^2}dF_c(\lambda), \quad \underline{m}^{(2)}(z) = \frac{1-c}{z^2} + cm^{(2)}(z), \qquad (3.3)$$

$$\begin{aligned}
V(c, m(z_1), m(z_2)) = c&\Big(m(z_1)\underline{m}(z_1) + z_1m(z_1)\underline{m}^{(2)}(z_1) + z_1m^{(2)}(z_1)\underline{m}(z_1)\Big) \\
&\times\Big(m(z_2)\underline{m}(z_2) + z_2m(z_2)\underline{m}^{(2)}(z_2) + z_2m^{(2)}(z_2)\underline{m}(z_2)\Big)
\end{aligned}$$

and $j, r = 1, 2, \ldots, k$. The contours in (3.1) and (3.2) are closed and are taken in the positive direction in the complex plane, each enclosing the support of F_c.

Based on Theorem 1, we can derive an asymptotic distribution for the proposed test statistic S_N as follows.

Theorem 2. *Under the assumptions of Theorem 1, the scaled statistic $N^2 S_N$ converges in distribution to*

$$R_0 = \int \Big(|V(\tau)|^2 + |Z(\tau)|^2\Big)^2 dU(\tau), \tag{3.4}$$

where $\big(V(\tau), Z(\tau)\big)$ is a Gaussian vector whose mean and variance are determined in (3.1) and (3.2) by taking $f_1(x)$ and $f_2(x)$ as $\sin(x)$ and $\cos(x)$, respectively.

We can evaluate the power of the statistic S_N for a class of local alternatives, although it is difficult to establish the asymptotic distribution for the test statistic under such a class of local alternative hypotheses.

Due to (2.12), the proposed statistic S_N can be written into the form as follows.

$$\begin{aligned} S_N &= \int \Big| \int \cos(\ell x) d\Big(F^{\hat{\mathbf{R}}_N}(x) - F_{c_N}(x)\Big)\Big|^2 dU(\ell) \\ &\quad + \int \Big| \int \sin(\ell x) d\Big(F^{\hat{\mathbf{R}}_N}(x) - F_{c_N}(x)\Big)\Big|^2 dU(\ell). \end{aligned}$$

Furthermore,

$$S_N = S_N^{\mathbb{H}_0} + \Delta_N, \tag{3.5}$$

where

$$\begin{aligned} S_N^{\mathbb{H}_0} &= \int \Big| \int \cos(\ell x) d\Big(F_{\mathbb{H}_0}^{\hat{\mathbf{R}}_N}(x) - F_{c_N}(x)\Big)\Big|^2 dU(\ell) \\ &\quad + \int \Big| \int \sin(\ell x) d\Big(F_{\mathbb{H}_0}^{\hat{\mathbf{R}}_N}(x) - F_{c_N}(x)\Big)\Big|^2 dU(\ell) \end{aligned}$$

and

$$\begin{aligned} \Delta_N &= \Big[\int \cos(\ell x) d\Big(F^{\hat{\mathbf{R}}_N}(x) - F_{\mathbb{H}_0}^{\hat{\mathbf{R}}_N}(x)\Big)\Big]^2 dU(\ell) + \Big[\int \sin(\ell x) d\Big(F^{\hat{\mathbf{R}}_N}(x) - F_{\mathbb{H}_0}^{\hat{\mathbf{R}}_N}(x)\Big)\Big]^2 dU(\ell) \\ &\quad +2\int \cos(\ell x) d\Big(F^{\hat{\mathbf{R}}_N}(x) - F_{\mathbb{H}_0}^{\hat{\mathbf{R}}_N}(x)\Big) dU(\ell) \int \cos(\ell x) d\Big(F_{\mathbb{H}_0}^{\hat{\mathbf{R}}_N}(x) - F_{c_N}(x)\Big) dU(\ell) \\ &\quad +2\int \sin(\ell x) d\Big(F^{\hat{\mathbf{R}}_N}(x) - F_{\mathbb{H}_0}^{\hat{\mathbf{R}}_N}(x)\Big) dU(\ell) \int \sin(\ell x) d\Big(F_{\mathbb{H}_0}^{\hat{\mathbf{R}}_N}(x) - F_{c_N}(x)\Big) dU(\ell). \end{aligned}$$

From (3.5), the power of the statistic S_N relies on the value of Δ_N.

Theorem 3. *In addition to Assumptions 1 and 2, let the following hold in probability,*

$$\limsup_{T\to\infty} N \Big| \int e^{i\ell x} d\Big(F_{\mathbb{H}_0}^{\hat{\mathbf{R}}_N}(x) - F_{\mathbb{H}_a}^{\hat{\mathbf{R}}_N}(x)\Big)\Big| = \infty, \tag{3.6}$$

where $F_{\mathbb{H}_0}^{\hat{\mathbf{R}}_N}$ stands for the ESD of $\hat{\mathbf{R}}_N$ under $\mathbb{H}_0$ and $F_{\mathbb{H}_a}^{\hat{\mathbf{R}}_N}$ is the ESD of $\hat{\mathbf{R}}_N$ under $\mathbb{H}_a$. Then

$$\lim_{T\to\infty} P(N^2 S_N > \gamma_\alpha | \mathbb{H}_a) = 1,$$

where γ_α is the critical value of $N^2 S_N$ under $\mathbb{H}_0$ corresponding to the significance level α.

Remark 1. *Note that if $F^{\hat{\mathbf{R}}_N}_{\mathbb{H}_0}$ and $F^{\hat{\mathbf{R}}_N}_{\mathbb{H}_a}$ have different limits in probability, then $\int e^{i\ell x} d\Big(F^{\hat{\mathbf{R}}_N}_{\mathbb{H}_0}(x) - F^{\hat{\mathbf{R}}_N}_{\mathbb{H}_a}(x)\Big)$ converges in probability to a nonzero constant depending on ℓ by Levy's continuity theorem. This ensures (3.6) is true. Most of the examples given in the subsequent sections satisfy (3.6).*

4 A local alternative hypothesis

It is well known that there are two commonly used cross-sectional dependent structures in panel data analysis: spatial models and factor models. In this section, we consider a simple factor model to describe cross–sectional dependence. An asymptotic theory is established as a consequence of our discussion.

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 $\mathbf{u}_1, \cdots, \mathbf{u}_N$, but the limit of the ESD associated with such vectors is the M-P law. Then a natural question is whether the statistic S_N works under this case.

We below investigate a local alternative hypothesis of the form:

$$\mathbb{H}_a: \quad u_{jt} = \varepsilon_{jt} + \frac{1}{\sqrt{T}} v_t, \; j = 1, \ldots, N; \; t = 1, \ldots, T, \tag{4.1}$$

where $\{\varepsilon_{jt}, j = 1, 2, \ldots, N; t = 1, 2, \ldots, T\}$ is a sequence of i.i.d. real random variables with $\mathbb{E}\varepsilon_{11} = 0$ and $\mathbb{E}\varepsilon_{11}^2 = 1$, and $\{v_t, t = 1, 2, \ldots, T\}$ is a sequence of i.i.d. real random variables, and is independent of $\{\varepsilon_{jt}, j = 1, 2, \ldots, N; t = 1, 2, \ldots, T\}$.

Model (4.1) can be written as the vector form

$$\mathbf{u}_j = \boldsymbol{\varepsilon}_j + \frac{1}{\sqrt{T}} \mathbf{v}, \quad j = 1, 2, \ldots, N, \tag{4.2}$$

or the matrix form

$$\mathbf{U} = \boldsymbol{\varepsilon} + \frac{1}{\sqrt{T}} \mathbf{v}\mathbf{e}^{\tau}, \tag{4.3}$$

where $\mathbf{U} = (\mathbf{u}_1, \ldots, \mathbf{u}_N)$, $\boldsymbol{\varepsilon} = (\boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_2, \ldots, \boldsymbol{\varepsilon}_N)$, $\mathbf{v} = (v_1, \ldots, v_T)^{\tau}$ and $\mathbf{e}$ is an $N \times 1$ vector with all elements being one.

Under the local alternative hypothesis (4.1), the residuals $u_{1t}, u_{2t}, \ldots, u_{Nt}$ are dependent due to the common factor $\frac{1}{\sqrt{T}} v_t$. This kind of dependence is rather weak in the sense of the covariance between u_{jt} and u_{kt} $(j \neq k)$ being $\frac{1}{T}$, which tends to 0 as T goes to infinity.

By the rank inequality (see Lemma 3.5 of Yin (1986)) and the fact that $rank(\mathbf{v}\mathbf{e}^T) \leq 1$, it can be concluded that the limit of the ESD of the matrix $\mathbf{R}_N$ is the same as that of the sample correlation matrix of $\{\varepsilon_{1t}, \varepsilon_{2t}, \ldots, \varepsilon_{Nt}\}$, i.e. the M-P law. Even so, we still would like to use the proposed statistic S_N to capture this kind of cross-sectional dependence.

Theorem 4. *Consider the local alternative hypothesis (4.1). In addition to Assumptions 1 and 2, suppose that $\{\varepsilon_{jt} : j = 1, 2, \ldots, N; t = 1, 2, \ldots, T\}$ are i.i.d with $\mathbb{E}\varepsilon_{jt} = 0$ and $\mathbb{E}\varepsilon_{jt}^4 < \infty$; $\{v_t : t =$*

$1, 2, \ldots, T\}$ *are i.i.d with* $\mathbb{E}v_t = 0$ *and* $\mathbb{E}v_t^4 < \infty$*; and moreover,* $\{\varepsilon_{jt} : j = 1, 2, \ldots, N; t = 1, 2, \ldots, T\}$ *and* $\{v_1, v_2, \ldots, v_T\}$ *are independent, and* $\mathbb{E}v_1^2 = \mathbb{E}\varepsilon_{11}^2$*. Let* $\kappa = \frac{\mathbb{E}(\varepsilon_{11}^4)}{\mathbb{E}(\varepsilon_{11}^2)}$*.*

Then, the proposed test statistic $N^2 S_N$ *converges in distribution to the random variable* R_2 *given by*

$$R_2 = \int \big(|W(\ell)|^2 + |Q(\ell)|^2\big) dU(\ell), \tag{4.4}$$

where $(W(\ell), Q(\ell))$ *is a Gaussian vector whose mean and covariance are specified below:*

$$\begin{aligned}
\mathbb{E}W(\ell) \quad = \quad & \frac{\kappa-1}{2\pi i}\oint_{\mathcal{C}} \cos(\ell z)\frac{c\underline{m}(z)\big(z(1+\underline{m}(z))+1-c\big)}{\big(\big(z(1+\underline{m}(z))-c\big)^2-c\big)\big(z(1+\underline{m}(z))-c\big)}dz \\
& -\frac{\kappa-3}{2\pi i}\oint_{\mathcal{C}} \cos(\ell z)\frac{cz\underline{m}(z)m^2(z)\big(1+\underline{m}(z)\big)\big(z(1+\underline{m}(z))+1-c\big)}{\big((z(1+\underline{m}(z))-c)^2-c\big)\big(1+cm(z)\big)}dz \\
& -\frac{1}{2\pi i}\oint_{\mathcal{C}} \cos(\ell z)\frac{c\underline{m}^{(2)}(z)\Big(z(1+\underline{m}(z))+1-c\Big)}{\underline{m}(z)\big(z+z\underline{m}(z)-c\big)\Big(\big(z(1+\underline{m}(z))-c\big)^2-c\Big)}dz \\
& +\frac{1}{2\pi i}\oint_{\mathcal{C}} \cos(\ell z)\frac{c\Big(1+z\underline{m}(z)-zm(z)\underline{m}(z)-z^2m(z)\underline{m}^2(z)\Big)\Big(1+\underline{m}(z)\Big)\Big(z(1+\underline{m}(z))+1-c\Big)}{z(1+cm(z))\big(z(1+\underline{m}(z))-c)^2-c\big)}dz \\
& +\frac{1}{2\pi i}\oint_{\mathcal{C}} \cos(\ell z)\Big(\frac{cm(z)}{z}-czm(z)\underline{m}^{(2)}(z)\Big)dz \\
& +\oint_{\mathcal{C}} \cos(\ell z)\frac{c\underline{m}^{(2)}(z)}{(1+\underline{m}(z))^2}dz \\
& -\oint_{\mathcal{C}} \cos(\ell z)\frac{c\underline{m}(z)\underline{m}^{(2)}(z)\,(\underline{m}(z)-2)}{(1+\underline{m}(z))\,(1+\underline{m}(z)-c\underline{m}^2(z))}dz \\
& -\oint_{\mathcal{C}} \cos(\ell z)\frac{c\underline{m}^{(2)}(z)\,\big[(1+2\underline{m}(z))\,\big(1+\underline{m}(z)-c\underline{m}^2(z)\big)-\underline{m}(z)\,(1+\underline{m}(z))\,(1-2c\underline{m}(z))\big]}{(1+\underline{m}(z)-c\underline{m}^2(z))\,(1+\underline{m}(z)+c\underline{m}(z))}dz
\end{aligned} \tag{4.5}$$

and covariance function

$$\begin{aligned}
Cov(W(\ell), Q(\ell)) \quad = \quad & -\frac{1}{2\pi^2}\oint_{\mathcal{C}_1}\oint_{\mathcal{C}_2} \cos(\ell z_1)\sin(\ell z_2)\frac{cm^{(2)}(z_1)m^{(2)}(z_2)}{\left(1+c\left(m\left(z_1\right)+m\left(z_2\right)\right)+c\left(c-1\right)m\left(z_1\right)m\left(z_2\right)\right)^2}dz_1dz_2 \\
& +\frac{\kappa-1}{4\pi^2}\oint_{\mathcal{C}_1}\oint_{\mathcal{C}_2} \cos(\ell z_1)\sin(\ell z_2)\frac{c\underline{m}^{(2)}(z_1)\underline{m}^{(2)}(z_2)}{(1+\underline{m}(z_1))^2(1+\underline{m}(z_2))^2}dz_1dz_2 \\
& -\frac{\kappa-3}{4\pi^2}\oint_{\mathcal{C}_1}\oint_{\mathcal{C}_2} \cos(\ell z_1)\sin(\ell z_2)V(c, m(z_1), m(z_2))dz_1dz_2,
\end{aligned} \tag{4.6}$$

where $m^{(2)}(z)$*,* $\underline{m}^{(2)}(z)$ *and* $V(c, m(z_1), m(z_2))$ *are defined in (3.3) and (3.4) respectively.*

Replacing cos(ℓz) *in* $\mathbb{E}[W(\ell)]$ *by sin*(ℓz) *yields the expression of* $\mathbb{E}[Q(\ell)]$*. The expressions of the covariances* $\mathrm{Cov}(W(\ell), W(\ell))$ *and* $\mathrm{Cov}(Q(\ell), Q(\ell))$ *are similar except replacing sin*(ℓz) *and cos*(ℓz) *by cos*(ℓz) *and sin*(ℓz)*, respectively. The contours in (4.5) and (4.6) both enclose the interval* $[(1-\sqrt{c})^2, (1+\sqrt{c})^2]$*. Moreover, the contours* γ_1 *and* γ_2 *are disjoint.*

In view of Theorem 4, we see that the proposed test statistic S_N still works mainly due to the involvement of the last term on the right–hand side of (4.5). Section 5 below employs the proposed test statistic to evaluate the finite–sample performance and the practical applicability of the proposed test.

5 Finite sample studies

We will present the empirical sizes and power values of the proposed test statistic under several scenarios.

5.1 Empirical sizes and power values

First, we introduce the method of calculating the empirical sizes and power values. Since the asymptotic distribution of the proposed modified test statistic $N^2 S_N$ is not a classical distribution, we calculate the critical value numerically. In detail, we generate K replications of Gaussian vector $\Big(W(\ell), Q(\ell)\Big)$ whose mean and covariance are given in (4.5) and (4.6) respectively. Put the K replications in an increasing order and let $\eta_{1-\alpha}$ be the $[K(1-\alpha)]$-th number. Meantime, we should generate K replications of the data set simulated under the null hypothesis and derive K values of the proposed test statistic. Then the empirical size can be calculated by

$$\hat{\alpha} = \frac{\{\sharp\ of\ N^2 S_N^{\mathbb{H}_0} \geq \eta_{1-\alpha}\}}{K}, \tag{5.1}$$

where $N^2 S_N^{\mathbb{H}_0}$ represents the value of the test statistic $N^2 S_N$ based on the data simulated under the null hypothesis.

In our simulation, we choose $K = 1000$ as the number of the replications. The significance level is $\alpha = 0.05$. Similarly, the empirical power is calculated by

$$\hat{\beta} = \frac{\{\sharp\ of\ N^2 S_N^{\mathbb{H}_a} \geq \eta_{1-\alpha}\}}{K}, \tag{5.2}$$

where $N^2 S_N^{\mathbb{H}_a}$ represents the value of the test statistic $N^2 S_N$ based on the data simulated under the alternative hypothesis.

5.2 Computational aspects

In the procedure of calculating both the empirical size and the empirical power in (5.1) and (5.2), respectively, we need to compute the asymptotic mean and variance derived in Theorem 1. Since the computation is relatively complicated, we provide a summary of the key steps to show how it is done. The code functions involved are displayed in Appendix A.

There are four key steps involved in computing the numerical values of the asymptotic mean and variance functions. They are summarised as follows.

Step 1. The LSD's $m(z)$ and $\underline{m}(z)$ are replaced by the estimators $\hat{m}(z) = \frac{1}{N} tr(\hat{\mathbf{R}}_N - z\mathbf{I}_N)^{-1}$ and $\underline{\hat{m}}(z) = \frac{1}{T} tr(\underline{\hat{\mathbf{R}}}_N - z\mathbf{I}_T)^{-1}$, respectively.

Step 2. The derivatives $m^{'}(z) = m^{(2)}(z)$ and $\underline{m}^{'}(z) = \underline{m}^{(2)}(z)$ are estimated by $\hat{m}^{(2)}(z) = \frac{1}{N}tr(\hat{\mathbf{R}}_N - z\mathbf{I}_N)^{-2}$ and $\underline{\hat{m}}^{(2)}(z) = \frac{1}{T}tr(\underline{\hat{\mathbf{R}}}_N - z\mathbf{I}_T)^{-2}$, respectively.

Step 3. For the asymptotic mean, we let $z = r \cdot e^{i\theta}$ by the polar coordinates transform and then replace the contour $\mathcal{C}$ by the circle $\{(r, \theta) : \theta \in [0, 2\pi]\}$, which involves the contour $\mathcal{C}$ inside. The integral in the asymptotic mean can be numerically computed by the MATLAB function named "*quad*". Similarly, for computing the asymptotic variance, the polar coordinates transforms $z_1 = r_1 \cdot e^{i\theta_1}$ and $z_2 = r_2 \cdot e^{i\theta_2}$ are utilized, and the two contours $\mathcal{C}_1$ and $\mathcal{C}_2$ are then replaced by circles $\mathcal{O}_1 := \{(r_1, \theta) : \theta \in [0, 2\pi]\}$ and $\mathcal{O}_2 := \{(r_2, \theta) : \theta \in [0, 2\pi]\}$, respectively, where $\mathcal{O}_1$ and $\mathcal{O}_2$ include $\mathcal{C}_1$ and $\mathcal{C}_2$, respectively. The double integral involved in the asymptotic variance can be simulated by the MATLAB function named "*dblquad*".

Step 4. The implementation of Steps 1–3 is realised in Section 5.3 by the code functions which are displayed in Appendix B.

5.3 Examples of implementation

The procedure proposed to calculate the empirical size and power values is stated as follows.

1. Data Generating Process (DGP): generate the data $Y_{jt} = \alpha_j + \mathbf{x}_{jt}^{\tau}\boldsymbol{\beta} + u_{jt}$ by following each example.

2. Calculate the statistic S_N defined in (2.12), where the $\hat{s}_N(\ell)$, $\hat{\mathbf{R}}_N$ and $\hat{\mathbf{u}}_j$ are defined in (2.11), (2.6) and (2.5) respectively.

3. Repeat K times of steps 1–2 and derive the number K statistic values $\{S_N^{(m)} : m = 1, 2, \ldots, K\}$.

4. The asymptotic mean and variance derived in Theorem 1 are calculated as follows. The LSD's $m(z)$ and $\underline{m}(z)$ are replaced by the estimators $\hat{m}(z) = \frac{1}{N}tr(\hat{\mathbf{R}}_N - z\mathbf{I}_N)^{-1}$ and $\underline{\hat{m}}(z) = \frac{1}{T}tr(\underline{\hat{\mathbf{R}}}_N - z\mathbf{I}_T)^{-1}$, respectively. The derivatives $m^{'}(z)$ and $\underline{m}^{'}(z)$ are estimated by $\hat{m}^{'}(z) = \frac{1}{N}tr(\hat{\mathbf{R}}_N - z\mathbf{I}_N)^{-2}$ and $\underline{\hat{m}}^{'}(z) = \frac{1}{T}tr(\underline{\hat{\mathbf{R}}}_N - z\mathbf{I}_T)^{-2}$, respectively.

   Since the integrand functions of the asymptotic mean and asymptotic variance are relatively complicated, we simply denote them by $g(z)$ and $h(z_1, z_2)$ respectively. Then by polar coordinates transforms, we have

$$\mathbb{E}X_f = \oint_{\mathcal{C}} g(z)dz = \int_0^{2\pi} g(r \cdot e^{i\theta}) r \cdot i \cdot e^{i\theta} d\theta$$

   and

$$Cov(X_{f_1}, X_{f_2}) = \oint_{\mathcal{C}} h(z_1, z_2)dz_1dz_2 = -\int_0^{2\pi}\int_0^{2\pi} h(r_1 \cdot e^{i\theta_1}, r_2 \cdot e^{i\theta_2}) r_1 r_2 \cdot e^{i(\theta_1+\theta_2)} d\theta_1 d\theta_2,$$

   where the circle $\{(r, \theta) : \theta \in [0, 2\pi]\}$ involves the contour $\mathcal{C}$ inside; moreover, the circles $\{(r_1, \theta) : \theta \in [0, 2\pi]\}$ and $\{(r_2, \theta) : \theta \in [0, 2\pi]\}$ include contours $\mathcal{C}_1$ and $\mathcal{C}_2$ inside respectively.

