

CLT for Largest Eigenvalues and Unit Root Tests for High-Dimensional Nonstationary Time Series

Bo Zhang*, Guangming Pan[†] and Jiti Gao[‡]

Abstract

Let $\{Z_{ij}\}$ be independent and identically distributed (i.i.d.) random variables with $EZ_{ij} = 0$, $E|Z_{ij}|^2 = 1$ and $E|Z_{ij}|^4 < \infty$. Define linear processes $Y_{tj} = \sum_{k=0}^{\infty} b_k Z_{t-k,j}$ with $\sum_{i=0}^{\infty} |b_i| < \infty$. Consider a p -dimensional time series model of the form: $\mathbf{x}_t = \mathbf{\Pi}\mathbf{x}_{t-1} + \Sigma^{1/2}\mathbf{y}_t$, $1 \leq t \leq T$ with $\mathbf{y}_t = (Y_{t1}, \dots, Y_{tp})'$ and $\Sigma^{1/2}$ be the square root of a symmetric positive definite matrix. Let $\mathbf{B} = (1/p)\mathbf{X}\mathbf{X}^*$ with $\mathbf{X} = (\mathbf{x}_1, \dots, \mathbf{x}_T)'$ and X^* be the conjugate transpose. This paper establishes both the convergence in probability and the asymptotic joint distribution of the first k largest eigenvalues of \mathbf{B} when \mathbf{x}_t is nonstationary. As an application, a new unit root test for a vector of high-dimensional time series is proposed and then studied both theoretically and numerically.

Keywords: Asymptotic normality, Largest eigenvalue, Linear process, Unit root test.

1 Introduction

There have been an increasing interest and significant developments on the theory and methodologies for handling high-dimensional data in recent years. Understanding high-dimensional sample covariance matrices, including its eigenvalues and eigenvectors, has proved to be extremely useful for such developments. Indeed, random matrix theory has provided useful estimation and testing procedures for high-dimensional data analysis. Recent discussions on this topic can be found in Johnstone [15], Paul and Aue [21] and Yao, Zheng and Bai [28].

Research towards understanding the eigenvalues of sample covariance matrices dates back to as early as the studies of Fisher [12], Hsu [13] and Roy [24], and has become increasingly active since the publication of the celebrated work of Marcenko and Pastur [18], in which the authors established a limiting spectral distribution (MP type distribution) for a sample covariance matrix for the case where p and n are comparable. More recent research has been devoted to establishing

*Bo Zhang, School of Physical and Mathematical Sciences, Nanyang Technological University, Singapore, 637371 (Email: bzhang007@e.ntu.edu.sg).

[†]Corresponding Author, Guangming Pan, School of Physical and Mathematical Sciences, Nanyang Technological University, Singapore, 637371 (Email: gmpan@ntu.edu.sg).

[‡]Jiti Gao, Department of Econometrics and Business Statistics, Monash University, VIC 3145, Australia (Email: jiti.gao@monash.edu).

asymptotic properties for the eigenvalues and eigenvectors of high-dimensional sample covariance matrices.

There are currently two main lines of research about asymptotic distributions of the largest eigenvalues of high-dimensional random matrices. The first line of research is concerned with the Tracy-Widom law of the largest eigenvalues of random matrices. It is well known that limiting distributions of the largest eigenvalues of high-dimensional random matrices, such as Wigner matrices, follow the Tracy-Widom law, which was originally discovered by Tracy and Widom in [26] and [27] for Gaussian Wigner ensembles. The largest eigenvalue of the Wishart matrix was investigated in Johnstone [14]. Several progresses for general sample covariance matrices have also been made, and we refer to [5] and [11] among others.

Empirical data from wireless communication, finance and speech recognition often suggest that some extreme eigenvalues of sample covariance matrices are well separated from the rest. This intrigues the second line of research about the spiked eigenvalues, which was first proposed in Johnstone [14]. Significant progresses have been made in recent years on the behaviour of these spiked eigenvalues. For instance, the CLTs of the largest eigenvalues of complex Gaussian sample covariance matrices with a spiked population were investigated in Baik et al. [3], which also reported an interesting phase transition phenomenon. Baik and Silverstein [4] further considered almost sure limits of the extreme sample eigenvalues of the general spiked population. Paul [20] established a CLT for the spiked eigenvalues under the Gaussian population and the population spikes being simple. The fluctuation of the extreme sample eigenvalues of the general spiked population with arbitrary multiplicity numbers was further reported in Bai and Yao [2].

Most of the above existing studies rely on the assumption that the observations of high dimensional data are independent, although dimensional correlation structure can be allowed. Observations of high-dimensional data in economics and finance, for example, are often highly dependent on time-dimension. In view of this, Zhang [29] investigated the empirical spectral distribution (ESD) of the sample covariance for the case where the data matrices are of the form $\mathbf{A}_1\mathbf{Z}\mathbf{A}_2$, where \mathbf{A}_1 and \mathbf{A}_2 are positive semidefinite matrices and \mathbf{Z} has independent entries satisfying some moment assumptions. This model is referred to as the separable covariance model and allows for some dependence among observations recorded over different time points. Liu, Aue and Paul [17] studied the ESD of sample covariance matrices and symmetrized sample autocovariance matrices constructed from a linear process. Note that their setting also accommodates dependence among observations due to the fact that linear processes are built from the same innovation vectors. However, the above two papers considered the ESD only.

To the best of our knowledge, there is no existing work available to deal with the largest eigenvalues of sample covariance matrices generated from high-dimensional nonstationary time series data. This paper belongs to the second line of research about the spiked eigenvalues. This paper establishes a joint asymptotic distribution for the first several largest eigenvalues of a large sample covariance matrix of high-dimensional nonstationary data. An additional contribution of this paper is that it develops a new unit root test that is devoted to testing nonstationarity

for high-dimensional dependent time series.

We conclude this section by giving its organization. Section 2 establishes an asymptotic distributional theory for the first several largest eigenvalues of the covariance matrix of a high-dimensional dependent time series. Section 3 proposes a new unit root test that is devoted to testing nonstationarity for high dimensional dependent data. Section 4 evaluates both the size and power properties of the proposed test. Section 5 gives some concluding remarks. Appendix A establishes some useful results for truncated versions of sample covariance matrices by truncating linear processes. Appendix B gives the full proofs of the main theorems in Section 3.

2 Asymptotic Theory

This section first introduces some necessary assumptions before we establish new asymptotic properties for the largest eigenvalues of the covariance matrix of a high-dimensional time series vector.

2.1 Matrix models

The paper is to consider high-dimensional covariance matrices for nonstationary time series. Specifically, define the following linear processes:

$$Y_{tj} = \sum_{k=0}^{\infty} b_k Z_{t-k,j} \quad (2.1)$$

with $\sum_{i=0}^{\infty} |b_i| < \infty$. Suppose that $\mathbf{y}_t = (Y_{t1}, \dots, Y_{tp})'$ is a p -dimensional time series where $\{Z_{ij}\}$ are independent and identically distributed (i.i.d.) random variables with $E Z_{ij} = 0$, $E|Z_{ij}|^2 = 1$ and $E|Z_{ij}|^4 < \infty$. Consider a p -dimensional time series model of the form:

$$\mathbf{x}_t = \mathbf{\Pi} \mathbf{x}_{t-1} + \mathbf{\Sigma}^{1/2} \mathbf{y}_t, \quad 1 \leq t \leq T, \quad (2.2)$$

where the spectral norm of the coefficient matrix $\mathbf{\Pi}$ is bounded by one ($0 \leq \|\mathbf{\Pi}\|_2 \leq 1$). Let

$$\mathbf{B} = \frac{1}{p} \mathbf{X} \mathbf{X}^* \quad (2.3)$$

with $\mathbf{X} = (\mathbf{x}_1, \dots, \mathbf{x}_T)'$. We also define a $T \times p$ matrix $\mathbf{X}_0 = (\mathbf{x}_0, \dots, \mathbf{x}_0)'$ consisting of the initial vector \mathbf{x}_0 of the time series. Here we would point out that when $\mathbf{\Pi} = \mathbf{0}$, $\mathbf{\Sigma}$ satisfies some conditions and Y_{tj} 's are i.i.d random variables, the Tracy-Widom distribution has been established for the large eigenvalue of \mathbf{B} in [5]. Also, when $\mathbf{\Pi} = \mathbf{0}$, $\mathbf{\Sigma}$ is a block matrix with spiked eigenvalues and Y_{tj} 's are i.i.d random variables, an asymptotic distribution (Gaussian distribution under some conditions) for the largest eigenvalues of \mathbf{B} has been discussed in [20] and [2]. It is not clear yet how the largest eigenvalues of \mathbf{B} may behave when Y_{tj} 's have some dependence structure. One case is that $\mathbf{\Pi} = \mathbf{0}$, but $\mathbf{\Sigma}$ is present in (2.1). When $\mathbf{\Pi} = \mathbf{I}$, (2.2) becomes nonstationary. The main motivation for considering such a model is the proposal of one unit root test to be discussed in the next section.

When $\mathbf{\Pi} = \mathbf{I}$, the sample covariance matrix \mathbf{B} can be rewritten as follows. Let $\mathbf{Y} = (\mathbf{y}_1, \dots, \mathbf{y}_T)'$ be a $T \times p$ random matrix. Define $\mathbf{C} = (C_{ij})_{1 \leq i, j \leq T}$ to be a $T \times T$ lower triangular matrix with

$$C_{ij} = 0 \text{ for } j > i \text{ and } C_{ij} = 1 \text{ for } 1 \leq j \leq i. \quad (2.4)$$

Now we may rewrite $\mathbf{X} = \mathbf{C}\mathbf{Y}\mathbf{\Sigma}^{1/2} + \mathbf{X}_0$ so that

$$\mathbf{B} = \frac{1}{p}\mathbf{X}\mathbf{X}^* = \frac{1}{p}\mathbf{C}\mathbf{Y}\mathbf{\Sigma}\mathbf{Y}^*\mathbf{C}^* + \frac{1}{p}\mathbf{C}\mathbf{Y}\mathbf{\Sigma}^{1/2}\mathbf{X}_0^* + \frac{1}{p}\mathbf{X}_0\mathbf{\Sigma}^{1/2}\mathbf{Y}^*\mathbf{C}^* + \frac{1}{p}\mathbf{X}_0\mathbf{X}_0^*. \quad (2.5)$$

This paper is to investigate the largest eigenvalues of \mathbf{B} for the case where $\mathbf{\Pi} = \mathbf{I}$ or $\|\mathbf{\Pi}\|_2 = \varphi < 1$. Throughout the paper, we make the following assumptions about the coefficients b_i and $\mathbf{\Sigma}$:

(A1) $\sum_{i=0}^{\infty} i|b_i| < \infty$.

(A2) $\sum_{i=0}^{\infty} b_i = s \neq 0$.

(A3) There exist two positive constants M_0 and M_1 such that $\|\mathbf{\Sigma}\|_2 \leq M_0$ and $\frac{\text{tr}(\mathbf{\Sigma})}{p} \geq M_1$.

(A4) $T \rightarrow \infty$ and $p \rightarrow \infty$ such that $\lim_{T \rightarrow \infty} \frac{\sqrt{p}}{T} = 0$.

Here $\|\cdot\|_2$ stands for either the spectral norm of a matrix or the Euclidean norm of a vector. The linear process includes MA(q) models and AR(1) models. Assumption A2 is easily satisfied. Note that we do not require p and T to be of the same order, which is being commonly used in the random matrix theory literature. We also need to make some assumptions about Z_{ij} and \mathbf{x}_0 .

(A5) $\{Z_{i,j}\}$ are i.i.d random variables with mean zero, variance one and bounded fourth moment.

Let $\mathbf{z}_t = (Z_{t1}, \dots, Z_{tp})'$, where t can be either positive or negative integer (for the purpose of introducing A7 below).

(A6) $E\|\mathbf{x}_0\|_2^2 = O(p)$.

(A7) $\mathbf{x}_0 = \sum_{k=0}^{\infty} \tilde{b}_k \mathbf{\Sigma}_1^{1/2} \mathbf{z}_{-k} + \tilde{b}_{-1} \mathbf{\Sigma}_2^{1/2} \tilde{\mathbf{z}} + \tilde{\mathbf{b}}_{-2}$, where $\|\mathbf{\Sigma}_1\|_2 \leq M_0$, $\|\mathbf{\Sigma}_2\|_2 \leq M_0$ and $\tilde{\mathbf{z}} = (\tilde{Z}_1, \dots, \tilde{Z}_p)'$ is independent of \mathbf{z}_t for any t , in which $\{\tilde{Z}_j\}$ are i.i.d random variables with mean zero, variance one and finite fourth moments. The coefficients satisfy $\sum_{k=0}^{\infty} |\tilde{b}_k| + |\tilde{b}_{-1}| < \infty$ and $\|\tilde{\mathbf{b}}_{-2}\|^2 = O(p)$.

We would like to remark that Assumption A7 implies Assumption A6.

2.2 Main results

To characterize the limits in probability of the eigenvalues of \mathbf{B} , define for $k = 1, \dots, T$,

$$\lambda_k = \frac{1}{2(1 + \cos \theta_k)} \text{ with } \theta_k = \frac{2(T+1-k)\pi}{2T+1}, \quad (2.6)$$

and

$$\gamma_k = \lambda_k \left(a_0 + 2 \sum_{j=1}^{\infty} a_j (-1)^j \cos(j\theta_k) \right), \quad (2.7)$$

where

$$a_i = \sum_{k=0}^{\infty} b_k b_{k+i}. \quad (2.8)$$

We first characterize the magnitude of λ_k and γ_k .

Proposition 1. *For any fixed constant $k \geq 1$, there is a constant c_k such that*

$$\lim_{T \rightarrow \infty} \frac{\gamma_k}{T^2} = c_k > 0 \quad (2.9)$$

and

$$\lim_{T \rightarrow \infty} \frac{\gamma_k}{\gamma_1} = \lim_{T \rightarrow \infty} \frac{\lambda_k}{\lambda_1} = \frac{1}{(2k-1)^2}. \quad (2.10)$$

We are now at a position to state the main results; their proofs are given in Appendix B. The first theorem develops an upper bound in probability for the spectral norm of \mathbf{B} for the stationary case. The second theorem gives a limit in probability for the first k largest eigenvalues of \mathbf{B} for nonstationary data. The third theorem establishes a joint distribution for the first k largest eigenvalues of \mathbf{B} for the nonstationary case.

Theorem 1. *Suppose that Assumptions A1-A6 hold. When $0 \leq \|\mathbf{\Pi}\|_2 = \varphi < 1$, we obtain*

$$\lim_{T \rightarrow \infty} P \left(\|\mathbf{B}\|_2 \leq \frac{8 \sum_{i \geq 0} |a_i|}{(1-\varphi)^2} M_0 \left(1 + \sqrt{\frac{T}{p}} \right)^2 \right) = 1, \quad (2.11)$$

where M_0 is a positive constant independent of p and T .

Theorem 2. *Suppose that Assumptions A1-A6 hold. Let ρ_k be the k th largest eigenvalue of \mathbf{B} . When $\mathbf{\Pi} = \mathbf{I}$ and k is fixed, we have*

$$\frac{\rho_k - \gamma_k \frac{\text{tr}(\mathbf{\Sigma})}{p}}{\gamma_1} \xrightarrow{i.p.} 0, \quad (2.12)$$

where *i.p.* means convergence in probability.

Theorem 3. *Suppose that Assumptions A1-A5 and A7 hold. Let ρ_k be the k th largest eigenvalue of \mathbf{B} . When $\mathbf{\Pi} = \mathbf{I}$ and k is fixed, the random vector*

$$\frac{\sqrt{p}}{\gamma_1} \left(\rho_1 - \gamma_1 \frac{\text{tr}(\mathbf{\Sigma})}{p}, \dots, \rho_k - \gamma_k \frac{\text{tr}(\mathbf{\Sigma})}{p} \right)' \quad (2.13)$$

converges weakly to a zero-mean Gaussian vector $\mathbf{w} = (w_1, \dots, w_k)'$ with the covariance function $\text{cov}(w_i, w_j) = 0$ for any $i \neq j$ and $\text{var}(w_i) = \frac{2\theta}{(2i-1)^4}$ with $\theta = \lim_{p \rightarrow \infty} \frac{\text{tr}(\mathbf{\Sigma}^2)}{p}$.

Remark 1. *The result holds for the complex case as well. In fact when \mathbf{Z} is complex, set*

$$\operatorname{Re}(Z_{jk}) = Z_{ij}^R, \quad \text{and} \quad \operatorname{Im}(Z_{jk}) = Z_{ij}^I. \quad (2.14)$$

Let Z_{ij}^R and Z_{ij}^I be independent. Then $\frac{\sqrt{p}}{\gamma_1} \left(\rho_1 - \gamma_1 \frac{\operatorname{tr}(\boldsymbol{\Sigma})}{p}, \dots, \rho_k - \gamma_k \frac{\operatorname{tr}(\boldsymbol{\Sigma})}{p} \right)'$ converges weakly to a zero-mean Gaussian vector $\mathbf{w} = (w_1, \dots, w_k)'$ with $\operatorname{var}(w_i) = \frac{2\theta}{(2i-1)^4} (1 - 2E(Z_{i1}^R)^2 E(Z_{i1}^I)^2)$, in which $\theta = \lim_{p \rightarrow \infty} \frac{\operatorname{tr}(\boldsymbol{\Sigma}^2)}{p}$. When $i \neq j$, $\operatorname{cov}(w_i, w_j) = 0$.

Remark 2. *If Assumption A7 does not hold but Assumption A6 is true, then Theorem 3 remains true under Assumptions A1-A3 and $\lim_{T \rightarrow \infty} \frac{p}{T} = 0$.*

Remark 3. *We would compare our results with those in [2]. [2] needs to assume that the observations are independent and that $\boldsymbol{\Sigma}$ has a spiked structure. In our paper, the observations are dependent. Furthermore, we don't need to assume a spiked structure of $\boldsymbol{\Sigma}$ since the spiked eigenvalues come naturally from the random walk structure.*

3 Unit Root Test

This section is to explore an application of the main results to the proposal of a new unit root test for a high-dimensional time series setting.

Unit root testing is a formal test to check whether time series data are stationary or non-stationary. Existing studies on this topic can be found in [10], [6] and [23]. In the past two decades, unit root testing in panels has received much attention. Many researchers (e.g. [9], [16] and [22]) consider the time series case where the error process is independent across individuals. There are also many results (e.g. [7]) that have investigated the case where the error process is cross-sectional dependent. In these papers, researchers often need to first estimate the covariance matrix about the cross-sectional dependence. However, when the dimensionality of the time series becomes large, it is hard to consistently estimate it without imposing some structure on the covariance matrix. We would like to propose one new test using the covariance matrix of a high-dimensional time series under consideration.

To this end, a key observation is that Theorem 2 indicates that the largest eigenvalue of \mathbf{B} is of order T^2 in probability (the order of γ_1 , which is given in Proposition 1), while Theorem 1 and Assumption (A4) imply that when $0 \leq \varphi < 1$, $\|\mathbf{B}\|_2 = o_p(T)$. This motivates us to propose a new unit root test based on the largest eigenvalues.

3.1 Test statistic

Theorem 3 states that when $\boldsymbol{\Pi} = \mathbf{I}$, the statistic $L_p = \frac{\sqrt{p}(\rho_1 - \gamma_1 \frac{\operatorname{tr}(\boldsymbol{\Sigma})}{p})}{\gamma_1 \sqrt{2\theta}}$ converges weakly to a standard normal variable. Note that $\gamma_1 \frac{\operatorname{tr}(\boldsymbol{\Sigma})}{p}$ and $\gamma_1 \sqrt{2\theta}$ are both unknown in practice. We would like to emphasize that γ_1 , $\frac{\operatorname{tr}(\boldsymbol{\Sigma})}{p}$ and θ can not be estimated individually. However it is possible to estimate their products as a whole. Specifically speaking, an estimator of $\gamma_1 \frac{\operatorname{tr}(\boldsymbol{\Sigma})}{p}$ is proposed below.

Theorem 4. Under the conditions of Theorem 3, we obtain

$$\frac{\lambda_1 \left(\sum_{i=2}^T \frac{(\mathbf{x}_i - \mathbf{x}_{i-1})'(\mathbf{x}_i - \mathbf{x}_{i-1})}{p^{(T-1)}} + 2 \sum_{j=1}^{m_1} \sum_{i=2}^{T-j} \frac{(\mathbf{x}_i - \mathbf{x}_{i-1})'(\mathbf{x}_{i+j} - \mathbf{x}_{i+j-1})}{p^{(T-j-1)}} \right) - \gamma_1 \frac{\text{tr}(\boldsymbol{\Sigma})}{p}}{\gamma_1} = o_p(p^{-1/2}), \quad (3.1)$$

when $m_1 = \lceil \sqrt{p} \rceil$, where λ_1 is given in (2.6).

