

Department of Econometrics and Business Statistics

<http://business.monash.edu/econometrics-and-business-statistics/research/publications>

Estimation and Testing for High– Dimensional Near Unit Root Time Series

Bo Zhang, Jiti Gao and Guangming Pan

April 2020

Working Paper 12/20

Estimation and Testing for High-Dimensional Near Unit Root Time Series

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

April 4, 2020

Abstract

This paper considers a n -dimensional time series model of the form

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

where $\mathbf{y}_t = (Y_{t1}, \dots, Y_{tn})^\top$ and $\mathbf{\Sigma}$ is the square root of a symmetric positive definite matrix. Here $\mathbf{\Pi}$ is a symmetric matrix which satisfies that $\|\mathbf{\Pi}\|_2 \leq 1$ and $T(1 - \|\mathbf{\Pi}\|_{\min})$ is bounded. The linear processes Y_{tj} is of the form $\sum_{k=0}^{\infty} b_k Z_{t-k,j}$ where $\sum_{i=0}^{\infty} |b_i| < \infty$ and $\{Z_{ij}\}$ are independent and identically distributed (i.i.d.) random variables with $E Z_{ij} = 0$, $E|Z_{ij}|^2 = 1$ and $E|Z_{ij}|^4 < \infty$. We first investigate the asymptotic behavior of the first k largest eigenvalues of the sample covariance matrices of the time series model. Then we propose a new estimator for the high-dimensional near unit root setting through using the largest eigenvalues of the sample covariance matrices and use it to test for near unit roots. Such an approach is theoretically novel and addresses some important estimation and testing issues in the high-dimensional near unit root setting. Simulations are also conducted to demonstrate the finite-sample performance of the proposed test statistic.

Keywords: Asymptotic normality, largest eigenvalue, linear process, near unit root test.

JEL Classification: C21, C32.

1 Introduction

There is a long literature about testing unit-root time series. Several testing methods have also been proposed for testing near unit root time series. The literature includes Phillips

*Bo Zhang, Department of Statistics and finance, University of Science and Technology of China, Hefei, 230026, China. Email: zhangbo890301@outlook.com

[†]*Corresponding Author:* Jiti Gao, Department of Econometrics and Business Statistics, Monash University, VIC 3145, Australia. Email: jiti.gao@monash.edu.

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

and Perron [24] for pioneering a highly popular unit-root test, Phillips [23] for establishing a novel estimation theory for a regression model associated with a unit root structure, Phillips and Xiao [25] for a survey of the relevant literature up to 1998, Phillips, Moon and Xiao [26] for estimating autoregressive roots near unity, and Moon and Phillips [17] for GMM estimation of panel data models with roots near unity. In the past two decades or so, there have been studies on testing unit roots and/or near unit roots in panel data settings. One key contribution is by Choi [9]. Some other contributions include Levin, Lin and Chu [16], Im, Pesaran, and Shin [13], Chang [7], Pesaran [20], Pesaran, Smith and Yamagata [22] and Pesaran [21].

The GMM estimation of panel data models with roots near unity in Moon and Phillips [17] is a classical method. However, it has some limitations. It has a convergence rate $\frac{1}{T\sqrt{n}}$ (n is the dimension of time series) when the root is smaller than 1 but the convergence rate is only $\frac{1}{Tn^{1/6}}$ when the root is 1. Moreover, it assumes that the error terms $\varepsilon_{it} \sim (0, \sigma^2)$ across i and t . In other words, it can't deal with the complicated cross-sectional dependence. As we know, the estimation of the complicated cross-sectional dependence is a persistent problem in high dimensional data. It leads to some difficulties for the tests and estimations when the cross-sectional dimension is proportional or greater than the time series sample size. So it's worthy exploring new methods for this case.

As demonstrated in Zhang, Pan and Gao [31], existing unit root tests available for the panel data setting are not directly applicable for the case where the cross-sectional dimension is proportional or greater than the time series sample size. Although there is complicated cross-sectional covariance matrix Σ in the model of Zhang, Pan and Gao [31], the paper shows that $tr(\Sigma)$, instead of Σ , affects the relation between unit root and the largest eigenvalues of sample covariance matrix. To deal with the near unit case, we establish some new relations between estimation and testing for near unit roots and the largest eigenvalues of the sample covariance matrix under consideration. Then we propose a new estimator for φ and establish a new statistic for testing near unit roots in the high-dimensional setting. One particular feature of the proposed test, which is quite different from those proposed in Zhang, Pan and Gao [31], is that the proposed test basically mimics the original version of the unit root test for the low-dimensional case (based on $\hat{\varphi} - \varphi$). Since the development of the theory for the proposed estimator and test relies on some asymptotic properties of the largest eigenvalues of high-dimensional sample covariance matrices, we will briefly discuss the relevant literature in the rest of this section.

In recent years, large sample properties for high-dimensional sample covariance matrices, including their eigenvalues and eigenvectors, have been proved to be extremely useful. In fact, random matrix theory provides many useful methods for estimation and testing procedures in high-dimensional data analysis. With respect to eigenvalues and eigenvectors, asymptotic properties for the largest eigenvalues are interesting and useful. 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. 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 [28] and [29] for Gaussian Wigner ensembles. The largest eigenvalue of the Wishart matrix was investigated in Johnstone [14]. Recent progress for general sample covariance matrices has also been made, and we refer to Bao, Pan and Zhou [5], Han, Pan and Zhang [12] and El Karoui [11] among others.

Empirical data from finance, speech recognition and wireless communication often implies that some extreme eigenvalues of sample covariance matrices are well separated from the rest. This introduces the second line of research about the spiked eigenvalues, which was first proposed in Johnstone [14]. There are some significant studies 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 [19] established a CLT for the spiked eigenvalues under the Gaussian population and the population spikes being simple. The asymptotic distribution 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 existing studies rely on the assumption that the observations of high dimensional data are independent. However, observations of high-dimensional data in economics and finance, for example, are often highly dependent. Zhang, Pan and Gao [31] may be the first paper to deal with the largest eigenvalues of sample covariance matrices generated from high-dimensional nonstationary time series data. Unfortunately, the result of [31] is also limited. It only considers the case that there is a unit root in the time series. As we know, there are many time series with near unit roots and our experience shows that the theory proposed in [31] is no longer applicable to near unit root settings. This therefore motivates us to develop some new estimation and testing procedures to deal with them.

This paper first proposes a new estimator for near unit roots. It then establishes some new asymptotic properties for the first k largest eigenvalues of the sample covariance matrices of the time series model with near unit roots. The paper then develops a new near unit root test for the high-dimensional case, which, as pointed out above, complements the approach proposed in [31]. As shown in Section 3 below, the proposed test naturally extends the well-known Phillips and Perron (PP) type (see [24]) test to the high-dimensional near unit root setting. Section 4 evaluates the finite-sample performance of the proposed test and then demonstrates that it outperforms its existing competitors.

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 2 also proposes a new near unit root test

that is devoted to testing nonstationarity for high dimensional dependent data. Section 3 evaluates both the size and power properties of the proposed test. Section 4 concludes the paper with some remarks. Appendix A establishes some useful lemmas for the largest eigenvalues of sample covariance matrices and the proof of the main theorems in section 2. Appendix B gives and proves the truncated versions of lemmas in Appendix A by truncating linear processes. Appendix C gives the proofs of lemmas in Appendix A.

2 Asymptotic Theory

2.1 Matrix models

The paper is to investigate high-dimensional sample covariance matrices for nonstationary time series. Let Y_{tj} be MA(q) with q being either fixed, or tending to infinity at a certain rate, or ∞ .

Define the linear processes Y_{tj} by

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

where $\sum_{i=0}^q |b_i| < \infty$, and $\{Z_{ij}\}$ are independent and identically distributed (i.i.d.) random variables with $E Z_{ij} = 0$, $E|Z_{ij}|^2 = 1$ and $E|Z_{ij}|^4 < \infty$. Suppose that $\mathbf{y}_t = (Y_{t1}, \dots, Y_{tp})^\top$ is a n -dimensional time series. 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 $\mathbf{\Pi}$ a $n \times n$ symmetric matrix.

Throughout the paper, we make the following assumptions about the coefficients b_i and $\mathbf{\Sigma}$:

Assumption 1 (The conditions about b_i). $\sum_{i=0}^q i|b_i| < \infty$ and $\sum_{i=0}^q b_i = s \neq 0$.

Assumption 2 (The conditions about $\mathbf{\Sigma}$). *There exist two positive constants M_0 and M_1 such that $\|\mathbf{\Sigma}\|_2 \leq M_0$ and $\text{tr}(\mathbf{\Sigma})/n \geq M_1$.*

We also need to impose an assumption on (T, n) .

Assumption 3 (The conditions about high-dimension). *Let $T \rightarrow \infty$ and $n \rightarrow \infty$ such that*

$$\lim_{T \rightarrow \infty} \frac{n^{1/2}}{T} = 0.$$

The first part of Assumption 1 implies that the linear process can include MA(q) models and AR(1) models. The second part of Assumption 1 is easily satisfied and plays an important role on the leading term of the largest eigenvalues. Assumption 2 is common. Assumption 3 shows that we do not require n and T to be of the same order, which is being commonly used in the random matrix theory literature.

We can also define a_i as follows.

$$a_i = \sum_{k=0}^q b_k b_{k+i}. \quad (2.3)$$

Then $a_i = EY_{tj}Y_{t+i,j}$.

Now we also need to make some assumptions about Z_{ij} and \mathbf{x}_0 .

Assumption 4 (The conditions about $\{Z_{ij}\}$). $\{Z_{ij}\}$ are independent and identically distributed random variables with mean zero, variance one and finite fourth moment. Let $\mathbf{z}_t = (Z_{t1}, \dots, Z_{tn})^\top$, where t can be either positive or negative integer (for the purpose of introducing Assumption 5 below).

Assumption 5 (The conditions about \mathbf{x}_0). $\mathbf{x}_0 = \sum_{k=0}^{\infty} \tilde{b}_k \Sigma_1^{1/2} \mathbf{z}_{-k} + \tilde{b}_{-1} \Sigma_2^{1/2} \tilde{\mathbf{z}} + \tilde{\mathbf{b}}_{-2}$, where $\|\Sigma_1\|_2 \leq M_0$, $\|\Sigma_2\|_2 \leq M_0$ and $\tilde{\mathbf{z}} = (\tilde{Z}_1, \dots, \tilde{Z}_n)^\top$ is independent of \mathbf{z}_t for any t , in which $\{\tilde{Z}_j\}$ are independent and identically distributed 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 $\|\tilde{\mathbf{b}}_{-2}\|^2 = O(n)$.

We would like to remark that Assumption 5 implies $E\|\mathbf{x}_0\|^2 = O(n)$. We next specify assumptions about $\mathbf{\Pi}$.

Assumption 6 (The conditions about $\mathbf{\Pi}$ and T). $\|\mathbf{\Pi}\|_2 \leq 1$,

$$\lim_{T \rightarrow \infty} T(1 - \|\mathbf{\Pi}\|_{\min}) = c \geq 0 \quad (2.4)$$

Definition 1.

$$\varphi = \frac{\text{tr}(\mathbf{\Pi}\Sigma)}{\text{tr}\Sigma}. \quad (2.5)$$

Assumption 7 (The conditions about $\mathbf{\Pi}$).

$$\frac{\text{tr}((\mathbf{\Pi} - \varphi \mathbf{I}_n)^2 \Sigma)}{\text{tr}\Sigma} = o\left(\frac{1}{T^2}\right). \quad (2.6)$$

Assumption 8 (The conditions about $\mathbf{\Pi}$).

$$\frac{\text{tr}((\mathbf{\Pi} - \varphi \mathbf{I}_n)^2 \Sigma)}{\text{tr}\Sigma} = o\left(\frac{1}{\sqrt{n}T^2}\right). \quad (2.7)$$

Remark 1. (2.2) is the same form as (2.2) in [31]. [31] considers two cases of (2.2): the CLT of largest eigenvalues for $\mathbf{\Pi} = \mathbf{I}$ and the upper bound of largest eigenvalues for $\|\mathbf{\Pi}\|_2 = \varphi < 1$ where φ is a constant. In this paper, the case is more complicated and challenging. Assumption 6 implies that $\mathbf{\Pi}$ is not a scale, so we can't immediately use the separable covariance matrix to deal with $\mathbf{X}\mathbf{X}^\top$. On the other hand, the eigenvalues of $\mathbf{\Pi}$ is close to 1, so \mathbf{x}_t is a non-stationary time series which is more complicated than the stationary ones.

2.2 The asymptotic theory for largest eigenvalues

Define the sample covariance matrix by

$$\mathbf{B} = \frac{1}{n} \mathbf{X} \mathbf{X}^\top. \quad (2.8)$$

In this section we show the asymptotic theory for largest eigenvalues of \mathbf{B} .

Definition 2. $\pi > \theta_{\varphi,1} > \cdots > \theta_{\varphi,T} > 0$ are the solutions to the equation

$$\varphi \sin T\theta + \sin(T+1)\theta = 0. \quad (2.9)$$

Definition 3. For $k = 1, \dots, T$,

$$\lambda_{\varphi,k} = \frac{1}{2\varphi(1 + \cos \theta_{\varphi,k}) + (1 - \varphi)^2} \quad (2.10)$$

and

$$\gamma_{\varphi,k} = \lambda_{\varphi,k} \left\{ a_0 + 2 \sum_{j=1}^{\infty} a_j (-1)^j \cos(j\theta_{\varphi,k}) \right\}, \quad (2.11)$$

where $\theta_{\varphi,k}$ is defined in (2.9).

Theorem 1. Suppose that Assumptions 1–7 hold. Let ρ_k be the k th largest eigenvalue of \mathbf{B} . When k is fixed,

$$\frac{\rho_k - \gamma_{\varphi,1} \frac{\text{tr} \boldsymbol{\Sigma}}{n}}{n} = o_p(1). \quad (2.12)$$

Moreover, when Assumption 8 also holds, the random vector

$$\frac{n}{[\text{tr}(2\boldsymbol{\Sigma}^2)]^{1/2}} \left(\frac{\rho_1 - \gamma_{\varphi,1} \frac{\text{tr} \boldsymbol{\Sigma}}{n}}{\gamma_{\varphi,1}}, \dots, \frac{\rho_k - \gamma_{\varphi,k} \frac{\text{tr} \boldsymbol{\Sigma}}{n}}{\gamma_{\varphi,k}} \right)^\top \quad (2.13)$$

converges weakly to a zero-mean Gaussian vector $\mathbf{w} = (w_1, \dots, w_k)^\top$ with the covariance function $\text{cov}(w_i, w_j) = 0$ for any $i \neq j$ and $\text{var}(w_i) = 1$.

3 Applications on near unit root

In this section, we give some applications of Theorem 1 on near unit root. Since the real data often contains some no random terms, we propose a new model with no random terms and near unit root at first.

3.1 Model and assumptions

Consider a n -dimensional time series model of the form:

$$\mathbf{x}_t - \delta_t = \boldsymbol{\Pi}(\mathbf{x}_{t-1} - \delta_{t-1}) + \boldsymbol{\Sigma}^{1/2} \mathbf{y}_t, \quad 1 \leq t \leq T, \quad \boldsymbol{\Pi} = \varphi \mathbf{I}_n, \quad (3.1)$$

where $\{\delta_t\}_{0 \leq t \leq T}$ is either a non-stochastic deterministic function or a stochastic trending function independent of $\{\mathbf{x}_t - \delta_t\}_{0 \leq t \leq T}$, and $0 < \varphi \leq 1$.