5. Generate K replications of the Gaussian vector $\Big(W(\ell), Q(\ell)\Big)$ whose mean and covariance are given in (4.5) and (4.6) respectively. Put the K replications in an increasing order and let $\eta_{1-\alpha}$ be the $[K(1-\alpha)]$-th number.

6. The empirical size or power is calculated as

$$\frac{\sum_{m=1}^{K} I_{\{S_N^{(m)} > \eta_{1-\alpha}\}}}{K}$$

5.3.1 Comparisons with the CD test

$$Y_{jt} = \alpha_j + X_{1,jt}\beta_1 + X_{2,jt}\beta_2 + u_{jt}, j = 1, 2, \ldots, N; t = 1, 2, \ldots, T, \tag{5.3}$$

where the regressors $X_{k,jt} \overset{i.i.d.}{\sim} \mathcal{N}(0, 1)$, their coefficients $\beta_1 = 0.6$, $\beta_2 = 0.8$ and the fixed effects are generated by $\alpha_j \overset{i.i.d.}{\sim} \mathcal{U}[-0.1, 0.1]$ for $j = 1, 2, \ldots, N-1$; $\alpha_N = -\sum_{j=1}^{N-1} \alpha_j$, with $\mathcal{N}$ and $\mathcal{U}$ standing for the normal distribution and the uniform distribution respectively. Under the null hypothesis, the error term $u_{jt} \overset{i.i.d.}{\sim} \mathcal{N}(0, 1)$. Note that $\{u_{jt}\}$ and $\{X_{k,jt} : k = 1, 2\}$ are generated independently.

Tables 1 and 2 show the empirical sizes of our proposed test and the CD test provided in Pesaran (2004) for (5.3) respectively. From the results, it can be seen that the proposed test statistic performs better than the CD test, in the sense of empirical sizes being close to the true size 0.05.

Table 1 and Table 2 near here

In the following sections, we consider several alternative hypotheses.

5.3.2 Spatial Models and Factor Models

In this part, we consider two types of cross-sectional dependent models: spatial models and factor models.

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

$$u_{jt} = \sum_{r=1}^{N} \omega_{jr}\varepsilon_{rt} + \varepsilon_{jt}, \quad j = 1, 2, \ldots, N; t = 1, 2, \ldots, T, \tag{5.4}$$

where $\varepsilon_{jt} \overset{i.i.d}{\sim} \mathcal{N}(0, 1), \forall j = 1, \ldots, N;\ t = 1, \ldots, T$. The coefficients $\omega_{jr} = \rho^{|j-r|}$ with $\rho = 0.2$, for any $j, r = 1, 2, \ldots, N$. In this case, the covariance between u_{jt} and u_{rt} $(j \le r)$ is

$$Cov(u_{jt}, u_{rt}) = \frac{\rho^{j+r} + \rho^{2N+2-(j+r)} - \rho^{r-j}(1+\rho^2)}{\rho^2 - 1} + \rho^{r-j}(r-j) + \omega_{jr} + \omega_{rj}. \tag{5.5}$$

Apply the proposed test statistic $N^2 S_N$ for the sample correlation matrix of $\mathbf{u} = (u_{1t}, u_{2t}, \ldots, u_{Nt})^{\tau}$. The empirical powers are illustrated in Table 3. These power values show that $N^2 S_N$ performs well numerically for capturing the cross–sectional dependence for SMA model.

Table 3 near here

Next, we consider a factor model as follows.

$$u_{jt} = \lambda_j f_t + \varepsilon_{jt}, \quad j = 1, 2, \ldots, N; \ t = 1, 2, \ldots, T, \tag{5.6}$$

where the idiosyncratic components $\varepsilon_{jt} \overset{i.i.d.}{\sim} \mathcal{N}(0, 1)$. The factors $f_t \overset{i.i.d.}{\sim} \mathcal{N}(0, 1)$. For the factor loadings, we consider two cases: (1) $\lambda_j \overset{i.i.d.}{\sim} U(0.1, 0.3)$; (2) $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \ldots, \lambda_N)^\tau \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}_1)$ with $\boldsymbol{\Sigma}_1 = (\sigma_{jr})$ and $\sigma_{jr} = \gamma^{|j-r|}$, $\gamma = 0.2$ and $\gamma = 0.5$.

The empirical powers in Table 4 and Table 5 show that, as the correlation between u_{jt} and u_{rt} (which is reflected in γ) increases, the power values also increase.

Table 4 and Table 5 near here

5.3.3 A Local Alternative Hypothesis

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

$$u_{jt} = \varepsilon_{jt} + \frac{1}{\sqrt{T}} v_t, \quad j = 1, 2, \ldots, N; \ t = 1, 2, \ldots, T, \tag{5.7}$$

where the idiosyncratic components $\varepsilon_{jt} \overset{i.i.d.}{\sim} \mathcal{N}(0, 1)$; $\mathbf{v} = (v_1, v_2, \ldots, v_T)^\tau \sim \mathcal{N}(\mathbf{0}, \Sigma_2)$, where $\Sigma_2 = (\eta_{ts})$ with $\eta_{ts} = \eta^{|t-s|}$ and $\eta = 0.2$. $\{v_t, t = 1, 2, \ldots, T\}$ are independent of $\{\varepsilon_{jt}, j = 1, 2, \ldots, N; t = 1, 2, \ldots, T\}$.

The simulation results in Table 6 show that the proposed test can capture the cross-sectional dependence in the residuals for the general panel data model (4.1).

Table 6 near here

5.3.4 Some Dependent but Uncorrelated Examples

Dependent structures of a set of random variables are often described by non-zero correlations among them. However, there are some data which are not independent but uncorrelated. We consider two examples and test their dependence by the proposed test statistic.

Nonlinear MA model

Consider nonlinear MA models of the form

$$u_{jt} = Z_{j-1,t} Z_{j-2,t} (Z_{j-2,t} + Z_{jt} + 1), \ t = 1, \ldots, T; \ j = 1, \ldots, N, \tag{5.8}$$

where $Z_{jt} \sim \mathcal{N}(0, 1)$. For any $j = 1, \ldots, N$, the correlation matrix of $\mathbf{u}_t = (u_{1t}, u_{2t}, \ldots, u_{Nt})^\tau$ is a diagonal matrix. This model is provided by Kuan and Lee (2004) which tests the martingale difference hypothesis. Our proposed cross-sectional independence test statistic can be applied to this nonlinear

MA model, and the powers in Table 7 show that this test statistic performs well numerically for this model.

From another aspect, this result also implies that the limit of the ESD 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.

Table 7 near here

Multiple ARCH(1) model

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

$$u_{jt} = Z_{jt}\sqrt{\alpha_0 + \alpha_1 u_{j-1,t}^2},\ t = 1, \dots, T;\ j = 1, \dots, N; \tag{5.9}$$

where $Z_{jt} \overset{i.i.d.}{\sim} \mathcal{N}(0,1)$ and $u_{0t} \overset{i.i.d.}{\sim} \mathcal{N}(0,1)$. $\alpha_0 = 2$ and $\alpha_1 = 0.5, 0.8$.

The power values are listed in Table 8. Although the power values are small comparably with those given for other examples, the results show that our proposed test statistic is effective for this model.

Table 8 near here

6 Conclusions

This paper has proposed a new statistic to test cross-sectional independence for a panel data model. This statistic is based on the characteristic function of the empirical spectral distribution of the sample correlation matrices. The asymptotic theory of a general class of linear spectral statistics for sample correlation matrices has been established, which is of significant interest in large dimensional random matrix theory. Our test statistic belongs to a general class of linear spectral statistics in the sense of covering some classical statistics. Furthermore, it can capture nonlinear dependence instead of just correlation. The nonlinear MA and ARCH(1) models used in the simulation part have demonstrated both the practical relevance and the applicability of the test proposed in this paper.

Acknowledgments

Guangming Pan was partially supported by the Ministry of Education, Tier 2 grant 2014-T2-2-060 and by a MOE Tier 1 Grant RG25/14 at the Nanyang Technological University, Singapore and he would also like to thank Prof. J. W. Silverstein for helpful discussions for the original version. The second and the third authors acknowledge comments by Farshid Vahid. The first three authors also thank the Australian Research Council Discovery Grants Program for its financial support under Grant Numbers: DP130104229 and DP150101012.

References

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

Bai, Z. D. and Silverstein, J. W. (2004). CLT for linear spectral statistics of large-dimensional sample covariance matrices. *Annals of Probability* **32(1A)** 553-605.

Bai, Z. D., Jiang, D. D., Yao, J. F. and Zheng, S. R. (2009). Corrections to LRT on large-dimensional covariance matrix by RMT. *Annals of Statistics* **37(6B)** 3822-3840.

Bai, Z. D. and Yao, J. F. (2008). Central limit theorems for eigenvalues in a spiked population model. *Annales de lÍnstitut Henri Poincaré* **44(3)** 447-474.

Bai, Z. D. and Yin, Y. Q. (1993). Limit of the smallest eigenvalue of a large dimensional sample covariance matrix. *Annals of Probability* **21(3)** 1275-1294.

Baltagi, B. H., Feng, Q. and Kao, C. H. (2012). A lagrange multiplier test for cross-sectional dependence in a fixed effects panel data model. *Journal of Econometrics* **170** 164-177.

Billingsley, P. (1999). Convergence of Probability Measures. *Wiley-Interscience Publication* **second edition**.

Breusch, T., and Pagan, A. (1980). The Lagrange multiplier test and its application to model specification in econometrics. *Review of Economic Studies* **47** 239-254.

Chen, J. Gao, J. and Li, D. (2012). A new diagnostic test for cross–sectional uncorrelatedness in nonparametric panel data models. *Econometric Theory* **28** 1144-1163.

Dozier, R. B. and Silverstein, J. W. (2007). On the empirical distribution of eigenvalues of large dimensional information-plus-noise type matrices. *Journal of Multivariate Analysis* **98(4)** 678-694.

Gao, J., Han, X., Pan, G. M. and Yang, Y. R. (2014). High dimensional correlation matrices: CLT and its applications. *http://arxiv.org/pdf/1411.0081.pdf* .

Guo, C. H. and Higham, N. J. (2006). A Schur-Newton method for the matrix p-th root and its inverse. *SIAM Journal on Matrix Analysis and Applications* **28(3)** 788–804.

Hsiao, C. (2003). Analysis of Panel Data. *Cambridge University Press* **2nd edition**.

Hsiao, C., Pesaran, M. H. and Pick, A. (2007). Diagnostic tests of cross section independence for nonlinear panel data models. *IZA discussion paper* **No.27**.

Horn, R. N., Johnson, C. R. (1999). Matrix Analysis. *Cambridge University Press*.

Jiang, T. F. (2004). The limiting distributions of eigenvalues of sample correlation matrices. *Sankhya* **66(1)** 35-48.

Kuan, C. M. and Lee, W. M. (2004). A new test for the martingale difference hypothesis. *Studies in Nonlinear Dynamics and Econometrics* **8(4)** 1-26.

Marcenko, V. A. and Pastur, L. A. (1967). Distribution for some sets of random matrices. *Mathematics of the USSR-Sbornik* **72(114)** 507-536.

Morrison, D. F. (2005). Multivariate Statistical Methods. *4th ed. Belmont* CA: Brooks/Cole.

Pan, G. M. and Zhou. W. (2011). Central limit theorem for Hotelling's T^2 statistics under large dimension. *Annals of Applied Probability* **21(5)** 1860-1910.

Pesaran, M. H. (2004). General diagnostic tests for cross–section dependence in panels. *Cambridge Working Paper in Economics* **No.0435**.

Pesaran, M. H., Ullah, A. and Yamagata, T. (2008). A bias-adjusted LM test of error cross-section independence. *Econometrics Journal* **11** 105-127.

Ryan, ∅., and Debbah, M. (2009). Asymptotic behavior of random vandermonde matrices with entries on the unit circle. *IEEE Transactions on Information Theory* **55(7)** 3115-3148.

Schott, J. (2005). Testing for complete independence in high dimensions. *Biometrika* **92** 951-956.

Silverstein, J. W. (1995). Strong convergence of the empirical distribution of eigenvalues of large dimensional random matrices. *Journal of Multivariate Analysis* **55** 331-339.

Su, L. and Ullah, A. (2009). Testing conditional uncorrelatednesss. *Journal of Business and Economic Statistics* **27** 18-29.

Xiao, H. and Zhou, W. (2010). On the limit of the smallest eigenvalue of some sample covariance matrix. *Journal of Theoretical Probability* **23** 1-20.

Yin, Y. Q. (1986). Limiting spectral distribution for a class of random matrices. *Journal of Multivariate Analysis* **20** 50-68.

Table 1: Sizes of the proposed test at the 5% significant level

T\ N	5	10	20	30	50	100
5	0.024	0.031	0.037	0.036	0.037	0.047
10	0.036	0.037	0.044	0.040	0.041	0.046
20	0.036	0.042	0.044	0.045	0.043	0.048
30	0.041	0.048	0.048	0.051	0.044	0.046
50	0.042	0.048	0.044	0.048	0.051	0.048
100	0.055	0.052	0.054	0.055	0.049	0.051

Table 2: Sizes of the CD test at the 5% significant level

T\ N	5	10	20	30	50	100
5	0.082	0.048	0.070	0.057	0.059	0.059
10	0.052	0.064	0.049	0.053	0.061	0.052
20	0.054	0.055	0.063	0.056	0.066	0.055
30	0.042	0.055	0.053	0.041	0.052	0.048
50	0.047	0.064	0.044	0.047	0.056	0.053
100	0.064	0.072	0.053	0.057	0.045	0.051

Table 3: Powers of the proposed test at the 5% significance level for SMA model

T\ N	5	10	20	30	50	100
5	0.130	0.138	0.264	0.253	0.494	0.792
10	0.167	0.289	0.482	0.649	0.669	0.844
20	0.230	0.434	0.646	0.652	0.748	0.855
30	0.324	0.531	0.661	0.760	0.791	0.869
50	0.439	0.604	0.699	0.803	0.899	0.924
100	0.937	0.951	0.983	0.992	1.000	1.000

Table 4: Powers of the proposed test at the 5% significance level for factor models (Case 1)

T\ N	5	10	20	30	50	100
5	0.122	0.149	0.217	0.298	0.521	0.742
10	0.138	0.213	0.308	0.389	0.572	0.809
20	0.179	0.237	0.378	0.452	0.599	0.853
30	0.236	0.362	0.419	0.503	0.604	0.886
50	0.621	0.637	0.701	0.722	0.782	0.898
100	0.833	0.849	0.903	0.958	0.993	1.000

7 Appendix A: Proofs of Main theorems

In this section, the proofs of Theorems 1–4 are provided. Before providing them, some useful lemmas are listed.

7.1 Some useful lemmas

Lemma 1 (Theorem 8.1 of Billingsley (1999))**.** *Let P_n and P be probability measures on a measurable space (C, φ), where C is a space and φ is a σ-algebra. 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 said to be 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 $E\{|X_n(t_2) - X_n(t_1)|^\gamma\} \leq |F(t_2) - F(t_1)|^\alpha$ holds for all t_1, t_2, and n.

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

$$X_n \xrightarrow{d} X \Rightarrow g(X_n) \xrightarrow{d} g(X). \tag{A.1}$$

Table 5: Powers of the proposed test at the 5% significance level for factor models (Case 2)

$\gamma = 0.2$						
T\ N	5	10	20	30	50	100
5	0.094	0.148	0.255	0.302	0.604	0.823
10	0.104	0.185	0.504	0.672	0.859	0.963
20	0.240	0.295	0.648	0.704	0.903	0.982
30	0.294	0.402	0.674	0.718	0.926	0.993
50	0.598	0.641	0.743	0.899	0.983	1.000
100	0.739	0.846	0.936	0.965	0.998	1.000
$\gamma = 0.9$						
T\ N	5	10	20	30	50	100
5	0.179	0.282	0.369	0.461	0.726	0.911
10	0.291	0.418	0.589	0.722	0.901	0.948
20	0.437	0.588	0.702	0.819	0.944	0.971
30	0.604	0.696	0.802	0.901	0.989	1.000
50	0.781	0.823	0.913	0.981	1.000	1.000
100	0.928	0.971	0.996	1.000	1.000	1.000

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

$$Re(\frac{f(a) - f(b)}{a - b}) = Re(f^{'}(u)),\ Im(\frac{f(a) - f(b)}{a - b}) = Im(f^{'}(v)), \tag{A.2}$$

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

Lemma 5 (Lemma 2 of Bai and Yin (1993)). *Let $\{W_{jk} : j, k = 1, 2, \ldots, N\}$ be a double array of i.i.d random variables and let $\alpha > 1/2$, $\beta \geq 0$ and $M > 0$ be constants. Then as $N \to \infty$,*

$$\max_{j \leq MN^{\beta}} \left| N^{-\alpha} \sum_{j=1}^{N} (W_{jk} - C) \right| \to 0, \quad a.s.,$$

if and only if $\mathbb{E}|W_{11}|^{(1+\beta)/\alpha} < \infty$, where

$$C = \begin{cases} \mathbb{E}W_{11}, & if \quad \alpha \leq 1, \\ any \quad number, & if \quad \alpha > 1. \end{cases}$$

Lemma 6 (Lemma 2.2 of Bai and Silverstein (2004)). *For $\mathbf{w} = (W_1, W_2, \ldots, W_N)^{\tau}$ i.i.d standardized (complex) entries, $\mathbf{C}$ $N \times N$ matrix (complex), we have, for any $p \geq 2$,*

$$\mathbb{E}|\mathbf{w}^*\mathbf{C}\mathbf{w} - tr\mathbf{C}|^p \leq M\Big[(\mathbb{E}|W_1|^4 \cdot tr\mathbf{C}\mathbf{C}^*)^{p/2} + \mathbb{E}|W_1|^{2p} tr(\mathbf{C}\mathbf{C}^*)^{p/2}\Big].$$

Table 6: Powers of the proposed test at the 5% significance level for the local alternative model

T\ N	5	10	20	30	50	100
5	0.143	0.140	0.189	0.247	0.398	0.529
10	0.151	0.201	0.220	0.301	0.387	0.799
20	0.276	0.371	0.403	0.438	0.504	0.801
30	0.466	0.599	0.671	0.711	0.782	0.904
50	0.573	0.698	0.892	0.930	0.978	0.977
100	0.600	0.802	0.933	0.975	0.983	0.995

Table 7: Powers of the proposed test at the 5% significance level for nonlinear MA model

T\ N	5	10	20	30	50	100
5	0.117	0.182	0.207	0.294	0.418	0.722
10	0.173	0.247	0.303	0.503	0.619	0.748
20	0.203	0.302	0.400	0.582	0.629	0.803
30	0.298	0.398	0.504	0.693	0.727	0.889
50	0.515	0.604	0.721	0.838	0.901	0.927
100	0.739	0.811	0.894	0.952	0.971	0.982

Lemma 7 (Corollary 7.3.8 of Horn and Johnson (1999))**.** *Let* $\mathbf{A} : m \times n$ *and* $\mathbf{B} : m \times n$ *be two deterministic matrices,* $\mathbf{E} = \mathbf{B} - \mathbf{A}$ *and* $q = \min\{m, n\}$*. If* $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_q$ *are the singular values of* $\mathbf{A}$ *and* $\tau_1 \geq \tau_2 \geq \cdots \geq \tau_q$ *are singular values of* $\mathbf{B}$*, then*

$$|\sigma_i - \tau_i| \leq ||\mathbf{E}||, \quad for \quad all \quad i = 1, 2, \ldots, q; \quad and \quad \Big[\sum_{i=1}^{q} (\sigma_i - \tau_i)^2\Big]^{1/2} \leq ||\mathbf{E}||.$$

Lemma 8. *Suppose a sequence of measures* $\mathbb{P}_n$ *on* $C[0, T]$ *satisfies the following conditions:*

1. *There exists* $a \geq 0$ *such that* $\lim_{n\to\infty} \mathbb{P}_n(|x(0)| \geq a) = 0$*;*
2. *For each* $\varepsilon > 0$, $\lim_{\delta\to 0} \limsup_{n\to\infty} \mathbb{P}_n(\{x : \omega_x(\delta) > \varepsilon\}) = 0$,

where $\omega_x(\delta) = \sup_{t\in[0,T]} |x(t + \delta) - x(t)|$*. Then the sequence* $\mathbb{P}_n$ *is tight.*

Lemma 9 (Lemma 2.3 of Bai and Silverstein (2004))**.** *Let* $f_1, f_2, \ldots$ *be analytic in* $\mathcal{D}$*, a connected open set in* $\mathbb{C}$*, satisfying* $|f_n(z)| \leq M$ *for every* n *and* $z \in \mathbb{D}$*, and* $f_n(z)$ *converges as* $n \to \infty$ *for each* z *that is in a subset of* $\mathcal{D}$ *having a limit point in* $\mathbb{D}$*. Then there exists a function* f*, analytic in* $\mathcal{D}$ *for which* $f_n(z) \to f(z)$ *and* $f_n'(z) \to f'(z)$ *for all* $z \in \mathcal{D}$*. Moreover, on any set bounded by a contour interior to* $\mathcal{D}$*, the convergence is uniform and* $\{f_n'(z)\}$ *is uniformly bounded by* $2M/\varepsilon$*, where* ε *is the distance between the contour and the boundary of* $\mathcal{D}$*.*

Table 8: Powers of the proposed test at the 5% significance level for multiple ARCH(1) model

$\alpha_1 = 0.5$						
T\ N	5	10	20	30	50	100
5	0.072	0.107	0.148	0.204	0.401	0.720
10	0.094	0.123	0.189	0.272	0.550	0.762
20	0.128	0.204	0.341	0.301	0.601	0.802
30	0.217	0.381	0.396	0.418	0.726	0.894
50	0.399	0.505	0.603	0.728	0.812	0.928
100	0.696	0.812	0.873	0.903	0.931	0.969
$\alpha_1 = 0.8$						
T\ N	5	10	20	30	50	100
5	0.124	0.183	0.238	0.370	0.634	0.829
10	0.184	0.195	0.281	0.471	0.752	0.896
20	0.218	0.299	0.447	0.693	0.802	0.938
30	0.329	0.432	0.575	0.738	0.926	0.995
50	0.698	0.742	0.845	0.903	0.942	0.992
100	0.828	0.878	0.939	0.941	0.998	0.998

7.2 Proof of Theorem 1

In order to simplify notation we use M to denote constants which may change from line to line. Recall from (2.4) in the main paper that the centralization of the original model is

$$\widetilde{\mathbf{Y}}_j = \widetilde{\mathbf{X}}_j\boldsymbol{\beta} + \widetilde{\mathbf{u}}_j, \quad j = 1, 2, \ldots, N.$$

Under the null hypothesis $\mathbb{H}_0$, it is well known that the convergence rate of the least–square estimator $\hat{\boldsymbol{\beta}}$ for the parameter $\boldsymbol{\beta}$ (see Hsiao (2003)) is

$$||\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}|| = O_P\Big(\frac{1}{\sqrt{NT}}\Big). \tag{A.3}$$

With the estimator $\hat{\boldsymbol{\beta}}$, we can decompose $\hat{\mathbf{u}}_j$ for the error component $\mathbf{u}_j$, i.e.

$$\hat{\mathbf{u}}_j = \widetilde{\mathbf{Y}}_j - \widetilde{\mathbf{X}}_j\hat{\boldsymbol{\beta}} = \widetilde{\mathbf{u}}_j + \widetilde{\mathbf{X}}_j(\boldsymbol{\beta} - \hat{\boldsymbol{\beta}}). \tag{A.4}$$

Define the matrix $\hat{\underline{\mathbf{R}}}_N$ by

$$\hat{\underline{\mathbf{R}}}_N = \sum_{j=1}^{N} \frac{\hat{\mathbf{u}}_j\hat{\mathbf{u}}_j^{\tau}}{||\hat{\mathbf{u}}_j||^2}. \tag{A.5}$$

Then the matrix $\hat{\underline{\mathbf{R}}}_N$ has the same non-zero eigenvalues as those of the sample correlation matrix $\hat{\mathbf{R}}_N$ other than $|N-T|$ zero eigenvalues. The main part of the proposed statistic is

$$\begin{aligned} N\Big(s_N(\ell)-s(\ell)\Big) &= N\int e^{i\ell\lambda}d\Big(F^{\hat{\mathbf{R}}_N}(\lambda)-F_{c_N}(\lambda)\Big) \\ &= -\frac{1}{2\pi i}\oint_{\mathcal{C}} e^{i\ell z}\Big(tr\big(\hat{\mathbf{R}}_N-z\mathbf{I}_N\big)^{-1}-Nm_{c_N}(z)\Big)dz \\ &= -\frac{1}{2\pi i}\oint_{\mathcal{C}} e^{i\ell z}\Big(tr\big(\hat{\underline{\mathbf{R}}}_N-z\mathbf{I}_T\big)^{-1}-T\underline{m}_{c_N}(z)\Big)dz, \end{aligned} \tag{A.6}$$

where the contour $\mathcal{C}$ is closed and is taken in the positive direction in the complex plane, enclosing the support of $F_c(\cdot)$.