However, finding a consistent estimator for $\gamma_1 \sqrt{2 \frac{\text{tr}(\boldsymbol{\Sigma}^2)}{p}}$ is challenging. Our strategy is to find an estimator for $a_i a_j \text{tr} \boldsymbol{\Sigma}^2$ first, which turns out to be $S_{\sigma^2, i, j}$ given below. To this end, let m_2 be a number to be specified later. For $0 \leq i \leq j \leq m_2$, define

$$S_{\sigma^2, i, j} = \frac{\sum_{f=2}^{\lfloor T/2 \rfloor - j} \sum_{g=f+\lfloor T/2 \rfloor}^{T-j} ((\mathbf{x}_f - \mathbf{x}_{f-1})'(\mathbf{x}_g - \mathbf{x}_{g-1}))((\mathbf{x}_{f+i} - \mathbf{x}_{f+i-1})'(\mathbf{x}_{g+j} - \mathbf{x}_{g+j-1}))}{(T - j/2 - \frac{3}{2}\lfloor T/2 \rfloor)(\lfloor T/2 \rfloor - j - 1)}. \quad (3.2)$$

We next approximate γ_1 by $\lambda_1 \left(a_0 + 2 \sum_{j=1}^{m_2} a_j \right)$. By carefully selecting the appropriate terms $S_{\sigma^2, i, j}$ and expanding the square of $\lambda_1 \left(a_0 + 2 \sum_{j=1}^{m_2} a_j \right)$, we may then construct S_{σ^2} , the estimator of $\gamma_1 \sqrt{2 \frac{\text{tr}(\boldsymbol{\Sigma}^2)}{p}}$, as follows: $S_{\sigma^2} = S_{\sigma^2, 0, 0} + 4 \sum_{i=1}^{m_2} S_{\sigma^2, i, i} + 4 \sum_{i=1}^{m_2} S_{\sigma^2, 0, i} + 8 \sum_{i=1}^{m_2-1} \sum_{j=i+1}^{m_2} S_{\sigma^2, i, j}$. To make it nonnegative we below add the absolute value sign for S_{σ^2} .

Theorem 5. Let m_2 tend to infinity such that $m_2^4 \max\{pT^{-2}, p^{-1}\} = o(1)$. Under the conditions of Theorem 3, we have

$$\frac{\lambda_1 \sqrt{2 \frac{|S_{\sigma^2}|}{p}}}{\gamma_1 \sqrt{2 \frac{\text{tr}(\boldsymbol{\Sigma}^2)}{p}}} = 1 + o_p(1), \quad (3.3)$$

where λ_1 is given in (2.6).

Once the two estimators are available, we can construct a test statistic, T_N , of the form:

$$T_N = \sqrt{p} \frac{\rho_1 - \lambda_1 \left(\sum_{i=2}^T \frac{(\mathbf{x}_i - \mathbf{x}_{i-1})'(\mathbf{x}_i - \mathbf{x}_{i-1})}{p^{(T-1)}} + 2 \sum_{j=1}^{m_1} \sum_{i=2}^{T-j} \frac{(\mathbf{x}_i - \mathbf{x}_{i-1})'(\mathbf{x}_{i+j} - \mathbf{x}_{i+j-1})}{p^{(T-j-1)}} \right)}{\lambda_1 \sqrt{2 \frac{|S_{\sigma^2}|}{p}}}, \quad (3.4)$$

where λ_1 is given in (2.6).

Theorem 6. Under the conditions of Theorem 5, we have

$$T_N \longrightarrow N(0, 1). \quad (3.5)$$

Remark 4. The conditions imposed on m_1 and m_2 can be further relaxed. For example, if there exists a positive integer s such that $b_i = 0$ for any $i > s$ in (2.1), we find $a_i = 0$ for any $i > s$ in (2.8). So one can choose $m_1 = \min\{s, \lceil \sqrt{p} \rceil\}$ in this case. Also, one can choose $m_2 = s$ as long as $s^4 \max\{pT^{-2}, p^{-1}\} = o(1)$. This point helps us to simplify the design and the verifications of the assumptions for the simulation in Section 4 below.

Now we investigate the power of T_N for the case where $\{Y_{tj}\}$ in (2.1) are i.i.d, a type of local alternatives under H_1 .

Theorem 7. *Let Assumptions A1–A5 and A7 hold with $b_i = 0$ for $i \geq 1$ and $\mathbf{\Pi} = \varphi \mathbf{I}$ with $0 \leq \varphi < 1$. Then under the case of $m_1 = m_2 = 0$, we have*

$$\lim_{T \rightarrow \infty} P(T_N > C_0 | H_1) = 1 \quad (3.6)$$

for some $C_0 > \ell_\alpha$, where ℓ_α is the α -level critical value of the limiting distribution of T_N .

The proofs of Theorems 4–7 are given in Appendix B.

Remark 5. *There are some well-known panel unit root tests (e.g. [9] and [16]). They considered the case of $\mathbf{\Pi} = \text{diag}(\varphi_1, \dots, \varphi_N)$ and used the estimators of φ_i to test whether $\mathbf{\Pi} = \mathbf{I}$. Moreover, when the covariance matrix $\mathbf{\Sigma}$ is involved, it has to be estimated in order to test whether $\mathbf{\Pi} = \mathbf{I}$ (e.g. [7]). So such existing tests may only work for the finite-dimensional case. By contrast, our test makes the best use of the properties of the largest eigenvalues of B instead of estimators of φ_i . In addition, we do not impose special structures, such as sparsity on the covariance matrix $\mathbf{\Sigma}$.*

The advantages of our test over existing tests will also be demonstrated by the finite-sample evaluation in Section 4 below.

4 Simulation

This section is to conduct some simulations to investigate the size and power of T_N . We first specify the initial vector x_0 and the coefficients b_i involved in the linear process. Let

$$b_i = \begin{cases} 1 & i = 0, \\ 0.3 & i = 1, \\ 0 & i \geq 2 \end{cases}, \quad \tilde{b}_i = \begin{cases} 0.6 & i = 0, \\ 0.8 & i = -1, \\ 0 & \text{otherwise} \end{cases}$$

and

$$(\mathbf{\Sigma}^{1/2})_{ij} = (\mathbf{\Sigma}_1^{1/2})_{ij} = (\mathbf{\Sigma}_2^{1/2})_{ij} = \begin{cases} 1 & i = j, \\ 0.2 & |i - j| = 1, \\ 0 & |i - j| \geq 2. \end{cases}$$

Let \tilde{Z}_j follow the uniform distribution over the interval $[-\sqrt{3}, \sqrt{3}]$. Let Z_{ij} be either the uniform distribution over the interval $[-\sqrt{3}, \sqrt{3}]$ or the standard normal distribution. We also consider the case where the matrix $\mathbf{\Pi}$ is not diagonal. The respective entries of two such matrices $\mathbf{\Pi}_1$ and $\mathbf{\Pi}_2$ are given as follows:

$$(\mathbf{\Pi}_1)_{ij} = \begin{cases} 0.599 & i = j, \\ 0.2 & |i - j| = 1, \\ 0 & |i - j| \geq 2, \end{cases}$$

$$(\mathbf{\Pi}_2)_{ij} = \begin{cases} 0.5 & i = j, \\ 0.2 & |i - j| = 1, \\ 0 & |i - j| \geq 2. \end{cases}$$

One can verify that the largest eigenvalue of $\mathbf{\Pi}_1$ is smaller than 0.999 and the smallest eigenvalue of $\mathbf{\Pi}_1$ is bigger than 0.199. Similarly, the largest eigenvalue of $\mathbf{\Pi}_2$ is smaller than 0.9 and the smallest eigenvalue of $\mathbf{\Pi}_2$ is bigger than 0.1. We choose $m_1 = m_2 = 1$ for such $\mathbf{\Pi}_1$ and $\mathbf{\Pi}_2$ as pointed out by Remark 4.

The results of the test statistic T_N are based on 1000 replications. The nominal size is set to be 0.05. We consider the case when both p and T are large (regardless of whether T is larger than p or smaller than p). Different values of p , T and $\mathbf{\Pi}$ are given in Tables 1–4. One can see that the power increases when $\mathbf{\Pi}$ moves away from the identity matrix.

Table 1: The empirical size for T_N

Z_{ij}	Normal			Uniform		
	40	60	80	40	60	80
40	0.057	0.051	0.052	0.068	0.058	0.051
60	0.065	0.035	0.052	0.056	0.058	0.048
80	0.061	0.052	0.045	0.062	0.046	0.063

Table 2: The empirical power for T_N when $\mathbf{\Pi} = 0.99\mathbf{I}$

Z_{ij}	Normal			Uniform		
	40	60	80	40	60	80
40	0.085	0.167	0.348	0.088	0.161	0.354
60	0.136	0.349	0.595	0.132	0.362	0.596
80	0.208	0.506	0.801	0.215	0.501	0.795

4.1 Comparison with some existing tests when p is large

There are several existing unit root tests available for panel data. The first type of statistics considered the case where the random variables Z_{it} and Z_{js} are independent when $i \neq j$. For example one may see [9]. In this case the p -dimensional test can be converted into p independent 1-dimensional tests. As a consequence, this idea can use the sample size of $O(pT)$. Unfortunately, when Σ is unknown, this method doesn't work.

Table 3: The empirical power for T_N when $\mathbf{\Pi} = 0.98\mathbf{I}$

Z_{ij}	Normal			Uniform		
	40	60	80	40	60	80
p \ T	40	60	80	40	60	80
40	0.327	0.658	0.948	0.319	0.662	0.946
60	0.520	0.922	0.999	0.531	0.917	1
80	0.717	0.986	1	0.709	0.990	1

Table 4: The empirical power for T_N when $\mathbf{\Pi} = \mathbf{\Pi}_1$

Z_{ij}	Normal			Uniform		
	40	60	80	40	60	80
p \ T	40	60	80	40	60	80
40	0.916	0.967	0.980	0.919	0.965	0.982
60	0.983	0.998	0.999	0.981	0.996	0.998
80	0.998	1	1	0.999	1	1

When $\mathbf{\Sigma}$ is unknown, an immediate idea is to estimate $\mathbf{\Sigma}$. [7] showed that the Bootstrap method with estimation of $\mathbf{\Sigma}$ performs better than the t-bar statistic for the case where p is fixed and T is large. [7] also stated that the Bootstrap-OLS performs better than Bootstrap-GLS when p is large. Furthermore, GLS doesn't work when $p \geq T$. So we will compare T_N with the t-statistic corresponding to Bootstrap-OLS t_{ols}^* and the F-statistic corresponding to Bootstrap-OLS F_{ols}^* .

In [7], when the error (\mathbf{y}_t in our paper) is an MA process, they used a finite order AR model to approximate it. Unfortunately, the approximation may perform poorly sometimes. For example, one may consider the case where \mathbf{y}_t is an MA(1) process and b_1 is near to 1. In contrast, T_N can perform well even though $b_1 = 0.99$. We use the same $\mathbf{\Sigma}$ as in Tables 1–4. The results are given in Tables 5–8. In this case, one can see that T_N performs well.

Table 5: The empirical size for T_N when $b_1 = 0.99$

Z_{ij}	Normal			Uniform		
	40	60	80	40	60	80
p \ T	40	60	80	40	60	80
40	0.041	0.039	0.043	0.049	0.057	0.048
60	0.052	0.046	0.039	0.059	0.048	0.049
80	0.044	0.047	0.054	0.055	0.049	0.043

Now we consider the case where \mathbf{y}_t is independent. In other words, we assume that $b_i = 0$

Table 6: The empirical power for T_N when $b_1 = 0.99$ and $\mathbf{\Pi} = 0.99\mathbf{I}$

Z_{ij}	Normal			Uniform		
	40	60	80	40	60	80
40	0.086	0.174	0.369	0.089	0.182	0.366
60	0.148	0.365	0.634	0.155	0.359	0.631
80	0.249	0.549	0.817	0.256	0.542	0.825

Table 7: The empirical power for T_N when $b_1 = 0.99$ and $\mathbf{\Pi} = 0.98\mathbf{I}$

Z_{ij}	Normal			Uniform		
	40	60	80	40	60	80
40	0.364	0.717	0.949	0.359	0.722	0.956
60	0.574	0.945	0.998	0.580	0.939	1
80	0.748	0.989	1	0.752	0.992	1

Table 8: The empirical power for T_N when $b_1 = 0.99$ and $\mathbf{\Pi} = \mathbf{\Pi}_1$

Z_{ij}	Normal			Uniform		
	40	60	80	40	60	80
40	0.956	0.982	0.991	0.962	0.979	0.994
60	0.993	0.997	0.999	0.992	0.998	1
80	0.996	1	1	0.998	1	1

when $i \geq 1$. The results of the three test statistics based on 1000 replications, 500 bootstrap replications and different values of p , T and $\mathbf{\Pi}$ are reported in Tables 9—11. The nominal size is set to be 0.05.

Table 9: The empirical size for three tests when $b_i = 0$ for $i > 0$.

test		T_N		t_{ols}^*		F_{ols}^*	
p	T	Normal	Uniform	Normal	Uniform	Normal	Uniform
40	40	0.038	0.052	0	0	0.001	0.002
40	60	0.044	0.055	0.001	0.001	0.004	0.003
40	80	0.053	0.046	0.004	0.008	0.012	0.015
60	40	0.046	0.049	0	0	0.006	0.001
60	60	0.052	0.047	0	0	0	0
60	80	0.050	0.044	0	0	0	0.002
80	40	0.056	0.052	0.001	0	0.004	0.003
80	60	0.041	0.048	0	0	0.001	0.001
80	80	0.048	0.053	0	0	0	0

One can find that the empirical sizes of t_{ols}^* and F_{ols}^* suffer from the size distortion when p is large. This indicates that their asymptotic distributions may not hold under the null hypothesis when p is large. One of the reasons is that when p is large and the population covariance matrix is not assumed to have some special structures, we can't find any consistent estimators for the population covariance matrix and other unknown parameters. As a consequence, their asymptotic distributions may fail to hold under the null. Meanwhile, their power is worse than T_N when p is large and $\mathbf{\Pi}$ is close to \mathbf{I} . In addition, T_N doesn't require bootstrap repetitions. This saves computational time.

Table 10: The empirical power for three tests when $b_i = 0$ for $i \geq 1$ and $\mathbf{\Pi} = 0.99\mathbf{I}$.

test		T_N		t_{ols}^*		F_{ols}^*	
p	T	Normal	Uniform	Normal	Uniform	Normal	Uniform
40	40	0.078	0.081	0.002	0	0.002	0
40	60	0.178	0.187	0.024	0.022	0.025	0.025
40	80	0.379	0.375	0.095	0.087	0.098	0.088
60	40	0.180	0.191	0.003	0.003	0.003	0.003
60	60	0.402	0.397	0.004	0.002	0.004	0.003
60	80	0.656	0.660	0.039	0.044	0.043	0.047
80	40	0.238	0.249	0.001	0.001	0.001	0.001
80	60	0.536	0.554	0.005	0.005	0.005	0.006
80	80	0.834	0.837	0.007	0.008	0.008	0.009

Table 11: The empirical power for three tests when $b_i = 0$ for $i \geq 1$ and $\mathbf{\Pi} = 0.98\mathbf{I}$.

test		T_N		t_{ols}^*		F_{ols}^*	
p	T	Normal	Uniform	Normal	Uniform	Normal	Uniform
40	40	0.378	0.382	0.012	0.012	0.013	0.012
40	60	0.768	0.777	0.145	0.152	0.153	0.156
40	80	0.947	0.953	0.381	0.386	0.386	0.388
60	40	0.634	0.641	0.019	0.019	0.019	0.022
60	60	0.955	0.957	0.064	0.059	0.065	0.061
60	80	0.998	0.999	0.363	0.347	0.373	0.355
80	40	0.820	0.814	0.017	0.016	0.018	0.017
80	60	0.995	0.996	0.096	0.117	0.104	0.121
80	80	1	1	0.226	0.251	0.236	0.263

5 Conclusions and Discussion

This paper has developed an asymptotic theory for the largest eigenvalues of the covariance matrix of a high-dimensional time series vector. As an application, a new unit root test developed for testing nonstationarity in high-dimensional time series vectors has been proposed and then discussed both theoretically and numerically. The small sample properties discussed in Section 4 have offered the support to the theory established in Sections 2 and 3.

6 Acknowledgments

We would like to thank Professor Yoosoon Chang for kindly providing us with the code of simulations for [7].

A Results for Truncated Matrices

This section is to consider the truncated version of the sample covariance matrix. To this end, define

$$Y_{ij,l} = \sum_{k=0}^l b_k Z_{i-k,j}$$

with $l = \max\{p, T\}$, a truncated version of Y_{tj} in (2.1). However, to simplify notation, we let $b_i = 0$ for all $i > l$ in this section, so that we still use Y_{ij} instead of $Y_{ij,l}$. In this way a_i defined in (2.8) and Y_{tj} in (2.1) respectively become

$$a_i = \sum_{k=0}^{l-i} b_k b_{k+i}, \quad Y_{tj} = \sum_{k=0}^l b_k Z_{t-k,j}.$$

Furthermore let $\mathbf{F} = (F_{ij})$ be a $T \times (T+l)$ matrix with

$$F_{ij} = \begin{cases} b_{l+i-j} & i \leq j \leq i+l, \\ 0 & \text{otherwise.} \end{cases} \quad (\text{A.1})$$

It follows that $\mathbf{Y} = \mathbf{F}\mathbf{Z}_{\mathbf{p}}$, where $\mathbf{Z}_{\mathbf{p}}$ is a $(T+l) \times p$ random matrix with $(\mathbf{Z}_{\mathbf{p}})_{i,j} = Z_{i-l,j}$. For the sake of notation simplicity, we below denote $\mathbf{Z}_{\mathbf{p}}$ by \mathbf{Z} and $(\mathbf{Z}_{\mathbf{p}})_{i,j}$ by Z_{ij} . Let $\mathbf{A} = (A_{ij})_{T \times T} = (a_{|i-j|})_{T \times T}$. We then have $\mathbf{A} = \mathbf{F}\mathbf{F}'$. We would remind readers that l depends on T so that $a_{|i-j|}$ depends on T .

We also assume that $\mathbf{x}_0 = \mathbf{0}$ in this section.

A.1 Upper bound of the spectral norm of \mathbf{B} from stationary data

This subsection is to investigate the upper bound of the spectral norm of \mathbf{B} from stationary data.