We also define two $T \times n$ matrices $\mathbf{\Delta} = (\delta_1, \dots, \delta_T)^\top$ and $\mathbf{\dot{X}}_0 = (\mathbf{x}_0 - \delta_0, \dots, \mathbf{x}_0 - \delta_0)^\top$ consisting of the initial vector \mathbf{x}_0 of the time series. Since there is δ_t in \mathbf{x}_t , we use a new matrix instead of \mathbf{B} .

$$\mathbf{\dot{B}} = \frac{1}{n} \mathbf{X}(\mathbf{I} - \mathbf{H})\mathbf{X}^\top, \quad (3.2)$$

where $\mathbf{H} = \frac{1}{n} \mathbf{1}\mathbf{1}'$ with $\mathbf{1}$ being a $1 \times n$ column vector. \mathbf{H} is introduced to remove the effect caused by δ_t , as commented in Remark 2.

Then we need some new assumptions.

Assumption 9 (The conditions about $\mathbf{x}_0 - \delta_0$). $\mathbf{x}_0 - \delta_0 = \sum_{k=0}^{\infty} \tilde{b}_k \mathbf{\Sigma}_1^{1/2} \mathbf{z}_{-k} + \tilde{b}_{-1} \mathbf{\Sigma}_2^{1/2} \tilde{\mathbf{z}} + \tilde{\mathbf{b}}_{-2}$, where $\|\mathbf{\Sigma}_1\|_2 \leq M_0$, $\|\mathbf{\Sigma}_2\|_2 \leq M_0$ and $\tilde{\mathbf{z}} = (\tilde{Z}_1, \dots, \tilde{Z}_n)^\top$ is independent of \mathbf{z}_t for any t , in which $\{\tilde{Z}_j\}$ are independent and identically distributed 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 $\|\tilde{\mathbf{b}}_{-2}\|^2 = O(n)$.

Assumption 10 (The conditions about δ_t). $\{\delta_t\}_{0 \leq t \leq T}$ is non-random or independent with $\{\mathbf{x}_t - \delta_t\}_{0 \leq t \leq T}$ and $\text{tr}\{\mathbf{\Delta}(\mathbf{I} - \mathbf{H})\mathbf{\Delta}^\top\} = \sum_{t=1}^T \delta_t^\top (\mathbf{I} - \mathbf{H}) \delta_t = o_p(n^{1/2}T^2)$.

Remark 2. The order of $\sum_{t=1}^T \delta_t^\top (\mathbf{I} - \mathbf{H}) \delta_t$ may be different from that of $\sum_{t=1}^T \delta_t^\top \delta_t$. For example, consider $\delta_t = g_t \mathbf{1} + \tilde{\delta}_t$ where g_t is large and $\|\tilde{\delta}_t\|$ is small. One may find that $\sum_{t=1}^T \delta_t^\top \delta_t$ is large but $\sum_{t=1}^T \delta_t^\top (\mathbf{I} - \mathbf{H}) \delta_t = \sum_{t=1}^T \tilde{\delta}_t^\top (\mathbf{I} - \mathbf{H}) \tilde{\delta}_t$ is small.

Assumption 11 (The conditions about φ , q and n).

$$\min\{q, n^{1/2}\} n^{1/2} (1 - \varphi) = O(1). \quad (3.3)$$

Assumption 12 (The conditions about $\delta_t - \delta_{t-1}$).

$$\frac{\min\{q, n^{1/2}\}}{n^{1/2}} \sum_{t=2}^T (\delta_t - \delta_{t-1})^\top (\mathbf{I} - \mathbf{H}) (\delta_t - \delta_{t-1}) = o_p(T)$$

3.2 Estimation and testing theory

With Theorem 1 and some calculations, we can get the following lemma.

Lemma 1. Suppose that Assumptions 1-6 and 9-10 hold. Let $\dot{\rho}_k$ be the k th largest eigenvalue of $\mathbf{\dot{B}}$. When k is fixed, the random vector

$$\frac{n}{[\text{tr}\{2(\mathbf{I} - \mathbf{H})\mathbf{\Sigma}(\mathbf{I} - \mathbf{H})\mathbf{\Sigma}\}]^{1/2}} \left(\frac{\dot{\rho}_1 - \gamma_{\varphi,1} \frac{\text{tr}\{\mathbf{\Sigma}^{1/2}(\mathbf{I} - \mathbf{H})\mathbf{\Sigma}^{1/2}\}}{n}}{\gamma_{\varphi,1}}, \dots, \frac{\dot{\rho}_k - \gamma_{\varphi,k} \frac{\text{tr}\{\mathbf{\Sigma}^{1/2}(\mathbf{I} - \mathbf{H})\mathbf{\Sigma}^{1/2}\}}{n}}{\gamma_{\varphi,k}} \right)^\top \quad (3.4)$$

converges weakly to a zero-mean Gaussian vector $\mathbf{w} = (w_1, \dots, w_k)^\top$ with the covariance function $\text{cov}(w_i, w_j) = 0$ for any $i \neq j$ and $\text{var}(w_i) = 1$.

We are now in a position to propose an estimator of φ and a statistic based on Lemma 1. The overall strategy is as follows. Since \mathbf{X} is observed one can calculate $\hat{\rho}_1$, the largest sample eigenvalue. Then a straightforward idea is to estimate $\varphi(\theta_{\varphi,1})$ by solving

$$\hat{\rho}_1 - \gamma_{\varphi,1} \frac{\text{tr}\{\boldsymbol{\Sigma}(\mathbf{I} - \mathbf{H})\}}{n} = 0 \quad (3.5)$$

which is inspired from Lemma 1. However, there are unknown a_i and $\boldsymbol{\Sigma}$ in the equation besides φ . Particularly it's hard to get a good estimator of $\boldsymbol{\Sigma}$ in the high-dimensional case. Fortunately, it is possible to estimate the products $a_i \text{tr}\{\boldsymbol{\Sigma}(\mathbf{I} - \mathbf{H})\}$ and $a_i [\text{tr}\{(\mathbf{I} - \mathbf{H})\boldsymbol{\Sigma}(\mathbf{I} - \mathbf{H})\boldsymbol{\Sigma}\}]^{1/2}$ involved in the asymptotic variance of $\hat{\rho}_1$. We then plug these estimators into (3.5) to get the estimators of $\theta_{\varphi,1}$ and φ .

We now start with estimating q defined in (2.1) as follow. Set

$$\zeta_j = \frac{\sum_{t=2}^{T-j} \dot{\mathbf{x}}_{t,t+j}}{n(T-j-1)}, \quad \dot{\mathbf{x}}_{t,s} = (\mathbf{x}_t - \mathbf{x}_{t-1})^\top (\mathbf{I} - \mathbf{H}) (\mathbf{x}_s - \mathbf{x}_{s-1}). \quad (3.6)$$

If q is bounded then its estimator is

$$\hat{q} = \min \left\{ 0 \leq i < \lfloor n^{1/2} \rfloor : |\zeta_j| < 2n^{-1/4}T^{-1/2}, i < j < \lfloor n^{1/2} \rfloor \right\}. \quad (3.7)$$

The estimator is based on the facts that the leading term of $E\zeta_j$ is $a_j \text{tr}\{\boldsymbol{\Sigma}^{1/2}(\mathbf{I} - \mathbf{H})\boldsymbol{\Sigma}^{1/2}\}$, the value of a_j is zero when $j > q$ and that the order of the standard deviation of ζ_j is smaller than $n^{-1/4}T^{-1/2}$. If q is not bounded then choose \hat{q} as the largest integer not larger than $n^{1/2}$. In order to simplify notation we below use m for \hat{q} .

The next step is to estimate $\gamma_{\varphi,1} \text{tr}\{\boldsymbol{\Sigma}(\mathbf{I} - \mathbf{H})\}$. Define

$$\dot{\mu}_m = \sum_{t=2}^T \frac{\dot{\mathbf{x}}_{t,t}}{n(T-1)} + 2 \sum_{j=1}^m \sum_{t=2}^{T-j} \frac{\dot{\mathbf{x}}_{t,t+j}}{n(T-j-1)}. \quad (3.8)$$

It turns out that $\frac{\text{tr}[\boldsymbol{\Sigma}(\mathbf{I} - \mathbf{H})]}{n} \kappa(\theta_{\varphi,1})(a_0 + 2 \sum_{j=1}^{\infty} a_j)$ is the leading term of $\dot{\mu}_m$ where

$$\kappa(\theta) = \left(-\frac{\sin(T+1)\theta}{\sin T\theta} \right)^m - m \frac{\left(-\frac{\sin(T+1)\theta}{\sin T\theta} \right)^3 - \left(-\frac{\sin(T+1)\theta}{\sin T\theta} \right)^{2T-m}}{2(T-1)}.$$

When $\varphi = 1$, $\kappa(\theta_{1,1}) = 1$. Otherwise, $\kappa(\theta_{\varphi,1}) = 1 + O(m(1-\varphi))$. Moreover the estimator of the term which contains $a_i [\text{tr}\{(\mathbf{I} - \mathbf{H})\boldsymbol{\Sigma}(\mathbf{I} - \mathbf{H})\boldsymbol{\Sigma}\}]^{1/2}$ can be constructed as follows. One may verify that $\text{Var}(\dot{\mathbf{x}}_{t,s}^2) = (a_{|t-s|}^2 + a_0^2) \text{tr}\{(\mathbf{I} - \mathbf{H})\boldsymbol{\Sigma}(\mathbf{I} - \mathbf{H})\boldsymbol{\Sigma}\} \{1 + o(1)\}$ and $a_{|t-s|}$ is negligible when $|t-s|$ is sufficiently large. Hence we define

$$\dot{S}_{\sigma^2,0} = \frac{\sum_{t=2}^{\lfloor T/2 \rfloor} \sum_{s=t+\lfloor T/2 \rfloor}^T \dot{\mathbf{x}}_{t,s}^2}{(T - \frac{3}{2}\lfloor T/2 \rfloor)(\lfloor T/2 \rfloor - 1)} \quad \text{and} \quad \dot{S}_{\sigma^2,m} = \frac{|\dot{\mu}_m| (2 \frac{\dot{S}_{\sigma^2,0}}{n})^{1/2}}{\sum_{i=2}^T \frac{\dot{\mathbf{x}}_{i,i}}{n(T-1)}}. \quad (3.9)$$

In addition, the reciprocal of $\lambda_{\varphi,1}$ can be rewritten as

$$g_2(\theta) = -2 \frac{\sin(T+1)\theta}{\sin T\theta} (1 + \cos \theta) + \left(1 + \frac{\sin(T+1)\theta}{\sin T\theta} \right)^2 = \frac{1 - \cos 2\theta}{1 - \cos 2T\theta}.$$

In view of (3.5) we define

$$h(\theta) = \dot{\rho}_1 g_2(\theta) \kappa(\theta) - \dot{\mu}_m \quad (3.10)$$

and obtain the largest solution $\hat{\theta}$ to $h(\theta) = 0$ on $(0, \pi)$. From (2.9), we can calculate $\hat{\varphi} = g_1(\hat{\theta})$, where $g_1(\theta) = -\frac{\sin(T+1)\theta}{\sin T\theta}$.

Then we have the following results.

Theorem 2. *Suppose that Assumptions 1–6 and 9–12 hold. Then,*

$$n^{1/2} \frac{\dot{\rho}_1 g_2(\theta_{\varphi,1}) \kappa(\theta_{\varphi,1}) - \dot{\mu}_m}{\dot{S}_{\sigma^2,m}} \rightarrow N(0, 1). \quad (3.11)$$

Proposition 1. *Under the condition of Theorem 2 and assume that q is finite, then*

$$\lim_{p, T \rightarrow \infty} P(\hat{q} \geq q) = 1. \quad (3.12)$$

Define $\tilde{T}_N = \frac{n^{1/2} \rho_1 f'(\hat{\theta})}{\dot{S}_{\sigma^2,m} g_1'(\hat{\theta})} (\hat{\varphi} - \varphi)$, where $f(\theta) = g_2(\theta) \kappa(\theta)$. We can establish our main results as follows.

Theorem 3. *Suppose that Assumptions 1–6 and 9–12 hold.*

(1) *Then the solution $\hat{\theta}$ is unique such that*

$$\lim_{p, T \rightarrow \infty} P \left(\text{there is } \hat{\theta} \in \left(\frac{T\pi}{T+1}, \frac{T\pi}{T+\frac{1}{4}} \right) \text{ so that } h(\hat{\theta}) = 0 \right) = 1. \quad (3.13)$$

(2) *As $n, T \rightarrow \infty$, we have*

$$n^{1/2} \frac{\rho_1 f'(\hat{\theta})}{\dot{S}_{\sigma^2,m}} (\hat{\theta} - \theta_{\varphi,1}) \rightarrow N(0, 1). \quad (3.14)$$

(3) *Moreover, we have as $n, T \rightarrow \infty$*

$$\tilde{T}_N \rightarrow N(0, 1). \quad (3.15)$$

Remark 3. *Equation (3.15) shows that the convergence rate of $\hat{\varphi}$ is $\frac{\dot{S}_{\sigma^2,m} g_1'(\hat{\theta})}{n^{1/2} \dot{\rho}_1 f'(\hat{\theta})} = O(T^{-1} n^{-1/2})$. We can compare it with the method in [17]. The paper by [17] considers an AR(1) process with the fixed term $\beta_i' g_{pt}$, where $g_{pt} = (t, t^2, \dots, t^p)'$. The dimension of time series in [17] is n , and p is just a fixed constant (i.e. $p = 1$ means linear trend). The error term (in [17]) is that $\varepsilon_{it} \sim N(0, \sigma^2)$ across i and t so that it doesn't allow for cross-sectional dependence. On the other hand, the method in [17] is based on two moment conditions $m_{1,it}(c)$ and $m_{2,it}(c)$. When $\varphi = 1 + c_0/T$ and $c_0 < 0$, the method in [17] has the same rate of convergence as that by us. However, when $c_0 = 0$ (or very near to 0, i.e. $\varphi = 1 - T^{-1} n^{-1}$), the first order information in the moment condition is asymptotically zero at the true parameter. Then it leads to the slower rate of convergence $O(T^{-1} n^{-1/6})$. More details are available in Section 4 of [17].*

Remark 4. \tilde{T}_N has the form $\frac{n^{1/2} \rho_1 f'(\hat{\theta})}{S_{\sigma^2, m, g_1'(\hat{\theta})}}(\hat{\varphi} - \varphi)$. A natural idea is to compare \tilde{T}_N with the PP type (see [24]) test. The main difference is that the asymptotic distribution of the proposed test is standard normal, while the PP type test has a non-Gaussian limiting distribution. The dimension n going to infinity plays an important role in the difference. Moreover, our model and testing theory accommodates a general cross-sectional dependence structure.

We would like to further comment that \tilde{T}_N is applicable to many different choices of φ involved in (2.4). For a general form of $\varphi = 1 - \chi/T$, we consider both unit-root and near unit-root scenarios: $\chi = 0$, $\chi = n^{-1}$, $\chi = n^{-1/2}$ or $\chi = 1$. As shown in the simulation studies in Section 4 below, \tilde{T}_N works well numerically.