First, we prove that

$$tr\Big(\hat{\underline{\mathbf{R}}}_N-z\mathbf{I}_T\Big)^{-1}=tr\Big(\widetilde{\underline{\mathbf{R}}}_N-z\mathbf{I}_T\Big)^{-1}+o_P(1), \tag{A.7}$$

where $\widetilde{\underline{\mathbf{R}}}_N=\sum_{j=1}^N\frac{\hat{\mathbf{u}}_j\hat{\mathbf{u}}_j^\tau}{||\widetilde{\mathbf{u}}_j||^2}$ and $o_P(1)$ holds uniformly for $z\in\mathcal{C}$.

To this end, we claim that

$$\sqrt{\hat{\lambda}_{\max}}-\sqrt{\lambda_{\max}}\to 0,\quad a.s.\quad and\quad \sqrt{\hat{\lambda}_{\min}}-\sqrt{\lambda_{\min}}\to 0,\quad a.s. \tag{A.8}$$

$$\sqrt{\widetilde{\lambda}_{\max}}-\sqrt{\lambda_{\max}}\to 0,\quad a.s.\quad and\quad \sqrt{\widetilde{\lambda}_{\min}}-\sqrt{\lambda_{\min}}\to 0,\quad a.s., \tag{A.9}$$

where $\hat{\lambda}_{\max},\hat{\lambda}_{\min},\widetilde{\lambda}_{\max},\widetilde{\lambda}_{\min},\lambda_{\max},\lambda_{\min}$, denote the largest and smallest non-zero eigenvalues of $\hat{\underline{\mathbf{R}}}_N,\widetilde{\underline{\mathbf{R}}}_N,\underline{\mathbf{R}}_N$, respectively, with $\underline{\mathbf{R}}_N=\sum_{j=1}^N\frac{\widetilde{\mathbf{u}}_j\widetilde{\mathbf{u}}_j^\tau}{||\widetilde{\mathbf{u}}_j||^2}$.

In fact, from Lemma 5 and (A.3) we conclude that

$$\begin{aligned} \max_{1\le j\le N}\Big|\frac{\widetilde{\mathbf{u}}_j^\tau\widetilde{\mathbf{X}}_j(\boldsymbol{\beta}-\hat{\boldsymbol{\beta}})}{T}\Big| &= \max_{1\le j\le N}\Big|\frac{\sum_{k=1}^T\widetilde{\mathbf{u}}_j^\tau\mathbf{e}_k\widetilde{\mathbf{x}}_{jk}^\tau(\boldsymbol{\beta}-\hat{\boldsymbol{\beta}})}{T}\Big| \\ &= \max_{1\le j\le N}\Big|\frac{\sum_{k=1}^T\widetilde{\mathbf{u}}_j^\tau\mathbf{e}_k\sum_{r=1}^p\widetilde{X}_{jk,r}(\beta_r-\hat{\beta}_r)}{T}\Big| \\ &\le \sum_{r=1}^p\max_{1\le j\le N}\Big|\frac{1}{T}\sum_{k=1}^T\widetilde{u}_{jk}\widetilde{X}_{jk,r}\Big|\cdot|\beta_r-\hat{\beta}_r|=o_P\Big(\frac{1}{T\sqrt{N}}\Big), \end{aligned} \tag{A.10}$$

$$\max_{1\le j\le N}\Big|\frac{(\boldsymbol{\beta}-\hat{\boldsymbol{\beta}})^\tau\widetilde{\mathbf{X}}_j^\tau\widetilde{\mathbf{X}}_j(\boldsymbol{\beta}-\hat{\boldsymbol{\beta}})}{T}\Big|\le||\boldsymbol{\beta}-\hat{\boldsymbol{\beta}}||^2\cdot\max_{1\le j\le N}\Big|\Big|\frac{\widetilde{\mathbf{X}}_j^\tau\widetilde{\mathbf{X}}_j}{T}\Big|\Big|=O_P\Big(\frac{1}{TN}\Big), \tag{A.11}$$

and

$$\max_{1\le j\le N}\frac{T}{||\widetilde{\mathbf{u}}_j||^2}=\frac{1}{\min_{1\le j\le N}\Big(\frac{||\widetilde{\mathbf{u}}_j||^2}{T}-E\frac{||\widetilde{\mathbf{u}}_1||^2}{T}\Big)+E\frac{||\widetilde{\mathbf{u}}_1||^2}{T}}=O_P(1), \tag{A.12}$$

where $\widetilde{\mathbf{x}}_{jk}=(\widetilde{X}_{jk,1},\widetilde{X}_{jk,2},\ldots,\widetilde{X}_{jk,p})^\tau$, $\boldsymbol{\beta}=(\beta_1,\cdots,\beta_p)^\tau$, and $\mathbf{e}_k$ is a $T\times 1$ vector with its k-th element being one and others zero. This, together with Lemma 7 and (A.4), yields

$$\begin{aligned}
&\Big|\sqrt{\hat{\lambda}_{\max}}-\sqrt{\lambda_{\max}}\Big| \le \Big|\Big|\underline{\widetilde{\mathbf{R}}}_N-\underline{\mathbf{R}}_N\Big|\Big| = \Big|\Big|\sum_{j=1}^{N}\frac{1}{||\widetilde{\mathbf{u}}_j||^2}\Big(\hat{\mathbf{u}}_j\hat{\mathbf{u}}_j^{\tau}-\widetilde{\mathbf{u}}_j\widetilde{\mathbf{u}}_j^{\tau}\Big)\Big|\Big| \\
\le\ & \Big|\Big|\sum_{j=1}^{N}\frac{1}{||\widetilde{\mathbf{u}}_j||^2}\Big(\hat{\mathbf{u}}_j-\widetilde{\mathbf{u}}_j\Big)\hat{\mathbf{u}}_j^{\tau}\Big|\Big| + \Big|\Big|\sum_{j=1}^{N}\frac{1}{||\widetilde{\mathbf{u}}_j||^2}\widetilde{\mathbf{u}}_j\Big(\hat{\mathbf{u}}_j^{\tau}-\widetilde{\mathbf{u}}_j^{\tau}\Big)\Big|\Big| \\
\le\ & N\cdot\max_{1\le j\le N}\Big|\frac{\hat{\mathbf{u}}_j^{\tau}(\hat{\mathbf{u}}_j-\widetilde{\mathbf{u}}_j)}{||\widetilde{\mathbf{u}}_j||^2}\Big| + N\cdot\max_{1\le j\le N}\Big|\frac{\widetilde{\mathbf{u}}_j^{\tau}(\hat{\mathbf{u}}_j-\widetilde{\mathbf{u}}_j)}{||\widetilde{\mathbf{u}}_j||^2}\Big| \\
\le\ & N\cdot\max_{1\le j\le N}\Big|\frac{\big(\widetilde{\mathbf{u}}_j+\widetilde{\mathbf{X}}_j(\boldsymbol{\beta}-\hat{\boldsymbol{\beta}})\big)^{\tau}\widetilde{\mathbf{X}}_j(\boldsymbol{\beta}-\hat{\boldsymbol{\beta}})}{T}\Big|\cdot\max_{1\le j\le N}\Big|\frac{T}{||\widetilde{\mathbf{u}}_j||^2}\Big| \\
&+N\cdot\max_{1\le j\le N}\Big|\frac{\widetilde{\mathbf{u}}_j^{\tau}\widetilde{\mathbf{X}}_j(\boldsymbol{\beta}-\hat{\boldsymbol{\beta}})}{T}\Big|\cdot\max_{1\le j\le N}\Big|\frac{T}{||\widetilde{\mathbf{u}}_j||^2}\Big| = O_P\Big(\frac{1}{\sqrt{T}}\Big). \qquad (A.13)
\end{aligned}$$

Thus the first part of (A.9) is proved. By Lemma 7 the second part of (A.9) can be similarly derived.

As in (A.13) it is easy to obtain

$$\begin{aligned}
&\Big|\sqrt{\hat{\lambda}_{\max}}-\sqrt{\widetilde{\lambda}_{\max}}\Big| \le \Big|\Big|\underline{\hat{\mathbf{R}}}_N-\underline{\widetilde{\mathbf{R}}}_N\Big|\Big| = \Big|\Big|\sum_{j=1}^{N}\hat{\mathbf{u}}_j\hat{\mathbf{u}}_j^{\tau}\Big(\frac{1}{||\hat{\mathbf{u}}_j||^2}-\frac{1}{||\widetilde{\mathbf{u}}_j||^2}\Big)\Big|\Big| \\
\le\ & N\cdot\max_{1\le j\le N}\Big|\frac{2(\boldsymbol{\beta}-\hat{\boldsymbol{\beta}})^{\tau}\widetilde{\mathbf{X}}_j^{\tau}\widetilde{\mathbf{u}}_j}{T}\Big|\cdot\max_{1\le j\le N}\Big|\frac{T}{||\widetilde{\mathbf{u}}_j||^2}\Big| \\
&+N\cdot\max_{1\le j\le N}\Big|\frac{(\boldsymbol{\beta}-\hat{\boldsymbol{\beta}})^{\tau}\widetilde{\mathbf{X}}_j^{\tau}\widetilde{\mathbf{X}}_j(\boldsymbol{\beta}-\hat{\boldsymbol{\beta}})}{T}\Big|\cdot\max_{1\le j\le N}\frac{T}{||\widetilde{\mathbf{u}}_j||^2} = o_P\Big(\frac{\sqrt{N}}{T}\Big) = o_P(1). \qquad (A.14)
\end{aligned}$$

Then (A.8) can be obtained from (A.13) and (A.14).

Now we introduce some formulas that will be frequently used in the proof. For any invertible matrices $\mathbf{A}$ and $\mathbf{B}$, vectors $\mathbf{r}$, $\mathbf{w}$ and a scalar q,

$$\mathbf{A}^{-1}-\mathbf{B}^{-1} = \mathbf{A}^{-1}(\mathbf{B}-\mathbf{A})\mathbf{B}^{-1}, \qquad (A.15)$$

$$\Big(\mathbf{A}+q\mathbf{r}\mathbf{w}^{\tau}\Big)^{-1} = \mathbf{A}^{-1}-\frac{q\mathbf{A}^{-1}\mathbf{r}\mathbf{w}^{\tau}\mathbf{A}^{-1}}{1+q\mathbf{w}^{\tau}\mathbf{A}^{-1}\mathbf{r}} \qquad (A.16)$$

and

$$(\mathbf{A}+q\mathbf{r}\mathbf{w}^{\tau})^{-1}\mathbf{r} = \frac{\mathbf{A}^{-1}\mathbf{r}}{1+q\mathbf{w}^{\tau}\mathbf{A}\mathbf{r}}. \qquad (A.17)$$

For $j=1,2,\cdots,N$, let $W_j(z)=\frac{\hat{\mathbf{u}}_j^{\tau}\big(\underline{\hat{\mathbf{R}}}_N-z\mathbf{I}_T\big)^{-1}\big(\underline{\widetilde{\mathbf{R}}}_N-z\mathbf{I}_T\big)^{-1}\hat{\mathbf{u}}_j}{||\hat{\mathbf{u}}_j||^2}$. From (A.15) we have

$$\begin{aligned}
&\Big|tr\Big(\underline{\hat{\mathbf{R}}}_N-z\mathbf{I}_T\Big)^{-1}-tr\Big(\underline{\widetilde{\mathbf{R}}}_N-z\mathbf{I}_N\Big)^{-1}\Big| \\
=\ & \Big|tr\Big(\underline{\widetilde{\mathbf{R}}}_N-z\mathbf{I}_N\Big)^{-1}\Big(\underline{\widetilde{\mathbf{R}}}_N-\underline{\hat{\mathbf{R}}}_N\Big)\Big(\underline{\hat{\mathbf{R}}}_N-z\mathbf{I}_T\Big)^{-1}\Big| \\
=\ & \Big|tr\Big[\Big(\underline{\widetilde{\mathbf{R}}}_N-z\mathbf{I}_T\Big)^{-1}\sum_{j=1}^{N}\hat{\mathbf{u}}_j\hat{\mathbf{u}}_j^{\tau}\Big(\frac{1}{||\widetilde{\mathbf{u}}_j||^2}-\frac{1}{||\hat{\mathbf{u}}_j||^2}\Big)\Big(\underline{\hat{\mathbf{R}}}_N-z\mathbf{I}_T\Big)^{-1}\Big]\Big| \\
=\ & \Big|\sum_{j=1}^{N}W_j(z)\frac{\widetilde{\mathbf{u}}_j^{\tau}\widetilde{\mathbf{X}}_j(\boldsymbol{\beta}-\hat{\boldsymbol{\beta}})+(\boldsymbol{\beta}-\hat{\boldsymbol{\beta}})^{\tau}\widetilde{\mathbf{X}}_j^{\tau}\widetilde{\mathbf{u}}_j+(\boldsymbol{\beta}-\hat{\boldsymbol{\beta}})^{\tau}\widetilde{\mathbf{X}}_j^{\tau}\widetilde{\mathbf{X}}_j(\boldsymbol{\beta}-\hat{\boldsymbol{\beta}})}{||\widetilde{\mathbf{u}}_j||^2}\Big| \\
\le\ & N\cdot\max_{1\le j\le N}\Big|W_j(z)\cdot\frac{\widetilde{\mathbf{u}}_j^{\tau}\widetilde{\mathbf{X}}_j(\boldsymbol{\beta}-\hat{\boldsymbol{\beta}})}{||\widetilde{\mathbf{u}}_j||^2}\Big| + N\cdot\max_{1\le j\le N}\Big|W_j(z)\cdot\frac{(\boldsymbol{\beta}-\hat{\boldsymbol{\beta}})^{\tau}\widetilde{\mathbf{X}}_j^{\tau}\widetilde{\mathbf{u}}_j}{||\widetilde{\mathbf{u}}_j||^2}\Big| \\
&+N\cdot\max_{1\le j\le N}\Big|W_j(z)\cdot\frac{(\boldsymbol{\beta}-\hat{\boldsymbol{\beta}})^{\tau}\widetilde{\mathbf{X}}_j^{\tau}\widetilde{\mathbf{X}}_j(\boldsymbol{\beta}-\hat{\boldsymbol{\beta}})}{||\widetilde{\mathbf{u}}_j||^2}\Big|, \qquad (A.18)
\end{aligned}$$

where the last equality uses (A.4).

We below consider the term $W_j(z)$. To this end, introduce a truncation version $\widehat{\hat{\underline{\mathbf{R}}}_N^{-1}}(z)$ for $\hat{\underline{\mathbf{R}}}_N^{-1}(z) = \Big(\hat{\underline{\mathbf{R}}}_N - z\mathbf{I}_T\Big)^{-1}$. Let $v_0 > 0$, w_r be any number greater than $(1+\sqrt{c})^2$ and w_ℓ be any negative number if $c \geq 1$. Otherwise we choose $w_\ell \in \big(0,(1-\sqrt{c})^2\big)$. Define $\mathcal{C}_r = \{w_r + v : v \in [T^{-1}\rho_T, v_0]\}$, where

$$\mathcal{C}_\ell = \begin{cases} \{w_\ell + iv : v \in [T^{-1}\rho_T, v_0]\}, & if \quad w_\ell > 0; \\ \{w_\ell + iv : v \in [0, v_0]\}, & if \quad w_\ell < 0, \end{cases}$$

in which ρ_T decreases to 0 as $T \to \infty$ and $\rho_T \geq T^{-\alpha}$ for some $\alpha \in (0,1)$. Let $\mathcal{C}_T^+ = \mathcal{C}_\ell \cup \mathcal{C}_w \cup \mathcal{C}_r$, where $\mathcal{C}_w = \{w + iv_0 : w \in [w_\ell, w_r]\}$ and $\mathcal{C}_T^-$ denotes the symmetric part of $\mathcal{C}_T^+$ with respect to the real axis. We now define the truncated process $\widehat{\hat{\underline{\mathbf{R}}}_N^{-1}}(z)$ of the process $\hat{\underline{\mathbf{R}}}_N^{-1}(z)$ by

$$\widehat{\hat{\underline{\mathbf{R}}}_N^{-1}}(z) = \begin{cases} \hat{\underline{\mathbf{R}}}_N^{-1}(z), & for \quad z \in \mathcal{C}_T = \mathcal{C}_T^+ \cup \mathcal{C}_T^-; \\ \frac{Tv+\rho_T}{2\rho_T}\hat{\underline{\mathbf{R}}}_N^{-1}(z_{r_1}) + \frac{\rho_T - Tv}{2\rho_T}\hat{\underline{\mathbf{R}}}_N^{-1}(z_{r_2}), & for \quad w = w_r, v \in [-T^{-1}\rho_T, T^{-1}\rho_T]; \\ \frac{Tv+\rho_T}{2\rho_T}\hat{\underline{\mathbf{R}}}_N^{-1}(z_{\ell_1}) + \frac{\rho_T - Tv}{2\rho_T}\hat{\underline{\mathbf{R}}}_N^{-1}(z_{\ell_2}), & for \quad w = w_\ell > 0, v \in [-T^{-1}\rho_T, T^{-1}\rho_T], \end{cases}$$

where $z_{r_1} = w_r - iT^{-1}\rho_T$, $z_{r_2} = w_r + iT^{-1}\rho_T$, $z_{\ell_1} = w_\ell - iT^{-1}\rho_T$ and $z_{\ell_2} = w_\ell + iT^{-1}\rho_T$. Similarly, we can define the truncation version $\widehat{\tilde{\underline{\mathbf{R}}}_N^{-1}}(z)$ of $\tilde{\underline{\mathbf{R}}}_N^{-1}(z) = \Big(\tilde{\underline{\mathbf{R}}}_N - z\mathbf{I}_T\Big)^{-1}$. Define $\widehat{W_j(z)}$, the analogue of $W_j(z)$, with replacing $\hat{\underline{\mathbf{R}}}_N^{-1}(z)$ and $\tilde{\underline{\mathbf{R}}}_N^{-1}(z)$ by $\widehat{\hat{\underline{\mathbf{R}}}_N^{-1}}(z)$ and $\widehat{\tilde{\underline{\mathbf{R}}}_N^{-1}}(z)$ respectively.