Proposition 2. *Suppose that Assumptions A1-A5 hold. When $0 \leq \|\mathbf{\Pi}\|_2 = \varphi < 1$,*

$$\lim_{T \rightarrow \infty} P(\|\mathbf{B}\|_2 \leq \frac{8 \sum_{i \geq 0} |a_i|}{(1-\varphi)^2} M_0 (1 + \sqrt{\frac{T}{p}})^2) = 1.$$

Proof of Proposition 2: By (2.2) we may write

$$\mathbf{x}_t = \sum_{k_1=0}^{t-1} \mathbf{\Pi}^{k_1} \mathbf{\Sigma}^{1/2} \mathbf{y}_{t-k_1}.$$

This, together with (2.3), implies that

$$\begin{aligned} \frac{1}{p} \mathbf{X}' \mathbf{X} &= \frac{1}{p} \sum_{t=1}^T \mathbf{x}_t \mathbf{x}_t' = \frac{1}{p} \sum_{t=1}^T \sum_{k_1=0}^{t-1} \sum_{k_2=0}^{t-1} \mathbf{\Pi}^{k_1} \mathbf{\Sigma}^{1/2} \mathbf{y}_{t-k_1} \mathbf{y}_{t-k_2}' \mathbf{\Sigma}^{1/2} \mathbf{\Pi}^{k_2} \\ &= \frac{1}{p} \sum_{k_1=0}^{T-1} \sum_{k_2=0}^{T-1} \mathbf{\Pi}^{k_1} \mathbf{\Sigma}^{1/2} \left(\sum_{t=\max(k_1, k_2)+1}^T \mathbf{y}_{t-k_1} \mathbf{y}_{t-k_2}' \right) \mathbf{\Sigma}^{1/2} \mathbf{\Pi}^{k_2}. \end{aligned}$$

Note that

$$\sum_{t=\max(k_1, k_2)+1}^T \mathbf{y}_{t-k_1} \mathbf{y}_{t-k_2}' = \mathbf{Y}' \tilde{\mathbf{C}}_{k_1}' \tilde{\mathbf{C}}_{k_2} \mathbf{Y}$$

where $\tilde{\mathbf{C}}_k$ is a $T \times T$ matrix with elements $\tilde{C}_{k,ij} = I(i - j = k)$. It's easy to know $\|\tilde{\mathbf{C}}_k\|_2 \leq 1$. We then conclude that

$$\begin{aligned} \left\| \frac{1}{p} \mathbf{X}' \mathbf{X} \right\|_2 &\leq \sum_{k_1=0}^{T-1} \sum_{k_2=0}^{T-1} \varphi^{k_1+k_2} \left\| \frac{1}{p} \mathbf{Y}' \mathbf{Y} \right\|_2 \|\mathbf{\Sigma}\|_2 \leq \frac{M_0}{(1-\varphi)^2} \left\| \frac{1}{p} \mathbf{Y}' \mathbf{Y} \right\|_2 \\ &\leq \frac{M_0}{(1-\varphi)^2} \left\| \frac{1}{p} \mathbf{Z}' \mathbf{Z} \right\|_2 \|\mathbf{A}\|_2. \end{aligned} \quad (\text{A.2})$$

As we know the matrix \mathbf{A} is a Toeplitz matrix so that $\|\mathbf{A}\|_2 \leq 2 \sum_{i \geq 0} |a_i|$ (see [19]). From the assumption (A4) and the results in [8] and [1] we have

$$\lim_{T \rightarrow \infty} P\left(\left\| \frac{1}{p} \mathbf{Z}' \mathbf{Z} \right\|_2 \leq 4\left(1 + \sqrt{\frac{T}{p}}\right)^2\right) = 1. \quad (\text{A.3})$$

Proposition then follows from Lemma 1, the assumption (A3) and (A.2). \square

A.2 Eigenvalues of \mathbf{CAC}^*

This subsection and the following subsections are to consider the case $\mathbf{\Pi} = \mathbf{I}$. In this case one has

$$\mathbf{B} = (1/p) \mathbf{X} \mathbf{X}' = (1/p) \mathbf{C} \mathbf{Y} \mathbf{\Sigma} \mathbf{Y}' \mathbf{C}' = (1/p) \mathbf{C} \mathbf{F} \mathbf{Z}_p \mathbf{\Sigma} \mathbf{Z}_p' \mathbf{F}' \mathbf{C}'. \quad (\text{A.4})$$

We below investigate the eigenvalues and eigenvectors of $\mathbf{C} \mathbf{F} \mathbf{F}' \mathbf{C}' = \mathbf{CAC}^*$ at first. These are crucial steps.

Since it's very hard to find the eigenvalues of \mathbf{CAC}^* directly we use the following strategy. At first, we note that the eigenvalues of \mathbf{CAC}^* and $\mathbf{AC}^* \mathbf{C}$ are the same. We obtain the eigenvalues and eigenvectors of $\mathbf{C}^* \mathbf{C}$ by first studying $(\mathbf{C}^* \mathbf{C})^{-1}$. The next key step is to construct a matrix \mathbf{A}_m defined in (A.27) below which has the same eigenvectors as $\mathbf{C}^* \mathbf{C}$. In the mean time, it is easier to find the eigenvalues of $\mathbf{A}_m \mathbf{C}^* \mathbf{C}$. We then use the eigenvalues of $\mathbf{CA}_m \mathbf{C}^*$ to approximate those of \mathbf{CAC}^* . Our results are summarized in the following series of Lemmas and Theorems.

The first two lemmas describe the eigenvalues of $\mathbf{C}^* \mathbf{C}$ and decide their limits.

Lemma 1. Let $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_T \geq 0$ be the eigenvalues of $\mathbf{C}^*\mathbf{C}$. We then have

$$\lambda_k = \frac{1}{2(1 + \cos \theta_k)}, \quad \theta_k = \frac{2(T+1-k)\pi}{2T+1}, \quad k = 1, \dots, T. \quad (\text{A.5})$$

Lemma 2. Using the notation in Lemma 1,

$$\lim_{T \rightarrow \infty} \frac{\lambda_k}{T^2} = \frac{4}{\pi^2(2k-1)^2} \quad (\text{A.6})$$

for any fixed k .

Lemma 3 below specifies the eigenvectors of $\mathbf{C}^*\mathbf{C}$.

Lemma 3. Let $\tilde{\mathbf{x}}_k = (x_{k,1}, \dots, x_{k,T})'$ be a $T \times 1$ vector with

$$x_{k,i} = (-1)^{T-i} \sin(T-i+1)\theta_k, \quad -l \leq i \leq T+l. \quad (\text{A.7})$$

Then $\{\tilde{\mathbf{x}}_k, 1 \leq k \leq T\}$ are orthogonal and satisfy for any k

$$\mathbf{C}^*\mathbf{C}\tilde{\mathbf{x}}_k = \lambda_k \tilde{\mathbf{x}}_k. \quad (\text{A.8})$$

Lemma 4 below specifies the eigenvalues of $\mathbf{A}_m \mathbf{C}^*\mathbf{C}$ and gives their approximation to those of $\mathbf{A} \mathbf{C}^*\mathbf{C}$.

Lemma 4. Define γ_k by

$$\gamma_k = \lambda_k(a_0 + 2 \sum_{1 \leq j \leq T-1} a_j (-1)^j \cos(j\theta_k)). \quad (\text{A.9})$$

For any fixed constant $k \geq 1$, there is a constant c_k such that

$$\lim_{T \rightarrow \infty} \frac{\gamma_k}{T^2} = c_k > 0 \quad (\text{A.10})$$

and

$$\lim_{T \rightarrow \infty} \frac{\gamma_k}{\gamma_1} = \lim_{T \rightarrow \infty} \frac{\lambda_k}{\lambda_1} = \frac{1}{(2k-1)^2}. \quad (\text{A.11})$$

Let $\beta_1 \geq \beta_2 \geq \dots \geq \beta_T$ be the eigenvalues of $\mathbf{A} \mathbf{C}^*\mathbf{C}$. If \mathbf{A} satisfies the assumptions (A1) and (A2), then for any fixed integers $i \geq 1$ and $j \geq 1$ the following holds

$$\left| \frac{\beta_i - \gamma_i}{\gamma_j} \right| = O(T^{-1}). \quad (\text{A.12})$$

For any $\epsilon > 0$ there exists T_0 and k_0 where k_0 is a fixed number independent of T such that when $T \geq T_0$ and $k \geq k_0$,

$$\left| \frac{\beta_k}{\gamma_1} \right| \leq \epsilon. \quad (\text{A.13})$$

Lemma 5. Suppose that \mathbf{A} satisfies the assumptions (A1) and (A2). Then

$$\text{tr}(\mathbf{A} \mathbf{C}^*\mathbf{C}) = a_0 \frac{(T+1)T}{2} + \sum_{1 \leq j \leq T-1} a_j (T-j+1)(T-j) \quad (\text{A.14})$$

and

$$\lim_{T \rightarrow \infty} \frac{\beta_k}{\text{tr}(\mathbf{A} \mathbf{C}^*\mathbf{C})} = \lim_{T \rightarrow \infty} \frac{\gamma_k}{\text{tr}(\mathbf{A} \mathbf{C}^*\mathbf{C})} = \frac{8}{\pi^2(2k-1)^2}. \quad (\text{A.15})$$

Lemma 6. Suppose that \mathbf{A} satisfies the assumptions (A1) and (A2). For any $\epsilon > 0$, we can find T_0 and k_0 , where k_0 is a finite number independent of T , such that when $T \geq T_0$,

$$\left| \frac{\sum_{k>k_0} \beta_k}{\gamma_1} \right| < \epsilon. \quad (\text{A.16})$$

Proof of Lemma 1:

Let $\mathbf{M}_T = (\mathbf{C}^* \mathbf{C})^{-1}$. Define the characteristic function of \mathbf{M}_T by $g_T(\lambda) = \det(\lambda \mathbf{I}_T - \mathbf{M}_T)$. We can verify that the entries of the inverse matrix \mathbf{C}^{-1} , a $T \times T$ lower triangular matrix, are as following

$$C_{ij}^{-1} = \begin{cases} 1 & i = j, \\ -1 & i = j + 1, \\ 0 & \text{otherwise.} \end{cases} \quad (\text{A.17})$$

It follows that $M_{i,j}$, the elements of $\mathbf{M}_T = (\mathbf{C}^* \mathbf{C})^{-1}$, satisfy

$$M_{ij} = \begin{cases} 1 & i = j = 1, \\ 2 & i = j > 1, \\ -1 & |i - j| = 1, \\ 0 & \text{otherwise.} \end{cases} \quad (\text{A.18})$$

By the cofactor expansion we obtain a recurrence relation as following

$$g_T(\lambda) = (\lambda - 2)g_{T-1}(\lambda) - g_{T-2}(\lambda). \quad (\text{A.19})$$

Consider $\lambda \in (0, 4)$ at first. Hence we may write $\lambda = \lambda(\theta) = 2 + 2 \cos \theta$. We can further solve (A.19) to get

$$g_T(\lambda) = \frac{\sin T\theta + \sin(T+1)\theta}{\sin \theta}. \quad (\text{A.20})$$

When $\sin \theta \neq 0$, $g_T(\lambda) = 0$ is equivalent to

$$\sin T\theta + \sin(T+1)\theta = 0. \quad (\text{A.21})$$

Let $h_T(\theta) = \sin T\theta + \sin(T+1)\theta = 2 \sin(T+1/2)\theta \cos \frac{\theta}{2}$. Note that (A.5) gives T different solutions which satisfy $h_T(\theta) = 0$ and $\sin \theta \neq 0$. On the other hand, observe that there are at most T solutions for $g_T(\lambda) = 0$. The proof of (A.5) is complete. □

Lemmas 2 and 3 can be verified with some straightforward computations and the simple fact that

$$\sin(k+j)\theta + \sin(k-j)\theta = 2 \sin k\theta \cos j\theta. \quad (\text{A.22})$$

We ignore the details here.

Proof of Lemma 4:

Let's prove (A.10) and (A.11) at first. Note that

$$\left| (a_0 + 2 \sum_{1 \leq j \leq T-1} a_j (-1)^j \cos(j\theta_k)) - (a_0 + 2 \sum_{1 \leq j \leq \infty} a_j) \right| \leq 2 \sum_{1 \leq j \leq T-1} |a_j| \left| \cos\left(\frac{j(2k-1)\pi}{2T+1}\right) - 1 \right| + 2 \sum_{T \leq j} |a_j|.$$

For a fixed k , we can find a j_k to satisfy $\frac{\pi}{3} \leq \frac{j_k(2k-1)\pi}{2T+1} \leq \frac{\pi}{2}$. It follows that

$$\begin{aligned} & 2 \sum_{1 \leq j \leq j_k} |a_j| \left| \cos\left(\frac{j(2k-1)\pi}{2T+1}\right) - 1 \right| \leq 2 \sum_{1 \leq j \leq j_k} |a_j| \left(\frac{j(2k-1)\pi}{2T+1}\right)^2 \\ & \leq \frac{2j_k(2k-1)^2\pi^2}{(2T+1)^2} \sum_{1 \leq j \leq j_k} j|a_j| \leq \frac{(2k-1)\pi^2}{(2T+1)} \sum_{1 \leq j \leq \infty} j|a_j| \end{aligned}$$

and that

$$\begin{aligned} & 2 \sum_{j_k < j \leq T-1} |a_j| \left| \cos\left(\frac{j(2k-1)\pi}{2T+1}\right) - 1 \right| + 2 \sum_{T \leq j} |a_j| \leq 4 \sum_{j \geq j_k} |a_j| \\ & \leq j_k^{-1} 4 \sum_{j \geq j_k} j|a_j| \leq \frac{3(2k-1)}{2T+1} \sum_{1 \leq j \leq \infty} j|a_j|. \end{aligned}$$

From the assumption (A2), (B.4) and truncation conditions we can find

$$\lim_{T \rightarrow \infty} (a_0 + 2 \sum_{1 \leq j \leq T-1} a_j (-1)^j \cos(j\theta_k)) = \lim_{T \rightarrow \infty} (a_0 + 2 \sum_{1 \leq j \leq \infty} a_j) = \left(\sum_{i=0}^{\infty} b_i\right)^2 = s^2 > 0. \quad (\text{A.23})$$

In view of (A.6), (A.9) and (A.23), we can prove (A.10) and (A.11).

Now we consider the eigenvalues of $\mathbf{A}\mathbf{C}^*\mathbf{C}$. From (A.5)

$$\sin(T-i)\theta_k = -\sin(T+i+1)\theta_k. \quad (\text{A.24})$$

In view of (A.7) and (A.24), we obtain

$$x_{k,i} = x_{k,1-i}, \quad -T \leq i \leq 0 \quad (\text{A.25})$$

and

$$x_{k,i} = -x_{k,2T+2-i}, \quad T+2 \leq i \leq 2T. \quad (\text{A.26})$$

We construct a new matrix \mathbf{A}_m whose s th row, $\mathbf{a}_{m,s}$, satisfies that

$$\mathbf{a}_{m,s}\tilde{\mathbf{x}}_k = a_0 x_{k,s} + \sum_{1 \leq j \leq T-1} a_j (x_{k,s-j} + x_{k,s+j}) = (a_0 + 2 \sum_{1 \leq j \leq T-1} a_j (-1)^j \cos j\theta_k) x_{k,s}. \quad (\text{A.27})$$

Let \mathbf{a}_s be the s th row of \mathbf{A} . We can find that

$$\mathbf{a}_s\tilde{\mathbf{x}}_k = a_0 x_{k,s} + \sum_{1 \leq j \leq s-1} a_j x_{k,s-j} + \sum_{1 \leq j \leq T-s} a_j x_{k,s+j}. \quad (\text{A.28})$$

We further define $T \times T$ matrix \mathbf{A}_1 by

$$(\mathbf{A}_1)_{ij} = \begin{cases} a_{i+j-1} & i+j \leq T, \\ -a_{2T-i-j+2} & i+j \geq T+3, \\ 0 & T+1 \leq i+j \leq T+2. \end{cases} \quad (\text{A.29})$$

Let

$$\mathbf{A}_m = \mathbf{A} + \mathbf{A}_1. \quad (\text{A.30})$$

One can verify

$$\mathbf{A}_m\tilde{\mathbf{x}}_k = (a_0 + 2 \sum_{1 \leq j \leq T-1} a_j (-1)^j \cos j\theta_k) \tilde{\mathbf{x}}_k. \quad (\text{A.31})$$

It follows that

$$\mathbf{A}_m \mathbf{C}^* \mathbf{C} \tilde{\mathbf{x}}_k = \lambda_k \mathbf{A}_m \tilde{\mathbf{x}}_k = \gamma_k \tilde{\mathbf{x}}_k, \quad (\text{A.32})$$

which implies that γ_k is the eigenvalues of $\mathbf{A}_m \mathbf{C}^* \mathbf{C}$.

Now we consider $\mathbf{C} \mathbf{A}_1 \mathbf{C}^*$. It is easily seen that,

$$\|\mathbf{C} \mathbf{A}_1 \mathbf{C}^*\|_2 \leq T \max_{i,j} \{ |(\mathbf{C} \mathbf{A}_1 \mathbf{C}^*)_{i,j}| \}. \quad (\text{A.33})$$

Recalling (A.29) we can find that

$$\max_{i,j} \{ |(\mathbf{C} \mathbf{A}_1 \mathbf{C}^*)_{i,j}| \} \leq 2 \sum_{i=1}^{T-1} i |a_i|. \quad (\text{A.34})$$

We conclude from (A.33) that

$$\|\mathbf{C} \mathbf{A}_1 \mathbf{C}^*\|_2 \leq 2 \sum_{i=0}^{T-1} i |a_i| T.$$

In view of (B.4), we have

$$\|\mathbf{C} \mathbf{A}_1 \mathbf{C}^*\|_2 \leq 2 \sum_{i=0}^{T-1} i |a_i| T = O(T). \quad (\text{A.35})$$

Let $\tilde{\gamma}_1 \geq \tilde{\gamma}_2 \geq \dots \geq \tilde{\gamma}_T$ be the eigenvalues of $\mathbf{C} \mathbf{A}_m \mathbf{C}^*$. For fixed integers i , β_i is the i th largest eigenvalue of $\mathbf{C} \mathbf{A} \mathbf{C}^*$. It follows that

$$\left| \frac{\beta_i - \tilde{\gamma}_i}{\gamma_j} \right| \leq \frac{\|\mathbf{C}(\mathbf{A}_m - \mathbf{A})\mathbf{C}^*\|_2}{\gamma_j} = \frac{\|\mathbf{C} \mathbf{A}_1 \mathbf{C}^*\|_2}{\gamma_j}. \quad (\text{A.36})$$

From (A.9) and (A.11) we can find T_i for any fixed i such that when $T > T_i$, $\tilde{\gamma}_i = \gamma_i$. By (A.10) and (A.35) we can prove (A.12).

(A.13) follows from Lemma 6 directly. □

Proof of Lemma 5:

One can verify (A.14) with some computation. Observe that

$$\left| a_0 + \sum_{1 \leq j \leq T-1} a_j \frac{(T-j+1)(T-j)}{\frac{(T+1)T}{2}} - (a_0 + 2 \sum_{1 \leq j \leq \infty} a_j) \right| = O(T^{-1}).$$

This, together with (A.6), (A.9), (A.23), (A.12) and the assumption (A2), implies (A.15). □

Proof of Lemma 6:

Observe that

$$\sum_{k=1}^{\infty} \frac{1}{(2k-1)^2} = \frac{3}{4} \sum_{k=1}^{\infty} \frac{1}{k^2} = \frac{\pi^2}{8}. \quad (\text{A.37})$$

For any $\epsilon > 0$, we can find k_0 such that

$$\left| \sum_{k=1}^{k_0} \frac{1}{(2k-1)^2} - \frac{\pi^2}{8} \right| < \frac{\epsilon}{3}. \quad (\text{A.38})$$

From (A.11), (A.12) and (A.15), we can also find T_0 such that when $T \geq T_0$,

$$\left| \sum_{k=1}^{k_0} \frac{1}{(2k-1)^2} - \frac{\sum_{k=1}^{k_0} \beta_k}{\gamma_1} \right| < \frac{\epsilon}{3}. \quad (\text{A.39})$$

and

$$\left| \frac{\text{tr}(\mathbf{A}\mathbf{C}^*\mathbf{C})}{\gamma_1} - \frac{\pi^2}{8} \right| < \frac{\epsilon}{3}. \quad (\text{A.40})$$

It follows from (A.38)-(A.40) that

$$\left| \frac{\sum_{k>k_0} \beta_k}{\gamma_1} \right| = \left| \frac{\text{tr}(\mathbf{A}\mathbf{C}^*\mathbf{C})}{\gamma_1} - \frac{\sum_{k=1}^{k_0} \beta_k}{\gamma_1} \right| < \epsilon. \quad (\text{A.41})$$

□

A.3 Eigenvectors of $\mathbf{C}\mathbf{A}\mathbf{C}^*$

This section is to investigate the eigenvectors of $\mathbf{C}\mathbf{A}\mathbf{C}^*$. At first we normalize $\{\tilde{\mathbf{x}}_{\mathbf{k}}\}_{1 \leq k \leq T}$ to get $\{\tilde{\mathbf{y}}_{\mathbf{k}}\}_{1 \leq k \leq T}$. Then we study the eigenvectors of $\mathbf{A}\mathbf{C}^*\mathbf{C}$ by representing them with $\{\tilde{\mathbf{y}}_{\mathbf{k}}\}_{1 \leq k \leq T}$. At last we give some result about the eigenvectors of $\mathbf{C}\mathbf{A}\mathbf{C}^*$ which is necessary for the future proof in Section 3.4. Our results are the following.