4 Simulation

4.1 The simulation for \tilde{T}_N

This subsection is to conduct some simulations to investigate the size and power of \tilde{T}_N . We now consider the setting where $\mathbf{y}_t = \psi \mathbf{z}_{t-1} + \mathbf{z}_t$, $\psi = 0.5$ and $\Sigma = (\Sigma_{i,j}) = (0.3^{|i-j|})$. We also set the elements of δ_t as $\delta_{it} = \cos\left(\frac{2\pi(i+t)}{T}\right)$.

Since \tilde{T}_N works for different φ , we rewrite $\varphi = 1 - \frac{\chi}{T}$. The bigger χ is, the further φ goes to 1. The size and power results of \tilde{T}_N based on 1000 replications and different values of p , T , χ for the different hypotheses are reported in Tables 1–4. The sizes and power values were all computed using asymptotic critical values from the standard normal distribution.

Tables 1–4 below provide us with the following findings.

- Table 1 shows that the test \tilde{T}_N has good sizes and power values even when n and T are as small as 10 and 20, respectively. Meanwhile, the power values increase when χ increases, and the sizes become stable even when n increases to 80. Table 1 particularly reveals some findings for the case where the null hypothesis is a unit root of the form: $H_0 : \varphi = 1$, and the alternative is a near unit root of the form: $H_1 : \varphi = 1 - \frac{1}{T}$. In this setting, the power values are quite substantial with 0.422 for the case of $(n, T) = (20, 20)$.
- Tables 2–4 have similar features to those observed in Table 1, but there are some differences. For instances, Table 2 shows that when $H_0 : \chi = \frac{1}{n}$ holds, the so-called power values under $H_1 : \chi = 0$ or $H_1 : \chi = \frac{1}{\sqrt{n}}$ look more like sizes. Under $H_1 : \chi = 1$, however, when the departure is substantial, the power values increase significantly, and they also increase when either n or T increases.

Table 1: Sizes and power values under $H_0 : \chi = 0$

n	T	$\chi = 0$ (size)	$\chi = \frac{1}{n}$ (power)	$\chi = \frac{1}{\sqrt{n}}$ (power)	$\chi = 1$ (power)
20	20	0.035	0.041	0.075	0.422
20	30	0.033	0.044	0.067	0.436
20	40	0.038	0.038	0.065	0.453
20	60	0.040	0.034	0.065	0.438
20	80	0.049	0.041	0.060	0.435
30	20	0.044	0.048	0.078	0.627
30	30	0.043	0.043	0.075	0.612
30	40	0.034	0.035	0.063	0.662
30	60	0.042	0.045	0.063	0.654
30	80	0.038	0.044	0.063	0.665
40	20	0.054	0.049	0.104	0.744
40	30	0.043	0.040	0.077	0.758
40	40	0.036	0.049	0.082	0.793
40	60	0.034	0.039	0.070	0.812
40	80	0.037	0.046	0.064	0.818
60	20	0.041	0.070	0.096	0.875
60	30	0.037	0.054	0.078	0.928
60	40	0.050	0.035	0.076	0.926
60	60	0.039	0.047	0.070	0.945
60	80	0.043	0.054	0.096	0.936
80	20	0.062	0.058	0.100	0.921
80	30	0.044	0.048	0.104	0.985
80	40	0.040	0.048	0.082	0.986
80	60	0.044	0.033	0.084	0.980
80	80	0.042	0.047	0.094	0.982

Table 2: Sizes and power values under $H_0 : \chi = \frac{1}{n}$

n	T	$\chi = 0$ (power)	$\chi = \frac{1}{n}$ (size)	$\chi = \frac{1}{\sqrt{n}}$ (power)	$\chi = 1$ (power)
20	20	0.035	0.036	0.061	0.376
20	30	0.033	0.037	0.059	0.394
20	40	0.038	0.035	0.060	0.406
20	60	0.040	0.033	0.055	0.405
20	80	0.053	0.037	0.046	0.397
30	20	0.039	0.042	0.068	0.589
30	30	0.039	0.041	0.057	0.587
30	40	0.035	0.034	0.052	0.628
30	60	0.046	0.043	0.051	0.621
30	80	0.042	0.041	0.055	0.635
40	20	0.049	0.041	0.094	0.719
40	30	0.040	0.036	0.070	0.737
40	40	0.031	0.046	0.069	0.769
40	60	0.035	0.043	0.064	0.794
40	80	0.038	0.043	0.057	0.799
60	20	0.034	0.065	0.087	0.865
60	30	0.034	0.050	0.073	0.916
60	40	0.048	0.033	0.072	0.915
60	60	0.040	0.049	0.060	0.937
60	80	0.041	0.052	0.081	0.926
80	20	0.058	0.054	0.092	0.920
80	30	0.043	0.045	0.095	0.982
80	40	0.038	0.047	0.072	0.985
80	60	0.044	0.040	0.076	0.977
80	80	0.040	0.050	0.085	0.978

Table 3: Sizes and power values under $H_0 : \chi = \frac{1}{\sqrt{n}}$

n	T	$\chi = 0$ (power)	$\chi = \frac{1}{n}$ (power)	$\chi = \frac{1}{\sqrt{n}}$ (size)	$\chi = 1$ (power)
20	20	0.027	0.026	0.038	0.263
20	30	0.043	0.034	0.032	0.260
20	40	0.041	0.040	0.038	0.272
20	60	0.048	0.045	0.039	0.270
20	80	0.061	0.053	0.044	0.249
30	20	0.038	0.033	0.040	0.433
30	30	0.053	0.044	0.032	0.441
30	40	0.056	0.044	0.034	0.459
30	60	0.066	0.061	0.037	0.465
30	80	0.060	0.059	0.034	0.455
40	20	0.037	0.023	0.058	0.585
40	30	0.044	0.041	0.040	0.598
40	40	0.053	0.050	0.042	0.605
40	60	0.047	0.065	0.047	0.631
40	80	0.060	0.069	0.035	0.648
60	20	0.029	0.036	0.045	0.774
60	30	0.045	0.031	0.046	0.845
60	40	0.064	0.041	0.053	0.822
60	60	0.055	0.064	0.037	0.838
60	80	0.057	0.074	0.044	0.840
80	20	0.023	0.034	0.045	0.882
80	30	0.033	0.034	0.052	0.948
80	40	0.052	0.050	0.046	0.957
80	60	0.058	0.049	0.037	0.946
80	80	0.063	0.069	0.042	0.955

Table 4: Sizes and power values under $H_0 : \chi = 1$

n	T	$\chi = 0$ (power)	$\chi = \frac{1}{n}$ (power)	$\chi = \frac{1}{\sqrt{n}}$ (power)	$\chi = 1$ (size)
20	20	0.348	0.323	0.226	0.042
20	30	0.470	0.414	0.282	0.041
20	40	0.497	0.437	0.317	0.037
20	60	0.522	0.500	0.333	0.043
20	80	0.555	0.494	0.365	0.041
30	20	0.504	0.478	0.344	0.033
30	30	0.613	0.610	0.418	0.040
30	40	0.680	0.637	0.477	0.036
30	60	0.686	0.660	0.542	0.034
30	80	0.713	0.692	0.526	0.036
40	20	0.632	0.607	0.448	0.044
40	30	0.775	0.745	0.589	0.044
40	40	0.782	0.747	0.604	0.053
40	60	0.828	0.807	0.673	0.045
40	80	0.839	0.803	0.673	0.046
60	20	0.807	0.749	0.646	0.048
60	30	0.899	0.876	0.769	0.044
60	40	0.911	0.912	0.823	0.048
60	60	0.930	0.919	0.825	0.034
60	80	0.937	0.931	0.837	0.036
80	20	0.872	0.878	0.779	0.045
80	30	0.952	0.947	0.887	0.038
80	40	0.965	0.955	0.915	0.041
80	60	0.980	0.984	0.934	0.052
80	80	0.982	0.978	0.940	0.050

4.2 The simulation for $\hat{\varphi} - \varphi$

Now we focus on the behavior of $\hat{\varphi} - \varphi$. In other words, we care the behavior of estimation. We use the same setting as the above section. Recalling Remark 3, $\hat{\varphi} - \varphi = O_p(\frac{1}{T\sqrt{n}})$. Then $T^2n(\hat{\varphi} - \varphi)^2 = O_p(1)$. So we calculate the mean of $T^2n(\hat{\varphi} - \varphi)^2/4$ based on 1000 replications and different values of n , T and $\chi(\varphi)$ in Table 5. Table 5 shows that $\hat{\varphi} - \varphi$ has a convergence rate $\frac{1}{T\sqrt{n}}$.

Table 5: the mean of $T^2n(\hat{\varphi} - \varphi)^2/4$

n	T	$\chi = 0$	$\chi = \frac{1}{n}$	$\chi = \frac{1}{\sqrt{n}}$	$\chi = 1$
20	20	1.3754	1.5352	1.6711	2.7702
20	30	1.1999	1.2610	1.4096	1.8408
20	40	1.0865	1.1684	1.2341	1.6764
20	60	1.0867	1.1076	1.2644	1.6051
20	80	1.1116	1.1295	1.1237	1.6379
30	20	1.4247	1.4883	1.5624	2.4028
30	30	1.2099	1.2381	1.2433	1.8688
30	40	1.0726	1.0669	1.1787	1.6307
30	60	1.0656	1.1450	1.0521	1.4518
30	80	1.0160	1.0830	1.0571	1.5252
40	20	1.4265	1.4588	2.1065	3.4832
40	30	1.0790	1.1512	1.2601	1.9609
40	40	1.0936	1.1726	1.2232	1.7571
40	60	0.9418	1.0686	1.1141	1.5655
40	80	0.9816	1.0869	0.9975	1.3137
60	20	1.3737	1.6416	1.5330	4.2121
60	30	1.1282	1.2343	1.1402	3.2031
60	40	1.1184	1.0110	1.2183	1.7469
60	60	1.0384	1.1015	1.0936	1.4613
60	80	0.9706	1.0873	1.1293	1.4726
80	20	1.5966	1.8976	2.1593	5.8800
80	30	1.1685	1.2091	1.2543	2.5119
80	40	1.0690	1.1528	1.1673	1.5459
80	60	1.0313	0.9233	1.0518	1.6845
80	80	0.9832	1.0022	1.0530	1.5762

5 Conclusions and discussion

This paper has proposed a new estimator $\hat{\varphi}$ for the case where the dimensionality of a vector of time series diverges along with the sample size. Then a new statistic based on $\hat{\varphi} - \varphi$ for near unit test is also proposed. The proposed test has been constructed in a similar fashion to that of the original Dickey–Fuller test, and the main difference is that the asymptotic distribution of the proposed test is standard normal, while the DF test has a nonstandard limiting distribution.

As shown in the simulation study, both the estimator and the proposed test work well in the finite–sample cases.

6 Acknowledgements

The authors would like to thank the Australian Research Council Discovery Grants Program for its support under Grant numbers: DP150101012 & DP170104421.

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. *Annales de l'Institut Henri 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. *Annals of Probability* **33**, 1643–1697.
- [4] BAIK, J. and SILVERSTEIN, J. W. (2006). Eigenvalues of large sample covariance matrices of spiked population models. *Journal of Multivariate Analysis* **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. *Annals of Statistics* **43**, 382–421.
- [6] CHAN, N.H. and WEI, C.Z. (1988). Limiting distributions of least squares estimates of unstable autoregressive processes. *Annals of Statistics* **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. *Journal of 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. *Journal of the American Statistical Association* **74**, 423–431.

- [11] EL KAROUÏ, N. (2007). Tracy-Widom limit for the largest eigenvalue of a large class of complex sample covariance matrices, *Annals of Probability* **35**, 663–714.
- [12] HAN, X., PAN, G. M. and ZHANG, B. (2016). The Tracy–Widom law for the largest eigenvalue of F type matrix, *Annals of Statistics* **44**, 1564–1592.
- [13] IM, K., PESARAN, M.H. and SHIN, Y. (2003). Testing for unit roots in heterogeneous panels. *Journal of Econometrics* **115**, 53–74.
- [14] JOHNSTONE, I.M. (2001). On the distribution of the largest eigenvalue in principal component analysis. *Annals of Statistics* **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] MOON H. R. and PHILLIPS, P, C. B. (2004). GMM Estimation of Autoregressive Roots Near Unity with Panel Data. *Econometrica* **72**, 467–522.
- [18] PAN, G. M., GAO, J. and YANG, Y. R. (2014). Testing independence among a large number of high dimensional random vectors. *Journal of the American Statistical Association* **109**, 600–612.
- [19] PAUL, D. (2007). Asymptotics of sample eigen–structure for a large dimensional spiked covariance model. *Statistica Sinica* **17**, 1617–1642.
- [20] PESARAN, M.H. (2007). A simple panel unit root test in the presence of cross–sectional dependence. *Journal of Applied Econometrics* **22**, 265–312.
- [21] PESARAN, M.H. (2015). *Time Series and Panel Data Econometrics*. Oxford University Press, Oxford.
- [22] PESARAN, M.H., SMITH, L. V. and YAMAGATA, T. (2013). Panel unit root tests in the presence of a multifactor error structure. *Journal of Econometrics* **175**, 94–115.
- [23] PHILLIPS, P, C. B. (1988). Regression theory for near integrated time series. *Econometrica* **56**, 1021–1044.
- [24] PHILLIPS, P, C. B. and PERRON, P. (1988). Testing for a unit root in time series regression. *Biometrika* **75**, 335–346.
- [25] PHILLIPS, P, C. B. and XIAO, Z. (1998). A Primer on unit root testing. *Journal of Economic Surveys* **12**, 423–470.
- [26] PHILLIPS, P, C. B., MOON H. R. and XIAO, Z. (2001). How to estimate autoregressive roots near unity ? *Econometric Theory* **17**, 26–69.
- [27] SOSHNIKOV, A. (2002). A note on universality of the distribution of the largest eigenvalues in certain sample covariance matrices. *Journal of Statistical Physics* **108**, 1033–1056.
- [28] TRACY, C. A. and WIDOM, H. (1994). Level-spacing distributions and the Airy kernel. *Communications in Mathematical Physics* **159**, 151–174.
- [29] TRACY, C. A. and WIDOM, H. (1996). On orthogonal and symplectic matrix ensembles. *Communications in Mathematical Physics* **177**, 727–754.

- [30] YAO, J.F., ZHENG, S. R. and BAI, Z.D. (2015). *Large Sample Covariance Matrices and High-Dimensional Data Analysis*. Cambridge University Press.
- [31] ZHANG, B , PAN, G. M, and GAO, J. (2018). CLT for Largest eigenvalues and unit root tests for high-dimensional nonstationary time series. *Annals of Statistics* **46**, 2186–2215.

A Proofs of the main results

We first will give some lemmas which are proved in Appendix B and C.

Lemma 2. *Suppose that Assumptions 1–6 hold and $\mathbf{\Pi} = \varphi \mathbf{I}_n$. Let ρ_k be the k th largest eigenvalue of \mathbf{B} . When k is fixed, the random vector*

$$\frac{n}{[\text{tr}(2\mathbf{\Sigma}^2)]^{1/2}} \left(\frac{\rho_1 - \gamma_{\varphi,1} \frac{\text{tr}\mathbf{\Sigma}}{n}}{\gamma_{\varphi,1}}, \dots, \frac{\rho_k - \gamma_{\varphi,k} \frac{\text{tr}\mathbf{\Sigma}}{n}}{\gamma_{\varphi,k}} \right)^\top$$

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 $\text{var}(w_i) = 1$.