From the above definition we conclude that

$$\Big|\Big|\widehat{\hat{\underline{\mathbf{R}}}_N^{-1}}(z)\Big| \leq \frac{M}{v_0} + \frac{M}{|\lambda_{\max}(\hat{\underline{\mathbf{R}}}_N) - w_r|} + \frac{M}{|\lambda_{\min}(\hat{\underline{\mathbf{R}}}_N) - w_\ell|}, \tag{A.19}$$

$$\Big|\Big|\widehat{\tilde{\underline{\mathbf{R}}}_N^{-1}}(z)\Big| \leq \frac{M}{v_0} + \frac{M}{|\lambda_{\max}(\tilde{\underline{\mathbf{R}}}_N) - w_r|} + \frac{M}{|\lambda_{\min}(\tilde{\underline{\mathbf{R}}}_N) - w_\ell|}, \tag{A.20}$$

and

$$\Big|\widehat{W_j(z)}\Big| < \frac{M}{v_0} + \frac{M}{|\lambda_{\max}(\hat{\underline{\mathbf{R}}}_N) - w_r|} + \frac{M}{|\lambda_{\min}(\hat{\underline{\mathbf{R}}}_N) - w_\ell|} + \frac{M}{|\lambda_{\max}(\hat{\underline{\mathbf{R}}}_N) - w_r|} + \frac{M}{|\lambda_{\min}(\hat{\underline{\mathbf{R}}}_N) - w_\ell|}. \tag{A.21}$$

It follows that for any analytic function $g(\cdot)$ on any region involving the contour $\mathcal{C}$

$$\begin{aligned} \Big|\int g(z)\Big(\widehat{W_j(z)} - W_j(z)\Big)dz\Big| \quad &\leq \Big|\int g(z)\frac{\hat{\mathbf{u}}_j^\tau\Big(\widehat{\hat{\underline{\mathbf{R}}}_N^{-1}}(z) - \hat{\underline{\mathbf{R}}}_N^{-1}(z)\Big)\widehat{\tilde{\underline{\mathbf{R}}}_N^{-1}}(z)\hat{\mathbf{u}}_j}{||\hat{\mathbf{u}}_j||^2}\Big| \\ &+\Big|\int g(z)\frac{\hat{\mathbf{u}}_j^\tau\hat{\underline{\mathbf{R}}}_N^{-1}(z)\Big(\widehat{\tilde{\underline{\mathbf{R}}}_N^{-1}}(z) - \tilde{\underline{\mathbf{R}}}_N^{-1}(z)\Big)\hat{\mathbf{u}}_j}{||\hat{\mathbf{u}}_j||^2}\Big| \\ \leq M\Big(\frac{1}{v_0} + \frac{1}{|\lambda_{\max}(\underline{\mathbf{R}}_N) - w_r|} + \frac{1}{|\lambda_{\min}(\underline{\mathbf{R}}_N) - w_\ell|}\Big) &+ \frac{M}{|\lambda_{\max}(\hat{\underline{\mathbf{R}}}_N) - w_r|} + \frac{M}{|\lambda_{\min}(\hat{\underline{\mathbf{R}}}_N) - w_\ell|}, \end{aligned} \tag{A.22}$$

which converges to zero in probability by (A.8), (A.9), Theorem 1 of Jiang (2004) and Theorem 1 of Xiao and Zhou (2010).

In view of the equivalence of $\widehat{W_j(z)}$ and $W_j(z)$ in the sense of (A.22) it is enough to consider the term $\widehat{W_j(z)}$ instead of $W_j(z)$. However for notational simplicity we still use $W_j(z)$, $\hat{\underline{\mathbf{R}}}_N^{-1}(z)$ and $\tilde{\underline{\mathbf{R}}}_N^{-1}(z)$ instead of their truncation versions. It follows from (A.10), (A.11), (A.12), (A.21) and (A.18) that

$$\Big|tr\Big(\hat{\underline{\mathbf{R}}}_N - z\mathbf{I}_T\Big)^{-1} - tr\Big(\tilde{\underline{\mathbf{R}}}_N - z\mathbf{I}_N\Big)^{-1}\Big| = O_P\Big(\frac{\sqrt{N}}{T}\Big) = o_P(1), \tag{A.23}$$

uniformly for $z \in \mathcal{C} = \mathcal{C}_T^- \cup \mathcal{C}_T^+$.

For simplicity, we only consider $z = w + iv_0 \in \mathcal{C}_w$ below and the remaining cases can be analyzed similarly.

Consider the Stieltjes transform $tr\big(\underline{\widetilde{\mathbf{R}}}_N - z\mathbf{I}_T\big)^{-1}$ below. From (A.4) we may write

$$tr\big(\underline{\widetilde{\mathbf{R}}}_N - z\mathbf{I}_T\big)^{-1} = tr\Big(\sum_{j=1}^N \mathbf{a}_j\mathbf{a}_j^\tau + \sum_{j=1}^N \mathbf{a}_j\mathbf{b}_j^\tau + \sum_{j=1}^N \mathbf{b}_j\mathbf{a}_j^\tau + \sum_{j=1}^N \mathbf{b}_j\mathbf{b}_j^\tau - z\mathbf{I}_T\Big)^{-1}, \tag{A.24}$$

where

$$\mathbf{a}_j = \frac{\widetilde{\mathbf{u}}_j}{||\widetilde{\mathbf{u}}_j||} \quad \text{and} \quad \mathbf{b}_j = \frac{\widetilde{\mathbf{X}}_j(\boldsymbol{\beta} - \hat{\boldsymbol{\beta}})}{||\widetilde{\mathbf{u}}_j||}. \tag{A.25}$$

Define $\mathbf{A}(z) = \sum_{j=1}^N \mathbf{a}_j\mathbf{a}_j^\tau - z\mathbf{I}_T$. We further obtain from (A.15) and (A.24) that

$$\Big|tr\Big(\underline{\widetilde{\mathbf{R}}}_N - z\mathbf{I}_T\Big)^{-1} - tr\mathbf{A}^{-1}(z)\Big| \le \Big|\sum_{j=1}^N A_j\Big| + \Big|\sum_{j=1}^N B_j\Big| + \Big|\sum_{j=1}^N C_j\Big|, \tag{A.26}$$

where $A_j = \mathbf{b}_j^\tau\mathbf{A}^{-1}(z)\underline{\widetilde{\mathbf{R}}}_N^{-1}(z)\mathbf{a}_j$, $B_j = \mathbf{a}_j^\tau\mathbf{A}^{-1}(z)\underline{\widetilde{\mathbf{R}}}_N^{-1}(z)\mathbf{b}_j$ and $C_j = \mathbf{b}_j^\tau\mathbf{A}^{-1}(z)\underline{\widetilde{\mathbf{R}}}_N^{-1}(z)\mathbf{b}_j$.

As will be seen, the three terms on the right hand of (A.26) converge to zero in probability. First, as in (A.11) and (A.12) one may verify that

$$\max_{1\le j\le N} ||\mathbf{b}_j|| = O_P\Big(\frac{1}{\sqrt{NT}}\Big), \tag{A.27}$$

which immediately implies $\Big|\sum_{j=1}^N C_j\Big| \xrightarrow{i.p.} 0$.

Since A_j and B_j are similar, we prove $\sum_{j=1}^N A_j \xrightarrow{i.p.} 0$ only and the proof of $\sum_{j=1}^N B_j \xrightarrow{i.p.} 0$ follows analogously. To this end, we first expand A_j. Let

$$\mathbf{A}_1^{(r)}(z) = \Big(\mathbf{A}(z) + \sum_{k=1}^N \mathbf{a}_k\mathbf{b}_k^\tau + \sum_{k=1}^N \mathbf{b}_k\mathbf{a}_k^\tau + \sum_{k=r}^N \mathbf{b}_k\mathbf{b}_k^\tau\Big)^{-1},$$

$$\mathbf{A}_2^{(r)}(z) = \Big(\mathbf{A}(z) + \sum_{k=1}^N \mathbf{a}_k\mathbf{b}_k^\tau + \sum_{k=r}^N \mathbf{b}_k\mathbf{a}_k^\tau\Big)^{-1}, \ \mathbf{A}_3^{(r)}(z) = \Big(\mathbf{A}(z) + \sum_{k=r}^N \mathbf{a}_k\mathbf{b}_k^\tau\Big)^{-1}.$$

We conclude from (A.16) that

$$\begin{aligned} A_j &= \mathbf{b}_j^\tau\mathbf{A}^{-1}(z)\mathbf{A}_1^{(1)}(z)\mathbf{a}_j \\ &= \mathbf{b}_j^\tau\mathbf{A}^{-1}(z)\mathbf{A}_1^{(2)}(z)\mathbf{a}_j - \frac{\mathbf{b}_j^\tau\mathbf{A}^{-1}(z)\mathbf{A}_1^{(2)}(z)\mathbf{b}_1\mathbf{b}_1^\tau\mathbf{A}_1^{(2)}(z)\mathbf{a}_j}{1 + \mathbf{b}_1^\tau\mathbf{A}_1^{(2)}(z)\mathbf{b}_1}. \end{aligned} \tag{A.28}$$

For the first term of (A.28), using the formula (A.16) again, we obtain

$$\mathbf{b}_j^\tau\mathbf{A}^{-1}(z)\mathbf{A}_1^{(2)}(z)\mathbf{a}_j = \mathbf{b}_j^\tau\mathbf{A}^{-1}(z)\mathbf{A}_1^{(3)}(z)\mathbf{a}_j - \frac{\mathbf{b}_j^\tau\mathbf{A}^{-1}(z)\mathbf{A}_1^{(3)}(z)\mathbf{b}_2\mathbf{b}_2^\tau\mathbf{A}_1^{(3)}(z)\mathbf{a}_j}{1 + \mathbf{b}_2^\tau\mathbf{A}_1^{(3)}(z)\mathbf{b}_2}.$$

Applying (A.16) repeatedly yields

$$A_j = \mathbf{b}_j^\tau\mathbf{A}^{-1}(z)\mathbf{A}_2^{(1)}(z)\mathbf{a}_j - \eta_1^{(j)}(z), \tag{A.29}$$

where $\eta_1^{(j)}(z) = \sum_{k=1}^N \frac{\mathbf{b}_j^\tau\mathbf{A}^{-1}(z)\mathbf{A}_1^{(k+1)}(z)\mathbf{b}_k\mathbf{b}_k^\tau\mathbf{A}_1^{(k+1)}(z)\mathbf{a}_j}{1+\mathbf{b}_k^\tau\mathbf{A}_1^{(k+1)}(z)\mathbf{b}_k}$ with $\mathbf{A}_1^{(N+1)}(z) = \mathbf{A}_2^{(1)}(z)$.

Similarly, for the first term on the right hand-side of (A.29), we have

$$\mathbf{b}_j^\tau \mathbf{A}^{-1}(z)\mathbf{A}_2^{(1)}(z)\mathbf{a}_j = \mathbf{b}_j^\tau \mathbf{A}^{-1}(z)\mathbf{A}_3^{(1)}(z)\mathbf{a}_j - \eta_2^{(j)}(z), \tag{A.30}$$

where $\eta_2^{(j)}(z) = \sum_{k=1}^N \frac{\mathbf{b}_j^\tau \mathbf{A}^{-1}(z)\mathbf{A}_2^{(k+1)}(z)\mathbf{b}_k\mathbf{a}_k^\tau\mathbf{A}_2^{(k+1)}(z)\mathbf{a}_j}{1+\mathbf{a}_k^\tau\mathbf{A}_2^{(k+1)}(z)\mathbf{b}_k}$ with $\mathbf{A}_2^{(N+1)}(z) = \mathbf{A}_3^{(1)}(z)$.

Using (A.16) again for the first term on the right hand-side of (A.30), we get

$$\mathbf{b}_j^\tau \mathbf{A}^{-1}(z)\mathbf{A}_3^{(1)}(z)\mathbf{a}_j = \mathbf{b}_j^\tau \mathbf{A}^{-2}(z)\mathbf{a}_j - \eta_3^{(j)}(z), \tag{A.31}$$

where $\eta_3^{(j)}(z) = \sum_{k=1}^N \frac{\mathbf{b}_j^\tau \mathbf{A}^{-1}(z)\mathbf{A}_3^{(k+1)}(z)\mathbf{a}_k\mathbf{b}_k^\tau\mathbf{A}_3^{(k+1)}(z)\mathbf{a}_j}{1+\mathbf{b}_k^\tau\mathbf{A}_3^{(k+1)}(z)\mathbf{a}_k}$ with $\mathbf{A}_3^{(N+1)}(z) = \mathbf{A}^{-1}(z)$.

Putting (A.29)-(A.31) together we have

$$A_j = \mathbf{b}_j^\tau \mathbf{A}^{-2}(z)\mathbf{a}_j - \eta_1^{(j)}(z) - \eta_2^{(j)}(z) - \eta_3^{(j)}(z). \tag{A.32}$$

We are now in a position to prove that $\sum_{j=1}^N A_j = o_P(1)$. To this end, write

$$\mathbf{b}_j = \frac{\widetilde{\mathbf{X}}_j(\boldsymbol{\beta} - \hat{\boldsymbol{\beta}})}{||\widetilde{\mathbf{u}}_j||} = \sum_{r=1}^p \frac{(\beta_r - \hat{\beta}_r)\dot{\widetilde{\mathbf{x}}}_{jr}}{||\widetilde{\mathbf{u}}_j||}. \tag{A.33}$$

It follows that

$$\begin{aligned} &\Big|\sum_{j=1}^N \mathbf{b}_j^\tau \mathbf{A}^{-2}(z)\mathbf{a}_j\Big| = \Big|\sum_{j=1}^N\sum_{r=1}^p \frac{(\beta_r - \hat{\beta}_r)\dot{\widetilde{\mathbf{x}}}_{jr}^\tau \mathbf{A}^{-2}(z)\widetilde{\mathbf{u}}_j}{||\widetilde{\mathbf{u}}_j||^2}\Big| \\ \le\ & \sum_{r=1}^p \Big|\beta_r - \hat{\beta}_r\Big| \cdot \max_{1\le j\le N}\Big|\frac{T}{||\widetilde{\mathbf{u}}_j||^2}\Big| \cdot \sum_{j=1}^N \Big|\frac{\dot{\widetilde{\mathbf{x}}}_{jr}^\tau \mathbf{A}^{-2}(z)\widetilde{\mathbf{u}}_j}{T}\Big| = O_P\Big(\frac{1}{\sqrt{N}}\Big), \end{aligned} \tag{A.34}$$

where $\dot{\widetilde{\mathbf{x}}}_{jr}$ is the r-th column of $\widetilde{\mathbf{X}}_j$, the last estimate uses (A.3), (A.12) and the fact that

$$\mathbb{E}\Big|\dot{\widetilde{\mathbf{x}}}_{jr}^\tau \mathbf{A}^{-2}(z)\widetilde{\mathbf{u}}_j\Big|^2 = \mathbb{E}\Big(\widetilde{\mathbf{u}}_j^\tau \mathbf{A}^{-2}(\bar{z})\boldsymbol{\Sigma}_x \mathbf{A}^{-2}(z)\widetilde{\mathbf{u}}_j\Big) \le \frac{M}{v^4}\|\boldsymbol{\Sigma}_x\|\mathbb{E}\Big(\widetilde{\mathbf{u}}_j^\tau\widetilde{\mathbf{u}}_j\Big) = O(T),$$

with $\boldsymbol{\Sigma}_x$ being a $T \times T$ matrix whose diagonal elements are $1 - \frac{1}{T}$ and off-diagonal elements are $-\frac{1}{T}$.

In order to deal with $\sum_{j=1}^N \eta_3^{(j)}(z)$, note that $\|\mathbf{A}^{-1}(z)\| \le 1/v_0$. But we have to prove that the spectral norm of $\mathbf{A}_3^{(r)}(z)$ is bounded. We conclude from the main theorems in Jiang (2004) and Xiao and Zhou (2010), and (A.12) that with probability one

$$\|\sum_{j=1}^N \frac{\dot{\widetilde{\mathbf{x}}}_{jr}\widetilde{\mathbf{u}}_j^\tau}{||\widetilde{\mathbf{u}}_j||^2}\| = \|\mathbf{S}_1\mathbf{S}_2\mathbf{S}_3\| \le \|\mathbf{S}_1\|\|\mathbf{S}_2\|\|\mathbf{S}_3\| \le M \tag{A.35}$$

where $\mathbf{S}_1 = \frac{1}{\sqrt{T}}(\dot{\widetilde{\mathbf{x}}}_{1r}, \cdots, \dot{\widetilde{\mathbf{x}}}_{Nr})$, $\mathbf{S}_2 = \mathrm{diag}\Big(\frac{\sqrt{T}}{||\widetilde{\mathbf{u}}_1||}, \cdots, \frac{\sqrt{T}}{||\widetilde{\mathbf{u}}_N||}\Big)$ and $\mathbf{S}_3 = \Big(\frac{\widetilde{\mathbf{u}}_1^\tau}{||\widetilde{\mathbf{u}}_1||}, \cdots, \frac{\widetilde{\mathbf{u}}_N^\tau}{||\widetilde{\mathbf{u}}_N||}\Big)$.

Using (A.33) we obtain from (A.3) and (A.35) that

$$\begin{aligned} &\Big|\Big|\sum_{j=1}^N \mathbf{b}_j\mathbf{a}_j^\tau\Big|\Big| = \Big|\Big|\sum_{j=1}^N\sum_{r=1}^p(\beta_r - \hat{\beta}_r)\sum_{j=1}^N \frac{\dot{\widetilde{\mathbf{x}}}_{jr}\widetilde{\mathbf{u}}_j^\tau}{||\widetilde{\mathbf{u}}_j||^2}\Big|\Big| \\ \le\ & \sum_{r=1}^p \Big|\beta_r - \hat{\beta}_r\Big| \cdot \Big|\Big|\sum_{j=1}^N \frac{\dot{\widetilde{\mathbf{x}}}_{jr}\widetilde{\mathbf{u}}_j^\tau}{||\widetilde{\mathbf{u}}_j||^2}\Big|\Big| = O_P\Big(\frac{1}{\sqrt{NT}}\Big). \end{aligned} \tag{A.36}$$

Similarly, One can show

$$\Big|\Big|\sum_{j=r}^N \mathbf{b}_j\mathbf{a}_j^\tau\Big|\Big| = O_P\Big(\frac{1}{\sqrt{NT}}\Big). \tag{A.37}$$

In view of (A.37) we have

$$\| \mathbf{A}_3^{(r)}(z) \| = \| \mathbf{A}^{-1}(z) - \mathbf{A}_3^{(r)}(z)\Big(\sum_{k=r}^{N} \mathbf{a}_k \mathbf{b}_k^{\tau}\Big)\mathbf{A}^{-1}(z) \| \le \frac{1}{v_0} + \frac{M\|\mathbf{A}_3^{(r)}(z)\|}{v_0} O_p\left(\frac{1}{\sqrt{NT}}\right)$$

which further implies that with probability one

$$\| \mathbf{A}_3^{(r)}(z) \| \le M. \tag{A.38}$$

It follows from (A.38), (A.27) and (A.25) that

$$\max_k |\mathbf{b}_k^{\tau} \mathbf{A}_3^{(k+1)}(z)\mathbf{a}_k| \le \max_k \| \mathbf{A}_3^{(k+1)} \| \max_k \|\mathbf{b}_k\|\|\mathbf{a}_k\| = O_P\Big(\frac{1}{\sqrt{NT}}\Big). \tag{A.39}$$

We then conclude from (A.27), (A.36), (A.38) and (A.39)

$$\begin{aligned} \Big|\sum_{j=1}^{N} \eta_3^{(j)}(z)\Big| &= \Big|\sum_{k=1}^{N} \frac{\mathbf{b}_k^{\tau}\mathbf{A}_3^{(k+1)}(z)\sum_{j=1}^{N}\mathbf{a}_j\mathbf{b}_j^{\tau}\mathbf{A}^{-1}(z)\mathbf{A}_3^{(k+1)}(z)\mathbf{a}_k}{1+\mathbf{b}_k^{\tau}\mathbf{A}_3^{(k+1)}(z)\mathbf{a}_k}\Big| \\ &\le M\sum_{k=1}^{N}\frac{1}{1-\max_k|\mathbf{b}_k^{\tau}\mathbf{A}_3^{(k+1)}(z)\mathbf{a}_k|}\Big|\Big|\mathbf{b}_k^{\tau}\mathbf{A}_3^{(k+1)}(z)\Big|\Big|\cdot\Big|\Big|\sum_{j=1}^{N}\mathbf{a}_j\mathbf{b}_j^{\tau}\Big|\Big|\cdot\Big|\Big|\mathbf{A}^{-1}(z)\mathbf{A}_3^{(k+1)}(z)\mathbf{a}_k\Big|\Big| \\ &= O_P\Big(\frac{1}{\sqrt{NT}}\Big). \end{aligned} \tag{A.40}$$

Similarly one can prove that

$$\Big|\sum_{j=1}^{N}\eta_2^{(j)}(z)\Big| = O_P\Big(\frac{1}{\sqrt{NT}}\Big), \quad \Big|\sum_{j=1}^{N}\eta_1^{(j)}(z)\Big| = O_P\Big(\frac{1}{\sqrt{NT}}\Big). \tag{A.41}$$

It follows from (A.34), (A.40) and (A.41) that

$$\Big|\sum_{j=1}^{N} A_j\Big| = O_P\Big(\frac{1}{\sqrt{N}}\Big). \tag{A.42}$$

Summarizing the above we have

$$\Big|tr\underline{\widetilde{\mathbf{R}}}_N^{-1}(z) - tr\mathbf{A}^{-1}(z)\Big| = o_P(1), \tag{A.43}$$

uniformly for z on the contour.

Combining (A.23), (A.43) with Theorem 1 of Gao, el. (2014), it follows that the CLT of $tr\underline{\widetilde{\mathbf{R}}}_N^{-1}(z)$ is the same as that of $tr\mathbf{A}^{-1}(z)$, which is provided in Theorem 1 of Gao, el. (2014).

7.3 Proof of Theorem 2

Recall that

$$s_N(\ell) = \int e^{i\ell\lambda}dF^{\hat{\mathbf{R}}_N}(\lambda) =: \phi_N(\lambda) + i\cdot\psi_N(\ell),$$

where $\phi_N(\ell) = \int\cos(\ell\lambda)dF^{\hat{\mathbf{R}}_N}(\lambda)$ and $\psi_N(\lambda) = \int\sin(\ell\lambda)dF^{\hat{\mathbf{B}}_N}(\lambda)$.

Let ℓ belong to a closed interval $I = [L_1, L_2]$. To establish Theorem 2, in view of Theorem 1, it suffices to prove the tightness of $\{\big(\phi_N(\ell), \psi_N(\ell)\big) : \ell \in I\}$. Thus it suffices to prove the tightness of $N\Big(s_N(\ell) - s(\ell)\Big)$.