Lemma 7. *Recall the eigenvectors $\tilde{\mathbf{x}}_{\mathbf{k}}$ defined in Lemma 3. Then*

$$\sum_{j=1}^T (x_{k,j})^2 = \frac{2T+1}{4}. \quad (\text{A.42})$$

Let

$$\tilde{\mathbf{y}}_{\mathbf{k}} = \frac{\tilde{\mathbf{x}}_{\mathbf{k}}}{\|\tilde{\mathbf{x}}_{\mathbf{k}}\|}. \quad (\text{A.43})$$

Then $\{\tilde{\mathbf{y}}_{\mathbf{k}}\}_{1 \leq k \leq T}$ are orthogonal and the j th element of $\tilde{\mathbf{y}}_{\mathbf{k}}$, $y_{k,j}$, satisfies

$$|y_{k,j}| = \frac{|x_{k,j}|}{\sqrt{\frac{2T+1}{4}}} \leq \frac{2}{\sqrt{2T+1}}. \quad (\text{A.44})$$

Lemma 8. *Let $\{\mathbf{u}_{\mathbf{k}}\}_{1 \leq k \leq T}$ be orthogonal and real vectors such that $\|\mathbf{u}_{\mathbf{k}}\| = 1$ and*

$$\mathbf{C}\mathbf{A}\mathbf{C}^*\mathbf{u}_{\mathbf{k}} = \beta_k \mathbf{u}_{\mathbf{k}}. \quad (\text{A.45})$$

Define $\mathbf{f}_{\mathbf{k}} = \frac{\mathbf{C}^{-1}\mathbf{u}_{\mathbf{k}}}{\|\mathbf{C}^{-1}\mathbf{u}_{\mathbf{k}}\|}$ such that

$$\mathbf{f}_{\mathbf{k}} = \sum_{j=1}^T \alpha_{kj} \mathbf{y}_j \quad (\text{A.46})$$

with

$$\sum_{j=1}^T \alpha_{kj}^2 = 1. \quad (\text{A.47})$$

Then when $k \geq 1$ is fixed,

$$\frac{\alpha_{kk}^2 \lambda_k}{\sum_{j=1}^T \alpha_{kj}^2 \lambda_j} = 1 + O(T^{-1}), \quad (\text{A.48})$$

where $\{\lambda_j\}$ are given in Lemma 1.

Lemma 9. *Let $(S_{k,1}, \dots, S_{k,T+l})' = \mathbf{s}_{\mathbf{k}} = \frac{\mathbf{F}^*\mathbf{C}^*\mathbf{u}_{\mathbf{k}}}{\sqrt{\gamma_1}}$. Then $\{\mathbf{s}_{\mathbf{k}}\}_{1 \leq k \leq T}$ are orthogonal and*

$$\sum_{j=1}^{T+l} S_{k,j}^4 = O(T^{-1}). \quad (\text{A.49})$$

Proof of Lemma 7: From (A.7) we obtain

$$|x_{k,j}| \leq 1.$$

Lemma 7 can be then proved with some straightforward computations. We ignore details here. □

Proof of Lemma 8: From $\mathbf{f}_k = \frac{\mathbf{C}^{-1}\mathbf{u}_k}{\|\mathbf{C}^{-1}\mathbf{u}_k\|}$ and (A.45), we have $\|\mathbf{f}_k\| = 1$ and

$$\mathbf{A}\mathbf{C}^*\mathbf{C}\mathbf{f}_k = \beta_k\mathbf{f}_k. \quad (\text{A.50})$$

From (A.30) and (A.50), we have

$$\beta_k = \frac{\mathbf{f}_k^*\mathbf{C}^*\mathbf{C}\mathbf{A}\mathbf{C}^*\mathbf{C}\mathbf{f}_k}{\|\mathbf{C}\mathbf{f}_k\|^2} = \frac{\mathbf{f}_k^*\mathbf{C}^*\mathbf{C}(\mathbf{A}_m - \mathbf{A}_1)\mathbf{C}^*\mathbf{C}\mathbf{f}_k}{\|\mathbf{C}\mathbf{f}_k\|^2}.$$

It follows that

$$\frac{|\mathbf{f}_k^*\mathbf{C}^*\mathbf{C}\mathbf{A}_m\mathbf{C}^*\mathbf{C}\mathbf{f}_k| - |\mathbf{f}_k^*\mathbf{C}^*\mathbf{C}\mathbf{A}_1\mathbf{C}^*\mathbf{C}\mathbf{f}_k|}{\|\mathbf{C}\mathbf{f}_k\|^2} \leq \beta_k \leq \frac{|\mathbf{f}_k^*\mathbf{C}^*\mathbf{C}\mathbf{A}_m\mathbf{C}^*\mathbf{C}\mathbf{f}_k| + |\mathbf{f}_k^*\mathbf{C}^*\mathbf{C}\mathbf{A}_1\mathbf{C}^*\mathbf{C}\mathbf{f}_k|}{\|\mathbf{C}\mathbf{f}_k\|^2}. \quad (\text{A.51})$$

By (A.8), (A.43) and (A.46), we have

$$\|\mathbf{C}\mathbf{f}_k\| = \sqrt{\sum_{j=1}^T \alpha_{kj}^2 \lambda_j}. \quad (\text{A.52})$$

(A.8), (A.32), (A.43) and (A.46) imply that

$$\mathbf{C}^*\mathbf{C}\mathbf{A}_m\mathbf{C}^*\mathbf{C}\mathbf{f}_k = \mathbf{C}^*\mathbf{C}\sum_{j=1}^T \alpha_{kj} \mathbf{A}_m \mathbf{C}^*\mathbf{C}\tilde{\mathbf{y}}_j = \mathbf{C}^*\mathbf{C}\sum_{j=1}^T \alpha_{kj} \gamma_j \tilde{\mathbf{y}}_j = \sum_{j=1}^T \alpha_{kj} \gamma_j \lambda_j \tilde{\mathbf{y}}_j.$$

This and (A.46) ensure

$$\mathbf{f}_k^*\mathbf{C}^*\mathbf{C}\mathbf{A}_m\mathbf{C}^*\mathbf{C}\mathbf{f}_k = \sum_{j=1}^T \alpha_{kj}^2 \gamma_j \lambda_j. \quad (\text{A.53})$$

From (A.35), we have

$$\frac{|\mathbf{f}_k^*\mathbf{C}^*\mathbf{C}\mathbf{A}_1\mathbf{C}^*\mathbf{C}\mathbf{f}_k|}{\|\mathbf{C}\mathbf{f}_k\|^2} \leq \|\mathbf{C}\mathbf{A}_1\mathbf{C}^*\|_2 = O(T).$$

This, together with (A.51)-(A.53), implies that

$$\frac{\sum_{j=1}^T \alpha_{kj}^2 \gamma_j \lambda_j}{\sum_{j=1}^T \alpha_{kj}^2 \lambda_j} - O(T) \leq \beta_k \leq \frac{\sum_{j=1}^T \alpha_{kj}^2 \gamma_j \lambda_j}{\sum_{j=1}^T \alpha_{kj}^2 \lambda_j} + O(T).$$

By Lemma 4, for any fixed k we have

$$\sum_{j=1}^T \frac{\alpha_{kj}^2 \lambda_j}{\sum_{j=1}^T \alpha_{kj}^2 \lambda_j} \frac{\gamma_j}{\beta_k} - O(T^{-1}) \leq 1 \leq \sum_{j=1}^T \frac{\alpha_{kj}^2 \lambda_j}{\sum_{j=1}^T \alpha_{kj}^2 \lambda_j} \frac{\gamma_j}{\beta_k} + O(T^{-1}). \quad (\text{A.54})$$

Note that $\{\mathbf{u}_k\}_{1 \leq k \leq T}$ are orthogonal and $\{\tilde{\mathbf{y}}_k\}_{1 \leq k \leq T}$ are orthogonal. When $k \neq m$, from (A.8), (A.43) and (A.46) we have

$$0 = U_k^* U_m = \frac{\mathbf{f}_k^*\mathbf{C}^*\mathbf{C}\mathbf{f}_m}{\|\mathbf{C}\mathbf{f}_k\| \|\mathbf{C}\mathbf{f}_m\|} = \frac{\sum_{j=1}^T \alpha_{kj} \alpha_{mj} \lambda_j}{\|\mathbf{C}\mathbf{f}_k\| \|\mathbf{C}\mathbf{f}_m\|}.$$

This implies that

$$\sum_{j=1}^T \alpha_{kj} \alpha_{mj} \lambda_j = 0. \quad (\text{A.55})$$

Moreover let $v_{kj} = \frac{\alpha_{kj} \sqrt{\lambda_j}}{\sqrt{\sum_{j=1}^T \alpha_{kj}^2 \lambda_j}}$. We have

$$\sum_{j=1}^T v_{kj}^2 = 1. \quad (\text{A.56})$$

Note that (A.54) is equivalent to

$$\sum_{j=1}^T v_{kj}^2 \frac{\gamma_j}{\beta_k} - O(T^{-1}) \leq 1 \leq \sum_{j=1}^T v_{kj}^2 \frac{\gamma_j}{\beta_k} + O(T^{-1}). \quad (\text{A.57})$$

Also (A.55) implies that

$$\sum_{j=1}^T v_{kj} v_{mj} = 0. \quad (\text{A.58})$$

We consider v_{kj} for fixed k below. When $k = 1$ and T is big enough, Lemma 4, (A.56) and (A.57) imply

$$O(T^{-1}) = |1 - \sum_{j=1}^T v_{1j}^2 \frac{\gamma_j}{\beta_1}| \geq (1 - v_{11}^2) \frac{\beta_1 - \gamma_2}{\beta_1} - v_{11}^2 \frac{|\beta_1 - \gamma_1|}{\beta_1}. \quad (\text{A.59})$$

In view of (A.10)-(A.12), we have $\frac{\beta_1 - \gamma_1}{\beta_1} = O(T^{-1})$ and $\frac{\beta_1 - \gamma_2}{\beta_1} = \frac{8}{9} + o(1)$. It follows that (A.59) implies that $v_{11}^2 = 1 + O(T^{-1})$ and $\sum_{j=2}^T v_{1j}^2 = O(T^{-1})$. From (A.58), for any $k \neq 1$ we have

$$|v_{k1} v_{11}| = |\sum_{j=2}^T v_{kj} v_{1j}| \leq \sqrt{\sum_{j=2}^T v_{kj}^2} \sqrt{\sum_{j=2}^T v_{1j}^2} = O(T^{-1/2}). \quad (\text{A.60})$$

This implies $v_{k1}^2 = O(T^{-1})$. It's similar to obtain that $v_{22}^2 = 1 + O(T^{-1})$ and $v_{k2}^2 = O(T^{-1})$ for any $k \neq 2$.

By repeating these steps we conclude that $v_{kk}^2 = 1 + O(T^{-1})$ for any fixed k . This implies (A.48). \square

Proof of Lemma 9: Note that $\{\mathbf{s}_k\}_{1 \leq k \leq T}$ are orthogonal and real due to orthogonality of $\{\mathbf{u}_k\}_{1 \leq k \leq T}$. We conclude from (A.8) and (A.46) that

$$\mathbf{s}_k = \frac{\mathbf{F}^* \mathbf{C}^* \mathbf{C} \mathbf{f}_k}{\sqrt{\gamma_1} \|\mathbf{C} \mathbf{f}_k\|} = \frac{1}{\sqrt{\gamma_1} \|\mathbf{C} \mathbf{f}_k\|} \sum_{j=1}^T \alpha_{kj} \lambda_j \mathbf{F}^* \tilde{\mathbf{y}}_j = \mathbf{s}_{k,M} + \mathbf{s}_{k,R}, \quad (\text{A.61})$$

where

$$\mathbf{s}_{k,M} = \frac{1}{\sqrt{\gamma_1} \|\mathbf{C} \mathbf{f}_k\|} \alpha_{kk} \lambda_k \mathbf{F}^* \tilde{\mathbf{y}}_k, \quad \mathbf{s}_{k,R} = \frac{1}{\sqrt{\gamma_1} \|\mathbf{C} \mathbf{f}_k\|} \sum_{j \neq k} \alpha_{kj} \lambda_j \mathbf{F}^* \tilde{\mathbf{y}}_j. \quad (\text{A.62})$$

By Hölder's inequality, we have

$$\|\mathbf{s}_{k,R}\| = \left\| \frac{1}{\sqrt{\gamma_1} \|\mathbf{C} \mathbf{f}_k\|} \sum_{j \neq k} \alpha_{kj} \lambda_j \mathbf{F}^* \tilde{\mathbf{y}}_j \right\| \leq \frac{1}{\sqrt{\gamma_1} \|\mathbf{C} \mathbf{f}_k\|} \|\mathbf{F}\|_2 \sqrt{\sum_{j \neq k} \alpha_{kj}^2 \lambda_j^2}.$$

Recalling $\mathbf{A} = \mathbf{F} \mathbf{F}^*$, we have

$$\|\mathbf{F}\|_2 = \sqrt{\|\mathbf{A}\|_2}.$$

Since A is a Hermitian Toeplitz matrix, from [19],

$$\|\mathbf{A}\|_2 \leq 2 \sum_{0 \leq k \leq l} |a_k|.$$

By (B.4) we can get

$$\|\mathbf{F}\|_2 = \sqrt{\|\mathbf{A}\|_2} < \infty.$$

From Lemma 2, (A.48), and (A.52) we can obtain that for any fixed k ,

$$\frac{\sqrt{\sum_{j \neq k} \alpha_{kj}^2 \lambda_j^2}}{\|\mathbf{C} \mathbf{f}_k\|} \leq \sqrt{\frac{\sum_{j \neq k} \alpha_{kj}^2 \lambda_j}{\sum_{j=1}^T \alpha_{kj}^2 \lambda_j}} \sqrt{\lambda_1} = O(T^{1/2}).$$

This, together with (A.10), implies that for any fixed k ,

$$\|\mathbf{s}_{k,R}\| = O(T^{-1/2}). \quad (\text{A.63})$$

Similarly, we can also obtain that $\frac{1}{\sqrt{\gamma_1}\|\mathbf{C}\mathbf{f}_k\|}\alpha_{kk}\lambda_k$ is bounded for any fixed k .

Let $S_{k,M,j}$ be the j th element of $\mathbf{s}_{k,M}$ and $S_{k,R,j}$ be the j th element of $\mathbf{s}_{k,R}$. From (A.44), (A.48) and (A.62) and the assumption (A1) we can obtain that for any fixed k ,

$$|S_{k,M,j}| \leq \frac{1}{\sqrt{\gamma_1}\|\mathbf{C}\mathbf{f}_k\|} |\alpha_{kk}\lambda_k| \frac{2}{\sqrt{2T+1}} \sum_{h=0}^l |b_h| = O(T^{-1/2}). \quad (\text{A.64})$$

It follows from (A.61), (A.62) and (A.64) that for any fixed k ,

$$\begin{aligned} \sum_{j=1}^{T+l} S_{k,j}^4 &\leq 8 \sum_{j=1}^{T+l} (S_{k,R,j}^4 + S_{k,M,j}^4) \\ &\leq 8 \sum_{j=1}^{T+l} S_{k,M,j}^4 + 8 \left(\sum_{j=1}^{T+l} S_{k,R,j}^2 \right)^2 = O(T^{-1}). \end{aligned} \quad (\text{A.65})$$

□

A.4 Convergence in Probability

This section is to establish convergence in probability of the spiked eigenvalues of a kind of separable sample covariance matrices, which is enough for our purpose.

Lemma 10. *Let $\mathbf{D} = \frac{1}{p}\mathbf{W}\mathbf{Z}\mathbf{\Sigma}\mathbf{Z}^*\mathbf{W}^*$, where \mathbf{W} is a $T \times (T+l)$ matrix, $\mathbf{\Sigma}$ is a $p \times p$ positive-definite matrix with $\|\mathbf{\Sigma}\|_2 \leq M_0$, and \mathbf{Z} is defined below (A.1). Order the eigenvalues of $\mathbf{W}\mathbf{W}^*$ as $\tau_1 \geq \dots \geq \tau_T$ with τ_1 being bounded. Suppose that $\{\tau_k\}_{1 \leq k \leq T}$ satisfy the following conditions.*

(C1) *For any fixed k , there is a constant $c_k > 0$ such that*

$$\lim_{T \rightarrow \infty} \tau_k = c_k. \quad (\text{A.66})$$

(C2) *For any $\epsilon > 0$ there exist T_0 and k_0 , where k_0 is a constant independent of n and T , such that when $T \geq T_0$,*

$$\left| \sum_{k > k_0} \tau_k \right| < \epsilon. \quad (\text{A.67})$$

For any fixed k , denote the first k largest eigenvalues of \mathbf{D} by $\rho_1 \geq \dots \geq \rho_k$. Then $\rho_j - c_j \frac{\text{tr}(\mathbf{\Sigma})}{p} \rightarrow 0$ in probability.

Proof. We can find \mathbf{V} such that

$$\begin{aligned} \mathbf{V}\mathbf{W}^*\mathbf{W}\mathbf{V}^* &= \text{diag}(\tau_1, \dots, \tau_{T+l}) = \mathbf{\Lambda}_{\mathbf{T}+1}, \\ \mathbf{V}\mathbf{V}^* &= \mathbf{V}^*\mathbf{V} = \mathbf{I}_{\mathbf{T}+1}, \end{aligned}$$

where $\tau_k = 0$ when $k > T$. So $\mathbf{D}^* = \frac{1}{p}\mathbf{V}\mathbf{Z}\mathbf{\Sigma}\mathbf{Z}^*\mathbf{V}^*\mathbf{\Lambda}_{\mathbf{T}+1}$ has the same nonzero eigenvalues as $\frac{1}{p}\mathbf{Z}\mathbf{\Sigma}\mathbf{Z}^*\mathbf{W}^*\mathbf{W}$.

To prove the lemma, it suffices to prove that for any $\delta > 0$ and fix number k ,

$$\lim_{T \rightarrow \infty} P\left(\left|\rho_k - c_k \frac{\text{tr}(\mathbf{\Sigma})}{p}\right| > \delta\right) = 0. \quad (\text{A.68})$$

In view of (A.67), we can find $k_0 > 0$ such that

$$\left| \sum_{k>k_0} \tau_k \right| < \frac{\delta}{4M_0}. \quad (\text{A.69})$$

Write $\mathbf{\Lambda}_{\mathbf{T}+1} = \mathbf{\Lambda}_{\mathbf{T}+1}^{\mathbf{M}} + \mathbf{\Lambda}_{\mathbf{T}+1}^{\mathbf{R}}$, where

$$\begin{aligned} \mathbf{\Lambda}_{\mathbf{T}+1}^{\mathbf{M}} &= \text{diag}\{\tau_1, \tau_2, \dots, \tau_{k_0}, 0, \dots, 0\} \\ \mathbf{\Lambda}_{\mathbf{T}+1}^{\mathbf{R}} &= \text{diag}\{0, \dots, 0, \tau_{k_0+1}, \tau_{k_0+2}, \dots, \tau_{T+l}\}. \end{aligned} \quad (\text{A.70})$$

Let $h = \text{tr}(\frac{1}{p} \mathbf{VZZ}^* \mathbf{V}^* \mathbf{\Lambda}_{\mathbf{T}+1}^{\mathbf{R}})$. Note that $E(\mathbf{VZZ}^* \mathbf{V}^*)_{kk} = p$ and $\frac{\text{Var}(\mathbf{VZZ}^* \mathbf{V}^*)_{kk}}{p} < \infty$. We can evaluate the mean and variance of h as follows

$$E(h) = \sum_{k=k_0+1}^{T+l} E\left(\frac{(\mathbf{VZZ}^* \mathbf{V}^*)_{kk} \tau_k}{p}\right) = \sum_{k>k_0} \tau_k < \frac{\delta}{4M_0}$$

and

$$\begin{aligned} \text{Var}(h) &= \text{Var}\left(\sum_{k=k_0+1}^T \left(\frac{(\mathbf{VZZ}^* \mathbf{V}^*)_{kk} \tau_k}{p}\right)\right) \\ &\leq \left(\sum_{k=k_0+1}^T \sqrt{\text{Var}\left(\frac{(\mathbf{VZZ}^* \mathbf{V}^*)_{kk} \tau_k}{p}\right)}\right)^2 = O\left(\frac{(\sum_{k>k_0} \tau_k)^2}{p}\right) = O\left(\frac{1}{p}\right). \end{aligned}$$

Since $\mathbf{\Sigma}$ is positive-definite, we have $\|\frac{1}{p} \mathbf{VZZ}^* \mathbf{V}^* \mathbf{\Lambda}_{\mathbf{T}+1}^{\mathbf{R}}\|_2 \leq \text{tr}(\frac{1}{p} \mathbf{VZZ}^* \mathbf{V}^* \mathbf{\Lambda}_{\mathbf{T}+1}^{\mathbf{R}}) \leq M_0 h$. It follows that

$$\lim_{T \rightarrow \infty} P\left(\|\frac{1}{p} \mathbf{VZZ}^* \mathbf{V}^* \mathbf{\Lambda}_{\mathbf{T}+1}^{\mathbf{R}}\|_2 > \frac{\delta}{2}\right) = 0. \quad (\text{A.71})$$