Lemma 3. *If the Assumptions 1–6 and 9–12 hold, then*

$$\kappa(\theta_{\varphi,1}) = 1 + O_p(m(1 - \varphi)) \quad (\text{A.1})$$

and

$$\begin{aligned} \dot{\mu}_m &= \frac{\text{tr}\{\mathbf{\Sigma}^{1/2}(\mathbf{I} - \mathbf{H})\mathbf{\Sigma}^{1/2}\}}{n} \kappa(\theta_{\varphi,1}) (a_0 + 2 \sum_{j=1}^{\infty} a_j) \\ &+ o_p\{m(1 - \varphi)\} + O_p(1 - \varphi) + o_p(n^{-1/2}). \end{aligned} \quad (\text{A.2})$$

Lemma 4. *Under the conditions specified in lemma 3,*

$$\frac{\dot{S}_{\sigma^2,0}}{n} = a_0^2 \frac{\text{tr}\{\mathbf{\Sigma}^{1/2}(\mathbf{I} - \mathbf{H})\mathbf{\Sigma}(\mathbf{I} - \mathbf{H})\mathbf{\Sigma}^{1/2}\}}{n} + o_p(1), \quad (\text{A.3})$$

$$\dot{S}_{\sigma^2,m} = \frac{|\dot{\mu}_{m_2}| (2 \frac{\dot{S}_{\sigma^2,0}}{n})^{1/2}}{\sum_{i=2}^T \frac{\mathbf{x}_{i,i}}{n(T-1)}} = (a_0 + 2 \sum_{i=1}^{\infty} a_i) \left[\frac{2}{n} \text{tr}\{(\mathbf{I} - \mathbf{H})\mathbf{\Sigma}(\mathbf{I} - \mathbf{H})\mathbf{\Sigma}\} \right]^{1/2} + o_p(1) \quad (\text{A.4})$$

and

$$\frac{\gamma_{\varphi,1} \left[\frac{2}{n} \text{tr}\{(\mathbf{I} - \mathbf{H})\mathbf{\Sigma}(\mathbf{I} - \mathbf{H})\mathbf{\Sigma}\} \right]^{1/2}}{\lambda_{\varphi,1} \dot{S}_{\sigma^2,m}} = 1 + o_p(1). \quad (\text{A.5})$$

A.1 The proof of Theorem 1

Proof of Theorem 1: Rewrite $\mathbf{\Pi}$ as

$$\mathbf{\Pi} = \mathbf{U}_{\mathbf{\Pi}} \mathbf{\Lambda}_{\mathbf{\Pi}} \mathbf{U}_{\mathbf{\Pi}}^*. \quad (\text{A.6})$$

Then

$$\mathbf{U}_{\mathbf{\Pi}}^* \mathbf{x}_t = \mathbf{\Lambda}_{\mathbf{\Pi}} \mathbf{U}_{\mathbf{\Pi}}^* \mathbf{x}_{t-1} + \mathbf{U}_{\mathbf{\Pi}}^* \mathbf{\Sigma}^{1/2} \mathbf{y}_t. \quad (\text{A.7})$$

Define $\ddot{\mathbf{x}}_t = \mathbf{U}_\Pi^* \mathbf{x}_t$, $\mathbf{\Omega} = \mathbf{U}_\Pi^* \mathbf{\Sigma} \mathbf{U}_\Pi$,

$$\ddot{\mathbf{x}}_t = \mathbf{\Lambda}_\Pi \ddot{\mathbf{x}}_{t-1} + \mathbf{\Omega}^{1/2} \mathbf{y}_t \quad (\text{A.8})$$

and

$$\tilde{\mathbf{x}}_t = \varphi \tilde{\mathbf{x}}_{t-1} + \mathbf{\Omega}^{1/2} \mathbf{y}_t. \quad (\text{A.9})$$

Note that $\text{tr} \mathbf{\Omega} = \text{tr} \mathbf{\Sigma}$, we can get CLT for the largest eigenvalues of $\frac{1}{n} \tilde{X} \tilde{X}^*$ from Lemma 2. So we only need to prove

$$\left\| \frac{1}{n} \ddot{X} \ddot{X}^\top - \frac{1}{n} \tilde{X} \tilde{X}^\top \right\|_2 = O_p \left(\frac{T^4}{n} \text{tr}((\mathbf{\Pi} - \varphi \mathbf{I}_n)^2 \mathbf{\Sigma}) \right). \quad (\text{A.10})$$

Without loss of generality, we set $\mathbf{x}_0 = \ddot{\mathbf{x}}_0 = \tilde{\mathbf{x}}_0 = \mathbf{0}$.

$$\ddot{\mathbf{x}}_t = \sum_{i=0}^{t-1} \mathbf{\Lambda}_\Pi^i \mathbf{\Omega}^{1/2} \mathbf{y}_{t-i} \quad (\text{A.11})$$

and

$$\tilde{\mathbf{x}}_t = \sum_{i=0}^{t-1} \varphi^i \mathbf{\Omega}^{1/2} \mathbf{y}_{t-i}. \quad (\text{A.12})$$

Define \ddot{x}_{tk} as the k th element of $\ddot{\mathbf{x}}_t$ and \tilde{x}_{tk} as the k th element of $\tilde{\mathbf{x}}_t$. $\mathbf{\Lambda}_\Pi = \text{diag}\{\varphi_1, \dots, \varphi_n\}$ and $\mathbf{\Omega}^{1/2} \mathbf{y}_t = (\ddot{y}_{t1}, \dots, \ddot{y}_{tn})^\top$.

$$\ddot{x}_{tk} = \sum_{i=0}^{t-1} \varphi_k^i \ddot{y}_{t-i,k} \quad (\text{A.13})$$

and

$$\tilde{x}_{tk} = \sum_{i=0}^{t-1} \varphi^i \ddot{y}_{t-i,k} \quad (\text{A.14})$$

Then the (t, s) entry of $\frac{1}{n} \ddot{X} \ddot{X}^\top$ is

$$\ddot{B}_{ts} = \frac{1}{n} \sum_{k=1}^n \ddot{x}_{tk} \ddot{x}_{sk}. \quad (\text{A.15})$$

The (t, s) entry of $\frac{1}{n} \tilde{X} \tilde{X}^\top$ is

$$\tilde{B}_{ts} = \frac{1}{n} \sum_{k=1}^n \tilde{x}_{tk} \tilde{x}_{sk}. \quad (\text{A.16})$$

(A.13)-(A.16) and Taylor expansion imply that

$$\begin{aligned} \ddot{B}_{ts} - \tilde{B}_{ts} &= \frac{1}{n} \sum_{k=1}^n \left\{ \left(\sum_{i=0}^{t-1} \varphi_k^i \ddot{y}_{t-i,k} \right) \left(\sum_{i=0}^{s-1} \varphi_k^i \ddot{y}_{s-i,k} \right) - \left(\sum_{i=0}^{t-1} \varphi^i \ddot{y}_{t-i,k} \right) \left(\sum_{i=0}^{s-1} \varphi^i \ddot{y}_{s-i,k} \right) \right\} \\ &= \frac{1}{n} \sum_{k=1}^n \sum_{i=0}^{t-1} \sum_{j=0}^{s-1} (\varphi_k^{i+j} - \varphi^{i+j}) \ddot{y}_{t-i,k} \ddot{y}_{s-j,k} \\ &= \frac{1}{n} \sum_{i=0}^{t-1} \sum_{j=0}^{s-1} \sum_{k=1}^n \left\{ (i+j)(\varphi_k - \varphi) + C_{i+j,k} \frac{(i+j)(i+j-1)(\varphi_k - \varphi)^2}{2} \right\} \ddot{y}_{t-i,k} \ddot{y}_{s-j,k}. \quad (\text{A.17}) \end{aligned}$$

With (2.5) and Assumption 1, we can find

$$\begin{aligned}
& E(\ddot{B}_{ts} - \tilde{B}_{ts}) \\
&= \frac{1}{n} \sum_{i=0}^{t-1} \sum_{j=0}^{s-1} \sum_{k=1}^n \left\{ (i+j)(\varphi_k - \varphi) + C_{i+j,k} \frac{(i+j)(i+j-1)(\varphi_k - \varphi)^2}{2} \right\} E\ddot{y}_{t-i,k} \ddot{y}_{s-j,k} \\
&= \frac{1}{n} \sum_{i=0}^{t-1} \sum_{j=0}^{s-1} \sum_{k=1}^n \left\{ (i+j)(\varphi_k - \varphi) + C_{i+j,k} \frac{(i+j)(i+j-1)(\varphi_k - \varphi)^2}{2} \right\} a_{|t-s-i+j|} \Omega_{kk} \\
&= \frac{1}{n} \sum_{i=0}^{t-1} \sum_{j=0}^{s-1} (i+j) a_{|t-s-i+j|} \{tr(\mathbf{\Lambda} \mathbf{\Pi} \mathbf{\Omega}) - \varphi tr(\mathbf{\Omega})\} \\
&+ \frac{1}{n} \sum_{i=0}^{t-1} \sum_{j=0}^{s-1} \sum_{k=1}^n C_{i+j,k} \frac{(i+j)(i+j-1)(\varphi_k - \varphi)^2}{2} a_{|t-s-i+j|} \Omega_{kk} \\
&= \frac{1}{n} \sum_{i=0}^{t-1} \sum_{j=0}^{s-1} (i+j) a_{|t-s-i+j|} \{tr(\mathbf{\Pi} \mathbf{\Sigma}) - \varphi tr(\mathbf{\Sigma})\} \\
&+ \frac{1}{n} \sum_{i=0}^{t-1} \sum_{j=0}^{s-1} \sum_{k=1}^n C_{i+j,k} \frac{(i+j)(i+j-1)(\varphi_k - \varphi)^2}{2} a_{|t-s-i+j|} \Omega_{kk} \\
&= \frac{1}{n} \sum_{i=0}^{t-1} \sum_{j=0}^{s-1} a_{|t-s-i+j|} \frac{(i+j)(i+j-1)}{2} \sum_{k=1}^n C_{i+j,k} (\varphi_k - \varphi)^2 \Omega_{kk} \\
&= O\left(\frac{\max\{s^3, t^3\}}{n} tr((\mathbf{\Pi} - \varphi \mathbf{I}_n)^2 \mathbf{\Sigma})\right). \tag{A.18}
\end{aligned}$$

Then

$$\left(\sum_{t=1}^T \sum_{s=1}^T E(\ddot{B}_{ts} - \tilde{B}_{ts})^2 \right)^{1/2} = O\left(\frac{T^4}{n} tr((\mathbf{\Pi} - \varphi \mathbf{I}_n)^2 \mathbf{\Sigma})\right). \tag{A.19}$$

Similarly, we can find

$$\left(\sum_{t=1}^T \sum_{s=1}^T Var(\ddot{B}_{ts} - \tilde{B}_{ts}) \right)^{1/2} = o\left(\frac{T^4}{n} tr((\mathbf{\Pi} - \varphi \mathbf{I}_n)^2 \mathbf{\Sigma})\right). \tag{A.20}$$

(A.19)-(A.20) conclude (A.10). \square

A.2 The proof of Lemma 1

$\mathbf{\Pi} = \varphi \mathbf{I}_n$ implies that Assumptions 7-8 hold. Assumptions 9-10 implies that the part of δ_t can be ignored. Then we can replace $\mathbf{\Sigma}$ by $\mathbf{\Sigma}^{1/2}(\mathbf{I} - \mathbf{H})\mathbf{\Sigma}^{1/2}$ and use Theorem 1 to complete the proof.

A.3 The proof of Proposition 1

At first by (3.1) we write $\mathbf{x}_t - \mathbf{x}_{t-1}$ as follows

$$\mathbf{x}_t - \mathbf{x}_{t-1} = \mathbf{\Sigma}^{1/2} [\mathbf{y}_t + (\varphi - 1) \sum_{k=0}^{t-2} \varphi^k \mathbf{y}_{t-k-1} + (\varphi - 1) \varphi^{t-1} \mathbf{x}_0] + \delta_t - \delta_{t-1}. \tag{A.21}$$

Recalling the Assumptions 6 and 9, one can find that $E\|(\varphi - 1) \varphi^{t-1} \mathbf{x}_0 - \delta_0\|^2 \rightarrow 0$. Assumption 12 implies the term of $\delta_t - \delta_{t-1}$ is small enough. Below for facilitating statements we ignore the terms and write

$$\mathbf{x}_t - \mathbf{x}_{t-1} = \mathbf{\Sigma}^{1/2} [\mathbf{y}_t + (\varphi - 1) \sum_{k=0}^{t-2} \varphi^k \mathbf{y}_{t-k-1}]. \tag{A.22}$$

Recalling the notation $\hat{x}_{f,g}$ below (3.9) we can write $E\hat{\mathbf{x}}_{t,t-j}$ when $j \geq 1$ as

$$\begin{aligned} E\hat{\mathbf{x}}_{t,t-j} &= E(\mathbf{x}_t - \mathbf{x}_{t-1})^\top (\mathbf{I} - \mathbf{H})(\mathbf{x}_{t-j} - \mathbf{x}_{t-j-1}) \\ &= E\left[\left\{\mathbf{y}_t + (\varphi - 1) \sum_{k=0}^{t-2} \varphi^k \mathbf{y}_{t-k-1}\right\}^\top \boldsymbol{\Sigma}^{1/2} (\mathbf{I} - \mathbf{H}) \boldsymbol{\Sigma}^{1/2} \left\{\mathbf{y}_{t-j} + (\varphi - 1) \sum_{k=0}^{t-j-2} \varphi^k \mathbf{y}_{t-j-k-1}\right\}\right] \\ &= \text{tr}\left\{\boldsymbol{\Sigma}^{1/2} (\mathbf{I} - \mathbf{H}) \boldsymbol{\Sigma}^{1/2}\right\} \left(a_j + A_{1tj} + A_{2tj} + A_{3tj} + A_{4tj} + A_{5tj} + A_{6tj}\right), \end{aligned} \quad (\text{A.23})$$

where

$$\begin{aligned} A_{1tj} &= (\varphi - 1) \sum_{k=0}^{j-1} \varphi^k a_{j-k-1}, \quad A_{2tj} = (\varphi - 1) \sum_{k=j}^{t-2} \varphi^k a_{k+1-j}, \quad A_{3tj} = (\varphi - 1) \sum_{k=0}^{t-j-2} \varphi^k a_{k+j+1}, \\ A_{4tj} &= (\varphi - 1)^2 \sum_{k_1=0}^{j-1} \sum_{k_2=0}^{t-j-2} \varphi^{k_1+k_2} a_{k_2+j-k_1}, \quad A_{5tj} = (\varphi - 1)^2 \sum_{k_1=j}^{t-2} \sum_{k_2=k_1-j+1}^{t-j-2} \varphi^{k_1+k_2} a_{k_2+j-k_1}, \end{aligned}$$

and

$$A_{6tj} = (\varphi - 1)^2 \sum_{k_1=j}^{t-2} \sum_{k_2=0}^{k_1-j} \varphi^{k_1+k_2} a_{k_1-k_2-j}.$$

When q is bounded, we conclude that

$$E\hat{\mathbf{x}}_{t,t-j} = \text{tr}\left[\boldsymbol{\Sigma}^{1/2} (\mathbf{I} - \mathbf{H}) \boldsymbol{\Sigma}^{1/2}\right] \{a_j + O(T^{-1})\}$$

and

$$E\check{\zeta}_j = \frac{\text{tr}\left[\boldsymbol{\Sigma}^{1/2} (\mathbf{I} - \mathbf{H}) \boldsymbol{\Sigma}^{1/2}\right]}{n} \{a_j + O(T^{-1})\}.$$

Similarly, we can find that $\text{var}(\check{\zeta}_j) = O(T^{-1}n^{-1})$. Then we can conclude Proposition 1.