From the proof of Theorem 1, it is known that the asymptotic distribution of $N\Big(s_N(\ell)-s(\ell)\Big)$ is the same as that of $N\Big(\widetilde{s}_N(\ell)-s(\ell)\Big)$, where $\widetilde{s}_N(\ell)$ is $s_N(\ell)$ with $\hat{\mathbf{R}}_N$ replaced by $\mathbf{R}_N$ and $\mathbf{R}_N$ is the sample correlation matrix $\hat{\mathbf{R}}_N$ with $\hat{\mathbf{u}}_j$, $j=1,2,\dots,N$ replaced by $\widetilde{\mathbf{u}}_j$, $j=1,2,\dots,N$ respectively. So it is enough to provide the tightness of $N\Big(\widetilde{s}_N(\ell)-s(\ell)\Big)$.

Repeating the same truncation and centralization steps as those in Gao, el. (2014), we can assume that

$$|U_{jt}|<\delta_N\sqrt{T},\ EU_{jt}=0,\ E|U_{jt}|^2=1,\ \text{and}\ E|U_{jt}|^4<\infty. \tag{A.44}$$

Set $M_N(z)=N[m_{F_N^{\mathbf{R}}}(z)-m_{F_{c_N}}(z)]$. By the Cauchy theorem

$$f(x)=-\frac{1}{2\pi i}\oint\frac{f(z)}{z-x}dz, \tag{A.45}$$

we have, with probability one, for N large enough,

$$\int e^{i\ell x}dp(F^{\mathbf{R}_N}(x)-F_{c_N}(x))=-\frac{1}{2\pi i}\oint_{\mathcal{C}}e^{i\ell z}M_N(z)dz. \tag{A.46}$$

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

$$\mathcal{C}_u=\{x+iv_0: x\in[x_l,x_r]\}, \tag{A.47}$$

where $v_0>0$, x_r is any number greater than $(1+\sqrt{c})^2$, x_l is any negative number if $c\geq 1$ and otherwise choose $x_l\in(0,(1-\sqrt{c})^2)$. 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

$$\mathcal{C}_+=\{x_l+iv: v\in[0,v_0]\}\cup\mathcal{C}_u\cup\{x_r+iv: v\in[0,v_0]\}. \tag{A.48}$$

From Theorem 1 in Gao, el. (2014), the argument regarding the equivalence in probability of $M_N(z)$ and its truncation version in the proof of Theorem 1 of Gao, el. (2014), and Lemma 3, we have

$$\oint_{\mathcal{C}}|M_N(z)||dz|\xrightarrow{D}\oint_{\mathcal{C}}|M(z)||dz|, \tag{A.49}$$

where $M(z)$ is a Gaussian process, i.e. the limit of $M_N(z)$.

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

$$\begin{aligned}
&\sup_{|\ell_1-\ell_2|<\delta,\ell_1,\ell_2\in I}\Big|\oint_{\mathcal{C}}(e^{i\ell_1 z}-e^{i\ell_2 z})M_N(z)dz\Big|\\
\leq\ &\sup_{|\ell_1-\ell_2|<\delta,\ell_1,\ell_2\in I}\Big|\oint_{\mathcal{C}}\sqrt{\big(Re(ize^{i\ell_3 z})\big)^2+\big(Im(ize^{i\ell_4 z})\big)^2}\delta|M_N(z)||dz|\Big|\\
\leq\ &K\delta\Big|\oint_{\mathcal{C}}|M_N(z)||dz|\Big|\xrightarrow{D}K\delta\Big|\oint_{\mathcal{C}}|M(z)||dz|\Big|,\quad\text{as}\quad N\to\infty,
\end{aligned} \tag{A.50}$$

where ℓ_3 and ℓ_4 lie in the interval $[L_1,L_2]$, the last inequality uses (A.49) and the fact that $Re(ize^{i\ell_3 z})$, $Im(ize^{i\ell_4 z})$ 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 (A.50), we have for any $\varepsilon>0$,

$$P\Big(\sup_{|\ell_1-\ell_2|<\delta,\ell_1,\ell_2\in[0,1]}\Big|\oint_{\mathcal{C}}(e^{i\ell_1 z}-e^{i\ell_2 z})M_N(z)dz\Big|\geq\varepsilon\Big)\leq P\Big(K\delta\Big|\oint_{\mathcal{C}}|M_N(z)||dz|\Big|\geq\varepsilon\Big) \tag{A.51}$$

and

$$\lim_{\delta\to 0}\limsup_{N\to\infty}P\Big(K\delta\Big|\oint_{\mathcal{C}}|M_N(z)||dz|\Big|\geq\varepsilon\Big)=\lim_{\delta\to 0}P\Big(K\delta\Big|\oint_{\mathcal{C}}|M(z)||dz|\Big|\geq\varepsilon\Big)=0. \tag{A.52}$$

Hence (A.51) and (A.52) imply that

$$\lim_{\delta\to 0}\limsup_{N\to\infty} P\Big(\sup_{|\ell_1-\ell_2|<\delta,\ell_1,\ell_2\in I}\Big|\oint_{\mathcal{C}}(e^{i\ell_1 z}-e^{i\ell_2 z})M_N(z)dz\Big|\geq\varepsilon\Big)=0. \tag{A.53}$$

By Theorem 7.3 of Billingsley (1999), $\int e^{i\ell x}dN(F^{\mathbf{R}_N}(x)-F_{c_N}(x))$ is tight.

7.4 Proof of Theorem 3

Consider N^2S_N under the alternative hypothesis $\mathbf{H}_a$ and rewrite it as follows.

$$N^2S_N=N^2\int_{L_1}^{L_2}|s_N(\ell)-s(\ell)|^2dU(\ell)=\int_{L_1}^{L_2}\Big[S_N^{\cos,\mathbf{H}_a}(\ell)\Big]^2dU(\ell)+\int_{L_1}^{L_2}\Big[S_N^{\sin,\mathbf{H}_a}(\ell)\Big]^2dU(\ell),$$

where

$$S_N^{\cos,\mathbf{H}_a}(\ell)=\int\cos(\ell x)dN(F_{\mathbf{H}_a}^{\hat{\mathbf{R}}_N}(x)-F_{c_N}(x)),\ \ S_N^{\sin,\mathbf{H}_a}(\ell)=\int\sin(\ell x)dN(F_{\mathbf{H}_a}^{\hat{\mathbf{R}}_N}(x)-F_{c_N}(x)).$$

We may further write

$$\Big[S_N^{\cos,\mathbf{H}_a}(\ell)\Big]^2=\Big[S_N^{\cos,\mathbf{H}_0}(\ell)\Big]^2+\Big[S_N^{\cos,\mathbf{H}_a,\mathbf{H}_0}(\ell)\Big]^2+2\Big[S_N^{\cos,\mathbf{H}_a,\mathbf{H}_0}(\ell)\Big]\Big[S_N^{\cos,\mathbf{H}_0}(\ell)\Big],$$

where $S_N^{\cos,\mathbf{H}_0}(\ell)$ is obtained from $S_N^{\cos,\mathbf{H}_a}(\ell)$ with $F_{\mathbf{H}_a}^{\hat{\mathbf{R}}_N}(x)$ replaced by $F_{\mathbf{H}_0}^{\hat{\mathbf{R}}_N}(x)$ and

$$S_N^{\cos,\mathbf{H}_a,\mathbf{H}_0}(\ell)=\int\cos(\ell x)dN(F_{\mathbf{H}_a}^{\hat{\mathbf{R}}_N}(x)-F_{\mathbf{H}_0}^{\hat{\mathbf{R}}_N}(x)).$$

By Holder's inequality we obtain

$$\Big|\int_{L_1}^{L_2}\Big[S_N^{\cos,\mathbf{H}_a,\mathbf{H}_0}(\ell)\Big]\Big[S_N^{\cos,\mathbf{H}_0}(\ell)\Big]dU(\ell)\Big|^2\leq\int_{L_1}^{L_2}\Big[S_N^{\cos,\mathbf{H}_a,\mathbf{H}_0}(\ell)\Big]^2dU(\ell)\int_{L_1}^{L_2}\Big[S_N^{\cos,\mathbf{H}_0}(\ell)\Big]^2dU(\ell)\Big|.$$

This, together with the proof of Theorem 1 in Gao, el. (2014), implies

$$\int_{L_1}^{L_2}\Big[S_N^{\cos,\mathbf{H}_a,\mathbf{H}_0}(\ell)\Big]\Big[S_N^{\cos,\mathbf{H}_0}(\ell)\Big]dU(\ell)=o_p\Big(\int_{L_1}^{L_2}(S_N^{\cos,\mathbf{H}_a,\mathbf{H}_0}(\ell))^2+(S_N^{\cos,\mathbf{H}_a,\mathbf{H}_0}(\ell))^2dU(\ell)\Big).$$

Similarly,

$$\Big[S_N^{\sin,\mathbf{H}_a}(\ell)\Big]^2=\Big[S_N^{\sin,\mathbf{H}_0}(\ell)\Big]^2+\Big[S_N^{\sin,\mathbf{H}_a,\mathbf{H}_0}(\ell)\Big]^2+2\Big[S_N^{\sin,\mathbf{H}_a,\mathbf{H}_0}(\ell)\Big]\Big[S_N^{\sin,\mathbf{H}_0}(\ell)\Big]$$

and

$$\int_{L_1}^{L_2}\Big[S_N^{\sin,\mathbf{H}_a,\mathbf{H}_0}(\ell)\Big]\Big[S_N^{\sin,\mathbf{H}_0}(\ell)\Big]dU(\ell)=o_p\Big(\int_{L_1}^{L_2}(S_N^{\sin,\mathbf{H}_a,\mathbf{H}_0}(\ell))^2+(S_N^{\sin,\mathbf{H}_a,\mathbf{H}_0}(\ell))^2dU(\ell)\Big),$$

where $S_N^{\sin,\mathbf{H}_0}(\ell)$ is defined similarly and $S_N^{\sin,\mathbf{H}_a,\mathbf{H}_0}(\ell)=\int\sin(\ell x)dN(F_{\mathbf{H}_a}^{\hat{\mathbf{R}}_N}(x)-F_{\mathbf{H}_0}^{\hat{\mathbf{R}}_N}(x))$.

Note that

$$\int_{L_1}^{L_2}(S_N^{\sin,\mathbf{H}_a,\mathbf{H}_0}(\ell))^2+(S_N^{\sin,\mathbf{H}_a,\mathbf{H}_0}(\ell))^2dU(\ell)=\int_{L_1}^{L_2}\Big|\int e^{i\ell x}dp(F_{\mathbf{H}_a}^{\hat{\mathbf{R}}_N}(x)-F_{\mathbf{H}_0}^{\hat{\mathbf{R}}_N}(x))\Big|^2dU(\ell).$$

Summarizing the above we have obtained

$$N^2S_N=\int_{L_1}^{L_2}(\Big[S_N^{\cos,\mathbf{H}_0}(\ell)\Big]^2+\Big[S_N^{\sin,\mathbf{H}_0}(\ell)\Big]^2)dU(\ell)+\int_{L_1}^{L_2}\Big|\int e^{i\ell x}dN(F_{\mathbf{H}_a}^{\hat{\mathbf{R}}_N}(x)-F_{\mathbf{H}_0}^{\hat{\mathbf{R}}_N}(x))\Big|^2dU(\ell)$$

$$+o_p\Big(\int_{L_1}^{L_2}\Big|\int e^{i\ell x}dN(F_{\mathbf{H}_a}^{\hat{\mathbf{R}}_N}(x)-F_{\mathbf{H}_0}^{\hat{\mathbf{R}}_N}(x))\Big|^2dU(\ell)\Big).$$

For notational simplicity, we adopt the same notation as what has been used in the proof of Theorem 1.

7.5 Proof of Theorem 4

Recall the original model $\widetilde{\mathbf{Y}}_j = \widetilde{\mathbf{X}}_j\boldsymbol{\beta} + \widetilde{\mathbf{u}}_j, \quad j = 1, 2, \ldots, N$. Consider the local alternative hypothesis: $\mathbb{H}_a$: $\mathbf{u}_j = \boldsymbol{\varepsilon}_j + \frac{1}{\sqrt{T}}\mathbf{v}, \quad j = 1, 2, \ldots, N$. The sample correlation matrix $\underline{\hat{\mathbf{R}}}_N$ under $\mathbb{H}_a$ can be written as

$$\underline{\hat{\mathbf{R}}}_N = \sum_{j=1}^{N} \frac{\hat{\mathbf{u}}_j\hat{\mathbf{u}}_j^{\tau}}{||\hat{\mathbf{u}}_j||^2}, \tag{A.54}$$

where $\hat{\mathbf{u}}_j = \widetilde{\mathbf{u}}_j + \widetilde{\mathbf{X}}_j(\boldsymbol{\beta} - \hat{\boldsymbol{\beta}})$ and $\widetilde{\mathbf{u}}_j = \widetilde{\boldsymbol{\varepsilon}}_j + \frac{1}{\sqrt{T}}\widetilde{\mathbf{v}}$, with $\widetilde{\mathbf{v}} = \mathbf{v} - \bar{v}\mathbf{e}$, $\bar{v} = \frac{1}{T}\sum_{t=1}^{T} v_t$, $\widetilde{\boldsymbol{\varepsilon}}_j = \boldsymbol{\varepsilon}_j - \bar{\varepsilon}_j\mathbf{e}$, $\bar{\varepsilon}_j = \frac{1}{T}\sum_{t=1}^{T}\varepsilon_{jt}$ and $\mathbf{e}$ is a $T \times 1$ vector consisting of all 1.

From (A.54), it is equivalent to considering the sample correlation matrix constructed by $\frac{\hat{\mathbf{u}}_j}{\sqrt{\mathbb{E}(\varepsilon_{11}^2)}}$, $j = 1, 2, \ldots, N$. Based on the assumption that the variances of ε_{11} and v_1 are equal, the variances of elements of $\frac{\hat{\boldsymbol{\varepsilon}}_j}{\sqrt{\mathbb{E}(\varepsilon_{11}^2)}}$ and $\frac{\hat{\mathbf{v}}}{\sqrt{\mathbb{E}(\varepsilon_{11}^2)}}$ are all equal to 1. For simplicity of notation, we still denote them by $\hat{\mathbf{u}}_j$, $\hat{\boldsymbol{\varepsilon}}_j$ and $\hat{\mathbf{v}}$, respectively.

As in Theorem 1, one can prove the following two results:

$$tr\Big(\underline{\hat{\mathbf{R}}}_N - z\mathbf{I}_T\Big)^{-1} = tr\Big(\underline{\widetilde{\mathbf{R}}}_N - z\mathbf{I}_T\Big)^{-1} + o_P(1) \tag{A.55}$$

and

$$tr\Big(\underline{\widetilde{\mathbf{R}}}_N - z\mathbf{I}_T\Big)^{-1} = tr\Big(\underline{\mathbf{R}}_N - z\mathbf{I}_T\Big)^{-1} + o_P(1), \tag{A.56}$$

where $\underline{\widetilde{\mathbf{R}}}_N = \sum_{j=1}^{N} \frac{\hat{\mathbf{u}}_j\hat{\mathbf{u}}_j^{\tau}}{||\widetilde{\mathbf{u}}_j||^2}$ and $\underline{\mathbf{R}}_N = \sum_{j=1}^{N} \frac{\widetilde{\mathbf{u}}_j\widetilde{\mathbf{u}}_j^{\tau}}{||\widetilde{\mathbf{u}}_j||^2}$; and $o_P(1)$ holds uniformly for $z \in \mathcal{C}$. Indeed, by carefully checking on the proof of Theorem 1, the differences between the proof of (A.55) and (A.56) under $\mathbb{H}_a$ and that of (A.7) and (A.43) lie in the proofs of (A.3), (A.10) and (A.12) under $\mathbb{H}_a$ which are listed below.

1. Consider (A.3) under $\mathbb{H}_a$. Note that the estimator $\hat{\boldsymbol{\beta}}$ under $\mathbb{H}_a$ is

$$\begin{aligned}
\Big|\Big|\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}_0\Big|\Big| &= \Big|\Big|\Big(\sum_{j=1}^{N}\widetilde{\mathbf{X}}_j^{\tau}\widetilde{\mathbf{X}}_j\Big)^{-1}\Big(\sum_{j=1}^{N}\widetilde{\mathbf{X}}_j^{\tau}\widetilde{\mathbf{u}}_j\Big)\Big|\Big| \\
&\leq \Big|\Big|\Big(\sum_{j=1}^{N}\widetilde{\mathbf{X}}_j^{\tau}\widetilde{\mathbf{X}}_j\Big)^{-1}\Big(\sum_{j=1}^{N}\widetilde{\mathbf{X}}_j^{\tau}\widetilde{\varepsilon}_j\Big)\Big|\Big| + \Big|\Big|\Big(\sum_{j=1}^{N}\widetilde{\mathbf{X}}_j^{\tau}\widetilde{\mathbf{X}}_j\Big)^{-1}\Big(\sum_{j=1}^{N}\frac{1}{\sqrt{T}}\widetilde{\mathbf{X}}_j^{\tau}\widetilde{\mathbf{v}}\Big)\Big|\Big| \\
&= O_P\Big(\frac{1}{\sqrt{NT}}\Big) + O_P\Big(\frac{1}{T\sqrt{N}}\Big) = O_P\Big(\frac{1}{\sqrt{NT}}\Big),
\end{aligned}$$

where the second last term uses the facts that $\frac{1}{NT}\sum_{j=1}^{N}\widetilde{\mathbf{X}}_j^{\tau}\widetilde{\mathbf{X}}_j = O_P(1)$ and that by recalling $\dot{\widetilde{\mathbf{x}}}_{jr}$ below (A.34),

$$\begin{aligned}
\mathbb{E}\Big|\Big|\frac{1}{NT^{3/2}}\sum_{j=1}^{N}\widetilde{\mathbf{X}}_j^{\tau}\widetilde{\mathbf{v}}\Big|\Big|^2 &= \frac{1}{N^2T^3}\sum_{r=1}^{p}\mathbb{E}\Big(\sum_{j=1}^{N}\dot{\widetilde{\mathbf{x}}}_{jr}^{\tau}\widetilde{\mathbf{v}}\Big)^2 = \frac{1}{T^3N^2}\sum_{r=1}^{p}\sum_{j_1,j_2=1}^{N}\mathbb{E}\Big(\dot{\widetilde{\mathbf{x}}}_{j_1r}^{\tau}\widetilde{\mathbf{v}}\widetilde{\mathbf{v}}^{\tau}\dot{\widetilde{\mathbf{x}}}_{j_2r}\Big) \\
&= \frac{1}{T^3N^2}\sum_{r=1}^{p}\sum_{j=1}^{N}\mathbb{E}\widetilde{\mathbf{v}}^{\tau}\dot{\widetilde{\mathbf{x}}}_{jr}\dot{\widetilde{\mathbf{x}}}_{jr}^{\tau}\widetilde{\mathbf{v}} = O\Big(\frac{1}{T^2N}\Big).
\end{aligned}$$

2. Consider (A.10) under $\mathbb{H}_a$. From Lemma 5 we have

$$\max_{1\leq j\leq N}\Big|\sum_{k=1}^{T}\frac{1}{T}\widetilde{u}_{jk}\widetilde{X}_{jk,r}\Big| \leq \max_{1\leq j\leq N}\Big|\sum_{k=1}^{T}\frac{1}{T}\widetilde{\varepsilon}_{jk}\widetilde{X}_{jk,r}\Big| + \max_{1\leq j\leq N}\Big|\sum_{k=1}^{T}\frac{1}{T\sqrt{T}}\widetilde{v}_k\widetilde{X}_{jk,r}\Big| = O_P\Big(\frac{1}{\sqrt{T}}\Big).$$

3. (A.12) under $\mathbb{H}_a$ also holds because by Lemma 5

$$\max_{1\le j\le N}\frac{T}{||\widetilde{\mathbf{u}}_j||^2}=\max_{1\le j\le N}\frac{T}{||\widetilde{\boldsymbol{\varepsilon}}_j+\frac{1}{\sqrt{T}}\widetilde{\mathbf{v}}||^2}=\max_{1\le j\le N}\frac{1}{\frac{||\widetilde{\boldsymbol{\varepsilon}}_j||^2}{T}+\frac{||\widetilde{\mathbf{v}}||^2}{T^2}+\frac{2\widetilde{\boldsymbol{\varepsilon}}^\tau\widetilde{\mathbf{v}}}{T\sqrt{T}}}=O_P(1). \tag{A.57}$$

For later use one can similarly prove

$$\max_{1\le j\le N}\frac{T}{||\widetilde{\varepsilon}_j||^2}=O_P(1)\quad\text{and}\quad\max_{1\le j\le N}\Big|\frac{||\widetilde{\varepsilon}_j||^2}{T}-1\Big|=O_P(1). \tag{A.58}$$

Next, we develop the central limit theorem for the term $tr\Big(\underline{\mathbf{R}}_N-z\mathbf{I}_T\Big)^{-1}$. Let

$$\mathbf{q}_N=\sum_{j=1}^N\frac{\widetilde{\boldsymbol{\varepsilon}}_j}{||\widetilde{\mathbf{u}}_j||^2},\quad\mathbf{h}_N=\frac{\widetilde{\mathbf{v}}}{\sqrt{T}},\quad d_N=\sum_{j=1}^N\frac{1}{||\widetilde{\mathbf{u}}_j||^2},\quad\underline{\mathbf{B}}_N=\sum_{j=1}^N\frac{\widetilde{\boldsymbol{\varepsilon}}_j\widetilde{\boldsymbol{\varepsilon}}_j^\tau}{||\widetilde{\mathbf{u}}_j||^2}.$$

Using the formula (A.16), we have

$$\begin{aligned}
tr\Big(\underline{\mathbf{R}}_N-z\mathbf{I}_T\Big)^{-1}&=tr\Big(\underline{\mathbf{B}}_N+\mathbf{q}_N\mathbf{h}_N^\tau+\mathbf{h}_N\mathbf{q}_N^\tau+d_N\mathbf{h}_N\mathbf{h}_N^\tau-z\mathbf{I}_T\Big)^{-1}\\
&=tr\Big(\underline{\mathbf{B}}_N-z\mathbf{I}_T+\mathbf{q}_N\mathbf{h}_N^\tau+\mathbf{h}_N\mathbf{q}_N^\tau\Big)^{-1}-\omega_3\\
&=tr(\underline{\mathbf{B}}_N-z\mathbf{I}_T+\mathbf{q}_N\mathbf{h}_N^\tau)^{-1}-\omega_2-\omega_3\\
&=tr(\underline{\mathbf{B}}_N-z\mathbf{I}_T)^{-1}-\omega_1-\omega_2-\omega_3,
\end{aligned}$$

where

$$\begin{aligned}
&\omega_1=\frac{\mathbf{h}_N^\tau(\underline{\mathbf{B}}_N-z\mathbf{I}_T)^{-2}\mathbf{q}_N}{1+\mathbf{h}_N^\tau(\underline{\mathbf{B}}_N-z\mathbf{I}_T)^{-1}\mathbf{q}_N},\quad\omega_2=\frac{\mathbf{q}_N^\tau(\underline{\mathbf{B}}_N-z\mathbf{I}_T+\mathbf{q}_N\mathbf{h}_N^\tau)^{-2}\mathbf{h}_N}{1+\mathbf{q}_N^\tau(\underline{\mathbf{B}}_N-z\mathbf{I}_T+\mathbf{q}_N\mathbf{h}_N^\tau)^{-1}\mathbf{h}_N},\\
&\omega_3=\frac{d_N\mathbf{h}_N^\tau(\underline{\mathbf{B}}_N-z\mathbf{I}_T+\mathbf{q}_N\mathbf{h}_N^\tau+\mathbf{h}_N\mathbf{q}_N^\tau)^{-2}\mathbf{h}_N}{1+d_N\mathbf{h}_N^\tau(\underline{\mathbf{B}}_N-z\mathbf{I}_T+\mathbf{q}_N\mathbf{h}_N^\tau+\mathbf{h}_N\mathbf{q}_N^\tau)^{-1}\mathbf{h}_N}.
\end{aligned} \tag{A.59}$$

We below investigate the terms $tr\Big(\underline{\mathbf{B}}_N-z\mathbf{I}_T\Big)^{-1}$, ω_1,ω_2 and ω_3 one by one. For simplicity we only consider $z=w+iv_0$ with $v_0>0$ as in Theorem 1.