Denote the k th largest eigenvalue of $\frac{1}{p} \mathbf{VZZ}^* \mathbf{V}^* \mathbf{\Lambda}_{\mathbf{T}+1}^{\mathbf{M}}$ by ρ_k^M . From the definition of $\mathbf{\Lambda}_{\mathbf{T}+1}^{\mathbf{M}}$, we conclude that $\frac{1}{n} \mathbf{VZZ}^* \mathbf{V}^* \mathbf{\Lambda}_{\mathbf{T}+1}^{\mathbf{M}}$ has the same nonzero eigenvalues as its upper left $k_0 \times k_0$ block. By Theorem 7.1 of [2] it's easy to prove that the limit of off-diagonal elements in the upper left $k_0 \times k_0$ block is 0 in probability. Recall that k_0 is a constant which doesn't depend on T . We conclude that the nonzero eigenvalues of $\frac{1}{p} \mathbf{VZZ}^* \mathbf{V}^* \mathbf{\Lambda}_{\mathbf{T}+1}^{\mathbf{M}}$ are diagonal elements of the upper left $k_0 \times k_0$ block. From Theorem 7.1 of [2], the limit of diagonal elements in the upper left $k_0 \times k_0$ block can be obtained as follow

$$\lim_{T \rightarrow \infty} P\left(|\rho_k^M - \tau_k \frac{\text{tr}(\mathbf{\Sigma})}{p}| > \frac{\delta}{2}\right) = 0. \quad (\text{A.72})$$

It follows from (A.64), (A.71) and (A.72) that

$$\begin{aligned} &\lim_{T \rightarrow \infty} P\left(|\rho_k - \lambda_k \frac{\text{tr}(\mathbf{\Sigma})}{p}| > \delta\right) \\ &\leq \lim_{T \rightarrow \infty} P\left(|\rho_k^M - \lambda_k \frac{\text{tr}(\mathbf{\Sigma})}{p}| + |\rho_k^M - \rho_k| > \delta\right) \\ &\leq \lim_{T \rightarrow \infty} P\left(|\rho_k^M - \lambda_k \frac{\text{tr}(\mathbf{\Sigma})}{p}| > \frac{\delta}{2}\right) + \lim_{T \rightarrow \infty} P\left(|\rho_k^M - \rho_k| > \frac{\delta}{2}\right) \\ &\leq \lim_{T \rightarrow \infty} P\left(|\rho_k^M - \lambda_k \frac{\text{tr}(\mathbf{\Sigma})}{p}| > \frac{\delta}{2}\right) + \lim_{T \rightarrow \infty} P\left(\|\frac{1}{p} \mathbf{VZZ}^* \mathbf{V}^* \mathbf{\Lambda}_{\mathbf{T}+1}^{\mathbf{R}}\|_2 > \frac{\delta}{2}\right) = 0. \end{aligned}$$

□

We apply Lemma 10 with $\mathbf{D} = \frac{\mathbf{B}}{\gamma_1}$ where \mathbf{B} is defined in (A.4). Lemma 4 and Lemma 6 ensure that conditions in Lemma 10 are satisfied so that Proposition 3 below holds.

Proposition 3. *Let ρ_k be the k th largest eigenvalue of \mathbf{B} . When $\varphi = 1$, $\frac{\rho_k - \gamma_k \frac{\text{tr}(\mathbf{\Sigma})}{p}}{\gamma_1} \rightarrow 0$ in probability.*

A.5 CLT of the first k largest eigenvalues

This section is to develop central limit theorems of the first k largest eigenvalues of \mathbf{B} .

Proposition 4. *Let ρ_k be the k th largest eigenvalue of \mathbf{B} . When $\varphi = 1$, $(\sqrt{p}\frac{\rho_1 - \gamma_1}{\gamma_1}, \dots, \sqrt{p}\frac{\rho_k - \gamma_k}{\gamma_1})'$ converges weakly to a zero-mean Gaussian vector $\mathbf{w} = (w_1, \dots, w_k)'$ with covariance $\text{cov}(w_i, w_j) = \delta_{ij} \frac{\theta}{(2i-1)^4} (2 - 4E(Z_{i1}^R)^2 E(Z_{i1}^I)^2)$ and $\theta = \lim_{p \rightarrow \infty} \frac{\text{tr}(\Sigma^2)}{p}$.*

Proof. Recalling the definitions of \mathbf{u}_k and \mathbf{s}_k in Section 3.2, we denote $(\mathbf{u}_1, \dots, \mathbf{u}_T)$ by \mathbf{U} and $\frac{\mathbf{F}^* \mathbf{C}^* \mathbf{U}}{\sqrt{\gamma_1}}$ by \mathbf{S} . Note that $\{\mathbf{u}_k\}_{1 \leq k \leq T}$ and $\{\mathbf{s}_k\}_{1 \leq k \leq T}$ are both orthogonal and real. Since $\mathbf{s}_j^* \mathbf{s}_i = 0$ for $i \neq j$ we have

$$\mathbf{S} \mathbf{S}^* = \mathbf{\Lambda} = \text{diag}\left\{\frac{\beta_1}{\gamma_1}, \dots, \frac{\beta_T}{\gamma_1}\right\}. \quad (\text{A.73})$$

In view of (A.4), let

$$\mathbf{D} = \frac{\mathbf{U}^* \mathbf{B} \mathbf{U}}{\gamma_1} = \frac{1}{p} \mathbf{S}^* \mathbf{Z} \Sigma \mathbf{Z}^* \mathbf{S}. \quad (\text{A.74})$$

The eigenvalues of \mathbf{D} are ordered as $\frac{\rho_T}{\gamma_1} \leq \dots \leq \frac{\rho_1}{\gamma_1}$.

To rewrite \mathbf{D} as a block matrix we first introduce the following notation. For a fixed number $k > 0$, let $\mathbf{z}_j = (Z_{1j}, \dots, Z_{(T+l)j})'$. Set $\mathbf{V}_1 = \frac{1}{\sqrt{p}}(\xi_1, \dots, \xi_p) = \frac{1}{\sqrt{p}} \mathbf{Q}_1 \mathbf{Z} = \frac{1}{\sqrt{p}}(\mathbf{s}_1^*, \dots, \mathbf{s}_k^*)' \mathbf{Z}$ and $\mathbf{V}_2 = \frac{1}{\sqrt{p}}(\eta_1, \dots, \eta_p) = \frac{1}{\sqrt{p}} \mathbf{Q}_2 \mathbf{Z} = \frac{1}{\sqrt{p}}(\mathbf{s}_{k+1}^*, \dots, \mathbf{s}_T^*)' \mathbf{Z}$ where $\mathbf{Q}_1 = (\mathbf{s}_1^*, \dots, \mathbf{s}_k^*)'$ and $\mathbf{Q}_2 = (\mathbf{s}_{k+1}^*, \dots, \mathbf{s}_T^*)'$. Then

$$\xi_j = (\xi_j(1), \dots, \xi_j(k))' = (\mathbf{s}_1^* \mathbf{z}_j, \dots, \mathbf{s}_k^* \mathbf{z}_j)' \quad (\text{A.75})$$

and

$$\eta_j = (\eta_j(k+1), \dots, \eta_j(p))' = (\mathbf{s}_{k+1}^* \mathbf{z}_j, \dots, \mathbf{s}_T^* \mathbf{z}_j)' \quad (\text{A.76})$$

Let $\mathbf{\Lambda}_1 = \text{cov}(\xi_j) = \mathbf{Q}_1 \mathbf{Q}_1^*$ and $\mathbf{\Lambda}_2 = \text{cov}(\eta_j) = \mathbf{Q}_2 \mathbf{Q}_2^*$. In view of (A.73), we have

$$\mathbf{\Lambda}_1 = \text{diag}\left\{\frac{\beta_1}{\gamma_1}, \dots, \frac{\beta_k}{\gamma_1}\right\}, \quad \mathbf{\Lambda}_2 = \text{diag}\left\{\frac{\beta_{k+1}}{\gamma_1}, \dots, \frac{\beta_T}{\gamma_1}\right\}. \quad (\text{A.77})$$

From Lemmas 4, 5 and (A.37) we can find a constant M_k such that

$$\lim_{T \rightarrow \infty} |\text{tr}(\mathbf{\Lambda}_2)| = \lim_{T \rightarrow \infty} |\text{tr}(\mathbf{\Lambda}) - \text{tr}(\mathbf{\Lambda}_1)| = \left| \frac{\pi^2}{8} - \sum_{j=1}^k \frac{1}{(2j-1)^2} \right| < M_k. \quad (\text{A.78})$$

In view of (A.74)-(A.76), we can rewrite \mathbf{D} as

$$\mathbf{D} = \begin{pmatrix} \mathbf{V}_1 \Sigma \mathbf{V}_1^* & \mathbf{V}_1 \Sigma \mathbf{V}_2^* \\ \mathbf{V}_2 \Sigma \mathbf{V}_1^* & \mathbf{V}_2 \Sigma \mathbf{V}_2^* \end{pmatrix} \triangleq \begin{pmatrix} \mathbf{W}_{11} & \mathbf{W}_{12} \\ \mathbf{W}_{21} & \mathbf{W}_{22} \end{pmatrix}. \quad (\text{A.79})$$

The characteristic polynomial of \mathbf{D} is

$$0 = |\lambda \mathbf{I}_T - \mathbf{D}| = |\lambda \mathbf{I}_{T-k} - \mathbf{W}_{22}| |\lambda \mathbf{I}_k - \mathbf{K}_p(\lambda)|, \quad (\text{A.80})$$

where

$$\mathbf{K}_p(\lambda) = \mathbf{W}_{11} + \mathbf{W}_{12}(\lambda \mathbf{I}_{T-k} - \mathbf{W}_{22})^{-1} \mathbf{W}_{21}. \quad (\text{A.81})$$

We conclude from Lemmas 4 and 6 that $\mathbf{W}_{22} = \mathbf{V}_2 \Sigma \mathbf{V}_2^* = \frac{1}{p} \mathbf{Q}_2 \mathbf{Z} \Sigma \mathbf{Z}^* \mathbf{Q}_2^*$ satisfies the conditions of Lemma 10. Lemma 10 immediately implies that the largest eigenvalue of \mathbf{W}_{22} , ρ , tends to $\frac{\gamma_{k+1} \frac{\text{tr}(\Sigma)}{p}}{\gamma_1}$

in probability. On the other hand, from Lemma 10, we also see that when $j \leq k$, $\frac{\rho_j - \gamma_j \frac{\text{tr}(\Sigma)}{p}}{\gamma_1} \rightarrow 0$ in probability. Since we want to study the first k largest eigenvalues, from Lemma 4 and (A.80), it's sufficient to consider the characteristic polynomial

$$0 = |\lambda \mathbf{I}_k - \mathbf{K}_p(\lambda)| = |\mathbf{G}(\lambda)|, \quad (\text{A.82})$$

where

$$\mathbf{G}(\lambda) = \{G_{ij}(\lambda)\}_{1 \leq i, j \leq k} = \lambda \mathbf{I}_k - \mathbf{K}_p(\lambda). \quad (\text{A.83})$$

From (A.81) we write

$$\begin{aligned} \mathbf{K}_p(\lambda) &= \mathbf{W}_{11} + \mathbf{W}_{12}(\lambda \mathbf{I}_{T-k} - \mathbf{W}_{22})^{-1} \mathbf{W}_{21} \\ &= \mathbf{V}_1 \Sigma \mathbf{V}_1^* + \mathbf{V}_1 \Sigma \mathbf{V}_2^* (\lambda \mathbf{I}_{T-k} - \mathbf{W}_{22})^{-1} \mathbf{V}_2 \Sigma \mathbf{V}_1^* \\ &= \mathbf{V}_1 (\Sigma + \mathbf{A}_p(\lambda)) \mathbf{V}_1^*, \end{aligned}$$

where

$$\mathbf{A}_p(\lambda) = \Sigma \mathbf{V}_2^* (\lambda \mathbf{I}_{T-k} - \mathbf{W}_{22})^{-1} \mathbf{V}_2 \Sigma. \quad (\text{A.84})$$

It follows that

$$\mathbf{K}_p(\lambda) = \frac{1}{\sqrt{p}} \mathbf{R}_p + \Lambda_1 \frac{\text{tr}(\Sigma)}{p} + \mathbf{V}_1 \mathbf{A}_p(\lambda) \mathbf{V}_1^*, \quad (\text{A.85})$$

where

$$\mathbf{R}_p = \{R_{ij}\}_{1 \leq i, j \leq k} = \sqrt{p} \mathbf{V}_1 \Sigma \mathbf{V}_1^* - \frac{\text{tr}(\Sigma)}{\sqrt{p}} \Lambda_1 = \sqrt{p} \mathbf{V}_1 \Sigma \mathbf{V}_1^* - \sqrt{p} \Lambda_1 \frac{\text{tr}(\Sigma)}{p}. \quad (\text{A.86})$$

Now we consider the Hermitian matrix $\mathbf{V}_1 \mathbf{A}_p(\lambda) \mathbf{V}_1^*$ in (A.85). When λ is a solution of (A.82), we have $\lambda > \|\mathbf{W}_{22}\|_2$ in probability due to Lemma 10. Hence the eigenvalues of $(\lambda \mathbf{I}_{T-k} - \mathbf{W}_{22})^{-1}$ and $\sqrt{p} \mathbf{V}_1 \mathbf{A}_p(\lambda) \mathbf{V}_1^*$ are non-negative in probability when λ is a solution of (A.82). Evidently we have

$$\|\sqrt{p} \mathbf{V}_1 \mathbf{A}_p(\lambda) \mathbf{V}_1^*\|_2 \leq \|\Sigma\|_2^2 \|(\lambda \mathbf{I}_{T-k} - \mathbf{W}_{22})^{-1}\|_2 \text{tr}(\sqrt{p} \mathbf{V}_1 \mathbf{V}_2^* \mathbf{V}_2 \mathbf{V}_1^*). \quad (\text{A.87})$$

Note that eigenvalues of $\sqrt{p} \mathbf{V}_1 \mathbf{V}_2^* \mathbf{V}_2 \mathbf{V}_1^*$ are also non-negative.

The next aim is to prove $E(\sqrt{p} \mathbf{V}_1 \mathbf{V}_2^* \mathbf{V}_2 \mathbf{V}_1^*) = o_p(1)$. Let $h_j = (\sqrt{p} \mathbf{V}_1 \mathbf{V}_2^* \mathbf{V}_2 \mathbf{V}_1^*)_{jj} \geq 0$. we can claim that $h_j = o_p(1)$. In fact

$$h_j = \frac{1}{p\sqrt{p}} \mathbf{s}_j^* \mathbf{Z} \mathbf{Z}^* \mathbf{Q}_2^* \mathbf{Q}_2 \mathbf{Z} \mathbf{s}_j. \quad (\text{A.88})$$

Let $E Z_{ij}^2 = y$, $E(Z_{ij}^*)^2 = z$ and $E|Z_{ij}|^4 = x + 1$. Write

$$\begin{aligned} E(h_j) &= \frac{1}{p\sqrt{p}} \mathbf{s}_j^* E(\mathbf{Z} \mathbf{Z}^* \mathbf{Q}_2^* \mathbf{Q}_2 \mathbf{Z}) \mathbf{s}_j \\ &= \frac{1}{p\sqrt{p}} \sum_{i=1}^p \sum_{m=1}^p \mathbf{s}_j^* E(\mathbf{z}_i \mathbf{z}_i^* \mathbf{Q}_2^* \mathbf{Q}_2 \mathbf{z}_m \mathbf{z}_m^*) \mathbf{s}_j \\ &= \frac{1}{p\sqrt{p}} \left(\sum_{i=1}^p \sum_{m \neq i} \mathbf{s}_j^* E(\mathbf{z}_i \mathbf{z}_i^* \mathbf{Q}_2^* \mathbf{Q}_2 \mathbf{z}_m \mathbf{z}_m^*) \mathbf{s}_j + \sum_{i=1}^p \mathbf{s}_j^* E(\mathbf{z}_i \mathbf{z}_i^* \mathbf{Q}_2^* \mathbf{Q}_2 \mathbf{z}_i \mathbf{z}_i^*) \mathbf{s}_j \right). \end{aligned} \quad (\text{A.89})$$

Consider the first term on the right hand of (A.89). When $i \neq m$,

$$\mathbf{s}_j^* E(\mathbf{z}_i \mathbf{z}_i^* \mathbf{Q}_2^* \mathbf{Q}_2 \mathbf{z}_m \mathbf{z}_m^*) \mathbf{s}_j = \mathbf{s}_j^* E(\mathbf{z}_i \mathbf{z}_i^*) \mathbf{Q}_2^* \mathbf{Q}_2 E(\mathbf{z}_m \mathbf{z}_m^*) \mathbf{s}_j = \mathbf{s}_j^* \mathbf{Q}_2^* \mathbf{Q}_2 \mathbf{s}_j.$$

Recall that $\{\mathbf{S}_k\}_{1 \leq k \leq T}$ are orthogonal and real. Since $j \leq k$, from the definition of \mathbf{Q}_2 ,

$$\mathbf{s}_j^* E(\mathbf{z}_i \mathbf{z}_i^* \mathbf{Q}_2^* \mathbf{Q}_2 \mathbf{z}_m \mathbf{z}_m^*) \mathbf{s}_j = \mathbf{s}_j^* \mathbf{Q}_2^* \mathbf{Q}_2 \mathbf{s}_j = 0, \quad (\text{A.90})$$

which implies the first term of on the right hand of (A.89) equals 0.

Consider the second term on the right hand of (A.89) now. Let $\mathbf{P} = (P_{rt})_{1 \leq r, t \leq T+l} = \mathbf{Q}_2^* \mathbf{Q}_2$ and $\mathbf{H}^v = (H_{im}^v)_{1 \leq i, m \leq T+l} = \mathbf{z}_v \mathbf{z}_v^* \mathbf{Q}_2^* \mathbf{Q}_2 \mathbf{z}_v \mathbf{z}_v^* = \mathbf{z}_v \mathbf{z}_v^* \mathbf{P} \mathbf{z}_v \mathbf{z}_v^*$. Then

$$\begin{aligned} E(H_{im}^v) &= E(Z_{iv} Z_{mv}^* \sum_{1 \leq r, t \leq T+l} P_{rt} Z_{rv}^* Z_{tv}) \\ &= \begin{cases} P_{ii} E|Z_{iv}|^4 + \sum_{r \neq i} P_{rr} E|Z_{iv}|^2 E|Z_{rv}|^2 & i = m, \\ P_{im} E|Z_{iv}|^2 E|Z_{mv}|^2 + P_{mi} E Z_{iv}^2 E(Z_{mv}^*)^2 & i \neq m. \end{cases} \\ &= \begin{cases} P_{ii} x + \sum_{r=1}^{T+l} P_{rr} & i = m, \\ P_{im} + P_{mi} yz & i \neq m. \end{cases} \end{aligned}$$

It follows that

$$\begin{aligned} & \sum_{v=1}^p \mathbf{s}_j^* E(\mathbf{z}_v \mathbf{z}_v^* \mathbf{P} \mathbf{z}_v \mathbf{z}_v^*) \mathbf{s}_j \\ &= \sum_{v=1}^p \mathbf{s}_j^* E(\mathbf{H}^v) \mathbf{s}_j \\ &= \sum_{v=1}^p \sum_{i=1}^{T+l} \sum_{m=1}^{T+l} S_{ji}^* S_{jm} E(H_{im}^v) \\ &= \sum_{v=1}^p \sum_{i=1}^{T+l} S_{ji}^* S_{ji} (P_{ii} x + \sum_{r=1}^{T+l} P_{rr}) + \sum_{v=1}^p \sum_{i=1}^{T+l} \sum_{m=1, m \neq i}^{T+l} S_{ji}^* S_{jm} (P_{im} + P_{mi} yz) \\ &= \sum_{v=1}^p \sum_{i=1}^{T+l} S_{ji}^* S_{ji} (P_{ii} (x - 1 - yz) + \sum_{r=1}^{T+l} P_{rr}) + \sum_{v=1}^p \sum_{i=1}^{T+l} \sum_{m=1}^{T+l} S_{ji}^* S_{jm} (P_{im} + P_{mi} yz) \\ &\leq \text{tr}(\mathbf{P}) (|x - 1 - yz| + 1) \sum_{v=1}^p \sum_{i=1}^{T+l} S_{ji}^* S_{ji} + \sum_{v=1}^p \sum_{i=1}^{T+l} \sum_{m=1}^{T+l} S_{ji}^* S_{jm} (P_{im} + P_{mi} yz) \\ &= \text{tr}(\mathbf{P}) (|x - 1 - yz| + 1) p \|\mathbf{s}_j\|^2 + p \sum_{i=1}^{T+l} \sum_{m=1}^{T+l} S_{ji}^* S_{jm} (P_{im} + P_{mi} yz) \\ &= \text{tr}(\mathbf{P}) (|x - 1 - yz| + 1) p \|\mathbf{s}_j\|^2 + p \mathbf{s}_j^* \mathbf{P} \mathbf{s}_j + p y z \mathbf{s}_j^* \mathbf{P} \mathbf{s}_j. \end{aligned}$$

By (A.90) and $\mathbf{P} = \mathbf{Q}_2^* \mathbf{Q}_2$, we have

$$\sum_{v=1}^p \mathbf{s}_j^* E(\mathbf{z}_v \mathbf{z}_v^* \mathbf{P} \mathbf{z}_v \mathbf{z}_v^*) \mathbf{s}_j \leq \text{tr}(\mathbf{P}) (|x - 1 - yz| + 1) p \|\mathbf{s}_j\|^2.$$

Also,

$$\text{tr}(\mathbf{P}) = \text{tr}(\mathbf{Q}_2^* \mathbf{Q}_2) = \text{tr}(\mathbf{\Lambda}_2).$$

From (A.78), (A.88) and (A.90) we can obtain

$$E(h_j) \leq \frac{1}{\sqrt{p}} \text{tr}(\mathbf{P}) (|x - 1 - yz| + 1) \|\mathbf{s}_j\|^2 \rightarrow 0, \quad (\text{A.91})$$

as claimed.