A.4 The proof of Theorem 2

proof of Theorem 2. We conclude from Lemmas 1, 3 and 4 that under the conditions in Theorem 2,

$$n^{1/2} \frac{\rho_1 - \frac{\lambda_{\varphi,1}}{\kappa(\theta_{\varphi,1})} \hat{\mu}_m}{\lambda_{\varphi,1} \hat{S}_{\sigma^2,m}} \rightarrow N(0, 1). \quad (\text{A.24})$$

Note that $\lambda_{\varphi,1} = \frac{1}{g_2(\theta_{\varphi,1})}$ and (A.1), (3.11) then follows. \square

A.5 The proof of Theorem 3

proof of Theorem 3. At first we prove (3.13). From Lemmas 1, 3 and 4 we have $\frac{\rho_1}{\hat{\mu}_m} = \lambda_{\varphi,1} \{1 + o_p(1)\}$. We can define $\check{\varphi} = \varphi - \frac{1}{T}$ and $\check{\theta} \in (\frac{T\pi}{T+1}, \pi)$ such that $g_1(\check{\theta}) = \check{\varphi}$. Then

$$\lim_{p, T \rightarrow \infty} P\left(h(\check{\theta}) > 0\right) = 1.$$

Similarly we obtain

$$\lim_{p, T \rightarrow \infty} P\left(h\left(\frac{T\pi}{T + \frac{1}{4}}\right) < 0\right) = 1.$$

Thus (3.13) follows from the continuity of $h(\theta)$.

We next prove that the solution is unique. It suffices to prove that $f(\theta)$ is monotone on the interval $\left(\frac{T\pi}{T+1}, \frac{T\pi}{T+\frac{1}{4}}\right)$. From the chain rule $f'(\theta) = g_2'(\theta)\kappa(\theta) + g_2(\theta)\kappa'(\theta)$. It is straightforward to verify that $\kappa(\theta) = 1 + o(1)$, $\kappa'(\theta) = O(m)$ and

$$g_2(\theta) = \frac{\sin^2 \theta}{\sin^2 T\theta}.$$

As for $g_2'(\theta)$ we have

$$g_2'(\theta) = \frac{\sin 2\theta \sin T\theta - 2T \cos T\theta \sin^2 \theta}{\sin^3 T\theta}. \quad (\text{A.25})$$

We rewrite $\theta = \frac{T\pi}{T+\iota}$ where $0 < \iota < 1$. Then

$$g_1(\theta) = -\frac{\sin(T+1)\theta}{\sin T\theta} = \frac{\sin \frac{T(1-\iota)\pi}{T+\iota}}{\sin \frac{T\iota\pi}{T+\iota}}. \quad (\text{A.26})$$

$$\begin{aligned} g_2'(\theta) &= \frac{\sin \frac{2T\pi}{T+\iota} \sin(T\pi - \frac{T\iota\pi}{T+\iota}) - 2T \cos(T\pi - \frac{T\iota\pi}{T+\iota}) \sin^2 \frac{T\pi}{T+\iota}}{\sin^3(T\pi - \frac{T\iota\pi}{T+\iota})} \\ &= 2 \sin \frac{T\pi}{T+\iota} \frac{\cos \frac{T\pi}{T+\iota} \sin(T\pi - \frac{T\iota\pi}{T+\iota}) - T \cos(T\pi - \frac{T\iota\pi}{T+\iota}) \sin \frac{T\pi}{T+\iota}}{\sin^3(T\pi - \frac{T\iota\pi}{T+\iota})} \\ &= 2 \sin \frac{T\pi}{T+\iota} \frac{\cos \frac{T\pi}{T+\iota} (-1)^{T+1} \sin \frac{T\iota\pi}{T+\iota} + T (-1)^{T+1} \cos \frac{T\iota\pi}{T+\iota} \sin \frac{T\pi}{T+\iota}}{(-1)^{T+1} \sin^3 \frac{T\iota\pi}{T+\iota}} \\ &= 2 \sin \frac{\iota\pi}{T+\iota} \frac{-\cos \frac{\iota\pi}{T+\iota} \sin \frac{T\iota\pi}{T+\iota} + T \cos \frac{T\iota\pi}{T+\iota} \sin \frac{\iota\pi}{T+\iota}}{\sin^3 \frac{T\iota\pi}{T+\iota}}. \end{aligned} \quad (\text{A.27})$$

When $\frac{1}{4} < \iota < 1$, $-\cos \frac{\iota\pi}{T+\iota} \sin \frac{T\iota\pi}{T+\iota} + T \cos \frac{T\iota\pi}{T+\iota} \sin \frac{\iota\pi}{T+\iota} < 0$ has a constant order. It follows that there is a constant $j > 0$ such that

$$g_2'(\theta) < -j \left| \frac{\sin \theta}{\sin^3 T\theta} \right| \quad (\text{A.28})$$

when $\frac{1}{4} < \iota < 1$. This, together with $\kappa(\theta) = 1 + o(1)$, $\kappa'(\theta) = O(m)$ and

$$g_2(\theta) = \frac{\sin^2 \theta}{\sin^2 T\theta},$$

implies that there is a constant $j_1 > 0$ such that

$$Tf'(\theta) < -j_1. \quad (\text{A.29})$$

Then the solution is unique.

Now we prove (3.14). Since $\hat{\theta} \in (\frac{T\pi}{T+1}, \pi)$ is the solution of $h(\theta) = 0$, we can rewrite (3.11) as

$$n^{1/2} \frac{\rho_1 f(\theta_{\varphi,1}) - \rho_1 f(\hat{\theta})}{\dot{S}_{\sigma^2, m}} \rightarrow N(0, 1). \quad (\text{A.30})$$

This leads to $f(\theta_{\varphi,1}) - f(\hat{\theta}) = O_p(T^{-2}n^{-1/2})$. This, together with (A.29), implies that

$$\hat{\theta} - \theta_{\varphi,1} = O_p(T^{-1}n^{-1/2}). \quad (\text{A.31})$$

Then we use Taylor expansions. We only need to prove $f''(\tilde{\theta}) = O_p(1)$ for any $\tilde{\theta}$ between $\theta_{\varphi,1}$ and $\hat{\theta}$. Since (A.31) ensures that $\theta_{\varphi,1}$ and $\hat{\theta}$ are very close, it's not hard. The proof of (3.15) is similar. We ignore the details. \square

B Results for truncated matrices

This section is to consider the truncated version of the sample covariance matrix for Lemma 2. The overall strategy of the proof, which needs more delicate analysis, is similar to that of [31]. However for the completeness of the proof and for facilitating the readers we repeat some necessary steps in [31] and still use the same notation as in [31].

To this end, define

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

with $l = \max\{n, T\}$, 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 (2.3) 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

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

It follows that $\mathbf{Y} = \mathbf{F}\mathbf{Z}_n$, where \mathbf{Z}_n is a $(T+l) \times n$ random matrix with $(\mathbf{Z}_n)_{i,j} = Z_{i-l,j}$. For the sake of notation simplicity, we below denote \mathbf{Z}_n by \mathbf{Z} and $(\mathbf{Z}_n)_{i,j}$ by Z_{ij} . Let $\mathbf{A} = (A_{ij})_{T \times T} = (a_{|i-j|})_{T \times T}$. We then have $\mathbf{A} = \mathbf{F}\mathbf{F}^\top$. 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.

Note that we have studied the case $\varphi = 1$ in [31]. Below we deal with with the case of $\varphi < 1$.

B.1 Eigenvalues of $\mathbf{C}_\varphi \mathbf{A} \mathbf{C}_\varphi^\top$

Write

$$\begin{aligned} \mathbf{B} &= (1/n) \mathbf{X} \mathbf{X}^\top = (1/n) \mathbf{C}_\varphi \mathbf{Y} \mathbf{\Sigma} \mathbf{Y}^\top \mathbf{C}_\varphi^\top \\ &= (1/n) \mathbf{C}_\varphi \mathbf{F} \mathbf{Z}_n \mathbf{\Sigma} \mathbf{Z}_n^\top \mathbf{F}^\top \mathbf{C}_\varphi^\top. \end{aligned} \quad (\text{B.2})$$

We below investigate the eigenvalues and eigenvectors of $\mathbf{C}_\varphi \mathbf{F} \mathbf{F}^\top \mathbf{C}_\varphi^\top = \mathbf{C}_\varphi \mathbf{A} \mathbf{C}_\varphi^\top$ at first. These are crucial steps.

Since it's very hard to find the eigenvalues of $\mathbf{C}_\varphi \mathbf{A} \mathbf{C}_\varphi^\top$ directly we use the following strategy. At first, we note that the eigenvalues of $\mathbf{C}_\varphi \mathbf{A} \mathbf{C}_\varphi^\top$ and $\mathbf{A} \mathbf{C}_\varphi^\top \mathbf{C}_\varphi$ are the same. We obtain the eigenvalues and eigenvectors of $\mathbf{C}_\varphi^\top \mathbf{C}_\varphi$ by first studying $(\mathbf{C}_\varphi^\top \mathbf{C}_\varphi)^{-1}$. The next key step is to approximate the eigenvalues of $\mathbf{C}_\varphi \mathbf{A} \mathbf{C}_\varphi^\top$. Our results are summarized in the following series of Lemmas and Theorems.

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

Lemma 5. *Let $\lambda_{\varphi,1} \geq \lambda_{\varphi,2} \geq \dots \geq \lambda_{\varphi,T} \geq 0$ be the eigenvalues of $\mathbf{C}_\varphi^\top \mathbf{C}_\varphi$. We then have*

$$\lambda_{\varphi,k} = \frac{1}{2\varphi(1 + \cos \theta_{\varphi,k}) + (1 - \varphi)^2} \quad \text{with} \quad \pi > \theta_{\varphi,1} > \dots > \theta_{\varphi,T} > 0, \quad (\text{B.3})$$

where $\pi > \theta_{\varphi,1} > \dots > \theta_{\varphi,T} > 0$ are the solutions of the equation

$$\varphi \sin T\theta + \sin(T+1)\theta = 0. \quad (\text{B.4})$$

Lemma 6. (Lemma C.2 in [31]) Using the notation in Lemma 5,

$$\lim_{T \rightarrow \infty} \frac{\lambda_{1,k}}{T^2} = \frac{4}{\pi^2(2k-1)^2} \quad (\text{B.5})$$

for any fixed k .

Lemma 7. Using the notation in Lemma 5, under the Assumption 6,

$$\lim_{T \rightarrow \infty} \frac{\lambda_{\varphi,k}}{T^2} > 0 \quad (\text{B.6})$$

for any fixed k . Furthermore, when T is big enough there exists a independent constant $\Upsilon > 0$ such that

$$\lim_{T \rightarrow \infty} \frac{\lambda_{1,k} - \lambda_{\varphi,k}}{\lambda_{1,k}} \leq \Upsilon T(1 - \varphi) \quad (\text{B.7})$$

for any k .

Lemma 8 below specifies the eigenvectors of $\mathbf{C}_\varphi^\top \mathbf{C}_\varphi$.

Lemma 8. Let $\tilde{\mathbf{x}}_{\mathbf{k}} = (x_{k,1}, \dots, x_{k,T})^\top$ be a $T \times 1$ vector with

$$x_{k,i} = (-1)^{T-i} \sin(T-i+1)\theta_{\varphi,k}, \quad -l \leq i \leq T+l. \quad (\text{B.8})$$

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

$$\mathbf{C}_\varphi^\top \mathbf{C}_\varphi \tilde{\mathbf{x}}_{\mathbf{k}} = \lambda_{\varphi,k} \tilde{\mathbf{x}}_{\mathbf{k}}. \quad (\text{B.9})$$

Lemma 9.

$$\sum_{j=1}^T (x_{k,j})^2 = \frac{T}{2} + o(T).$$

Let

$$\tilde{\mathbf{y}}_{\mathbf{k}} = \frac{\tilde{\mathbf{x}}_{\mathbf{k}}}{\|\tilde{\mathbf{x}}_{\mathbf{k}}\|}. \quad (\text{B.10})$$

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

$$|y_{k,j}| = \frac{|x_{k,j}|}{\|\tilde{\mathbf{x}}_{\mathbf{k}}\|} = O(T^{-1/2}). \quad (\text{B.11})$$

Lemma 10 gives the approximation to the eigenvalues of $\mathbf{A} \mathbf{C}_\varphi^\top \mathbf{C}_\varphi$.

Lemma 10. Define $\gamma_{\varphi,k}$ by

$$\gamma_{\varphi,k} = \lambda_{\varphi,k} [a_0 + 2 \sum_{1 \leq j \leq T-1} a_j (-1)^j \cos(j\theta_{\varphi,k})]. \quad (\text{B.12})$$

When $T(1 - \varphi)$ is bounded, for any fixed constant $k \geq 1$, there is a constant $c_{\varphi,k}$ such that

$$\lim_{T \rightarrow \infty} \frac{\gamma_{\varphi,k}}{T^2} = c_{\varphi,k} > 0. \quad (\text{B.13})$$

Let $\beta_{\varphi,1} \geq \beta_{\varphi,2} \geq \dots \geq \beta_{\varphi,T}$ be the eigenvalues of $\mathbf{A}\mathbf{C}_\varphi^\top\mathbf{C}_\varphi$. If \mathbf{A} satisfies Assumption 1, then for any fixed integers $i \geq 1$ and $j \geq 1$ the following holds

$$\left| \frac{\beta_{\varphi,i} - \gamma_{\varphi,i}}{\gamma_{\varphi,j}} \right| = O(T^{-1}). \quad (\text{B.14})$$

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 $k \geq k_0$,

$$\left| \frac{\beta_{\varphi,k}}{\gamma_{\varphi,1}} \right| \leq \epsilon. \quad (\text{B.15})$$

Lemma 11. (Lemma C.5 in [31]) Suppose that \mathbf{A} satisfies Assumption 1. Then

$$\text{tr}(\mathbf{A}\mathbf{C}_1^\top\mathbf{C}_1) = a_0 \frac{(T+1)T}{2} + \sum_{1 \leq j \leq T-1} a_j (T-j+1)(T-j)$$

and

$$\lim_{T \rightarrow \infty} \frac{\beta_{1,k}}{\text{tr}(\mathbf{A}\mathbf{C}_1^\top\mathbf{C}_1)} = \lim_{T \rightarrow \infty} \frac{\gamma_{1,k}}{\text{tr}(\mathbf{A}\mathbf{C}_1^\top\mathbf{C}_1)} = \frac{8}{\pi^2(2k-1)^2}.$$

Lemma 12. (Lemma C.6 in [31]) Suppose that \mathbf{A} satisfies Assumption 1. 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 $T \geq T_0$,

$$\left| \frac{\sum_{k > k_0} \beta_{1,k}}{\gamma_{1,1}} \right| < \epsilon. \quad (\text{B.16})$$

Lemma 13. Suppose that \mathbf{A} satisfies Assumption 1. 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 $T \geq T_0$,

$$\left| \frac{\sum_{k > k_0} \beta_{\varphi,k}}{\gamma_{\varphi,1}} \right| < \epsilon. \quad (\text{B.17})$$

of Lemma 5.