First, we establish the central limit theorem of $tr\Big(\underline{\mathbf{B}}_N-z\mathbf{I}_T\Big)^{-1}$. Let $\underline{\mathbf{A}}_N=\sum_{j=1}^N\frac{\widetilde{\boldsymbol{\varepsilon}}_j\widetilde{\boldsymbol{\varepsilon}}_j^\tau}{||\widetilde{\boldsymbol{\varepsilon}}_j||^2}$. Write

$$\underline{\mathbf{B}}_N-\underline{\mathbf{A}}_N=-\Big[\sum_{j=1}^N\frac{\widetilde{\boldsymbol{\varepsilon}}_j\widetilde{\boldsymbol{\varepsilon}}_j^\tau}{||\widetilde{\boldsymbol{\varepsilon}}_j||^2}\cdot\frac{\frac{2}{\sqrt{T}}\widetilde{\mathbf{v}}^\tau\widetilde{\boldsymbol{\varepsilon}}_j}{||\widetilde{\mathbf{u}}_j||^2}+\sum_{j=1}^N\frac{\widetilde{\boldsymbol{\varepsilon}}_j\widetilde{\boldsymbol{\varepsilon}}_j^\tau}{||\widetilde{\boldsymbol{\varepsilon}}_j||^2}\cdot\frac{\frac{1}{T}\widetilde{\mathbf{v}}^\tau\widetilde{\mathbf{v}}}{||\widetilde{\mathbf{u}}_j||^2}\Big]. \tag{A.60}$$

This, together with (A.15), yields,

$$\begin{aligned}
&tr\Big(\underline{\mathbf{A}}_N-z\mathbf{I}_T\Big)^{-1}-tr\Big(\underline{\mathbf{B}}_N-z\mathbf{I}_T\Big)^{-1}\\
=\;&tr\Big(\underline{\mathbf{A}}_N-z\mathbf{I}_T\Big)^{-1}\Big(\underline{\mathbf{B}}_N-\underline{\mathbf{A}}_N\Big)\Big(\underline{\mathbf{B}}_N-z\mathbf{I}_T\Big)^{-1}\\
=\;&:\omega_4+\omega_5,
\end{aligned} \tag{A.61}$$

where

$$\begin{aligned}
\omega_4&=-\sum_{j=1}^N\frac{\frac{2}{\sqrt{T}}\widetilde{\mathbf{v}}^\tau\widetilde{\boldsymbol{\varepsilon}}_j}{||\widetilde{\mathbf{u}}_j||^2}\frac{1}{||\widetilde{\boldsymbol{\varepsilon}}_j||^2}\widetilde{\boldsymbol{\varepsilon}}_j^\tau\Big(\underline{\mathbf{B}}_N-z\mathbf{I}_T\Big)^{-1}\Big(\underline{\mathbf{A}}_N-z\mathbf{I}_T\Big)^{-1}\widetilde{\boldsymbol{\varepsilon}}_j,\\
\omega_5&=-\sum_{j=1}^N\frac{\frac{1}{T}\widetilde{\mathbf{v}}^\tau\widetilde{\mathbf{v}}}{||\widetilde{\mathbf{u}}_j||^2}\frac{1}{||\widetilde{\boldsymbol{\varepsilon}}_j||^2}\widetilde{\boldsymbol{\varepsilon}}_j^\tau\Big(\underline{\mathbf{B}}_N-z\mathbf{I}_T\Big)^{-1}\Big(\underline{\mathbf{A}}_N-z\mathbf{I}_T\Big)^{-1}\widetilde{\boldsymbol{\varepsilon}}_j.
\end{aligned}$$

Before studying ω_4 and ω_5 we specify some notation.

$$\underline{\mathbf{A}}_N^{(j)} = \underline{\mathbf{A}}_N - \frac{\widetilde{\varepsilon}_j \widetilde{\varepsilon}_j^\tau}{||\widetilde{\varepsilon}_j||^2}, \quad \zeta_j = \frac{\widetilde{\varepsilon}_j^\tau}{||\widetilde{\varepsilon}_j||}\Big(\underline{\mathbf{A}}_N^{(j)} - z\mathbf{I}_T\Big)^{-1}\frac{\widetilde{\varepsilon}_j}{||\widetilde{\varepsilon}_j||}, \quad \zeta = \frac{1}{T}tr\Big(\underline{\mathbf{A}}_N - z\mathbf{I}_T\Big)^{-1},$$
$$\underline{\mathbf{B}}_N^{(j)} = \underline{\mathbf{B}}_N - \frac{\widetilde{\varepsilon}_j \widetilde{\varepsilon}_j^\tau}{||\widetilde{\mathbf{u}}_j||^2}, \quad \alpha_j = \frac{\widetilde{\varepsilon}_j^\tau}{||\widetilde{\mathbf{u}}_j||}\Big(\underline{\mathbf{B}}_N^{(j)} - z\mathbf{I}_T\Big)^{-1}\frac{\widetilde{\varepsilon}_j}{||\widetilde{\mathbf{u}}_j||}, \quad \alpha = \frac{1}{T}tr\left(\underline{\mathbf{B}}_N - z\mathbf{I}_T\right)^{-1},$$
$$\gamma_j = \frac{\widetilde{\varepsilon}_j^\tau}{||\widetilde{\varepsilon}_j||}\Big(\underline{\mathbf{B}}_N^{(j)} - z\mathbf{I}_T\Big)^{-1}\Big(\underline{\mathbf{A}}_N^{(j)} - z\mathbf{I}_T\Big)^{-1}\frac{\widetilde{\varepsilon}_j}{||\widetilde{\varepsilon}_j||}, \quad \gamma = \frac{1}{T}tr\Big(\underline{\mathbf{B}}_N - z\mathbf{I}_T\Big)^{-1}\Big(\underline{\mathbf{A}}_N - z\mathbf{I}_T\Big)^{-1}.$$

Moreover, for simplifying notation denote $\Big(\underline{\mathbf{A}}_N - z\mathbf{I}_T\Big)^{-1}, \Big(\underline{\mathbf{A}}_N^{(j)} - z\mathbf{I}_T\Big)^{-1}, \Big(\underline{\mathbf{B}}_N - z\mathbf{I}_T\Big)^{-1}, \Big(\underline{\mathbf{B}}_N^{(j)} - z\mathbf{I}_T\Big)^{-1}$ by $\underline{\mathbf{A}}_N^{-1}(z), \underline{\mathbf{A}}_{Nj}^{-1}(z), \underline{\mathbf{B}}_N^{-1}(z), \underline{\mathbf{B}}_{Nj}^{-1}(z)$ respectively.

From Lemma 5 of Gao, el. (2014) and (A.16) we have, for $j = 1, 2, \ldots, N$,

$$\mathbb{E}|\gamma_j - \gamma|^2 = O\left(\frac{1}{T}\right), \quad \mathbb{E}|\zeta_j - \zeta|^2 = O\left(\frac{1}{T}\right), \quad \mathbb{E}|\alpha_j - \alpha|^2 = \left(\frac{1}{T^2}\right). \tag{A.62}$$

The last estimate in (A.62) follows from Lemma 5 of Gao, el. (2014), (A.16) and the fact that

$$\mathbb{E}\Big|\alpha_j - \frac{\widetilde{\varepsilon}_j^\tau}{||\widetilde{\varepsilon}_j||}\Big(\underline{\mathbf{B}}_N^{(j)} - z\mathbf{I}_T\Big)^{-1}\frac{\widetilde{\varepsilon}_j}{||\widetilde{\varepsilon}_j||}\Big|^2 = \mathbb{E}\Big|\Big(\frac{\|\widetilde{\varepsilon}_j\|^2}{||\widetilde{\mathbf{u}}_j||^2} - 1\Big)\frac{\widetilde{\varepsilon}_j^\tau}{||\widetilde{\varepsilon}_j||}\underline{\mathbf{B}}_{Nj}^{-1}(z)\frac{\widetilde{\varepsilon}_j}{||\widetilde{\varepsilon}_j||}\Big|^2$$
$$\leq \frac{1}{v_0^2}\mathbb{E}\Big(\frac{\|\widetilde{\varepsilon}_j\|^2}{||\widetilde{\mathbf{u}}_j||^2} - 1\Big)^2 = O(\frac{1}{T}),$$

which can be proved as in Lemma 5 of Gao, el. (2014) by introducing an event to control the denominator $||\widetilde{\mathbf{u}}_j||$.

Note that

$$|\gamma_j| \leq 1/v_0, \quad |1/(1+\zeta_j)| \leq |z|/v_0, \quad |1/(1+\alpha_j)| \leq |z|/v_0. \tag{A.63}$$

In view of (A.17), (A.57), (A.62), (A.63) and Lemma 5 we have

$$\begin{aligned} |\omega_4| &= \Big|\sum_{j=1}^N \frac{\frac{2}{\sqrt{T}}\widetilde{\mathbf{v}}^\tau \widetilde{\varepsilon}_j}{T}\frac{T}{||\widetilde{\mathbf{u}}_j||^2}\frac{\gamma_j}{(1+\zeta_j)(1+\alpha_j)}\Big| \\ &= \Big|\frac{1}{T}\sum_{j=1}^N \frac{2}{\sqrt{T}}\widetilde{\mathbf{v}}^\tau \widetilde{\varepsilon}_j \frac{\gamma}{(1+\zeta)(1+\alpha)}\Big| + O_P(\frac{1}{\sqrt{T}}) = O_p(\frac{1}{\sqrt{T}}) \end{aligned} \tag{A.64}$$

where we also use the fact that

$$\begin{aligned} &\mathbb{E}\Big|\frac{1}{T}\sum_{j=1}^N \frac{1}{\sqrt{T}}\widetilde{\mathbf{v}}^\tau \widetilde{\varepsilon}_j\Big|^2 \leq \mathbb{E}\Big|\frac{1}{T}\sum_{j=1}^N \frac{1}{\sqrt{T}}\widetilde{\mathbf{v}}^\tau \widetilde{\varepsilon}_j\Big|^2 \\ &\leq \frac{1}{T^3}\sum_{j_1,j_2=1}^N \sum_{k_1,k_2=1}^T \mathbb{E}\big(\widetilde{v}_{k_1}\widetilde{\varepsilon}_{j_1k_1}\widetilde{\varepsilon}_{j_2k_2}\widetilde{v}_{k_2}\big) \leq \frac{M}{T}. \end{aligned} \tag{A.65}$$

Likewise, by the formula (A.17), ω_5 can be written as

$$\omega_5 = -\sum_{j=1}^N \frac{\frac{1}{T}\widetilde{\mathbf{v}}^\tau \widetilde{\mathbf{v}}}{||\widetilde{\mathbf{u}}_j||^2}\frac{\gamma_j}{(1+\alpha_j)(1+\zeta_j)} = -\frac{N}{T}\frac{\gamma}{(1+\zeta)(1+\alpha)} + O_p(\frac{1}{\sqrt{T}}).$$

Observe from (A.60), (A.15), (A.65) and (A.57) that

$$\begin{aligned} &\Big|\frac{1}{T}tr\underline{\mathbf{B}}_N^{-1}(z)\underline{\mathbf{A}}_N^{-1}(z) - \frac{1}{T}tr\underline{\mathbf{A}}_N^{-2}(z)\Big| \\ = &\Big|\frac{1}{T}\sum_{j=1}^N \frac{\frac{2}{\sqrt{T}}\widetilde{\mathbf{v}}^\tau \widetilde{\varepsilon}_j + \frac{1}{T}\widetilde{\mathbf{v}}^\tau \widetilde{\mathbf{v}}}{||\widetilde{\mathbf{u}}_j||^2}\frac{\widetilde{\varepsilon}_j}{||\widetilde{\varepsilon}_j||}\underline{\mathbf{A}}_N^{-2}(z)\underline{\mathbf{B}}_N^{-1}(z)\frac{\widetilde{\varepsilon}_j}{||\widetilde{\varepsilon}_j||}\Big| = O_P\left(\frac{1}{T}\right) \end{aligned} \tag{A.66}$$

and from (9) that

$$\frac{1}{T}tr\underline{\mathbf{A}}_N^{-2}(z) \xrightarrow{i.p.} \underline{m}^{(2)}(z), \tag{A.67}$$

where $\underline{m}^{(2)}(z)$ is defined in (3.3) of the main paper. These imply

$$\gamma \xrightarrow{i.p.} \underline{m}^{(2)}(z). \tag{A.68}$$

One can similarly prove that

$$\alpha \xrightarrow{i.p.} \underline{m}(z). \tag{A.69}$$

It follows that

$$\omega_5 \xrightarrow{i.p.} -\frac{c\underline{m}^{(2)}(z)}{(1+\underline{m}(z))^2}. \tag{A.70}$$

From (A.61), (A.64), (A.70) and Theorem 1 of Gao, el. (2014) which provides the CLT of $tr\underline{\mathbf{A}}_N^{-1}(z)$, it follows that

$$tr\underline{\mathbf{B}}_N^{-1}(z) - T\underline{m}_{c_N}(z) \xrightarrow{i.d.} \mathcal{N}\left(\mu + \frac{c\underline{m}^{(2)}(z)}{(1+\underline{m}(z))^2}, \sigma^2\right),$$

where μ and σ^2 are the asymptotic mean and variance that are derived in Theorem 1 of Gao, el. (2014).

We next determine the limits of ω_1, ω_2 and ω_3. As for the numerator of ω_1 defined in (A.59), note that the following relation: $\frac{\mathbf{h}_N^\tau}{\|\mathbf{h}_N^\tau\|}\underline{\mathbf{B}}_N^{-2}(z)\frac{\mathbf{q}_N}{\|\mathbf{q}_N\|} = \left(\frac{\mathbf{h}_N^\tau}{\|\mathbf{h}_N^\tau\|}\underline{\mathbf{B}}_N^{-1}(z)\frac{\mathbf{q}_N}{\|\mathbf{q}_N\|}\right)'$, where $(\cdot)'$ denotes the first derivative of $(\cdot)$ in the bracket with respect to z. In view of Lemma 9 it is sufficient to consider the limit of $\frac{\mathbf{h}_N^\tau}{\|\mathbf{h}_N^\tau\|}\left(\underline{\mathbf{B}}_N - z\mathbf{I}_T\right)^{-1}\frac{\mathbf{q}_N}{\|\mathbf{q}_N\|}$ in probability. From the formula (A.15), (A.17) and (A.60), it follows that

$$\frac{\mathbf{h}_N^\tau}{\|\mathbf{h}_N^\tau\|}\underline{\mathbf{B}}_N^{-1}(z)\frac{\mathbf{q}_N}{\|\mathbf{q}_N\|} = \frac{\mathbf{h}_N^\tau}{\|\mathbf{h}_N^\tau\|}\underline{\mathbf{A}}_N^{-1}(z)\frac{\mathbf{q}_N}{\|\mathbf{q}_N\|} + \sum_{j=1}^N \frac{\mathbf{h}_N^\tau}{\|\mathbf{h}_N^\tau\|}\underline{\mathbf{B}}_{Nj}^{-1}(z)\frac{\widetilde{\boldsymbol{\varepsilon}}_j}{T^{3/2}}Q_j, \tag{A.71}$$

where $Q_j = \frac{T^{3/2}\frac{\widetilde{\boldsymbol{\varepsilon}}_j^\tau}{\|\widetilde{\boldsymbol{\varepsilon}}_j\|}\underline{\mathbf{A}}_N^{-1}(z)\frac{\mathbf{q}_N}{\|\mathbf{q}_N\|}}{1+\alpha_j} \cdot \frac{\frac{2}{\sqrt{T}}\widetilde{\mathbf{v}}^\tau\widetilde{\boldsymbol{\varepsilon}}_j + \frac{1}{T}\widetilde{\mathbf{v}}^\tau\widetilde{\mathbf{v}}}{\|\widetilde{\boldsymbol{\varepsilon}}_j\|\cdot\|\widetilde{\mathbf{u}}_j\|^2}$. By (A.63), $|Q_j| \le MT^{3/2}\left|\frac{\frac{2}{\sqrt{T}}\widetilde{\mathbf{v}}^\tau\widetilde{\boldsymbol{\varepsilon}}_j + \frac{1}{T}\widetilde{\mathbf{v}}^\tau\widetilde{\mathbf{v}}}{\|\widetilde{\boldsymbol{\varepsilon}}_j\|\cdot\|\widetilde{\mathbf{u}}_j\|^2}\right|$ and also by the mutual independence between $\widetilde{\boldsymbol{\varepsilon}}_j$ and $\mathbf{h}_N^\tau\underline{\mathbf{B}}_{Nj}^{-1}(z)$, we have

$$\mathbb{E}\left|\mathbf{h}_N^\tau\underline{\mathbf{B}}_{Nj}^{-1}(z)\frac{\widetilde{\boldsymbol{\varepsilon}}_j}{T^{3/2}}\right|^2 = \frac{1}{T^3}\mathbb{E}\left(\mathbf{h}_N^\tau\underline{\mathbf{B}}_{Nj}^{-1}(z)\boldsymbol{\Sigma}_x\underline{\mathbf{B}}_{Nj}^{-1}(\bar{z})\mathbf{h}_N\right) = O\left(\frac{1}{T^3}\right), \tag{A.72}$$

where $\boldsymbol{\Sigma}_x$ is defined below (A.34) and $\underline{\mathbf{B}}_{Nj}^{-1}(\bar{z})$ denotes the complex transpose of $\underline{\mathbf{B}}_{Nj}^{-1}(z)$ (we below drop $\boldsymbol{\Sigma}_x$ whenever coming across similar calculations). It is straightforward to verify that

$$E\left|\frac{2}{\sqrt{T}}\widetilde{\mathbf{v}}^\tau\widetilde{\boldsymbol{\varepsilon}}_j + \frac{1}{T}\widetilde{\mathbf{v}}^\tau\widetilde{\mathbf{v}}\right|^2 = O(1). \tag{A.73}$$

These estimates, together with (A.57) and (A.58), imply

$$\sum_{j=1}^N \frac{\mathbf{h}_N^\tau}{\|\mathbf{h}_N^\tau\|}\underline{\mathbf{B}}_{Nj}^{-1}(z)\frac{\widetilde{\boldsymbol{\varepsilon}}_j}{T^{3/2}}Q_j = O_P(\frac{1}{\sqrt{T}}). \tag{A.74}$$

For the first term of (A.71), we use (A.17), (A.57), (A.72) and (A.62) to obtain

$$\begin{aligned}\mathbf{h}_N^\tau\underline{\mathbf{A}}_N^{-1}(z)\mathbf{q}_N &= \sum_{j=1}^N \mathbf{h}_N^\tau\underline{\mathbf{A}}_{Nj}^{-1}(z)\frac{\widetilde{\boldsymbol{\varepsilon}}_j}{T}\cdot\frac{T}{\|\widetilde{\mathbf{u}}_j\|^2}\cdot\frac{1}{1+\zeta_j} \\ &= \sum_{j=1}^N \mathbf{h}_N^\tau\underline{\mathbf{A}}_{Nj}^{-1}(z)\frac{\widetilde{\boldsymbol{\varepsilon}}_j}{T}\cdot\frac{1}{1+\zeta} + O_p(\frac{1}{\sqrt{T}}).\end{aligned} \tag{A.75}$$