Since k is a fix number, we can obtain

$$E(\text{tr}(\sqrt{p}\mathbf{V}_1\mathbf{V}_2^*\mathbf{V}_2\mathbf{V}_1^*)) = \sum_{j=1}^k E(h_j) = o_p(1). \quad (\text{A.92})$$

It follows from $\|(\lambda\mathbf{I}_{\mathbf{T}-\mathbf{k}} - \mathbf{W}_{22})^{-1}\|_2 = O_p(1)$, (A.87) and (A.92) that

$$\|\mathbf{V}_1\mathbf{A}_p(\lambda)\mathbf{V}_1^*\|_2 = o_p(p^{-1/2}). \quad (\text{A.93})$$

Now we consider $\frac{1}{\sqrt{p}}\mathbf{R}_p = \frac{1}{\sqrt{p}}(R_{ij})$ in (A.85). From (A.86) and the definition of \mathbf{V}_1 , \mathbf{Q}_1 and $\mathbf{\Lambda}_1$,

$$R_{ij} = \frac{1}{\sqrt{p}}(\mathbf{s}_i^*\mathbf{Z}\mathbf{\Sigma}\mathbf{Z}^*\mathbf{s}_j - \mathbf{s}_i^*\mathbf{s}_j\text{tr}(\mathbf{\Sigma})). \quad (\text{A.94})$$

Note that $E(\mathbf{s}_i^*\mathbf{z}_1\mathbf{z}_1^*\mathbf{s}_j) = \mathbf{s}_i^*\mathbf{s}_j$. With Theorem 7.1 of [2], we can prove that R_{ij} converges weakly to a zero-mean Gaussian variable r_{ij} with bounded variance. It follows that

$$\frac{1}{\sqrt{p}}R_{ij} = O_p(p^{-1/2}). \quad (\text{A.95})$$

Note that $\mathbf{\Lambda}_1 \frac{\text{tr}(\mathbf{\Sigma})}{p}$ is a diagonal matrix and hence we can find that any off-diagonal element of $\lambda\mathbf{I}_k - \mathbf{K}_n(\lambda)$ is $O_p(p^{-1/2})$. This, together with (A.83), implies that for any $i \neq j$, $G_{ij}(\lambda) = O_p(p^{-1/2})$ with λ satisfying $|\mathbf{G}(\lambda)| = 0$. Similarly, $G_{ii}(\lambda) = \lambda - \frac{\text{tr}(\mathbf{\Sigma})\beta_i}{p\gamma_1} + O_p(p^{-1/2})$ is absolutely bounded with λ satisfying $|\mathbf{G}(\lambda)| = 0$.

Denote by Ω_k all permutations σ of the set $\{1, 2, \dots, k\}$. By the Laplace formula of a determinant we have

$$\begin{aligned} 0 &= |\mathbf{G}(\lambda)| = \sum_{\sigma \in \Omega_k} \text{sgn}(\sigma) \prod_{j=1}^k G_{j,\sigma_j}(\lambda) \\ &= \sum_{\sigma \in \Omega_k, \sigma \neq [1, 2, \dots, k]} \text{sgn}(\sigma) \prod_{j=1}^k G_{j,\sigma_j}(\lambda) + \prod_{j=1}^k G_{jj}(\lambda). \end{aligned}$$

Recall that any off-diagonal element of $\mathbf{G}(\lambda)$ is $O_p(p^{-1/2})$. We conclude that when $\sigma \neq [1, 2, \dots, k]$, $\prod_{j=1}^k G_{j,\sigma_j}(\lambda) = O_p(p^{-1})$ since there are at least two different j_1 and j_2 such that $\sigma_{j_1} \neq j_1$ and $\sigma_{j_2} \neq j_2$. Since k is fixed, we have

$$\prod_{1 \leq j \leq k} G_{jj}(\lambda) = - \sum_{\sigma \in \Omega_k, \sigma \neq [1, 2, \dots, k]} \text{sgn}(\sigma) \prod_{j=1}^k G_{j,\sigma_j}(\lambda) = O_p(p^{-1}). \quad (\text{A.96})$$

Then when λ satisfies $|\mathbf{G}(\lambda)| = 0$, there exists j (not bigger than k) such that $|G_{jj}(\lambda)| = o(1)$. When $i \neq j$, from (A.85), (A.93) and (A.95)

$$\begin{aligned} &|G_{jj}(\lambda) - G_{ii}(\lambda)| \\ &\geq \frac{\text{tr}(\mathbf{\Sigma})}{p} |(\Lambda_1)_{jj} - (\Lambda_1)_{ii}| - \left| \frac{1}{\sqrt{p}}R_{jj} - \frac{1}{\sqrt{p}}R_{ii} \right| - |(\mathbf{V}_1\mathbf{A}_p(\lambda)\mathbf{V}_1^*)_{jj} - (\mathbf{V}_1\mathbf{A}_p(\lambda)\mathbf{V}_1^*)_{ii}| \\ &= \frac{\text{tr}(\mathbf{\Sigma})|\beta_j - \beta_i|}{p\gamma_1} + O_p(p^{-1/2}). \end{aligned}$$

By Lemma 4 we can obtain that for any $i \neq j$, $|G_{ii}(\lambda)| \geq \frac{\text{tr}(\mathbf{\Sigma})}{p} (|\frac{1}{(2i-1)^2} - \frac{1}{(2j-1)^2}|) + o_p(1)$. This, together with (A.96), implies that $|G_{jj}(\lambda)| = O_p(p^{-1})$. Hence $|G_{jj}(\lambda_j)| = O_p(p^{-1})$ for any $\lambda_1 > \lambda_2 > \dots > \lambda_k$ satisfying $|\mathbf{G}(\lambda_j)| = 0$ for $1 \leq j \leq k$. Write

$$(\lambda_1, \lambda_2, \dots, \lambda_k) = (\lambda_1 - G_{11}(\lambda_1) + O_p(p^{-1}), \dots, \lambda_k - G_{kk}(\lambda_k) + O_p(p^{-1})). \quad (\text{A.97})$$

It follows that

$$\begin{aligned} & (\sqrt{p}(\lambda_1 - \frac{\gamma_1 \text{tr}(\boldsymbol{\Sigma})}{\gamma_1 p}), \dots, \sqrt{p}(\lambda_k - \frac{\gamma_k \text{tr}(\boldsymbol{\Sigma})}{\gamma_1 p})) = \\ & (\sqrt{p}(\lambda_1 - \frac{\gamma_1 \text{tr}(\boldsymbol{\Sigma})}{\gamma_1 p} - G_{11}(\lambda_1) + O_p(p^{-1})), \dots, \sqrt{p}(\lambda_k - \frac{\gamma_k \text{tr}(\boldsymbol{\Sigma})}{\gamma_1 p} - G_{kk}(\lambda_k) + O_p(p^{-1}))). \end{aligned}$$

Via this, by (A.83), (A.85), (A.77), (A.86), (A.93) and Lemma 4 we further obtain

$$\begin{aligned} & \sqrt{p}[\lambda_j - \frac{\gamma_j \text{tr}(\boldsymbol{\Sigma})}{\gamma_1 p} - G_{jj}(\lambda_j) + O_p(p^{-1})] \\ &= \sqrt{p}[\lambda_j - \frac{\gamma_j \text{tr}(\boldsymbol{\Sigma})}{\gamma_1 p} - \lambda_j + \frac{\beta_j \text{tr}(\boldsymbol{\Sigma})}{\gamma_1 p} + \frac{1}{\sqrt{p}}R_{jj} + (\mathbf{V}_1 \mathbf{A}_n(\lambda_j) \mathbf{V}_1^*)_{jj} + O_p(p^{-1})] \\ &= R_{jj} + O(\sqrt{p}T^{-1}) + o_p(1). \end{aligned} \tag{A.98}$$

Recalling (A.94), we have

$$(R_{11}, \dots, R_{kk})' = (\frac{1}{\sqrt{p}}(\mathbf{s}_1^* \mathbf{Z} \boldsymbol{\Sigma} \mathbf{Z}^* \mathbf{s}_1 - \mathbf{s}_1^* \mathbf{s}_1 \text{tr}(\boldsymbol{\Sigma})), \dots, \frac{1}{\sqrt{p}}(\mathbf{s}_k^* \mathbf{Z} \boldsymbol{\Sigma} \mathbf{Z}^* \mathbf{s}_k - \mathbf{s}_k^* \mathbf{s}_k \text{tr}(\boldsymbol{\Sigma})))'.$$

Note that $E(\mathbf{s}_i^* \mathbf{z}_1 \mathbf{z}_1^* \mathbf{s}_j) = \mathbf{s}_i^* \mathbf{s}_j$. From Theorem 7.1 of [2], we can find that $(R_{11}, \dots, R_{kk})'$ converges weakly to a zero-mean Gaussian vector $\mathbf{w} = (w_1, \dots, w_k)'$.

We next determine the covariance between w_i and w_j for the complex case and the real case in a unified way. To this end, let $\omega = \lim_{p \rightarrow \infty} \frac{\sum_{1 \leq i \leq p} \Sigma_{ii}^2}{p}$ and $\theta = \lim_{p \rightarrow \infty} \frac{\text{tr}(\Sigma^2)}{p} = \lim_{p \rightarrow \infty} \frac{\text{tr}(\Sigma \Sigma')}{p}$. When $i \neq j$, from Theorem 7.1 of [2], we have

$$\begin{aligned} & \text{cov}(w_i, w_j) \\ &= \lim_{T \rightarrow \infty} \omega (E | \xi_1(i) |^2 | \xi_1(j) |^2 - (E | \xi_1(i) |^2)(E | \xi_1(j) |^2)) \\ &+ \lim_{T \rightarrow \infty} (\theta - \omega) (E \bar{\xi}_1(i) \xi_1(j)) (E \bar{\xi}_1(j) \xi_1(i)) + \lim_{T \rightarrow \infty} (\theta - \omega) (E \xi_1(i) \xi_1(j)) (E \bar{\xi}_1(i) \bar{\xi}_1(j)). \end{aligned}$$

Recalling (A.75), we have

$$\begin{aligned} & E(| \xi_1(i) |^2 | \xi_1(j) |^2) - E(| \xi_1(i) |^2) E(| \xi_1(j) |^2) \\ &= E(\mathbf{s}_i^* \mathbf{Z}_1 \mathbf{Z}_1^* \mathbf{s}_i \mathbf{s}_j^* \mathbf{Z}_1 \mathbf{Z}_1^* \mathbf{s}_j) - E(\mathbf{s}_i^* \mathbf{Z}_1 \mathbf{Z}_1^* \mathbf{s}_i) E(\mathbf{s}_j^* \mathbf{Z}_1 \mathbf{Z}_1^* \mathbf{s}_j). \end{aligned} \tag{A.99}$$

Recall that $\{\mathbf{s}_k\}_{1 \leq k \leq T}$ are orthogonal and real. We obtain

$$\begin{aligned} & E(\mathbf{s}_i^* \mathbf{z}_1 \mathbf{z}_1^* \mathbf{s}_i \mathbf{s}_j^* \mathbf{z}_1 \mathbf{z}_1^* \mathbf{s}_j) \\ &= E(\sum_{f_1=1}^{T+l} \sum_{f_2=1}^{T+l} S_{if_1} S_{if_2} Z_{f_1 1} Z_{f_2 1}^*) (\sum_{f_1=1}^{T+l} \sum_{f_2=1}^{T+l} S_{jf_1} S_{jf_2} Z_{f_1 1} Z_{f_2 1}^*) \\ &= \sum_{f_1=1}^{T+l} \sum_{f_2 \neq f_1}^{T+l} S_{if_1} S_{if_2} S_{jf_1} S_{jf_2} (E Z_{f_1 1}^2 (Z_{f_2 1}^*)^2 + E |Z_{f_1 1}|^2 |Z_{f_2 1}|^2) + E(\sum_{f_1=1}^{T+l} S_{if_1}^2 |Z_{f_1 1}|^2) (\sum_{f_2=1}^{T+l} S_{jf_2}^2 |Z_{f_2 1}|^2) \\ &= (yz + 1) \sum_{f_1=1}^{T+l} \sum_{f_2 \neq f_1}^{T+l} S_{if_1} S_{if_2} S_{jf_1} S_{jf_2} + E(\sum_{f_1=1}^{T+l} S_{if_1}^2 |Z_{f_1 1}|^2) (\sum_{f_2=1}^{T+l} S_{jf_2}^2 |Z_{f_2 1}|^2). \end{aligned} \tag{A.100}$$

Since $\{\mathbf{s}_k\}_{1 \leq k \leq T}$ are orthogonal, we conclude from (A.49) that

$$\begin{aligned} & (yz + 1) \sum_{f_1=1}^{T+l} \sum_{f_2 \neq f_1}^{T+l} S_{if_1} S_{if_2} S_{jf_1} S_{jf_2} \\ &= (yz + 1) \sum_{f_1=1}^{T+l} S_{if_1} S_{jf_1} \sum_{f_2=1}^{T+l} S_{if_2} S_{jf_2} - (yz + 1) \sum_{f_1=1}^{T+l} S_{if_1}^2 S_{jf_1}^2 \\ &= -(yz + 1) \sum_{f_1=1}^{T+l} S_{if_1}^2 S_{jf_1}^2 = O(T^{-1}) \end{aligned} \tag{A.101}$$

and

$$\begin{aligned}
& E\left(\sum_{f_1=1}^{T+l} S_{if_1}^2 |Z_{f_1}|^2\right) \left(\sum_{f_2=1}^{T+l} S_{jf_2}^2 |Z_{f_2}|^2\right) \\
&= \sum_{f_1=1}^{T+l} S_{if_1}^2 S_{jf_1}^2 (E|Z_{f_1}|^4 - 1) + E(\mathbf{s}_i^* \mathbf{z}_1 \mathbf{z}_1^* \mathbf{s}_i) E(\mathbf{s}_j^* \mathbf{z}_1 \mathbf{z}_1^* \mathbf{s}_j) \\
&= E(\mathbf{s}_i^* \mathbf{z}_1 \mathbf{z}_1^* \mathbf{s}_i) E(\mathbf{s}_j^* \mathbf{z}_1 \mathbf{z}_1^* \mathbf{s}_j) + O(T^{-1}).
\end{aligned} \tag{A.102}$$

Summarizing (A.99), (A.100), (A.101) and (A.102), we conclude that

$$\lim_{T \rightarrow \infty} \omega(E | \xi_1(i) |^2 | \xi_1(j) |^2 - (E | \xi_1(i) |^2)(E | \xi_1(j) |^2)) = 0.$$

Since $\{\mathbf{s}_k\}_{1 \leq k \leq T}$ are orthogonal and real, we also have

$$E\bar{\xi}_1(i)\xi_1(j) = 0, E\bar{\xi}_1(j)\xi_1(i) = 0$$

and

$$E\xi_1(i)\xi_1(j) = 0, E\bar{\xi}_1(i)\bar{\xi}_1(j) = 0.$$

This implies that

$$\text{cov}(w_i, w_j) = 0. \tag{A.103}$$

By (2.14) and (A.75) we can get

$$\begin{aligned}
& \text{Var}(w_i) \\
&= \omega \lim_{T \rightarrow \infty} \{E | \xi_1(i) |^4 - 2(E | \xi_1(i) |^2)^2 - (E\xi_1(i)^2)(E\bar{\xi}_1(i)^2)\} \\
&\quad + \theta \lim_{T \rightarrow \infty} (E | \xi_1(i) |^2)^2 + \theta \lim_{T \rightarrow \infty} (E\xi_1(i)^2)(E\bar{\xi}_1(i)^2) \\
&= \omega \lim_{T \rightarrow \infty} \{E | \xi_1(i) |^4 - 2(E | \xi_1(i) |^2)^2 - (E\xi_1(i)^2)(E\bar{\xi}_1(i)^2)\} \\
&\quad + \theta \frac{1}{(2i-1)^4} + \theta \frac{1}{(2i-1)^4} (1 - 4E(Z_{i1}^R)^2 E(Z_{i1}^I)^2).
\end{aligned}$$

From Lemma 4

$$\begin{aligned}
& \lim_{T \rightarrow \infty} \{E | \xi_1(i) |^4 - 2(E | \xi_1(i) |^2)^2 - (E\xi_1(i)^2)(E\bar{\xi}_1(i)^2)\} \\
&= \lim_{T \rightarrow \infty} \{E | \sum_{j=1}^{T+l} S_{ij} Z_{j1} |^4 - 2\left(\frac{\beta_i}{\gamma_1}\right)^2 - \left(\frac{\beta_i}{\gamma_1}\right)^2 (E(Z_{i1}^R)^2 - E(Z_{i1}^I)^2)^2\} \\
&= \lim_{T \rightarrow \infty} \{E\left(\sum_{j=1}^{T+l} S_{ij} Z_{j1}^R\right)^4 + E\left(\sum_{j=1}^{T+l} S_{ij} Z_{j1}^I\right)^4 + \\
&\quad 2E\left(\sum_{j=1}^{T+l} S_{ij} Z_{j1}^R\right)^2 \left(\sum_{j=1}^{T+l} S_{ij} Z_{j1}^I\right)^2\} - \frac{1}{(2i-1)^4} (2 + (E(Z_{i1}^R)^2 - E(Z_{i1}^I)^2)^2).
\end{aligned}$$

Recalling Z_{j1}^R and Z_{j1}^I are independent, we have

$$2E\left(\sum_{j=1}^{T+l} S_{ij} Z_{j1}^R\right)^2 \left(\sum_{j=1}^{T+l} S_{ij} Z_{j1}^I\right)^2 = 2E\left(\sum_{j=1}^{T+l} S_{ij}^2 (Z_{j1}^R)^2\right) E\left(\sum_{j=1}^{T+l} S_{ij}^2 (Z_{j1}^I)^2\right)$$

and

$$\begin{aligned}
& E\left(\sum_{j=1}^{T+l} S_{ij} Z_{j1}^R\right)^4 + E\left(\sum_{j=1}^{T+l} S_{ij} Z_{j1}^I\right)^4 + 2E\left(\sum_{j=1}^{T+l} S_{ij} Z_{j1}^R\right)^2 \left(\sum_{j=1}^{T+l} S_{ij} Z_{j1}^I\right)^2 \\
= & 3\left[\left(\sum_{j=1}^{T+l} S_{ij}^2 E(Z_{j1}^R)^2\right)^2 + \left(\sum_{j=1}^{T+l} S_{ij}^2 E(Z_{j1}^I)^2\right)^2\right] + \\
& \sum_{j=1}^{T+l} S_{ij}^4 [E(Z_{j1}^R)^4 + E(Z_{j1}^I)^4 - 3(E((Z_{j1}^R)^2)^2 + (E(Z_{j1}^I)^2)^2)] \\
& + 2E\left(\sum_{j=1}^{T+l} S_{ij}^2 (Z_{j1}^R)^2\right) E\left(\sum_{j=1}^{T+l} S_{ij}^2 (Z_{j1}^I)^2\right) \\
= & 3\left[\sum_{j=1}^{T+l} S_{ij}^2 E(Z_{j1}^R)^2 + \sum_{j=1}^{T+l} S_{ij}^2 E(Z_{j1}^I)^2\right]^2 \\
& + \sum_{j=1}^{T+l} S_{ij}^4 [E(Z_{j1}^R)^4 + E(Z_{j1}^I)^4 - 3((E(Z_{j1}^R)^2)^2 + (E(Z_{j1}^I)^2)^2)] - \\
& 4E\left(\sum_{j=1}^{T+l} S_{ij}^2 (Z_{j1}^R)^2\right) E\left(\sum_{j=1}^{T+l} S_{ij}^2 (Z_{j1}^I)^2\right) \\
= & (2 + (E(Z_{i1}^R)^2 - E(Z_{i1}^I)^2)^2) \left(\sum_{j=1}^{T+l} S_{ij}^2\right)^2 + \sum_{j=1}^{T+l} S_{ij}^4 [E(Z_{j1}^R)^4 + E(Z_{j1}^I)^4 - 3((E(Z_{j1}^R)^2)^2 + (E(Z_{j1}^I)^2)^2)].
\end{aligned}$$

In view of Lemma 4, (A.49) and (A.73), we have

$$(2 + (E(Z_{i1}^R)^2 - E(Z_{i1}^I)^2)^2) \left(\sum_{j=1}^{T+l} S_{ij}^2\right)^2 = (2 + (E(Z_{i1}^R)^2 - E(Z_{i1}^I)^2)^2) \left(\frac{1}{(2i-1)^4} + O(T^{-1})\right)$$

and

$$\sum_{j=1}^{T+l} S_{ij}^4 [E(Z_{j1}^R)^4 + E(Z_{j1}^I)^4 - 3((E(Z_{j1}^R)^2)^2 + (E(Z_{j1}^I)^2)^2)] = O(T^{-1}).$$

So we can obtain

$$\lim_{T \rightarrow \infty} \{E|\xi_1(i)|^4 - 2(E|\xi_1(i)|^2)^2 - (E\xi_1(i)^2)(E\bar{\xi}_1(i)^2)\} = 0.$$

It follows that

$$\text{Var}(w_i) = \theta \frac{1}{(2i-1)^4} (2 - 4E(Z_{i1}^R)^2 E(Z_{i1}^I)^2).$$

This, together with (A.98), (A.103) and the assumption (A4), implies Theorem 4. □

B The Proof of the Main Results

This section is to prove that the results obtained in Section 4 still hold for the general linear processes (without the truncation step performed there) and the general initial vector \mathbf{x}_0 (not necessarily zero).