Let $\mathbf{M}_T = (\mathbf{C}_\varphi^\top\mathbf{C}_\varphi)^{-1}$. 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}_φ^{-1} , a $T \times T$ lower triangular matrix, are of the form

$$C_{\varphi,ij}^{-1} = \begin{cases} 1 & i = j, \\ -\varphi & i = j + 1, \\ 0 & \text{otherwise.} \end{cases}$$

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

$$M_{ij} = \begin{cases} 1 & i = j = 1, \\ 1 + \varphi^2 & i = j > 1, \\ -\varphi & |i - j| = 1, \\ 0 & \text{otherwise.} \end{cases}$$

By the cofactor expansion we obtain a recurrence relation as following

$$g_T(\lambda) = (\lambda - 1 - \varphi^2)g_{T-1}(\lambda) - \varphi^2g_{T-2}(\lambda). \quad (\text{B.18})$$

Consider $\lambda \in (0, 4)$ at first. Hence we may write $\lambda = \lambda(\theta) = 1 + \varphi^2 + 2\varphi \cos \theta$. We can further solve (B.18) to get

$$g_T(\lambda) = \frac{\varphi^{T+1} \sin T\theta + \varphi^T \sin(T+1)\theta}{\sin \theta}.$$

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

$$\varphi \sin T\theta + \sin(T+1)\theta = 0.$$

Let $h_T(\theta) = \varphi \sin T\theta + \sin(T+1)\theta$. For any $1 \leq k \leq T$,

$$h_T\left(\frac{k\pi}{T+1/2}\right) = \varphi \sin \frac{kT\pi}{T+1/2} + \sin \frac{k(T+1)\pi}{T+1/2} = (-1)^k (1-\varphi) \sin \frac{2k\pi}{2T+1} \quad (\text{B.19})$$

and

$$h_T\left(\frac{k\pi}{T+1}\right) = \varphi \sin \frac{kT\pi}{T+1} = \varphi (-1)^{k-1} \sin \frac{(T+1-k)\pi}{T+1}. \quad (\text{B.20})$$

Then there is $\theta_{\varphi, T+1-k} \in [\frac{k\pi}{T+1}, \frac{k\pi}{T+1/2}]$ such that $h_T(\theta_{\varphi, k}) = 0$. So (B.3) 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 (B.3) is complete. \square

proof of Lemma 7. Recalling the proof of Lemma 5, $\theta_{\varphi, k} \in \left[\frac{(T+1-k)\pi}{T+1}, \frac{(T+1-k)\pi}{T+1/2}\right]$. Then we can find that both $T^2(1 + \cos \theta_{\varphi, k})$ and $T^2(1 - \varphi)^2$ are bounded for any fixed k . So we can proof (B.6).

We next prove (B.7), Write

$$\begin{aligned} \lambda_{1,k} - \lambda_{\varphi,k} &= \frac{1}{2 + 2 \cos \theta_{1,k}} - \frac{1}{1 + \varphi^2 + 2\varphi \cos \theta_{\varphi,k}} \\ &= \frac{(1 - \varphi)^2 + 2(\varphi - 1)(1 + \cos \theta_{\varphi,k}) + 2(\cos \theta_{\varphi,k} - \cos \theta_{1,k})}{(2 + 2 \cos \theta_{1,k})(1 + \varphi^2 + 2\varphi \cos \theta_{\varphi,k})}. \end{aligned} \quad (\text{B.21})$$

For any k , we define

$$\varpi_k = \theta_{1,k} - \theta_{\varphi,k} = \frac{(T+1-k)\pi}{T+1/2} - \theta_{\varphi,k}. \quad (\text{B.22})$$

Then $0 < \varpi_{T-k+1} \leq \frac{k\pi}{T(2T+1)}$,

$$\frac{\lambda_{1,k} - \lambda_{\varphi,k}}{\lambda_{1,k}} = \frac{(1 - \varphi)^2 + 2(\varphi - 1)(1 + \cos \theta_{\varphi,k}) + 4 \sin(\theta_{1,k} - \frac{\varpi_k}{2}) \sin \frac{\varpi_k}{2}}{1 + \varphi^2 + 2\varphi \cos \theta_{\varphi,k}}. \quad (\text{B.23})$$

Note that

$$0 < \frac{(1 - \varphi)^2}{1 + \varphi^2 + 2\varphi \cos \theta_{\varphi,k}} \leq \frac{\lambda_{1,1}}{T^2} T^2 (1 - \varphi)^2, \quad (\text{B.24})$$

and

$$\left| \frac{(\varphi - 1)(1 + \cos \theta_{\varphi,k})}{1 + \varphi^2 + 2\varphi \cos \theta_{\varphi,k}} \right| = (1 - \varphi) \left| \frac{(1 + \cos \theta_{\varphi,k})}{(1 - \varphi)^2 + 2\varphi(1 + \cos \theta_{\varphi,k})} \right| \leq \frac{1 - \varphi}{2\varphi}. \quad (\text{B.25})$$

Then we focus on the term $\sin(\theta_{1,k} - \frac{\varpi_k}{2}) \sin \frac{\varpi_k}{2}$. We need to study the relation between ϖ_k and φ . From (B.4), we can find that

$$(1 - \varphi) \sin T\theta_{\varphi,k} = \sin T\theta_{\varphi,k} + \sin(T+1)\theta_{\varphi,k} = 2 \sin(T+1/2)\theta_{\varphi,k} \cos \frac{\theta_{\varphi,k}}{2}.$$

It follows that

$$(1 - \varphi) \sin \left\{ \frac{T(T+1-k)\pi}{T+1/2} - T\varpi_k \right\} = 2 \sin \left\{ (T+1-k)\pi - (T+1/2)\varpi_k \right\} \cos \frac{\theta_{\varphi,k}}{2}.$$

Then

$$(1 - \varphi) \sin \left\{ (T+1-k)\pi - \frac{T+1-k}{2T+1}\pi - T\varpi_k \right\} = 2 \sin \left\{ (T+1-k)\pi - (T+1/2)\varpi_k \right\} \cos \frac{\theta_{\varphi,k}}{2}.$$

$$(1 - \varphi) \sin \left(\frac{T+1-k}{2T+1}\pi + T\varpi_k \right) = 2 \sin \left\{ (T+1/2)\varpi_k \right\} \cos \frac{\theta_{\varphi,k}}{2}.$$

Then we can find that

$$(1 - \varphi) \sin\left(\frac{T+1-k}{2T+1}\pi\right) \cos T\varpi_k + (1 - \varphi) \sin T\varpi_k \cos\left(\frac{T+1-k}{2T+1}\pi\right) = 2 \sin(T+1/2)\varpi_k \cos \frac{\theta_{\varphi,k}}{2}.$$

Recall that $0 < \varpi_{T-k+1} \leq k\pi/(2T^2 + T)$ and $\theta_{\varphi,k} \in [\frac{(T+1-k)\pi}{T+1}, \frac{(T+1-k)\pi}{T+1/2}]$. Then

$$\sin\left\{(T+1/2)\varpi_k\right\} \cos \frac{\theta_{\varphi,k}}{2} \neq 0.$$

It follows that

$$(1 - \varphi) \frac{\sin \frac{T+1-k}{2T+1}\pi \cos T\varpi_k}{2 \sin\left\{(T+1/2)\varpi_k\right\} \cos \frac{\theta_{\varphi,k}}{2}} + (1 - \varphi) \frac{\sin T\varpi_k \cos \frac{T+1-k}{2T+1}\pi}{2 \sin\left\{(T+1/2)\varpi_k\right\} \cos \frac{\theta_{\varphi,k}}{2}} = 1. \quad (\text{B.26})$$

Note that $0 < T\varpi_k < (T+1/2)\varpi_k \leq \frac{T+1-k\pi}{2T} \leq \frac{\pi}{2}$, and

$$0 < \frac{\sin T\varpi_k}{\sin\left\{(T+1/2)\varpi_k\right\}} < 1. \quad (\text{B.27})$$

Similarly, since $\theta_{\varphi,k} \in [\frac{(T+1-k)\pi}{T+1}, \frac{(T+1-k)\pi}{T+1/2}]$, we have

$$0 < \frac{\cos \frac{T+1-k}{2T+1}\pi}{\cos \frac{(T+1-k)\pi}{2T+2}} < \frac{\cos \frac{T+1-k}{2T+1}\pi}{\cos \frac{\theta_{\varphi,k}}{2}} < \frac{\cos \frac{T+1-k}{2T+1}\pi}{\cos \frac{(T+1-k)\pi}{2T+1}} = 1. \quad (\text{B.28})$$

(B.27)-(B.28) imply that

$$0 < (1 - \varphi) \frac{\sin T\varpi_k \cos \frac{T+1-k}{2T+1}\pi}{2 \sin\left\{(T+1/2)\varpi_k\right\} \cos \frac{\theta_{\varphi,k}}{2}} < \frac{1 - \varphi}{2}.$$

This, together with (B.26), implies that

$$\left|1 - (1 - \varphi) \frac{\sin \frac{T+1-k}{2T+1}\pi \cos T\varpi_k}{2 \sin\left\{(T+1/2)\varpi_k\right\} \cos \frac{\theta_{\varphi,k}}{2}}\right| < \frac{1 - \varphi}{2}. \quad (\text{B.29})$$

$$\frac{\sin\left\{(T+1/2)\varpi_k\right\}}{\cos(T\varpi_k)} = \tan(T\varpi_k) \cos \frac{\varpi_k}{2} + \sin \frac{\varpi_k}{2}$$

and

$$\frac{\sin \frac{T+1-k}{2T+1}\pi}{2 \cos \frac{\theta_{\varphi,k}}{2}} > 0.$$

Then (B.29) implies

$$\frac{2 - 2\varphi}{3 - \varphi} \frac{\sin \frac{T+1-k}{2T+1}\pi}{2 \cos \frac{\theta_{\varphi,k}}{2}} < \tan(T\varpi_k) \cos \frac{\varpi_k}{2} + \sin \frac{\varpi_k}{2} < \frac{2 - 2\varphi}{1 + \varphi} \frac{\sin \frac{T+1-k}{2T+1}\pi}{2 \cos \frac{\theta_{\varphi,k}}{2}} \quad (\text{B.30})$$

Since $\theta_{\varphi,k} \in [\frac{(T+1-k)\pi}{T+1}, \frac{(T+1-k)\pi}{T+1/2}]$, we conclude that

$$0 < \frac{\sin \frac{T+1-k}{2T+1}\pi}{2 \frac{T}{k} \cos \frac{\theta_{\varphi,k}}{2}} < \frac{1}{2 \frac{T}{k} \cos \frac{(T+1-k)\pi}{2T+1}} < \frac{\Upsilon_0}{2}, \quad (\text{B.31})$$

Here Υ_0 is a constant which doesn't depend on T and k .

Moreover, since $0 < \varpi_k \leq \frac{(T-k+1)\pi}{T(2T+1)} < \frac{\pi}{2T}$,

$$\cos \frac{\varpi_k}{2} = 1 + O\left(\frac{(T-k+1)^2}{T^4}\right) \quad (\text{B.32})$$

and

$$\sin \frac{\varpi_k}{2} = O\left(\frac{\tan T\varpi_k}{T}\right). \quad (\text{B.33})$$

(B.30)-(B.33) imply that there is a constant Υ_1 which doesn't depend on T and k such that

$$|\tan(T\varpi_k)| \leq \frac{T}{k}(1-\varphi)\Upsilon_1. \quad (\text{B.34})$$

It follows that there is a constant Υ_2 which doesn't depend on T and k such that

$$|\varpi_k| \leq \frac{1-\varphi}{k}\Upsilon_2 \quad (\text{B.35})$$

and

$$|\sin(\theta_{1,k} - \frac{\varpi_k}{2})\sin \frac{\varpi_k}{2}| \leq \frac{1-\varphi}{T}\Upsilon_2. \quad (\text{B.36})$$

(B.23)-(B.25), together with Assumption 6, (B.36) and (B.6), imply that when T is big enough there exists a constant $\Upsilon > 0$ such that

$$\frac{\lambda_{1,k} - \lambda_{\varphi,k}}{\lambda_{1,k}} \leq \Upsilon T(1-\varphi) \quad (\text{B.37})$$

for any k . □

Lemmas 8 can be verified with some straightforward computations and the simple fact that

$$\sin(k+j)\theta + \sin(k-j)\theta = 2\sin k\theta \cos j\theta. \quad (\text{B.38})$$

We ignore the details here.

proof of Lemma 9. From (B.8) we obtain

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

Lemma 9 can be then proved with some straightforward computations. We ignore details here. □

proof of Lemma 10.

Let's prove (B.13) at first. Note that

$$\begin{aligned} & \left| [a_0 + 2 \sum_{1 \leq j \leq T-1} a_j (-1)^j \cos(j\theta_{\varphi,k})] - (a_0 + 2 \sum_{1 \leq j \leq \infty} a_j) \right| \\ & \leq 2 \sum_{1 \leq j \leq T-1} |a_j| |\cos\{j(\pi - \theta_{\varphi,k})\} - 1| + 2 \sum_{T \leq j} |a_j|. \end{aligned} \quad (\text{B.39})$$

For a fixed k , we can find a j_k to satisfy $\pi/3 \leq j(\pi - \theta_{\varphi,k}) \leq \pi/2$. It follows that

$$\begin{aligned} 2 \sum_{1 \leq j \leq j_k} |a_j| |\cos\{j(\pi - \theta_{\varphi,k})\} - 1| & \leq \frac{1}{2} \sum_{1 \leq j \leq j_k} |a_j| j^2 (\pi - \theta_{\varphi,k})^2 \\ & \leq \frac{j_k (\pi - \theta_{\varphi,k})^2}{2} \sum_{1 \leq j \leq j_k} j |a_j| \\ & \leq \frac{k\pi^2}{4(T+1)} \sum_{1 \leq j \leq \infty} j |a_j| \end{aligned} \quad (\text{B.40})$$

and that

$$\begin{aligned} & 2 \sum_{j_k < j \leq T-1} |a_j| |\cos\{j(\pi - \theta_{\varphi,k})\} - 1| + 2 \sum_{T \leq j} |a_j| \leq 4 \sum_{j \geq j_k} |a_j| \\ & \leq j_k^{-1} 4 \sum_{j \geq j_k} j |a_j| \leq \frac{3(2k-1)}{2T+1} \sum_{1 \leq j \leq \infty} j |a_j|. \end{aligned} \quad (\text{B.41})$$

Assumption 1 implies that

$$\sum_{i=0}^{\infty} i |a_i| < \infty. \quad (\text{B.42})$$

From Assumption 1, (B.42) and truncation conditions we can find

$$\lim_{T \rightarrow \infty} [a_0 + 2 \sum_{1 \leq j \leq T-1} a_j (-1)^j \cos(j\theta_{\varphi,k})] = \lim_{T \rightarrow \infty} (a_0 + 2 \sum_{1 \leq j \leq \infty} a_j) = (\sum_{i=0}^{\infty} b_i)^2 = s^2 > 0. \quad (\text{B.43})$$

In view of (B.6), (B.12) and (B.43), we have (B.13).