We next consider the first term on the right hand of the second equality of (A.75). Let

$$a_j = \frac{\widetilde{\boldsymbol{\varepsilon}}_j^\tau}{\sqrt{T}} \underline{\mathbf{A}}_{Nj}^{-1}(z) \frac{\widetilde{\boldsymbol{\varepsilon}}_j}{\sqrt{T}}, \quad \zeta^{(2)} = \frac{1}{T} tr \underline{\mathbf{A}}_{Nj}^{-2}(z),$$
$$a_{j_1 j_2}^{(i)(k)} = \frac{\widetilde{\boldsymbol{\varepsilon}}_{j_i}^\tau}{\sqrt{T}} \underline{\mathbf{A}}_{Nj_1j_2}^{-k}(z) \frac{\widetilde{\boldsymbol{\varepsilon}}_{j_i}}{\sqrt{T}}, \quad a_{j_1 j_2}^{(1,2)(k)} = \frac{\widetilde{\boldsymbol{\varepsilon}}_{j_1}^\tau}{\sqrt{T}} \underline{\mathbf{A}}_{Nj_1j_2}^{-k}(z) \frac{\widetilde{\boldsymbol{\varepsilon}}_{j_2}}{\sqrt{T}},$$

with $j, j_1, j_2 = 1, 2, \ldots, N$; $i = 1, 2$, $k = 1, 2$ and $\underline{\mathbf{A}}_{Nj_1j_2}^{-k}(z) = \left(\underline{\mathbf{A}}_N - z\mathbf{I}_T - \frac{\widetilde{\boldsymbol{\varepsilon}}_{j_1}\widetilde{\boldsymbol{\varepsilon}}_{j_1}^\tau}{||\widetilde{\boldsymbol{\varepsilon}}_{j_1}||^2} - \frac{\widetilde{\boldsymbol{\varepsilon}}_{j_2}\widetilde{\boldsymbol{\varepsilon}}_{j_2}^\tau}{||\widetilde{\boldsymbol{\varepsilon}}_{j_2}||^2}\right)^{-k}$. From the formula (A.17), it follows that

$$\begin{aligned}
&\mathbb{E}\left|\sum_{j=1}^N \mathbf{h}_N^\tau \underline{\mathbf{A}}_{Nj}^{-1}(z) \frac{\widetilde{\boldsymbol{\varepsilon}}_j}{T}\right|^2 = \frac{1}{T^2}\mathbb{E}\left(\sum_{j_1,j_2=1}^N \frac{\widetilde{\boldsymbol{\varepsilon}}_{j_1}^\tau}{\sqrt{T}} \underline{\mathbf{A}}_{Nj_1}^{-1}(z) \underline{\mathbf{A}}_{Nj_2}^{-1}(\bar{z}) \frac{\widetilde{\boldsymbol{\varepsilon}}_{j_2}}{\sqrt{T}}\right) = \\
&\frac{1}{T^2}\mathbb{E}\left(\sum_{j=1}^N \frac{\widetilde{\boldsymbol{\varepsilon}}_j^\tau}{\sqrt{T}} \underline{\mathbf{A}}_{Nj_1}^{-2}(z) \frac{\widetilde{\boldsymbol{\varepsilon}}_j}{\sqrt{T}}\right) + \frac{1}{T^2}\mathbb{E}\Big[\sum_{j_1 \neq j_2}^N \frac{\widetilde{\boldsymbol{\varepsilon}}_{j_1}^\tau}{\sqrt{T}} \left(\underline{\mathbf{A}}_{Nj_1j_2}^{-1}(z) - \frac{\underline{\mathbf{A}}_{Nj_1j_2}^{-1}(z)\frac{\widetilde{\boldsymbol{\varepsilon}}_{j_2}}{||\widetilde{\boldsymbol{\varepsilon}}_{j_2}||} \cdot \frac{\widetilde{\boldsymbol{\varepsilon}}_{j_2}^\tau}{||\widetilde{\boldsymbol{\varepsilon}}_{j_2}||}\underline{\mathbf{A}}_{Nj_1j_2}^{-1}(z)}{1 + \frac{\widetilde{\boldsymbol{\varepsilon}}_{j_2}^\tau}{||\widetilde{\boldsymbol{\varepsilon}}_{j_2}||}\underline{\mathbf{A}}_{Nj_1j_2}^{-1}(z)\frac{\widetilde{\boldsymbol{\varepsilon}}_{j_2}^\tau}{||\widetilde{\boldsymbol{\varepsilon}}_{j_2}||}}\right) \\
&\qquad \cdot \left(\underline{\mathbf{A}}_{Nj_1j_2}^{-1}(z) - \frac{\underline{\mathbf{A}}_{Nj_1j_2}^{-1}(z)\frac{\widetilde{\boldsymbol{\varepsilon}}_{j_1}}{||\widetilde{\boldsymbol{\varepsilon}}_{j_1}||} \cdot \frac{\widetilde{\boldsymbol{\varepsilon}}_{j_1}^\tau}{||\widetilde{\boldsymbol{\varepsilon}}_{j_1}||}\underline{\mathbf{A}}_{Nj_1j_2}^{-1}(z)}{1 + \frac{\widetilde{\boldsymbol{\varepsilon}}_{j_1}^\tau}{||\widetilde{\boldsymbol{\varepsilon}}_{j_1}||}\underline{\mathbf{A}}_{Nj_1j_2}^{-1}(z)\frac{\widetilde{\boldsymbol{\varepsilon}}_{j_1}^\tau}{||\widetilde{\boldsymbol{\varepsilon}}_{j_1}||}}\right) \frac{\widetilde{\boldsymbol{\varepsilon}}_{j_2}}{\sqrt{T}}\Big] \\
= \quad & \frac{1}{T^2}\sum_{h=1}^5 \mathbb{E}\left(C_h\right), \qquad \text{(A.76)}
\end{aligned}$$

where

$$C_1 = \sum_{j=1}^N a_j, \quad C_2 = \sum_{j_1 \neq j_2}^N a_{j_1j_2}^{(1,2)(2)}, \quad C_3 = -\sum_{j_1 \neq j_2}^N a_{j_1j_2}^{(1)(2)} a_{j_1j_2}^{(1,2)(1)} \frac{||\widetilde{\boldsymbol{\varepsilon}}_{j_1}||^2}{T\left(1 + a_{j_1j_2}^{(1)(1)}\frac{T}{||\widetilde{\boldsymbol{\varepsilon}}_{j_1}||^2}\right)},$$
$$C_4 = -\sum_{j_1 \neq j_2}^N \frac{a_{j_1j_2}^{(1,2)(1)} a_{j_1j_2}^{(2)(2)} ||\widetilde{\boldsymbol{\varepsilon}}_{j_2}||^2}{T\left(1 + a_{j_1j_2}^{(2)(1)}\frac{T}{||\widetilde{\boldsymbol{\varepsilon}}_{j_2}||^2}\right)}, \quad C_5 = \sum_{j_1 \neq j_2}^N \frac{\left(a_{j_1j_2}^{(1,2)(1)}\right)^2 a_{j_1j_2}^{(1,2)(2)} ||\widetilde{\boldsymbol{\varepsilon}}_{j_1}||^2 \cdot ||\widetilde{\boldsymbol{\varepsilon}}_{j_2}||^2}{T^2\left(1 + a_{j_1j_2}^{(1)(1)}\frac{T}{||\widetilde{\boldsymbol{\varepsilon}}_{j_1}||^2}\right)\left(1 + a_{j_1j_2}^{(2)(1)}\frac{T}{||\widetilde{\boldsymbol{\varepsilon}}_{j_2}||^2}\right)}.$$

Here we also write $\underline{\mathbf{A}}_{Nj_2}^{-1}(\bar{z})$ as $\underline{\mathbf{A}}_{Nj_2}^{-1}(z)$ in order to simplifying notation. It is easy to verify that $\frac{1}{T^2}\mathbb{E}C_1 = O\left(\frac{1}{T}\right)$ and that by Lemma (6), $E|a_{j_1j_2}^{(1,2)(k)}|^2 = O\left(\frac{1}{T}\right)$.

The above estimates, together with (A.57 and (A.58), ensure $\mathbb{E}\left|\sum_{j=1}^N \mathbf{h}_N^\tau \underline{\mathbf{A}}_{Nj}^{-1}(z)\frac{\widetilde{\boldsymbol{\varepsilon}}_j}{T}\right|^2 = O\left(\frac{1}{T}\right)$, which implies, together with (A.75), that

$$\mathbf{h}_N^\tau \underline{\mathbf{A}}_N^{-1}(z)\mathbf{q}_N = O_P(\frac{1}{\sqrt{T}}). \qquad \text{(A.77)}$$

Moreover, let $\widetilde{\mathbf{q}}_N = \sum_{j=1}^N \frac{\widetilde{\boldsymbol{\varepsilon}}_j}{||\widetilde{\boldsymbol{\varepsilon}}_j||^2}$ and $\bar{\mathbf{q}}_N = \sum_{j=1}^N \frac{\widetilde{\boldsymbol{\varepsilon}}_j}{T}$.

It follows from (A.57) and (A.73) that

$$||\mathbf{q}_N - \widetilde{\mathbf{q}}_N|| = \left|\left|\sum_{j=1}^N \widetilde{\boldsymbol{\varepsilon}}_j \frac{\frac{2}{\sqrt{T}}\widetilde{\mathbf{v}}^\tau\widetilde{\boldsymbol{\varepsilon}}_j + \frac{1}{T}\widetilde{\mathbf{v}}^\tau\widetilde{\mathbf{v}}}{||\widetilde{\mathbf{u}}_j||^2 \cdot ||\widetilde{\boldsymbol{\varepsilon}}_j||^2}\right|\right| = O_P\left(\frac{1}{\sqrt{T}}\right).$$

As in (A.35), by (A.58)

$$\|\widetilde{\mathbf{q}}_N - \bar{\mathbf{q}}_N\| = \|\sum_{j=1}^N \frac{\widetilde{\boldsymbol{\varepsilon}}_j(T - ||\widetilde{\boldsymbol{\varepsilon}}_j||^2)}{T||\widetilde{\boldsymbol{\varepsilon}}_j||^2}\| = \|\boldsymbol{\mathcal{C}}\mathbf{D}_1\frac{\mathbf{e}}{\sqrt{T}}\| \le \|\boldsymbol{\mathcal{C}}\| \max_j \frac{\sqrt{T}(1 - \frac{\|\widetilde{\boldsymbol{\varepsilon}}_j\|^2}{T})}{\|\widetilde{\boldsymbol{\varepsilon}}_j\|} = o_p(1), \qquad \text{(A.78)}$$

where $\boldsymbol{\mathcal{C}} = \frac{1}{\|\widetilde{\boldsymbol{\varepsilon}}_j\|}(\widetilde{\boldsymbol{\varepsilon}}_1, \widetilde{\boldsymbol{\varepsilon}}_2, \cdots, \widetilde{\boldsymbol{\varepsilon}}_N)$ and $\mathbf{D}_1 = diag(\frac{\sqrt{T}(1 - \frac{\|\widetilde{\boldsymbol{\varepsilon}}_1\|^2}{T})}{\|\widetilde{\boldsymbol{\varepsilon}}_1\|}, \cdots, \frac{\sqrt{T}(1 - \frac{\|\widetilde{\boldsymbol{\varepsilon}}_N\|^2}{T})}{\|\widetilde{\boldsymbol{\varepsilon}}_N\|})$.

Therefore, $|\bar{\mathbf{q}}_N\|^2 \xrightarrow{i.p.} c$ by Lemma 1 of Pan and Zhou (2011). This, together with (A.78) and (A.78), implies that

$$\|\mathbf{q}_N\|^2 \xrightarrow{i.p.} c. \tag{A.79}$$

Also, it is easily seen that

$$\|\mathbf{h}_N\|^2 \xrightarrow{i.p.} 1. \tag{A.80}$$

It follows from (A.74), (A.77), (A.79), (A.80) and (A.71) that

$$\frac{\mathbf{h}_N^\tau}{\|\mathbf{h}_N^\tau\|}\underline{\mathbf{B}}_N^{-1}(z)\frac{\mathbf{q}_N}{\|\mathbf{q}_N\|} = o_P(1), \quad \omega_1 = o_P(1). \tag{A.81}$$

Next, we consider ω_2 defined in (A.59). Note that

$$\mathbf{q}_N^\tau\left(\underline{\mathbf{B}}_N - z\mathbf{I}_T + \mathbf{q}_N\mathbf{h}_N^\tau\right)^{-2}\mathbf{h}_N = \left[\mathbf{q}_N^\tau\left(\underline{\mathbf{B}}_N - z\mathbf{I}_T + \mathbf{q}_N\mathbf{h}_N^\tau\right)^{-1}\mathbf{h}_N\right]'. \tag{A.82}$$

As before, it is sufficient to consider $\mathbf{q}_N^\tau\left(\underline{\mathbf{B}}_N - z\mathbf{I}_T + \mathbf{q}_N\mathbf{h}_N^\tau\right)^{-1}\mathbf{h}_N$. Via (A.16) write

$$\begin{aligned} & 1+\mathbf{q}_N^\tau(\underline{\mathbf{B}}_N - z\mathbf{I}_T + \mathbf{q}_N\mathbf{h}_N^\tau)^{-1}\mathbf{h}_N \\ = \quad & 1+\mathbf{q}_N^\tau\underline{\mathbf{B}}_N^{-1}(z)\mathbf{h}_N - \frac{\mathbf{q}_N^\tau\underline{\mathbf{B}}_N^{-1}(z)\mathbf{q}_N\mathbf{h}_N^\tau\underline{\mathbf{B}}_N^{-1}(z)\mathbf{h}_N}{1+\mathbf{h}_N^\tau\underline{\mathbf{B}}_N^{-1}(z)\mathbf{q}_N}. \end{aligned} \tag{A.83}$$

In view of (A.79)-(A.81), it is sufficient to find the following terms: $\mathbf{q}_N^\tau\left(\underline{\mathbf{B}}_N - z\mathbf{I}_T\right)^{-1}\mathbf{q}_N$, $\mathbf{h}_N^\tau\left(\underline{\mathbf{B}}_N - z\mathbf{I}_T\right)^{-1}\mathbf{q}_N$ and $\mathbf{h}_N^\tau\left(\underline{\mathbf{B}}_N - z\mathbf{I}_T\right)^{-1}\mathbf{h}_N$.

We conclude from Lemma 6 that

$$\mathbf{h}_N^\tau\left(\underline{\mathbf{B}} - z\mathbf{I}_T\right)^{-1}\mathbf{h}_N \xrightarrow{i.p.} \underline{m}(z), \qquad as \quad T, N \to \infty. \tag{A.84}$$

We next consider $\mathbf{q}_N^\tau\left(\underline{\mathbf{B}}_N - z\mathbf{I}_T\right)^{-1}\mathbf{q}_N$. As in (A.35), by (A.57), (A.79) and (A.60) one can prove that

$$\left|\mathbf{q}_N^\tau\underline{\mathbf{B}}_N^{-1}(z)\mathbf{q}_N - \mathbf{q}_N^\tau\underline{\mathbf{A}}_N^{-1}(z)\mathbf{q}_N\right| = O_P\left(\frac{1}{T}\right). \tag{A.85}$$

Also, by an argument similar to (A.35), via (A.58) and (A.79) we have

$$\left|\mathbf{q}_N^\tau\underline{\mathbf{A}}_N^{-1}(z)\mathbf{q}_N - \mathbf{q}_N^\tau\mathbf{S}_{1N}^{-1}(z)\mathbf{q}_N\right| = o_P\left(1\right), \tag{A.86}$$

where $\mathbf{S}_{1N}^{-1}(z) = \left(\frac{1}{T}\sum_{j=1}^{N}\widetilde{\boldsymbol{\varepsilon}}_j\widetilde{\boldsymbol{\varepsilon}}_j^\tau - z\mathbf{I}\right)^{-1}$.

Moreover from (A.78) and (A.58)

$$\begin{aligned} & \left|\left|\mathbf{q}_N^\tau\mathbf{S}_{1N}^{-1}(z)\mathbf{q}_N - \bar{\mathbf{q}}_N^\tau\mathbf{S}_{1N}^{-1}(z)\bar{\mathbf{q}}_N\right|\right| \\ \le \quad & ||\mathbf{q}_N - \bar{\mathbf{q}}_N|| \cdot \left|\left|\mathbf{S}_{1N}^{-1}(z)\right|\right| \cdot ||\mathbf{q}_N|| + ||\widetilde{\mathbf{q}}_N|| \cdot \left|\left|\mathbf{S}_{1N}^{-1}(z)\right|\right| \cdot ||\mathbf{q}_N - \bar{\mathbf{q}}_N|| = O_P\left(\frac{1}{\sqrt{T}}\right). \end{aligned}$$

From the result in Pan and Zhou (2011)

$$\bar{\mathbf{q}}_N^\tau\mathbf{S}_{1N}^{-1}(z)\bar{\mathbf{q}}_N \xrightarrow{i.p.} \frac{c\underline{m}(z)}{1+\underline{m}(z)}, \quad as \quad T, N \to \infty \tag{A.87}$$

(one can see $M_n(z)$ below (1.13) of Pan and Zhou (2011) or one may prove (A.87) by rewriting it as a martingale as in Sections 3 and 4 in Pan and Zhou (2011)).

It follows that

$$\mathbf{q}_N^\tau\left(\underline{\mathbf{B}}_N - z\mathbf{I}_T\right)^{-1}\mathbf{q}_N \xrightarrow{i.p.} \frac{c\underline{m}(z)}{1+\underline{m}(z)}. \tag{A.88}$$

We then conclude from (A.82), (A.79), (A.80), (A.81), (A.83), (A.84), (A.88) that the limit of the numerator of ω_2 is

$$\begin{aligned} & \mathbf{q}_N^\tau \left(\underline{\mathbf{B}}_N - z\mathbf{I}_T + \mathbf{q}_N\mathbf{h}_N^\tau\right)^{-2}\mathbf{h}_N \\ = & \left[\mathbf{q}_N^\tau \left(\underline{\mathbf{B}}_N - z\mathbf{I}_T + \mathbf{q}_N\mathbf{h}_N^\tau\right)^{-1}\mathbf{h}_N\right]' \\ = & \left[\mathbf{q}_N^\tau \underline{\mathbf{B}}_N^{-1}(z)\mathbf{h}_N - \frac{\mathbf{q}_N^\tau \underline{\mathbf{B}}_N^{-1}(z)\mathbf{q}_N \cdot \mathbf{h}_N^\tau \underline{\mathbf{B}}_N^{-1}(z)\mathbf{h}_N}{1+\mathbf{h}_N^\tau \underline{\mathbf{B}}_N^{-1}(z)\mathbf{q}_N}\right]' \\ \xrightarrow{i.p.} & \left[-\frac{c\underline{m}^2(z)}{1+\underline{m}(z)}\right]' = \frac{c\underline{m}^2(z)\underline{m}^{(2)}(z) - 2c\underline{m}(z)\underline{m}^{(2)}(z)}{\left(1+\underline{m}(z)\right)^2}, \end{aligned} \tag{A.89}$$

where $m^{(2)}(z)$ is defined in (3.3) of the main paper. Meanwhile, for the denominator of ω_2 we have

$$1 + \mathbf{q}_N^\tau \left(\underline{\mathbf{B}}_N - z\mathbf{I}_T + \mathbf{q}_N\mathbf{h}_N^\tau\right)^{-1}\mathbf{h}_N \xrightarrow{i.p.} 1 - \tfrac{c\underline{m}^2(z)}{1+\underline{m}(z)}. \tag{A.90}$$

It follows from (A.89) and (A.90) that

$$\omega_2 \xrightarrow{i.p.} \frac{c\underline{m}(z)\underline{m}^{(2)}(z)\left(\underline{m}(z)-2\right)}{\left(1+\underline{m}(z)\right)\left(1+\underline{m}(z)-c\underline{m}^2(z)\right)}. \tag{A.91}$$

For ω_3 defined in (A.59), its numerator can be written as

$$\mathbf{h}_N^\tau \left(\underline{\mathbf{B}}_N - z\mathbf{I}_T + \mathbf{q}_N\mathbf{h}_N^\tau + \mathbf{h}_N\mathbf{q}_N^\tau\right)^{-2}\mathbf{h}_N = \left[\mathbf{h}_N^\tau \left(\underline{\mathbf{B}}_N - z\mathbf{I}_T + \mathbf{q}_N\mathbf{h}_N^\tau + \mathbf{h}_N\mathbf{q}_N^\tau\right)^{-1}\mathbf{h}_N\right]'.$$

As before it is sufficient to consider $\mathbf{h}_N^\tau \left(\underline{\mathbf{B}}_N - z\mathbf{I}_T + \mathbf{q}_N\mathbf{h}_N^\tau + \mathbf{h}_N\mathbf{q}_N^\tau\right)^{-1}\mathbf{h}_N$.