Lemma 11. Recall the definitions of \mathbf{Y} , λ_k and γ_k in Section 1. Let $l = \max\{p, T\}$ and \mathbf{Y}_1 be the truncated matrix of \mathbf{Y} in Section 4. Define

$$\gamma_{k,l} = \lambda_k(a_{0,l} + 2 \sum_{1 \leq j \leq T-1} a_{j,l}(-1)^j \cos(j\theta_k))$$

where

$$a_{j,l} = \sum_{j \leq k \leq l} b_k b_{k-j}. \quad (\text{B.1})$$

Then when $\mathbf{\Pi} = \mathbf{I}$,

$$\left\| \frac{(1/p)\mathbf{C}(\mathbf{Y}\Sigma\mathbf{Y}^* - \mathbf{Y}_1\Sigma\mathbf{Y}_1^*)\mathbf{C}^*}{\gamma_{1,l}} \right\|_2 = o_p(p^{-1/2}) \quad (\text{B.2})$$

and

$$\frac{|\gamma_{k,l} - \gamma_k|}{\gamma_{1,l}} = o(1). \quad (\text{B.3})$$

Proof of Lemma 11: We consider (B.3) first. To this end, observe that the assumption (A1) implies that

$$\sum_{i=0}^{\infty} i|a_i| < \infty, \quad (\text{B.4})$$

because

$$\sum_{i=0}^{\infty} i|a_i| \leq \sum_{i=0}^{\infty} i \sum_{k=0}^{\infty} |b_k| |b_{k+i}| = \sum_{k=0}^{\infty} |b_k| \left(\sum_{i=0}^{\infty} i|b_{k+i}| \right) \leq \sum_{k=0}^{\infty} |b_k| \left(\sum_{i=0}^{\infty} i|b_i| \right).$$

Write

$$\begin{aligned} \frac{|\gamma_{k,l} - \gamma_k|}{\gamma_{1,l}} &\leq \frac{\lambda_k}{\gamma_{1,l}} \left(\sum_{k>l} b_k^2 + 2 \sum_{j=1}^{T-1} \sum_{k>l} |b_k| |b_{k-j}| + 2 \sum_{j \geq T} |a_j| \right) \\ &\leq \frac{\lambda_k}{\gamma_{1,l}} \left(\sum_{k>l} b_k^2 + 2 \sum_{j=1}^{\infty} |b_j| \sum_{k>l} |b_k| + 2 \sum_{j \geq T} |a_j| \right). \end{aligned}$$

From the assumption (A1) and (B.4), we obtain that

$$\sum_{k>l} b_k^2 + 2 \sum_{j=1}^{\infty} |b_j| \sum_{k>l} |b_k| + 2 \sum_{j \geq T} |a_j| = o(1).$$

Moreover, Lemma 2 and assumption (A1) (or (A.9)) imply that $\frac{\lambda_k}{\gamma_{1,l}}$ is bounded. So we conclude (B.3).

Now we consider (B.2). Via Lemma 1, observe that

$$\begin{aligned} \left\| \frac{(1/p)\mathbf{C}(\mathbf{Y}\Sigma\mathbf{Y}^* - \mathbf{Y}_1\Sigma\mathbf{Y}_1^*)\mathbf{C}^*}{\gamma_{1,l}} \right\|_2 &\leq \frac{\|\mathbf{C}\|_2^2}{\gamma_{1,l}} \left\| (1/p)(\mathbf{Y}\Sigma\mathbf{Y}^* - \mathbf{Y}_1\Sigma\mathbf{Y}_1^*) \right\|_2 \\ &= \frac{\lambda_1}{\gamma_{1,l}} \left\| (1/p)(\mathbf{Y}\Sigma\mathbf{Y}^* - \mathbf{Y}_1\Sigma\mathbf{Y}_1^*) \right\|_2. \end{aligned}$$

As before $\frac{\lambda_1}{\gamma_{1,l}}$ is bounded. So we just need to consider $\left\| (1/p)(\mathbf{Y}\Sigma\mathbf{Y}^* - \mathbf{Y}_1\Sigma\mathbf{Y}_1^*) \right\|_2$. Let $\mathbf{K} = (K_{ij})_{1 \leq i \leq T, 1 \leq j \leq p} = \mathbf{Y} - \mathbf{Y}_1$. We can obtain that $K_{ij} = \sum_{k=l+1}^{\infty} b_k Z_{i-k,j}$ and

$$E|K_{ij}|^2 = \sum_{k=l+1}^{\infty} b_k^2.$$

By the assumption (A1) we can get

$$E|K_{ij}|^2 = \sum_{k=l+1}^{\infty} b_k^2 \leq l^{-2} \sum_{k=l+1}^{\infty} k^2 |b_k|^2 = o(l^{-2}),$$

which ensures that

$$E \left\| \frac{1}{\sqrt{p}} \mathbf{K} \right\|_F^2 = o(Tl^{-2}).$$

This, together with (A.3), implies that

$$\begin{aligned} & \left\| (1/p)(\mathbf{Y}\Sigma\mathbf{Y}^* - \mathbf{Y}_1\Sigma\mathbf{Y}_1^*) \right\|_2 = \left\| (1/p)(\mathbf{K}\Sigma\mathbf{Y}_1^* + \mathbf{Y}_1\Sigma\mathbf{K}^* + \mathbf{K}\Sigma\mathbf{K}^*) \right\|_2 \\ & \leq 2 \left\| \frac{1}{\sqrt{p}} \mathbf{K} \right\|_F \|\Sigma\|_2 \left\| \frac{1}{\sqrt{p}} \mathbf{Y}_1 \right\|_2 + \left\| \frac{1}{\sqrt{p}} \mathbf{K} \right\|_F^2 \|\Sigma\|_2 = o_p(p^{-1/2}). \end{aligned} \quad (\text{B.5})$$

So we can conclude (B.2). \square

Proof of Theorems 2: Recalling (2.5),

$$\mathbf{B} = \frac{1}{p} \mathbf{X}\mathbf{X}^* = \frac{1}{p} \mathbf{C}\mathbf{Y}\Sigma\mathbf{Y}^*\mathbf{C}^* + \frac{1}{p} \mathbf{C}\mathbf{Y}\Sigma^{1/2}\mathbf{X}_0^* + \frac{1}{p} \mathbf{X}_0\Sigma^{1/2}\mathbf{Y}^*\mathbf{C}^* + \frac{1}{p} \mathbf{X}_0\mathbf{X}_0^*.$$

The assumption A6 implies that

$$\left\| \frac{1}{p} \mathbf{X}_0\mathbf{X}_0^* \right\|_2 = O_p(T) \quad (\text{B.6})$$

and that

$$\left\| \frac{1}{p} \mathbf{C}\mathbf{Y}\Sigma^{1/2}\mathbf{X}_0^* \right\|_2 = O_p(T^{1/2} \left\| \frac{1}{p} \mathbf{C}\mathbf{Y}\Sigma\mathbf{Y}^*\mathbf{C}^* \right\|_2^{1/2}). \quad (\text{B.7})$$

We can write $\frac{(1/p)\mathbf{C}\mathbf{Y}\Sigma\mathbf{Y}^*\mathbf{C}^*}{\gamma_1}$ as

$$\begin{aligned} & \frac{(1/p)\mathbf{C}\mathbf{Y}\Sigma\mathbf{Y}^*\mathbf{C}^*}{\gamma_1} = \frac{\gamma_{1,l}}{\gamma_1} \frac{(1/p)\mathbf{C}\mathbf{Y}\Sigma\mathbf{Y}^*\mathbf{C}^*}{\gamma_{1,l}} \\ & = \frac{\gamma_{1,l}}{\gamma_1} \frac{(1/p)\mathbf{C}\mathbf{Y}_1\Sigma\mathbf{Y}_1^*\mathbf{C}^*}{\gamma_{1,l}} + \frac{\gamma_{1,l}}{\gamma_1} \frac{(1/p)\mathbf{C}(\mathbf{Y}\Sigma\mathbf{Y}^* - \mathbf{Y}_1\Sigma\mathbf{Y}_1^*)\mathbf{C}^*}{\gamma_{1,l}}. \end{aligned} \quad (\text{B.8})$$

From (B.3) we have $\lim_{T \rightarrow \infty} \frac{\gamma_{1,l}}{\gamma_1} = 1$. This, together with (2.5), Proposition 3, Lemma 4, (B.2), (B.6) and (B.7), implies Theorem 2. \square

Proof of Theorem 4: Let $m = \lceil \sqrt{p} \rceil$. From (2.6) and (B.4) we have

$$\left| (a_0 + 2 \sum_{1 \leq j \leq m} a_j (-1)^j \cos(j\theta_1)) - (a_0 + 2 \sum_{1 \leq j \leq \infty} a_j (-1)^j \cos(j\theta_1)) \right| \leq 2 \sum_{1+m \leq j \leq \infty} |a_j| = o(p^{-1/2}) \quad (\text{B.9})$$

and

$$\begin{aligned} & \left| (a_0 + 2 \sum_{1 \leq j \leq m} a_j (-1)^j \cos(j\theta_1)) - (a_0 + 2 \sum_{1 \leq j \leq m} a_j) \right| \\ & \leq 2 \sum_{1 \leq j \leq m} |a_j| (1 - \cos \frac{j\pi}{2T+1}) = O(p^{1/2} T^{-2}) = o(p^{-1/2}). \end{aligned} \quad (\text{B.10})$$

In view of (2.7) it suffices to prove that

$$\left| \frac{1}{T-1} \sum_{2 \leq i \leq T} \frac{\mathbf{y}_i' \Sigma \mathbf{y}_i}{p} + 2 \sum_{1 \leq j \leq m} \frac{1}{T-j-1} \sum_{2 \leq i \leq T-j} \frac{\mathbf{y}_i' \Sigma \mathbf{y}_{i+j}}{p} - (a_0 + 2 \sum_{1 \leq j \leq m} a_j) \frac{\text{tr}(\Sigma)}{p} \right| = o_p(p^{-1/2}). \quad (\text{B.11})$$

A direct calculation shows the following mean and variance

$$E \left(\frac{1}{T-1} \sum_{2 \leq i \leq T} \frac{\mathbf{y}_i' \Sigma \mathbf{y}_i}{p} + 2 \sum_{1 \leq j \leq m} \frac{1}{T-j-1} \sum_{2 \leq i \leq T-j} \frac{\mathbf{y}_i' \Sigma \mathbf{y}_{i+j}}{p} \right) - (a_0 + 2 \sum_{1 \leq j \leq m} a_j) \frac{\text{tr}(\Sigma)}{p} = 0, \quad (\text{B.12})$$

$$\begin{aligned}
& \text{Var}\left(\sum_{1 \leq j \leq m} \frac{1}{T-j-1} \sum_{2 \leq i \leq T-j} \frac{\mathbf{y}'_i \boldsymbol{\Sigma} \mathbf{y}_{i+j}}{p}\right) \\
&= \sum_{1 \leq i, j \leq m} \sum_{2 \leq f \leq T-i} \sum_{2 \leq g \leq T-j} \frac{\text{Cov}\left(\frac{\mathbf{y}'_f \boldsymbol{\Sigma} \mathbf{y}_{f+i}}{p}, \frac{\mathbf{y}'_g \boldsymbol{\Sigma} \mathbf{y}_{g+j}}{p}\right)}{(T-i-1)(T-j-1)}.
\end{aligned}$$

Moreover

$$\begin{aligned}
\text{Cov}\left(\frac{\mathbf{y}'_f \boldsymbol{\Sigma} \mathbf{y}_{f+i}}{p}, \frac{\mathbf{y}'_g \boldsymbol{\Sigma} \mathbf{y}_{g+j}}{p}\right) &= \frac{1}{p} \left[\frac{\sum_{i=1}^p \Sigma_{ii}^2}{p} E|Z_{ij}|^4 \sum_{k=0}^{\infty} b_k b_{k+i} b_{k+g-f} b_{k+g-f+j} 1_{(k+g-f \geq 0)} \right. \\
&\quad \left. + \frac{\text{tr}(\boldsymbol{\Sigma}^2)}{p} E|Z_{ij}|^2 (a_{|f-g|} a_{|f+i-g-j|} + a_{|f+i-g|} a_{|f-g-j|}) \right].
\end{aligned}$$

From the above, the assumption (A1) and (B.4) we conclude that

$$\text{Var}\left(\frac{1}{T-1} \sum_{2 \leq i \leq T} \frac{\mathbf{y}'_i \boldsymbol{\Sigma} \mathbf{y}_i}{p} + 2 \sum_{1 \leq j \leq m} \frac{1}{T-j-1} \sum_{2 \leq i \leq T-j} \frac{\mathbf{y}'_i \boldsymbol{\Sigma} \mathbf{y}_{i+j}}{p}\right) = O(p^{-1} m T^{-1}) = o(p^{-1}). \quad (\text{B.13})$$

(B.12) and (B.13) imply (B.11). \square

Proof of Theorem 5: Throughout the proof of this theorem, in order to simplify notation we use m to replace m_2 . In view of (2.6) and (2.7) it suffices to show that

$$\frac{S_{\sigma^2}}{(a_0 + 2 \sum_{j=1}^{\infty} a_j (-1)^j \cos(j\theta_1))^2 \text{tr}(\boldsymbol{\Sigma}^2)} = 1 + o_p(1).$$

(B.9) and (B.10) imply

$$\left| (a_0 + 2 \sum_{1 \leq j \leq m} a_j) - (a_0 + 2 \sum_{j=1}^{\infty} a_j (-1)^j \cos(j\theta_1)) \right| = o(m^{-1}) = o(1). \quad (\text{B.14})$$

It then suffices to prove that

$$\frac{S_{\sigma^2} - (a_0 + 2 \sum_{1 \leq j \leq m} a_j)^2 \text{tr}(\boldsymbol{\Sigma}^2)}{(a_0 + 2 \sum_{1 \leq j \leq m} a_j)^2 \text{tr}(\boldsymbol{\Sigma}^2)} = o_p(1).$$

Recall that $S_{\sigma^2} = S_{\sigma^2,0,0} + 4 \sum_{i=1}^m S_{\sigma^2,i,i} + 4 \sum_{i=1}^m S_{\sigma^2,0,i} + 8 \sum_{i=1}^{m-1} \sum_{j=i+1}^m S_{\sigma^2,i,j}$. Let $\tilde{S}_{\sigma^2,i,j} = S_{\sigma^2,i,j} - a_i a_j \text{tr}(\boldsymbol{\Sigma}^2)$. It is then sufficient to show that

$$\frac{\tilde{S}_{\sigma^2}}{(a_0 + 2 \sum_{1 \leq j \leq m} a_j)^2 \text{tr}(\boldsymbol{\Sigma}^2)} = \frac{\tilde{S}_{\sigma^2,0,0} + 4 \sum_{i=1}^m \tilde{S}_{\sigma^2,i,i} + 4 \sum_{i=1}^m \tilde{S}_{\sigma^2,0,i} + 8 \sum_{i=1}^{m-1} \sum_{j=i+1}^m \tilde{S}_{\sigma^2,i,j}}{(a_0 + 2 \sum_{1 \leq j \leq m} a_j)^2 \text{tr}(\boldsymbol{\Sigma}^2)} = o_p(1).$$

From the assumptions A2, A3 and (B.14) we have for big enough T

$$(a_0 + 2 \sum_{1 \leq j \leq m} a_j)^2 \text{tr}(\boldsymbol{\Sigma}^2) \geq \frac{1}{2} s^2 M_1^2 p, \quad (\text{B.15})$$

where we use the fact that

$$\text{tr}(\boldsymbol{\Sigma}^2) \geq \frac{(\text{tr} \boldsymbol{\Sigma})^2}{p}.$$

When T is big enough, for i and j not bigger than m ,

$$(T - j/2 - \frac{3}{2}[T/2])([T/2] - j - 1) \geq \frac{T^2}{9}. \quad (\text{B.16})$$

We next expand $\tilde{S}_{\sigma^2, i, j}$ in terms of Z_{ij} and write it a sum of the terms involving the high order of Z_{ij} and the terms involving the low order of Z_{ij} . Specifically write $\tilde{S}_{\sigma^2, i, j} = \tilde{S}_{\sigma^2, i, j, h} + \tilde{S}_{\sigma^2, i, j, l}$, where

$$\begin{aligned} \tilde{S}_{\sigma^2, i, j, h} &= \frac{1}{(T - j/2 - \frac{3}{2}[T/2])([T/2] - j - 1)} \sum_{f=2}^{[T/2]-j} \sum_{g=f+[T/2]}^{T-j} \left(\sum_{i_1, i_2=1}^p \Sigma_{i_1 i_1} \Sigma_{i_1 i_2} \sum_{s_1, s_2=-\infty}^T Z_{s_1 i_1}^3 Z_{s_2 i_2} \right. \\ &\quad (b_{f-s_1} b_{g-s_1} b_{f+i-s_1} b_{g+j-s_2} + b_{f-s_1} b_{g-s_1} b_{f+i-s_2} b_{g+j-s_1} + b_{f-s_1} b_{g-s_2} b_{f+i-s_2} b_{g+j-s_1} \\ &\quad \left. + b_{f-s_2} b_{g-s_1} b_{f+i-s_1} b_{g+j-s_1}) - 3 \sum_{i_1=1}^p \Sigma_{i_1 i_1}^2 \sum_{s_1=-\infty}^T Z_{s_1 i_1}^4 b_{f-s_1} b_{g-s_1} b_{f+i-s_1} b_{g+j-s_1} \right). \end{aligned} \quad (\text{B.17})$$

Note that $b_k = 0$ when $k < 0$. We then can conclude from the assumption A1 and (B.14) that

$$E|\tilde{S}_{\sigma^2, i, j, h}| = o(p^2 T^{-2}). \quad (\text{B.18})$$

Moreover write $\tilde{S}_{\sigma^2} = \tilde{S}_{\sigma^2, h} + \tilde{S}_{\sigma^2, l}$, where

$$\tilde{S}_{\sigma^2, h} = \tilde{S}_{\sigma^2, 0, 0, h} + 4 \sum_{i=1}^m \tilde{S}_{\sigma^2, i, i, h} + 4 \sum_{i=1}^m \tilde{S}_{\sigma^2, 0, i, h} + 8 \sum_{i=1}^{m-1} \sum_{j=i+1}^m \tilde{S}_{\sigma^2, i, j, h}.$$

(B.15) and (B.18) imply that

$$\frac{E|\tilde{S}_{\sigma^2, h}|}{(a_0 + 2 \sum_{1 \leq j \leq m} a_j)^2 \text{tr}(\boldsymbol{\Sigma}^2)} = o(m^2 p T^{-2}) = o(1). \quad (\text{B.19})$$

A direct calculation shows that

$$\begin{aligned} &(T - j/2 - \frac{3}{2}[T/2])([T/2] - j - 1) E \tilde{S}_{\sigma^2, i, j, l} \\ &= \sum_{f=2}^{[T/2]-j} \sum_{g=f+[T/2]}^{T-j} (a_{g+j-f} a_{g-f-i} \text{tr}(\boldsymbol{\Sigma}^2) + a_{g-f} a_{g+j-f-i} (\text{tr}(\boldsymbol{\Sigma}))^2) = o(p^2 T^{-1}). \end{aligned} \quad (\text{B.20})$$

This, together with (B.15) and (B.16), implies that

$$\frac{E \tilde{S}_{\sigma^2, l}}{(a_0 + 2 \sum_{1 \leq j \leq m} a_j)^2 \text{tr}(\boldsymbol{\Sigma}^2)} = o(m^2 p T^{-3}) = o(1). \quad (\text{B.21})$$

Via (B.15), (B.16), the assumption A1 and (B.14) one can also verify that

$$\text{Var}\left(\frac{\tilde{S}_{\sigma^2, l}}{(a_0 + 2 \sum_{1 \leq j \leq m} a_j)^2 \text{tr}(\boldsymbol{\Sigma}^2)}\right) = o(m^4 (p T^{-2} + p^{-1})) = o(1). \quad (\text{B.22})$$

This, together with (B.19) and (B.21), completes the proof. \square

Proof of Theorems 3 : We prove Remark 2 at first. Recalling the proof of Theorem 2, we just need to prove

$$\left\| \frac{1}{p} \mathbf{C} \mathbf{Y} \boldsymbol{\Sigma}^{1/2} \mathbf{X}_0^* \right\|_2 = o_p(p^{-1/2} T^2). \quad (\text{B.23})$$

(B.7) implies that $\left\| \frac{1}{p} \mathbf{C} \mathbf{Y} \boldsymbol{\Sigma}^{1/2} \mathbf{X}_0^* \right\|_2 = O_p(T^{3/2})$. Remark 2 then follows.