Now we consider the eigenvalues of $\mathbf{A}\mathbf{C}_{\varphi}^{\top}\mathbf{C}_{\varphi}$. From Lemmas 5-9, we can write $\mathbf{C}_{\varphi}^{\top}\mathbf{C}_{\varphi}$ as $\mathbf{V}_{\varphi}^{\top}\mathbf{\Lambda}_{\varphi}^{1/2}\mathbf{\Lambda}_{\varphi}^{1/2}\mathbf{V}_{\varphi}$, where

$$\mathbf{\Lambda}_{\varphi} = \text{diag}\{\lambda_{\varphi,1}, \dots, \lambda_{\varphi,T}\}, \quad \mathbf{V}_{\varphi}^{\top} = (\tilde{\mathbf{y}}_1, \dots, \tilde{\mathbf{y}}_T). \quad (\text{B.44})$$

Then we just need to consider the eigenvalues of $\mathbf{M}_{\mathbf{a},\varphi} = \mathbf{\Lambda}_{\varphi}^{1/2}\mathbf{V}_{\varphi}\mathbf{A}\mathbf{V}_{\varphi}^{\top}\mathbf{\Lambda}_{\varphi}^{1/2}$. Define the (i,j) entry of $\mathbf{M}_{\mathbf{a},\varphi}$ as $\dot{M}_{i,j}$. Then

$$\dot{M}_{i,j} = \lambda_{\varphi,i}^{1/2} \lambda_{\varphi,j}^{1/2} \tilde{\mathbf{y}}_i^{\top} \mathbf{A} \tilde{\mathbf{y}}_j = \frac{\lambda_{\varphi,i}^{1/2} \lambda_{\varphi,j}^{1/2}}{\|\tilde{\mathbf{x}}_i\| \|\tilde{\mathbf{x}}_j\|} \tilde{\mathbf{x}}_i^{\top} \mathbf{A} \tilde{\mathbf{x}}_j,$$

and

$$\tilde{\mathbf{x}}_i^{\top} \mathbf{A} \tilde{\mathbf{x}}_j = a_0 \sum_{h=1}^T x_{ih} x_{jh} + \sum_{h=1}^{T-1} a_h \left(\sum_{f=1}^{T-h} x_{if} x_{j,f+h} + \sum_{f=h+1}^T x_{if} x_{j,f-h} \right).$$

Recall $x_{j,f-h} + x_{j,f+h} = 2(-1)^h x_{jf} \cos(h\theta_{\varphi,j})$. Then

$$\begin{aligned} & \tilde{\mathbf{x}}_i^{\top} \mathbf{A} \tilde{\mathbf{x}}_j \\ = & \left[a_0 + 2 \sum_{h=1}^{T-1} a_h (-1)^h \cos(h\theta_{\varphi,j}) \right] \left(\sum_{h=1}^T x_{ih} x_{jh} \right) - \sum_{h=1}^{T-1} a_h \left(\sum_{f=T-h+1}^T x_{if} x_{j,f+h} + \sum_{f=1}^h x_{if} x_{j,f-h} \right). \end{aligned} \quad (\text{B.45})$$

$$\sum_{h=1}^T x_{ih} x_{jh} = 1\{i=j\} \|\tilde{\mathbf{x}}_i\| \|\tilde{\mathbf{x}}_j\|$$

and

$$\left| \sum_{h=1}^{T-1} a_h \left(\sum_{f=T-h+1}^T x_{if} x_{j,f+h} + \sum_{f=1}^h x_{if} x_{j,f-h} \right) \right| \leq 2 \sum_{h=1}^{T-1} h |a_h|.$$

Then we can write

$$\mathbf{M}_{\mathbf{a},\varphi} = \mathbf{M}_{\mathbf{a},\varphi,\mathbf{R}} + \text{diag}\{\gamma_{\varphi,1}, \dots, \gamma_{\varphi,T}\} = \mathbf{M}_{\mathbf{a},\varphi,\mathbf{R}} + \mathbf{\Lambda}_{\mathbf{a},\varphi,\mathbf{M}}. \quad (\text{B.46})$$

$$\|\mathbf{M}_{\mathbf{a},\varphi,\mathbf{R}}\|_2^2 \leq \sum_{i=1}^T \sum_{j=1}^T \frac{\lambda_{\varphi,i} \lambda_{\varphi,j}}{\|\tilde{\mathbf{x}}_i\|^2 \|\tilde{\mathbf{x}}_j\|^2} \left(2 \sum_{h=1}^{T-1} h |a_h| \right)^2 = O\left(\frac{\lambda_{\varphi,1}^2}{T^2}\right) = O(T^2). \quad (\text{B.47})$$

This, together with (B.13), leads to (B.14). \square

proof of Lemma 13. From (B.7) we can find that

$$\left| \sum_{k>k_0} \beta_{\varphi,k} \right| < \{1 + \Upsilon T(1 - \varphi)\} \sum_{k>k_0} \beta_{1,k}. \quad (\text{B.48})$$

This, together with (B.6) and (B.16), implies (B.17). \square

B.2 Eigenvectors of $\mathbf{C}_\varphi \mathbf{A} \mathbf{C}_\varphi^\top$

This section is to investigate the eigenvectors of $\mathbf{C}_\varphi \mathbf{A} \mathbf{C}_\varphi^\top$. Recalling (B.10), we normalize $\{\tilde{\mathbf{x}}_k\}_{1 \leq k \leq T}$ to get $\{\tilde{\mathbf{y}}_k\}_{1 \leq k \leq T}$. Then we study the eigenvectors of $\mathbf{A} \mathbf{C}_\varphi^\top \mathbf{C}_\varphi$ by representing them with $\{\tilde{\mathbf{y}}_k\}_{1 \leq k \leq T}$. At last we give some result about the eigenvectors of $\mathbf{C}_\varphi \mathbf{A} \mathbf{C}_\varphi^\top$ which is necessary for the future proof. Our results are the following.

Lemma 14. *Let $\{\mathbf{u}_k\}_{1 \leq k \leq T}$ be orthogonal and real vectors such that $\|\mathbf{u}_k\| = 1$ and*

$$\mathbf{C}_\varphi \mathbf{A} \mathbf{C}_\varphi^\top \mathbf{u}_k = \beta_{\varphi,k} \mathbf{u}_k. \quad (\text{B.49})$$

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

$$\mathbf{f}_k = \sum_{j=1}^T \alpha_{kj} \mathbf{y}_j \quad (\text{B.50})$$

with

$$\sum_{j=1}^T \alpha_{kj}^2 = 1. \quad (\text{B.51})$$

Then when $k \geq 1$ is fixed,

$$\frac{\alpha_{kk}^2 \lambda_{\varphi,k}}{\sum_{j=1}^T \alpha_{kj}^2 \lambda_{\varphi,j}} = 1 + O(T^{-1}), \quad (\text{B.52})$$

where $\{\lambda_{\varphi,j}\}$ are given in Lemma 5.

Lemma 15. *Let*

$$(S_{k,1}, \dots, S_{k,T+l})^\top = \mathbf{s}_k = \frac{\mathbf{F}^\top \mathbf{C}_\varphi^\top \mathbf{u}_k}{\gamma_{\varphi,1}^{1/2}}.$$

Then $\{\mathbf{s}_k\}_{1 \leq k \leq T}$ are orthogonal and

$$\sum_{j=1}^{T+l} S_{k,j}^4 = O(T^{-1}). \quad (\text{B.53})$$

proof of Lemma 14. From $\mathbf{f}_k = \frac{\mathbf{C}_\varphi^{-1} \mathbf{u}_k}{\|\mathbf{C}_\varphi^{-1} \mathbf{u}_k\|}$ and (B.49), we have $\|\mathbf{f}_k\| = 1$ and

$$\mathbf{A} \mathbf{C}_\varphi^\top \mathbf{C}_\varphi \mathbf{f}_k = \beta_{\varphi,k} \mathbf{f}_k. \quad (\text{B.54})$$

From (B.46) and (B.54), we have

$$\beta_{\varphi,k} = \frac{\mathbf{f}_k^\top \mathbf{C}_\varphi^\top \mathbf{C}_\varphi \mathbf{A} \mathbf{C}_\varphi^\top \mathbf{C}_\varphi \mathbf{f}_k}{\|\mathbf{C}_\varphi \mathbf{f}_k\|^2} = \frac{\mathbf{f}_k^\top \mathbf{V}_\varphi^\top \Lambda_\varphi^{1/2} (\Lambda_{\mathbf{a},\varphi,\mathbf{M}} + \mathbf{M}_{\mathbf{a},\varphi,\mathbf{R}}) \Lambda_\varphi^{1/2} \mathbf{V}_\varphi \mathbf{f}_k}{\|\mathbf{C}_\varphi \mathbf{f}_k\|^2}.$$

It follows that

$$\frac{|\mathbf{f}_k^\top \mathbf{V}_\varphi^\top \Lambda_\varphi^{1/2} \Lambda_{\mathbf{a},\varphi,\mathbf{M}} \Lambda_\varphi^{1/2} \mathbf{V}_\varphi \mathbf{f}_k| - |\mathbf{f}_k^\top \mathbf{V}_\varphi^\top \Lambda_\varphi^{1/2} \mathbf{M}_{\mathbf{a},\varphi,\mathbf{R}} \Lambda_\varphi^{1/2} \mathbf{V}_\varphi \mathbf{f}_k|}{\|\mathbf{C}_\varphi \mathbf{f}_k\|^2} \leq \beta_{\varphi,k} \quad (\text{B.55})$$

and

$$\beta_{\varphi,k} \leq \frac{|\mathbf{f}_k^\top \mathbf{V}_\varphi^\top \Lambda_\varphi^{1/2} \Lambda_{\mathbf{a},\varphi,\mathbf{M}} \Lambda_\varphi^{1/2} \mathbf{V}_\varphi \mathbf{f}_k| + |\mathbf{f}_k^\top \mathbf{V}_\varphi^\top \Lambda_\varphi^{1/2} \mathbf{M}_{\mathbf{a},\varphi,\mathbf{R}} \Lambda_\varphi^{1/2} \mathbf{V}_\varphi \mathbf{f}_k|}{\|\mathbf{C}_\varphi \mathbf{f}_k\|^2}. \quad (\text{B.56})$$

By (B.9), (B.10), (B.44) and (B.50), we have

$$\|\mathbf{C}_\varphi \mathbf{f}_k\|^2 = \sum_{j=1}^T \alpha_{kj}^2 \lambda_{\varphi,j}. \quad (\text{B.57})$$

(B.9), (B.10), (B.46) and (B.50) imply that

$$\mathbf{f}_k^\top \mathbf{V}_\varphi^\top \Lambda_\varphi^{1/2} \Lambda_{\mathbf{a},\varphi, \mathbf{M}} \Lambda_\varphi^{1/2} \mathbf{V}_\varphi \mathbf{f}_k = \sum_{j=1}^T \alpha_{kj}^2 \gamma_{\varphi,j} \lambda_{\varphi,j}. \quad (\text{B.58})$$

From (B.47), we have

$$\frac{|\mathbf{f}_k^\top \mathbf{V}_\varphi^\top \Lambda_\varphi^{1/2} \mathbf{M}_{\mathbf{a},\varphi, \mathbf{R}} \Lambda_\varphi^{1/2} \mathbf{V}_\varphi \mathbf{f}_k|}{\|\mathbf{C}_\varphi \mathbf{f}_k\|^2} \leq \|\mathbf{M}_{\mathbf{a},\varphi, \mathbf{R}}\|_2 = O(T).$$

This, together with (B.55)-(B.58), implies that

$$\frac{\sum_{j=1}^T \alpha_{kj}^2 \gamma_{\varphi,j} \lambda_{\varphi,j}}{\sum_{j=1}^T \alpha_{kj}^2 \lambda_{\varphi,j}} - O(T) \leq \beta_{\varphi,k} \leq \frac{\sum_{j=1}^T \alpha_{kj}^2 \gamma_{\varphi,j} \lambda_{\varphi,j}}{\sum_{j=1}^T \alpha_{kj}^2 \lambda_{\varphi,j}} + O(T).$$

By Lemma 10, for any fixed k we have

$$\sum_{j=1}^T \frac{\alpha_{kj}^2 \lambda_{\varphi,j}}{\sum_{j=1}^T \alpha_{kj}^2 \lambda_{\varphi,j}} \frac{\gamma_{\varphi,j}}{\beta_{\varphi,k}} - O(T^{-1}) \leq 1 \leq \sum_{j=1}^T \frac{\alpha_{kj}^2 \lambda_{\varphi,j}}{\sum_{j=1}^T \alpha_{kj}^2 \lambda_{\varphi,j}} \frac{\gamma_{\varphi,j}}{\beta_{\varphi,k}} + O(T^{-1}). \quad (\text{B.59})$$

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 (B.9), (B.10) and (B.50) we have

$$0 = \mathbf{u}_k^\top \mathbf{u}_m = \frac{\mathbf{f}_k^\top \mathbf{C}_\varphi^\top \mathbf{C}_\varphi \mathbf{f}_m}{\|\mathbf{C}_\varphi \mathbf{f}_k\| \|\mathbf{C}_\varphi \mathbf{f}_m\|} = \frac{\sum_{j=1}^T \alpha_{kj} \alpha_{mj} \lambda_{\varphi,j}}{\|\mathbf{C}_\varphi \mathbf{f}_k\| \|\mathbf{C}_\varphi \mathbf{f}_m\|}.$$

This implies that

$$\sum_{j=1}^T \alpha_{kj} \alpha_{mj} \lambda_{\varphi,j} = 0. \quad (\text{B.60})$$

Moreover let

$$v_{kj} = \frac{\alpha_{kj} \lambda_{\varphi,j}^{1/2}}{(\sum_{j=1}^T \alpha_{kj}^2 \lambda_{\varphi,j})^{1/2}}.$$

We have

$$\sum_{j=1}^T v_{kj}^2 = 1. \quad (\text{B.61})$$

Note that (B.59) is equivalent to

$$\sum_{j=1}^T v_{kj}^2 \frac{\gamma_{\varphi,j}}{\beta_{\varphi,k}} - O(T^{-1}) \leq 1 \leq \sum_{j=1}^T v_{kj}^2 \frac{\gamma_{\varphi,j}}{\beta_{\varphi,k}} + O(T^{-1}). \quad (\text{B.62})$$

Also (B.60) implies that

$$\sum_{j=1}^T v_{kj} v_{mj} = 0. \quad (\text{B.63})$$

We consider v_{kj} for fixed k below. When $k = 1$ and T is big enough, Lemma 10, (B.61) and (B.62) imply

$$O(T^{-1}) = |1 - \sum_{j=1}^T v_{1j}^2 \frac{\gamma_{\varphi,j}}{\beta_{\varphi,1}}| \geq (1 - v_{11}^2) \frac{\beta_{\varphi,1} - \gamma_{\varphi,2}}{\beta_{\varphi,1}} - v_{11}^2 \frac{|\beta_{\varphi,1} - \gamma_{\varphi,1}|}{\beta_{\varphi,1}}. \quad (\text{B.64})$$

In view of (B.13)-(B.14), we have $\frac{\beta_{\varphi,1} - \gamma_{\varphi,1}}{\beta_{\varphi,1}} = O(T^{-1})$. Recalling that $\theta_{\varphi,k} \in [\frac{(T+1-k)\pi}{T+1}, \frac{(T+1-k)\pi}{T+1/2}]$, we can find that $|\frac{\beta_{\varphi,1}}{\beta_{\varphi,1} - \gamma_{\varphi,2}}|$ is bounded. It follows that (B.64) implies that $v_{11}^2 = 1 + O(T^{-1})$ and $\sum_{j=2}^T v_{1j}^2 = O(T^{-1})$. From (B.63), for any $k \neq 1$ we have

$$|v_{k1} v_{11}| = |\sum_{j=2}^T v_{kj} v_{1j}| \leq (\sum_{j=2}^T v_{kj}^2)^{1/2} (\sum_{j=2}^T v_{1j}^2)^{1/2} = O(T^{-1/2}). \quad (\text{B.65})$$