By the formula (A.17) we have

$$\begin{aligned} & d_N\mathbf{h}_N^\tau\Big(\underline{\mathbf{B}}_N - z\mathbf{I}_T + \mathbf{q}_N\mathbf{h}_N^\tau + \mathbf{h}_N\mathbf{q}_N^\tau\Big)^{-1}\mathbf{h}_N = \frac{d_N\mathbf{h}_N^\tau(\underline{\mathbf{B}}_N - z\mathbf{I}_T + \mathbf{q}_N\mathbf{h}_N^\tau)^{-1}\mathbf{h}_N}{1+\mathbf{q}_N^\tau(\underline{\mathbf{B}}_N - z\mathbf{I}_T + \mathbf{q}_N\mathbf{h}_N^\tau)^{-1}\mathbf{h}_N} \\ = & \frac{d_N\mathbf{h}_N^\tau\underline{\mathbf{B}}_N^{-1}(z)\mathbf{h}_N}{\Big(1+\mathbf{h}_N^\tau\underline{\mathbf{B}}_N^{-1}(z)\mathbf{q}_N\Big)\Big(1+\mathbf{q}_N^\tau(\underline{\mathbf{B}}_N - z\mathbf{I}_T + \mathbf{q}_N\mathbf{h}_N^\tau)^{-1}\mathbf{h}_N\Big)} \\ = & \frac{d_N\mathbf{h}_N^\tau\underline{\mathbf{B}}_N^{-1}(z)\mathbf{h}_N}{\Big(1+\mathbf{h}_N^\tau\underline{\mathbf{B}}_N^{-1}(z)\mathbf{q}_N\Big)\Big(1+\mathbf{q}_N^\tau\underline{\mathbf{B}}_N^{-1}(z)\mathbf{h}_N - \frac{\mathbf{q}_N^\tau\underline{\mathbf{B}}_N^{-1}(z)\mathbf{q}_N\mathbf{h}_N^\tau\underline{\mathbf{B}}_N^{-1}(z)\mathbf{h}_N}{1+\mathbf{h}_N^\tau\underline{\mathbf{B}}_N^{-1}(z)\mathbf{q}_N}\Big)}. \end{aligned} \tag{A.92}$$

Based on (A.84), (A.84), (A.81), (A.80), (A.88) and the fact that

$$d_N = \frac{1}{T}\sum_{j=1}^N \frac{T}{||\widetilde{\mathbf{u}}_j||^2} = \frac{1}{T}\sum_{j=1}^N \frac{T}{||\widetilde{\boldsymbol{\varepsilon}}_j||^2 + \frac{1}{T}||\widetilde{\mathbf{v}}||^2 + \frac{2}{\sqrt{T}}\widetilde{\mathbf{v}}^\tau\widetilde{\boldsymbol{\varepsilon}}_j} \xrightarrow{i.p.} c, \tag{A.93}$$

for the numerator of ω_3 we have

$$\begin{aligned} & d_N\mathbf{h}_N^\tau \left(\underline{\mathbf{B}}_N - z\mathbf{I}_T + \mathbf{q}_N\mathbf{h}_N^\tau + \mathbf{h}_N\mathbf{q}_N^\tau\right)^{-2}\mathbf{h}_N \\ \xrightarrow{i.p.} & \left[\frac{c\underline{m}(z)\left(1+\underline{m}(z)\right)}{1+\underline{m}(z)-c\underline{m}^2(z)}\right]' \\ = & \frac{c\underline{m}^{(2)}(z)\left[\left(1+2\underline{m}(z)\right)\left(1+\underline{m}(z)-c\underline{m}^2(z)\right) - \underline{m}(z)\left(1+\underline{m}(z)\right)\left(1-2c\underline{m}(z)\right)\right]}{\left(1+\underline{m}(z)-c\underline{m}^2(z)\right)^2}. \end{aligned} \tag{A.94}$$

Similarly, the limit of the denominator of ω_3 is

$$1 + d_N\mathbf{h}_N^\tau \left(\underline{\mathbf{B}}_N - z\mathbf{I}_T + \mathbf{q}_N\mathbf{h}_N^\tau + \mathbf{h}_N\mathbf{q}_N^\tau\right)^{-1}\mathbf{h}_N \xrightarrow{i.p.} 1 + \frac{c\underline{m}(z)\left(1+\underline{m}(z)\right)}{1+\underline{m}(z)-c\underline{m}^2(z)}. \tag{A.95}$$

Combining (A.94) with (A.95) we obtain

$$\omega_3 \xrightarrow{i.p.} \frac{c\underline{m}^{(2)}(z)\left[(1+2\underline{m}(z))\left(1+\underline{m}(z)-c\underline{m}^2(z)\right)-\underline{m}(z)\left(1+\underline{m}(z)\right)\left(1-2c\underline{m}(z)\right)\right]}{\left(1+\underline{m}(z)-c\underline{m}^2(z)\right)\left(1+\underline{m}(z)+c\underline{m}(z)\right)}. \tag{A.96}$$

In summary, we have

$$\begin{aligned} & tr\left(\hat{\underline{\mathbf{R}}}_N - z\mathbf{I}_T\right)^{-1} - T\underline{m}_{c_N}(z) \\ = \quad & tr\left(\underline{\mathbf{A}}_N - z\mathbf{I}_T\right)^{-1} - T\underline{m}_{c_N}(z) - \omega_1 - \omega_2 - \omega_3 - \omega_4 - \omega_5 + o_P(1). \end{aligned}$$

The asymptotic distribution of $tr\left(\underline{\mathbf{A}}_N - z\mathbf{I}_T\right)^{-1} - T\underline{m}_{c_N}(z)$ is provided in Theorem 1 of Gao, el. (2014) and the limit of ω_i, $i = 1, 2, 3, 4, 5$ in probability are given in (A.81), (A.91), (A.95), (A.64) and (A.70). The proof of this theorem is completed by Slutsky's theorem.

8 Appendix B: Computational code

In this section, we will provide the code functions for the calculation of the empirical size for independence test. The code functions for other examples are similar and so omitted.

The main code function is displayed as follows.

```
function hatalpha=reg_clt_size(T,N,tl,tr,mm)
%%%%%%%%%%%%%%%%%%main function%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%Some Parameters%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
K=1000; %repeated times
c=N/T; %ratio of dimension over size
alpha=0.05; %significant level
a=(1-sqrt(c))^2; %left bound of mp law
b=(1+sqrt(c))^2; %right bound of mp law
standmu1=zeros(1,mm); %expectation of asymptotic distribution of sintx and costx
standmu2=zeros(1,mm);
standva1=zeros(1,mm);
standva2=zeros(1,mm);
standva3=zeros(1,mm); %variance of asymptotic distribution of sintx and costxvi=zeros(2,1);
myintegral=zeros(1,K); %limit of constructed test statictic
sr=zeros(1,K);
count=0; %number of random values that are large than quantile value
Test=zeros(1,K); %values of test statistic
TOL=0.0001; %precision of numeric integral
for k=1:K
    ii=1;
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%data generating process%%%%%%%%%%%%%%%%%%%%
beta=[1:0.1:2]';% a 2 by 1 vector;
kk=length(beta);
u=mvnrnd(zeros(1,N),eye(N),T);% error terms;
Y=zeros(T,N);% T by N
for i=1:N
XX(:,:,i)=mvnrnd(zeros(1,kk),eye(kk),T);
```

```
Y(:,i)=XX(:,:,i)*beta+u(:,i); % regression model
end
%%%%%%%%%%%%%%%%%%%%ols estimator of beta%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
invinner=zeros(kk,kk);
outmultiplier=zeros(kk,1);
for i=1:N
    invinner=invinner+XX(:,:,i)'*XX(:,:,i); % k by k;
    outmultiplier=outmultiplier+XX(:,:,i)'*Y(:,i); % k by 1;
end
beta_ols=inv(invinner)*outmultiplier;
%%%%%%%%%%%%%%%%%%%%%%%%%%%estimated error terms%%%%%%%%%%%%%%%%%%%%%%%%%%%
hat_u=zeros(T,N);
for i=1:N
    hat_u(:,i)=Y(:,i)-XX(:,:,i)*beta_ols;
end
u_reg=hat_u; % T by N;
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
for j=1:N
    r_reg(:,j)=u_reg(:,j)./norm(u_reg(:,j));
end
Rn_reg=r_reg*r_reg'; % T by T
A=Rn_reg;
eigvalue=eig(A);
for t=tl:((tr-tl)/mm):(tr-(tr-tl)/mm)
standmu1(1,ii)=1./(pi.*2i).*quad(@(theta)intelimmean1(T,N,eigenvale,r1,theta),0,2*pi);
standmu2(1,ii)=1./(pi.*2i).*quad(@(theta)intelimmean2(T,N,eigenvalue,r1,theta),0,2*pi);
standva1(1,ii)=-1./(4.*(pi.^2))
.*dblquad(@(theta1,theta2)intelimcov1(T,N,eigenvalue,r1,theta1,r2,theta2,t),0,2*pi,0,2*pi);
standva2(1,ii)=-1./(4.*(pi.^2))
.*dblquad(@(theta1,theta2)intelimcov2(T,N,eigenvalue,r1,theta1,r2,theta2,t),0,2*pi,0,2*pi);
standva3(1,ii)=-1./(4.*(pi.^2))
.*dblquad(@(theta1,theta2)intelimcov3(T,N,eigenvalue,r1,theta1,r2,theta2,t),0,2*pi,0,2*pi);
ii=ii+1;
%%%%%%%%%%%%%%%%%%%%%%calculation of proposed test statistic%%%%%%%%%%%%%%%
out11=quad(@(x)inf1(t,x,a,b,c),a,b,TOL);
out22=quad(@(x)inf2(t,x,a,b,c),a,b,TOL);
re=0;
rea=0;
rec=0;
red=0;
im=0;
ima=0;
imc=0;
imd=0;
for j=1:N
    re1=cos(t.*eigvalue(j));
```

```
    re=re+re1;
    re1a=cos(t.*eigvalueB(j));
    rea=rea+re1a;
    re1c=cos(t.*eigvalueC(j));
    rec=rec+re1c;
    re1d=cos(t.*eigvalueD(j));
    red=red+re1d;
    im1=sin(t.*eigvalue(j));
    im=im+im1;
    im1a=sin(t.*eigvalueB(j));
    ima=ima+im1a;
    im1c=sin(t.*eigvalueC(j));
    imc=imc+im1c;
    im1d=sin(t.*eigvalueD(j));
    imd=imd+im1d;
end
Test(1,k)=Test(1,k)+((re-N.*out11).^2+(im-N.*out22).^2).*(1/(tr-tl)).*((tr-tl)/mm);
standmu=[standmu1(1,ii),standmu2(1,ii)]';
standva=[standva1(1,ii),standva3(1,ii);standva3(1,ii),standva2(1,ii)];
vi(:,1)=mvnrnd(standmu,standva);
myintegral(1,k)=myintegral(1,k)+(vi(1,1).^2+vi(2,1)^2).*(1/(tr-tl)).*((tr-tl)/mm);
ii=ii+1;
end
end
sr(1,:)=sort(myintegral(1,:));
criticalp=sr(1,K*(1-alpha));
for k=1:K
    if Test(1,k)>=criticalp
        count=count+1;
    end
end
hatalpha=count./K;

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%Definition of inf1%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
function gun=inf1(t,x,a,b,c)
gun=(cos(t.*x)).*1./(2.*pi.*c).*sqrt((b-x).*(x-a))./x;

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%Definition of inf2%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
function gun=inf2(t,x,a,b,c)
gun=(sin(t.*x)).*1./(2.*pi.*c).*sqrt((b-x).*(x-a))./x;

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%Definition of intelimmean1%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
function meanint=intelimmean1(T,N,eigensamples,r,theta,t)
c=N/T;% ratio of N over T;
```

```
z=r.*exp(1i.*theta); % z values;
%%%%%%%%%%%%%%%%calculations of sz1,sz2 and derisz1,derisz2%%%%%%%%%%%%%%
stieltjes=0;% stieltjes transform;
for j=1:N
    stieltjes=stieltjes+1./(eigensamples(j)-z);
end
destieltjes=0;
for j=1:N
    destieltjes=destieltjes+1./(eigensamples(j)-z).^2;
end
stieltjes=stieltjes./N;% stieltjes transform;
destieltjes=destieltjes./N;% derivative of stieltjes transform
ulstj=-(1-c)./z+c.*stieltjes;% underline_m(z);
deulstj=(1-c)./z.^2+c.*destieltjes;% derivative of underline_m(z);
%%%%%%%%%%%%%%%%%%%%%%%%%some moment parameters%%%%%%%%%%%%%%%%%%%%%%%%%%%
mu_mean=0;
kappa=0;
for inum=1:N
    nume=0;
    deme=0;
    for jnum=1:T
        nume=nume+(X(jnum,inum)-mu_mean(inum)).^4;
        deme=deme+(X(jnum,inum)-mu_mean(inum)).^2;
    end
    nume=nume./T;
    deme=deme./T;
    kappa=kappa+nume./deme;
end
kappa=kappa./N;
%%%%%%%%%%%%%%%%%%%%%%%%%%%%limitvar%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
term=(kappa-3).*c.*z.*(1+ulstj).*ulstj.*stieltjes.^2.*(z+z.*ulstj+1-c)
     ./(((z+z.*ulstj-c).^2-c).*(1+c.*stieltjes));
meanint=sin(t.*z).*((kappa-1).*c.*ulstj.*(z+z.*ulstj+1-c)./(((z+z.*ulstj-c).^2-c)
        .*(z+z.*ulstj-c))-term-c.*deulstj.*(z+z.*ulstj+1-c)./(ulstj.*(z+z.*ulstj-c)
        .*((z+z.*ulstj-c).^2-c))+c.*(1+z.*ulstj-z.*ulstj.*stieltjes-z.^2.*stieltjes.*ulstj.^2)
        .*(1+ulstj).*(z+z.*ulstj+1-c)./(z.*(1+c.*stieltjes).*((z+z.*ulstj-c).^2-c))
        +c.*stieltjes./z-c.*z.*stieltjes.*deulstj).*z.*1i;

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%Definition of intelimmean2%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
function meanint=intelimmean2(T,N,eigensamples,r,theta,t)
c=N/T;% ratio of N over T;
z=r.*exp(1i.*theta); % z values;
%%%%%%%%%%%%%%%%calculations of sz1,sz2 and derisz1,derisz2%%%%%%%%%%%%%%
stieltjes=0;% stieltjes transform;
for j=1:N
    stieltjes=stieltjes+1./(eigensamples(j)-z);
```

```
end
destieltjes=0;
for j=1:N
    destieltjes=destieltjes+1./(eigensamples(j)-z).^2;
end
stieltjes=stieltjes./N;% stieltjes transform;
destieltjes=destieltjes./N;% derivative of stieltjes transform
ulstj=-(1-c)./z+c.*stieltjes;% underline_m(z);
deulstj=(1-c)./z.^2+c.*destieltjes;% derivative of underline_m(z);
%%%%%%%%%%%%%%%%%%%%%%%%some moment parameters%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
mu_mean=0;
kappa=0;
for inum=1:N
    nume=0;
    deme=0;
    for jnum=1:T
        nume=nume+(X(jnum,inum)-mu_mean(inum)).^4;
        deme=deme+(X(jnum,inum)-mu_mean(inum)).^2;
    end
    nume=nume./T;
    deme=deme./T;
    kappa=kappa+nume./deme;
end
kappa=kappa./N;
%%%%%%%%%%%%%%%%%%%%%%%%%%%%limitvar%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
term=(kappa-3).*c.*z.*(1+ulstj).*ulstj.*stieltjes.^2.*(z+z.*ulstj+1-c)
     ./(((z+z.*ulstj-c).^2-c).*(1+c.*stieltjes));
meanint=cos(t.*z).*((kappa-1).*c.*ulstj.*(z+z.*ulstj+1-c)./(((z+z.*ulstj-c).^2-c)
        .*(z+z.*ulstj-c))-term-c.*deulstj.*(z+z.*ulstj+1-c)./(ulstj.*(z+z.*ulstj-c)
        .*((z+z.*ulstj-c).^2-c))+c.*(1+z.*ulstj-z.*ulstj.*stieltjes-z.^2.*stieltjes.*ulstj.^2)
        .*(1+ulstj).*(z+z.*ulstj+1-c)./(z.*(1+c.*stieltjes).*((z+z.*ulstj-c).^2-c))
        +c.*stieltjes./z-c.*z.*stieltjes.*deulstj).*z.*1i;

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%Definition of intelimcov1%%%%%%%%%%%%%%%%%%%%%%
function covint=intelimcov1(T,N,eigensamples,r1,theta1,r2,theta2,t)
c=N/T;
z1=r1.*exp(1i.*theta1);
z2=r2.*exp(1i.*theta2);
%%%%%%%%%%%%%%%%%calculations of sz1,sz2 and derisz1,derisz2%%%%%%%%%%%%%%%
stieltjes1=0;
stieltjes2=0;
deristieltjes1=0;
deristieltjes2=0;
for j=1:N
    stieltjes1=stieltjes1+1./(eigensamples(j)-z1);
    stieltjes2=stieltjes2+1./(eigensamples(j)-z2);
```

```
    deristieltjes1=deristieltjes1+1./((eigensamples(j)-z1).^2);
    deristieltjes2=deristieltjes2+1./((eigensamples(j)-z2).^2);
end
stieltjes1=stieltjes1./N;
stieltjes2=stieltjes2./N;
ulstj1=-(1-c)./z1+c.*stieltjes1;
ulstj2=-(1-c)./z2+c.*stieltjes2;
deristieltjes1=deristieltjes1./N;
deristieltjes2=deristieltjes2./N;
deulstj1=(1-c)./z1.^2+c.*deristieltjes1;
deulstj2=(1-c)./z2.^2+c.*deristieltjes2;
%%%%%%%%%%%%%%%%%%%%%%%%some moment parameters%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
mu_mean=0;
kappa=0;
for inum=1:N
    nume=0;
    deme=0;
    for jnum=1:T
        nume=nume+(X(jnum,inum)-mu_mean(inum)).^4;
        deme=deme+(X(jnum,inum)-mu_mean(inum)).^2;
    end
    nume=nume./T;
    deme=deme./T;
    kappa=kappa+nume./deme;
end
kappa=kappa./N;
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%limitvar%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
term2=(kappa-3).*c.*(stieltjes1.*ulstj1+z1.*stieltjes1.*deulstj1+z1.*ulstj1.*deristieltjes1)
       .*(stieltjes2.*ulstj2+z2.*stieltjes2.*deulstj2+z2.*ulstj2.*deristieltjes2);
covint=-sin(t.z1).*sin(t.z2).*(2*c.*deristieltjes1*deristieltjes2./(1+c*stieltjes1+c*stieltjes2
       +c.*(c-1)*stieltjes1.*stieltjes2).^2
       -(kappa-1).*c.*deulstj1.*deulstj2./((1+ulstj1).*(1+ulstj2)).^2+term2).*z1.*z2;

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%Definition of intelimcov2%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
function covint=intelimcov2(T,N,eigensamples,r1,theta1,r2,theta2,t)
c=N/T;
z1=r1.*exp(1i.*theta1);
z2=r2.*exp(1i.*theta2);
%%%%%%%%%%%%%%%%%calculations of sz1,sz2 and derisz1,derisz2%%%%%%%%%%%%%%%
stieltjes1=0;
stieltjes2=0;
deristieltjes1=0;
deristieltjes2=0;
for j=1:N
    stieltjes1=stieltjes1+1./(eigensamples(j)-z1);
    stieltjes2=stieltjes2+1./(eigensamples(j)-z2);
```

```
    deristieltjes1=deristieltjes1+1./((eigensamples(j)-z1).^2);
    deristieltjes2=deristieltjes2+1./((eigensamples(j)-z2).^2);
end
stieltjes1=stieltjes1./N;
stieltjes2=stieltjes2./N;
ulstj1=-(1-c)./z1+c.*stieltjes1;
ulstj2=-(1-c)./z2+c.*stieltjes2;
deristieltjes1=deristieltjes1./N;
deristieltjes2=deristieltjes2./N;
deulstj1=(1-c)./z1.^2+c.*deristieltjes1;
deulstj2=(1-c)./z2.^2+c.*deristieltjes2;
%%%%%%%%%%%%%%%%%%%%%%%some moment parameters%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
mu_mean=0;
kappa=0;
for inum=1:N
    nume=0;
    deme=0;
    for jnum=1:T
        nume=nume+(X(jnum,inum)-mu_mean(inum)).^4;
        deme=deme+(X(jnum,inum)-mu_mean(inum)).^2;
    end
    nume=nume./T;
    deme=deme./T;
    kappa=kappa+nume./deme;
end
kappa=kappa./N;
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%limitvar%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
term2=(kappa-3).*c.*(stieltjes1.*ulstj1+z1.*stieltjes1.*deulstj1+z1.*ulstj1.*deristieltjes1)
       .*(stieltjes2.*ulstj2+z2.*stieltjes2.*deulstj2+z2.*ulstj2.*deristieltjes2);
covint=-cos(t.z1).*cos(t.z2).*(2*c.*deristieltjes1*deristieltjes2./(1+c*stieltjes1+c*stieltjes2
       +c.*(c-1)*stieltjes1.*stieltjes2).^2
       -(kappa-1).*c.*deulstj1.*deulstj2./((1+ulstj1).*(1+ulstj2)).^2+term2).*z1.*z2;

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%Definition of intelimcov3%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
function covint=intelimcov3(T,N,eigensamples,r1,theta1,r2,theta2,t)
c=N/T;
z1=r1.*exp(1i.*theta1);
z2=r2.*exp(1i.*theta2);
%%%%%%%%%%%%%%%%%calculations of sz1,sz2 and derisz1,derisz2%%%%%%%%%%%%%%%
stieltjes1=0;
stieltjes2=0;
deristieltjes1=0;
deristieltjes2=0;
for j=1:N
    stieltjes1=stieltjes1+1./(eigensamples(j)-z1);
    stieltjes2=stieltjes2+1./(eigensamples(j)-z2);
```

```
    deristieltjes1=deristieltjes1+1./((eigensamples(j)-z1).^2);
    deristieltjes2=deristieltjes2+1./((eigensamples(j)-z2).^2);
end
stieltjes1=stieltjes1./N;
stieltjes2=stieltjes2./N;
ulstj1=-(1-c)./z1+c.*stieltjes1;
ulstj2=-(1-c)./z2+c.*stieltjes2;
deristieltjes1=deristieltjes1./N;
deristieltjes2=deristieltjes2./N;
deulstj1=(1-c)./z1.^2+c.*deristieltjes1;
deulstj2=(1-c)./z2.^2+c.*deristieltjes2;
%%%%%%%%%%%%%%%%%%%%%%%%some moment parameters%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
mu_mean=0;
kappa=0;
for inum=1:N
    nume=0;
    deme=0;
    for jnum=1:T
        nume=nume+(X(jnum,inum)-mu_mean(inum)).^4;
        deme=deme+(X(jnum,inum)-mu_mean(inum)).^2;
    end
    nume=nume./T;
    deme=deme./T;
    kappa=kappa+nume./deme;
end
kappa=kappa./N;
%%%%%%%%%%%%%%%%%%%%%%%%%%%%limitvar%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
term2=(kappa-3).*c.*(stieltjes1.*ulstj1+z1.*stieltjes1.*deulstj1+z1.*ulstj1.*deristieltjes1)
       .*(stieltjes2.*ulstj2+z2.*stieltjes2.*deulstj2+z2.*ulstj2.*deristieltjes2);
covint=-sin(t.z1).*cos(t.z2).*(2*c.*deristieltjes1*deristieltjes2./(1+c*stieltjes1+c*stieltjes2
       +c.*(c-1)*stieltjes1.*stieltjes2).^2
       -(kappa-1).*c.*deulstj1.*deulstj2./((1+ulstj1).*(1+ulstj2)).^2+term2).*z1.*z2;
```