We now prove Theorem 3. The assumption A7 implies that

$$\left\| \frac{1}{p} \mathbf{X}_0 \mathbf{X}_0^* \right\|_2 = O_p(T). \quad (\text{B.24})$$

Our aim is to prove (B.23). Note that $\text{rank}(\mathbf{C}\mathbf{Y}\boldsymbol{\Sigma}^{1/2}\mathbf{X}_0^*) = 1$. Recalling the assumption A7 we can then find that

$$\left\| \frac{1}{p} \mathbf{C}\mathbf{Y}\boldsymbol{\Sigma}^{1/2}\mathbf{X}_0^* \right\|_2 = \frac{\sqrt{T}}{p} \sqrt{\sum_{t=1}^T \left(\sum_{i=1}^t \mathbf{y}'_i \boldsymbol{\Sigma}^{1/2} \mathbf{x}_0 \right)^2} \quad (\text{B.25})$$

$$\sum_{i=1}^t \mathbf{y}'_i \boldsymbol{\Sigma}^{1/2} \mathbf{x}_0 = \sum_{i=1}^t \mathbf{y}'_i \boldsymbol{\Sigma}^{1/2} \sum_{k=0}^{\infty} \tilde{b}_k \boldsymbol{\Sigma}_1^{1/2} \mathbf{z}_{-k} + \sum_{i=1}^t \mathbf{y}'_i \boldsymbol{\Sigma}^{1/2} \tilde{b}_{-1} \boldsymbol{\Sigma}_2^{1/2} \tilde{\mathbf{z}} + \sum_{i=1}^t \mathbf{y}'_i \boldsymbol{\Sigma}^{1/2} \tilde{\mathbf{b}}_{-2}. \quad (\text{B.26})$$

By (2.1) and a variable change we may write

$$\sum_{i=1}^t \mathbf{y}'_i = \sum_{j=1}^t \mathbf{z}'_j \left(\sum_{i=j}^t b_{i-j} \right) + \sum_{j=-\infty}^0 \mathbf{z}'_j \left(\sum_{i=1}^t b_{i-j} \right). \quad (\text{B.27})$$

Let $(\tilde{c}_{-2,1}, \dots, \tilde{c}_{-2,p})' = \tilde{\mathbf{c}}_{-2} = \boldsymbol{\Sigma}^{1/2} \tilde{\mathbf{b}}_{-2}$. The assumptions A3 and A7 imply $\|\tilde{\mathbf{c}}_{-2}\|^2 = O(p)$. Then

$$\sum_{i=1}^t \mathbf{y}'_i \boldsymbol{\Sigma}^{1/2} \tilde{\mathbf{b}}_{-2} = \sum_{i=1}^t \mathbf{y}'_i \tilde{\mathbf{c}}_{-2}.$$

It follows that

$$E \left(\sum_{i=1}^t \mathbf{y}'_i \boldsymbol{\Sigma}^{1/2} \tilde{\mathbf{b}}_{-2} \right) = 0 \quad (\text{B.28})$$

and

$$\text{Var} \left(\sum_{i=1}^t \mathbf{y}'_i \boldsymbol{\Sigma}^{1/2} \tilde{\mathbf{b}}_{-2} \right) = \|\tilde{\mathbf{c}}_{-2}\|^2 \left(\sum_{j=1}^t \left(\sum_{i=j}^t b_{i-j} \right)^2 + \sum_{j=-\infty}^0 \left(\sum_{i=1}^t b_{i-j} \right)^2 \right) = O(pt). \quad (\text{B.29})$$

These imply that

$$\sum_{i=1}^t \mathbf{y}'_i \boldsymbol{\Sigma}^{1/2} \tilde{\mathbf{b}}_{-2} = O_p(p^{1/2} t^{1/2}). \quad (\text{B.30})$$

As in (B.27) write

$$\sum_{i=1}^t \mathbf{y}'_i \boldsymbol{\Sigma}^{1/2} \tilde{b}_{-1} \boldsymbol{\Sigma}_2^{1/2} \tilde{\mathbf{z}} = \tilde{b}_{-1} \left(\sum_{j=1}^t \mathbf{z}'_j \boldsymbol{\Sigma}^{1/2} \boldsymbol{\Sigma}_2^{1/2} \tilde{\mathbf{z}} \left(\sum_{i=j}^t b_{i-j} \right) + \sum_{j=-\infty}^0 \mathbf{z}'_j \boldsymbol{\Sigma}^{1/2} \boldsymbol{\Sigma}_2^{1/2} \tilde{\mathbf{z}} \left(\sum_{i=1}^t b_{i-j} \right) \right).$$

The assumption A7 implies that $\tilde{\mathbf{z}}$ is independent of \mathbf{z}_t and that \tilde{b}_{-1} is bounded. It follows that

$$\sum_{i=1}^t \mathbf{y}'_i \boldsymbol{\Sigma}^{1/2} \tilde{b}_{-1} \boldsymbol{\Sigma}_2^{1/2} \tilde{\mathbf{z}} = O_p(p^{1/2} t^{1/2}). \quad (\text{B.31})$$

Now we consider the first term of the right hand of (B.26). From (B.27), write

$$\begin{aligned} & \sum_{i=1}^t \mathbf{y}'_i \boldsymbol{\Sigma}^{1/2} \sum_{k=0}^{\infty} \tilde{b}_k \boldsymbol{\Sigma}_1^{1/2} \mathbf{z}_{-k} \\ = & \sum_{j=1}^t \sum_{k=0}^{\infty} \mathbf{z}'_j \boldsymbol{\Sigma}^{1/2} \boldsymbol{\Sigma}_1^{1/2} \mathbf{z}_{-k} \tilde{b}_k \left(\sum_{i=j}^t b_{i-j} \right) + \sum_{j=-\infty}^0 \sum_{k=0}^{\infty} \mathbf{z}'_j \boldsymbol{\Sigma}^{1/2} \boldsymbol{\Sigma}_1^{1/2} \mathbf{z}_{-k} \tilde{b}_k \left(\sum_{i=1}^t b_{i-j} \right). \end{aligned}$$

Direct calculations indicate that

$$E\left(\sum_{i=1}^t \mathbf{y}_i' \boldsymbol{\Sigma}^{1/2} \sum_{k=0}^{\infty} \tilde{b}_k \boldsymbol{\Sigma}_1^{1/2} \mathbf{z}_{-k}\right) = \sum_{k=0}^{\infty} \text{tr}(\boldsymbol{\Sigma}^{1/2} \boldsymbol{\Sigma}_1^{1/2}) \tilde{b}_k \left(\sum_{i=1}^t b_{i+k}\right) = O(p) \quad (\text{B.32})$$

and

$$\text{Var}\left(\sum_{i=1}^t \mathbf{y}_i' \boldsymbol{\Sigma}^{1/2} \sum_{k=0}^{\infty} \tilde{b}_k \boldsymbol{\Sigma}_1^{1/2} \mathbf{z}_{-k}\right) = O(pt). \quad (\text{B.33})$$

(B.30)-(B.33) and the assumption A4 imply that

$$\left\| \frac{1}{p} \mathbf{C} \mathbf{Y} \boldsymbol{\Sigma}^{1/2} \mathbf{X}_0^* \right\|_2 = O_p(\max(p^{-1/2} T^{3/2}, T)) = o_p(p^{-1/2} T^2). \quad (\text{B.34})$$

The proof of Theorem 3 is complete. □

Proof of Theorem 1: Define $\mathbf{X}_{0\Pi} = (\Pi \mathbf{x}_0, \dots, \Pi^T \mathbf{x}_0)'$ and $\mathbf{X}_{1\Pi} = \mathbf{X} - \mathbf{X}_{0\Pi}$. Write

$$\begin{aligned} \mathbf{B} &= (1/p) \mathbf{X} \mathbf{X}^* \\ &= (1/p) \mathbf{X}_{1\Pi} \mathbf{X}_{1\Pi}^* + (1/p) \mathbf{X}_{1\Pi} \mathbf{X}_{0\Pi}^* + (1/p) \mathbf{X}_{0\Pi} \mathbf{X}_{1\Pi}^* + (1/p) \mathbf{X}_{0\Pi} \mathbf{X}_{0\Pi}^*. \end{aligned} \quad (\text{B.35})$$

Observe that

$$\|(1/p) \mathbf{X}_{0\Pi}^* \mathbf{X}_{0\Pi}\|_2 = \|(1/p) \sum_{t=1}^T \Pi^t \mathbf{x}_0 \mathbf{x}_0' \Pi^{t'}\|_2 \leq \frac{1}{p(1-\varphi^2)} \|\mathbf{x}_0\|_2^2. \quad (\text{B.36})$$

This, together with the assumption A6, implies

$$\|(1/p) \mathbf{X}_{0\Pi}^* \mathbf{X}_{0\Pi}\|_2 = O_p(1). \quad (\text{B.37})$$

Recalling (A.2) we have

$$\|(1/p) \mathbf{X}_{1\Pi}^* \mathbf{X}_{1\Pi}\|_2 \leq \frac{M_0}{(1-\varphi)^2} \|(1/p) \mathbf{Y}^* \mathbf{Y}\|_2.$$

We then conclude from (A.2), (A.3) and (B.5) that

$$\lim_{T \rightarrow \infty} P(\|(1/p) \mathbf{X}_{1\Pi}^* \mathbf{X}_{1\Pi}\|_2 \leq \frac{2 \sum_{i \geq 0} |a_i|}{(1-\varphi)^2} M_0 (1 + \sqrt{\frac{T}{p}})^2) = 1. \quad (\text{B.38})$$

By Holder's inequality

$$\|(1/p) \mathbf{X}_{0\Pi} \mathbf{X}_{1\Pi}^*\|_2 \leq \sqrt{\|(1/p) \mathbf{X}_{0\Pi}^* \mathbf{X}_{0\Pi}\|_2 \|(1/p) \mathbf{X}_{1\Pi}^* \mathbf{X}_{1\Pi}\|_2}. \quad (\text{B.39})$$

(B.37)-(B.39) ensure Theorem 1. □

Proof of Theorem 7: We claim that

$$\sum_{i=2}^T \frac{(\mathbf{x}_i - \mathbf{x}_{i-1})' (\mathbf{x}_i - \mathbf{x}_{i-1})}{p(T-1)} - \frac{2a_0 \text{tr}(\boldsymbol{\Sigma})}{p(1+\varphi)} \xrightarrow{i.p.} 0 \quad (\text{B.40})$$

and

$$\frac{|S_{\sigma^2}|}{p} - \frac{4a_0^2 \text{tr}(\boldsymbol{\Sigma}^2)}{p(1+\varphi)^2} \xrightarrow{i.p.} 0. \quad (\text{B.41})$$

Indeed, the proofs of (B.40) and (B.41) are similar to those of Theorem 4 and 5 (replacing $m_1 = m_2$ there by 0). Moreover from Theorem 1 we have $\rho_1 = o_p(T)$. This, together with (B.40) and (B.41), ensures that

$$\frac{T_N}{\sqrt{p}} + \frac{\frac{\text{tr}(\Sigma)}{p}}{\sqrt{\frac{2\text{tr}(\Sigma^2)}{p}}} \xrightarrow{i.p.} 0, \quad (\text{B.42})$$

which implies (3.6). □

C The simulation for the traditional case

We investigate the performance of T_N for the case where p is small and T is large.

Table 12: The results of the test T_N for the case where $Z_{ij} \sim N(0, 1)$

p	T	I(size)	0.99I (power)	0.97I (power)	0.95I (power)	0.9I (power)	Π_2 (power)
10	40	0.061	0.018	0.040	0.107	0.272	0.331
10	80	0.058	0.007	0.044	0.171	0.393	0.274
10	160	0.047	0.004	0.102	0.289	0.485	0.200
10	300	0.046	0.007	0.270	0.468	0.586	0.147
20	40	0.066	0.043	0.218	0.546	0.918	0.921
20	80	0.048	0.077	0.742	0.987	1	0.997
20	160	0.038	0.393	1	1	1	1
20	300	0.068	0.964	1	1	1	1

Table 13: The results of the test T_N when Z_{ij} follows a uniform distribution over $[-\sqrt{3}, \sqrt{3}]$

p	T	I(size)	0.99I (power)	0.97I (power)	0.95I (power)	0.9I (power)	Π_2 (power)
10	40	0.053	0.023	0.035	0.101	0.291	0.322
10	80	0.062	0.007	0.040	0.154	0.383	0.266
10	160	0.051	0.007	0.119	0.272	0.492	0.222
10	300	0.042	0.009	0.263	0.459	0.603	0.143
20	40	0.047	0.027	0.229	0.524	0.931	0.911
20	80	0.052	0.060	0.778	0.989	1	0.998
20	160	0.039	0.377	1	1	1	1
20	300	0.041	0.963	1	1	1	1

Different values of p , T and $\mathbf{\Pi}$ are given in Tables 12 and 13, which indicate that the size is approximately 0.05 even if p is as small as 10 and T is as small as 40. Its power increases as the sample size increases.

References

- [1] BAI, Z. D. and SILVERSTEIN, J. W. (2006). *Spectral Analysis of Large Dimensional Random Matrices*. 2nd Edition, Springer, New York.
- [2] BAI, Z.D. and YAO, J.F (2008). Central limit theorems for eigenvalues in a spiked population model. *Ann. Inst. H. Poincaré* **44**, 447-474.
- [3] BAIK, J., BEN AROUS, G. and PÉCHÉ, S. (2005). Phase transition of the largest eigenvalue for non-null complex sample covariance matrices. *Ann. Probab.* **33**, 1643–1697.
- [4] BAIK, J. and SILVERSTEIN, J. W. (2006). Eigenvalues of large sample covariance matrices of spiked population models. *J. Multivariate Anal.* **97**, 1382–1408.
- [5] BAO, Z. G., PAN, G. M. and ZHOU, W. (2015). Universality for the largest eigenvalue of sample covariance matrices with general population. *Ann. Statist.* **43**, 382–421.
- [6] CHAN, N.H. and WEI, C.Z. (1988). Limiting distributions of least squares estimates of unstable autoregressive processes. *Ann. Statist.* **16**, 367–401.
- [7] CHANG, Y. (2004). Bootstrap unit root tests in panels with cross sectional dependency. *Journal of Econometrics* **120**, 263–293.
- [8] CHEN, B. B. and PAN, G. M. (2012). Convergence of the largest eigenvalue of normalized sample covariance matrices when p and n both tend to infinity with their ratio converging to zero. *Bernoulli* **18**, 1405–1420.
- [9] CHOI, IN. (2001). Unit root tests for panel data. *J. International Money and Finance* **20**, 249–272.
- [10] DICKEY, D.A. and FULLER, W.A. (1979). Distribution of the estimators for autoregressive time series with an unit root. *J. Amer. Statist. Assoc.* **74**, 423–431.
- [11] EL KAROUI, N. (2007). Tracy-Widom limit for the largest eigenvalue of a large class of complex sample covariance matrices, *Ann. Probab.* **35**, 663–714.
- [12] FISHER, R. A. (1937). The sampling distribution of some statistics obtained from non-linear equations, *Ann. Eugenics.* **9**, 238–249.
- [13] HSU, P. L.(1939). On the distribution of roots of certain determinant equations, *Ann. Eugenics.* **9**, 250–258.
- [14] JOHNSTONE, I.M. (2001). On the distribution of the largest eigenvalue in principal component analysis. *Ann. Statist.* **29**, 295–327.
- [15] JOHNSTONE, I.M. (2007). High dimensional statistical inference and random matrices. *In International Congress of Mathematicians I*, 307–333.
- [16] LEVIN, A., LIN, C.F. and CHU, C.S.J.(2002). Unit root tests in panel data: asymptotic and finite-sample properties. *Journal of Econometrics* **108**, 1–24.

- [17] LIU, H. Y., AUE. A. and PAUL. D.(2015) - On the Marcenko-Pastur law for linear time series. *Ann. Statist.* **43**, 675–712.
- [18] MARČENKO, V. A. and PASTUR, L. A.(1967). Distribution for some sets of random matrices, *Math. USSR-Sb.* **1**, 457–483.
- [19] PAN, G. M , GAO, J and YANG, Y. R.(2014). Testing independence among a large number of high dimensional random vectors. *J. Amer. Statist. Assoc.* **109**, 600–612.
- [20] PAUL, D. (2007) . Asymptotics of sample eigen–structure for a large dimensional spiked covariance model.*Statist. Sinica.* **17**, 1617–1642.
- [21] PAUL, D. and AUE, A.(2014). Random matrix theory in statistics: a review. *J. Statist. Plan. & Infer.* **150**, 1–29.
- [22] IM, K., PESARAN, M.H. and SHIN, Y.(2003). Testing for unit roots in heterogeneous panels. *Journal of Econometrics* **115**, 53–74.
- [23] PHILLIPS, P and PERRON, P(1988). Testing for a unit root in time series regression. *Biometrika* **75**, 335–346.
- [24] ROY, S. N.(1939). P –statistics and some generalizations in analysis of variance appropriate to multivariate problems, *Sankhyā.* **4**, 381–396.
- [25] SOSHIKOV, A.(2002). A note on universality of the distribution of the largest eigenvalues in certain sample covariance matrices. *J. Stat. Phys.* **108**, 1033–1056.
- [26] TRACY, C. A. and WIDOM, H.(1994). Level-spacing distributions and the Airy kernel. *Comm. Math. Phys.* **159**, 151–174.
- [27] TRACY, C. A. and WIDOM, H.(1996). On orthogonal and symplectic matrix ensembles. *Comm. Math. Phys.* **177**, 727–754.
- [28] YAO, J.F., ZHENG, S. R. and BAI, Z.D. (2015). *Large Sample Covariance Matrices and High-Dimensional Data Analysis*. Cambridge University Press.
- [29] ZHANG, L. (2006). Spectral analysis of large dimensional random matrices. Ph.D. thesis, National University of Singapore, Singapore.