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 (B.52). \square

proof of Lemma 15. 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 (B.9) and (B.50) that

$$\mathbf{s}_k = \frac{\mathbf{F}^\top \mathbf{C}_\varphi^\top \mathbf{C}_\varphi \mathbf{f}_k}{\gamma_{\varphi,1}^{1/2} \|\mathbf{C}_\varphi \mathbf{f}_k\|} = \frac{1}{\gamma_{\varphi,1}^{1/2} \|\mathbf{C}_\varphi \mathbf{f}_k\|} \sum_{j=1}^T \alpha_{kj} \lambda_{\varphi,j} \mathbf{F}^\top \tilde{\mathbf{y}}_j = \mathbf{s}_{k,M} + \mathbf{s}_{k,R}, \quad (\text{B.66})$$

where

$$\mathbf{s}_{k,M} = \frac{1}{\gamma_{\varphi,1}^{1/2} \|\mathbf{C}_\varphi \mathbf{f}_k\|} \alpha_{kk} \lambda_{\varphi,k} \mathbf{F}^\top \tilde{\mathbf{y}}_k, \quad \mathbf{s}_{k,R} = \frac{1}{\gamma_{\varphi,1}^{1/2} \|\mathbf{C}_\varphi \mathbf{f}_k\|} \sum_{j \neq k} \alpha_{kj} \lambda_{\varphi,j} \mathbf{F}^\top \tilde{\mathbf{y}}_j. \quad (\text{B.67})$$

By Hölder's inequality, we have

$$\|\mathbf{s}_{k,R}\| = \left\| \frac{1}{\gamma_{\varphi,1}^{1/2} \|\mathbf{C}_\varphi \mathbf{f}_k\|} \sum_{j \neq k} \alpha_{kj} \lambda_{\varphi,j} \mathbf{F}^\top \tilde{\mathbf{y}}_j \right\| \leq \frac{1}{\gamma_{\varphi,1}^{1/2} \|\mathbf{C}_\varphi \mathbf{f}_k\|} \|\mathbf{F}\|_2 (\sum_{j \neq k} \alpha_{kj}^2 \lambda_{\varphi,j}^2)^{1/2}.$$

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

$$\|\mathbf{F}\|_2 = \|\mathbf{A}\|_2^{1/2}.$$

Since A is a Hermitian Toeplitz matrix, from [18],

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

By (B.42) we can get

$$\|\mathbf{F}\|_2 = \|\mathbf{A}\|_2^{1/2} < \infty.$$

From Lemma 6, (B.52), and (B.57) we can obtain that for any fixed k ,

$$\frac{(\sum_{j \neq k} \alpha_{kj}^2 \lambda_{\varphi,j}^2)^{1/2}}{\|\mathbf{C}_\varphi \mathbf{f}_k\|} \leq \left(\frac{\lambda_{\varphi,1} \sum_{j \neq k} \alpha_{kj}^2 \lambda_{\varphi,j}}{\sum_{j=1}^T \alpha_{kj}^2 \lambda_{\varphi,j}} \right)^{1/2} = O(T^{1/2}).$$

This, together with (B.13), implies that for any fixed k ,

$$\|\mathbf{s}_{k,R}\| = O(T^{-1/2}). \quad (\text{B.68})$$

Similarly, we can also obtain that $\frac{1}{\gamma_{\varphi,1}^{1/2} \|\mathbf{C}_\varphi \mathbf{f}_k\|} \alpha_{kk} \lambda_k$ 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 (B.11), (B.52) and (B.67) and Assumption 1 we can obtain that for any fixed k ,

$$|S_{k,M,j}| \leq \frac{1}{\gamma_{\varphi,1}^{1/2} \|\mathbf{C}_\varphi \mathbf{f}_k\|} |\alpha_{kk}| \lambda_k \frac{2}{(2T+1)^{1/2}} \sum_{h=0}^l |b_h| = O(T^{-1/2}). \quad (\text{B.69})$$

It follows from (B.66)-(B.69) that for any fixed k ,

$$\begin{aligned} \sum_{j=1}^{T+l} S_{k,j}^4 &\leq 8 \sum_{j=1}^{T+l} (S_{k,R,j}^4 + S_{k,M,j}^4) \\ &\leq 8 \sum_{j=1}^{T+l} S_{k,M,j}^4 + 8 \left(\sum_{j=1}^{T+l} S_{k,R,j}^2 \right)^2 = O(T^{-1}). \end{aligned} \quad (\text{B.70})$$

□

B.3 The proof of Lemma 2

The proof of the remaining part of the Lemma 2 for truncated matrices is the same as that of the part C.2.3-C.2.4 in [31]. Moreover, we can prove Lemma 2 with the result for truncated matrices and the same method as that of Appendix B in [31].

C Appendix C: The proof of Lemmas 3 and 4

C.1 The proof of Lemma 3

At first we consider the impact of δ_t . Recalling the definition of $\dot{\mu}_m$, we write

$$\begin{aligned}
\dot{\mu}_m &= \sum_{i=2}^T \frac{\dot{\mathbf{x}}_{i,i}}{n(T-1)} + 2 \sum_{j=1}^m \sum_{i=2}^{T-j} \frac{\dot{\mathbf{x}}_{i,i+j}}{n(T-j-1)} \\
&= \sum_{i=2}^T \frac{(\mathbf{x}_i - \delta_i - \mathbf{x}_{i-1} + \delta_{i-1})^\top (\mathbf{I} - \mathbf{H})(\mathbf{x}_i - \delta_i - \mathbf{x}_{i-1} + \delta_{i-1})}{n(T-1)} \\
&\quad + 2 \sum_{j=1}^m \sum_{i=2}^{T-j} \frac{(\mathbf{x}_i - \delta_i - \mathbf{x}_{i-1} + \delta_{i-1})^\top (\mathbf{I} - \mathbf{H})(\mathbf{x}_{i+j} - \delta_{i+j} - \mathbf{x}_{i+j-1} + \delta_{i+j-1})}{n(T-j-1)} \\
&\quad + \sum_{i=2}^T \frac{2(\mathbf{x}_i - \delta_i - \mathbf{x}_{i-1} + \delta_{i-1})^\top (\mathbf{I} - \mathbf{H})(\delta_i - \delta_{i-1})}{n(T-1)} \\
&\quad + 2 \sum_{j=1}^m \sum_{i=2}^{T-j} \frac{(\delta_i - \delta_{i-1})^\top (\mathbf{I} - \mathbf{H})(\mathbf{x}_{i+j} - \delta_{i+j} - \mathbf{x}_{i+j-1} + \delta_{i+j-1})}{n(T-j-1)} \\
&\quad + 2 \sum_{j=1}^m \sum_{i=2}^{T-j} \frac{(\mathbf{x}_i - \delta_i - \mathbf{x}_{i-1} + \delta_{i-1})^\top (\mathbf{I} - \mathbf{H})(\delta_{i+j} - \delta_{i+j-1})}{n(T-j-1)} \\
&\quad + \sum_{i=2}^T \frac{(\delta_i - \delta_{i-1})^\top (\mathbf{I} - \mathbf{H})(\delta_i - \delta_{i-1})}{n(T-1)} \\
&\quad + 2 \sum_{j=1}^m \sum_{i=2}^{T-j} \frac{(\delta_i - \delta_{i-1})^\top (\mathbf{I} - \mathbf{H})(\delta_{i+j} - \delta_{i+j-1})}{n(T-j-1)}.
\end{aligned} \tag{C.1}$$

Recall the definition of m , $m \leq n^{1/2}$. Recall Proposition 1, when q is bounded,

$$\lim_{n, T \rightarrow \infty} P(q \leq \hat{q}) = 1. \tag{C.2}$$

Moreover

$$\begin{aligned}
& \left| \sum_{i=2}^T \frac{(\delta_i - \delta_{i-1})(\mathbf{I} - \mathbf{H})^\top (\delta_i - \delta_{i-1})}{n(T-1)} + 2 \sum_{j=1}^m \sum_{i=2}^{T-j} \frac{(\delta_i - \delta_{i-1})(\mathbf{I} - \mathbf{H})^\top (\delta_{i+j} - \delta_{i+j-1})}{n(T-j-1)} \right| \\
& \leq (2m+1) \sum_{i=2}^T \frac{(\delta_i - \delta_{i-1})^\top (\mathbf{I} - \mathbf{H})(\delta_i - \delta_{i-1})}{n(T-m-1)} \\
& = \frac{1}{n^{1/2}} O \left(\sum_{i=2}^T \frac{\min\{q, n^{1/2}\} (\delta_i - \delta_{i-1})^\top (\mathbf{I} - \mathbf{H})(\delta_i - \delta_{i-1})}{n^{1/2} T} \right) = o_p(n^{-1/2}),
\end{aligned} \tag{C.3}$$

which is ensured by Assumption 12. The other parts involving $(\delta_i - \delta_{i-1})$ are similar. Note that Assumptions 6 and 10-12 ensure that δ_t does not change the asymptotic properties of the estimators. In the following part of this subsection we only consider the case when $\delta_t = \mathbf{0}$.

proof of Lemma 3. Recall (A.23),

$$\begin{aligned} & \frac{1}{n(T-j-1)} \sum_{t=j+2}^T E \dot{\mathbf{x}}_{t,t-j} \\ &= \frac{\text{tr}(\boldsymbol{\Sigma}^{1/2}(\mathbf{I} - \mathbf{H})\boldsymbol{\Sigma}^{1/2})}{n} \frac{1}{(T-j-1)} \sum_{t=j+2}^T (a_j + A_{1tj} + A_{2tj} + A_{3tj} + A_{4tj} + A_{5tj} + A_{6tj}). \end{aligned} \quad (\text{C.4})$$

We next analyze these term by term.

The terms A_{ktj} , $k = 1, \dots, 6$ can be rewritten as follows:

$$\begin{aligned} \sum_{t=j+2}^T A_{2tj} &= \sum_{t=j+2}^T (\varphi - 1) \sum_{k=1}^{t-1-j} \varphi^{k+j-1} a_k \\ &= (\varphi - 1) \sum_{k=1}^{T-1-j} \varphi^{k+j-1} \sum_{t=j+1+k}^T a_k = (\varphi - 1) \sum_{k=1}^{T-1-j} \varphi^{k+j-1} (T-j-k) a_k; \end{aligned} \quad (\text{C.5})$$

$$\sum_{t=j+2}^T A_{3tj} = \sum_{t=j+2}^T (\varphi - 1) \sum_{k=j+1}^{t-1} \varphi^{k-j-1} a_k = (\varphi - 1) \sum_{k=j+1}^{T-1} \varphi^{k-j-1} (T-k) a_k; \quad (\text{C.6})$$

$$\begin{aligned} \sum_{t=j+2}^T A_{4tj} &= \frac{1-\varphi}{1+\varphi} \sum_{k_2=1}^{T-2} a_{k_2} \varphi^{|j-k_2|} (T-1 - \max\{k_2, j\}) \\ &\quad - \frac{1}{(1+\varphi)^2} \sum_{k_2=1}^{T-2} a_{k_2} (\varphi^{2+|k_2-j|} - \varphi^{\min\{2+k_2+j, 2T-k_2-j\}}) \end{aligned} \quad (\text{C.7})$$

$$- \frac{1-\varphi}{1+\varphi} \sum_{k_2=1}^{T-2} a_{k_2} \varphi^{k_2+j} \max\{T-k_2-j-1, 0\}; \quad (\text{C.8})$$

$$\sum_{t=j+2}^T A_{5tj} = \frac{1-\varphi}{1+\varphi} \sum_{k_2=1}^{T-j-2} a_{k_2} \varphi^{j+k_2} (T-1-j-k_2) \quad (\text{C.9})$$

$$- \frac{1}{(1+\varphi)^2} \sum_{k_2=1}^{T-j-2} a_{k_2} (\varphi^{2+k_2+j} - \varphi^{2T-k_2-j}); \quad (\text{C.10})$$

and

$$\sum_{t=j+2}^T A_{6tj} = \frac{1-\varphi}{1+\varphi} \sum_{k_2=0}^{T-j-2} a_{k_2} \varphi^{j+k_2} (T-1-j-k_2) \quad (\text{C.11})$$

$$- \frac{1}{(1+\varphi)^2} \sum_{k_2=0}^{T-j-2} a_{k_2} (\varphi^{2+k_2+j} - \varphi^{2T-k_2-j}). \quad (\text{C.12})$$

From (C.4)-(C.11) we conclude that

$$\begin{aligned}
& \frac{1}{(T-j-1)} \sum_{t=j+2}^T \frac{E\dot{\mathbf{x}}_{t,t-j}}{\text{tr}(\boldsymbol{\Sigma}^{1/2}(\mathbf{I}-\mathbf{H})\boldsymbol{\Sigma}^{1/2})} \\
&= \sum_{k=0}^{T-j-2} a_k (1\{j=k\}) \left(1 + \frac{1-\varphi}{1+\varphi} - \frac{1}{(1+\varphi)^2} \frac{\varphi^2 - \varphi^{2T-2k}}{T-j-1} \right) \\
&+ 1\{k < j\} \left((\varphi-1)\varphi^{j-1-k} + \frac{1-\varphi}{1+\varphi} \varphi^{j-k} - \frac{1}{(1+\varphi)^2} \frac{\varphi^{2+j-k} - \varphi^{2T-k-j}}{T-j-1} \right) \\
&+ 1\{k > j\} \left((\varphi-1)\varphi^{k-1-j} \frac{T-k}{T-1-j} + \frac{1-\varphi}{1+\varphi} \varphi^{k-j} \frac{T-k-1}{T-1-j} - \frac{1}{(1+\varphi)^2} \frac{\varphi^{2+k-j} - \varphi^{2T-k-j}}{T-j-1} \right) \\
&+ (\varphi-1)\varphi^{k-1+j} \frac{T-k-j}{T-1-j} + \frac{1-\varphi}{1+\varphi} \varphi^{k+j} \frac{T-k-1-j}{T-1-j} - \frac{1}{(1+\varphi)^2} \frac{\varphi^{2+k+j} - \varphi^{2T-k-j}}{T-j-1} \\
&+ O\left((\varphi-1) \sum_{k=T-j-1}^{T-1} |a_k| \right). \tag{C.13}
\end{aligned}$$

Similarly $E\dot{\mathbf{x}}_{t,t}$ can be written as

$$\begin{aligned}
E\dot{\mathbf{x}}_{t,t} &= E(\mathbf{x}_t - \mathbf{x}_{t-1})^\top (\mathbf{I} - \mathbf{H})(\mathbf{x}_t - \mathbf{x}_{t-1}) \\
&= E[\mathbf{y}_t + (\varphi-1) \sum_{k=0}^{t-2} \varphi^k \mathbf{y}_{t-k-1}]^\top \boldsymbol{\Sigma}^{1/2}(\mathbf{I}-\mathbf{H})\boldsymbol{\Sigma}^{1/2}[\mathbf{y}_{t-j} + (\varphi-1) \sum_{k=0}^{t-2} \varphi^k \mathbf{y}_{t-k-1}] \\
&= \text{tr}[\boldsymbol{\Sigma}^{1/2}(\mathbf{I}-\mathbf{H})\boldsymbol{\Sigma}^{1/2}] [a_0 + 2(\varphi-1) \sum_{k=0}^{t-2} \varphi^k a_{k+1} + 2\frac{1-\varphi}{1+\varphi} \sum_{k=1}^{t-2} a_k (\varphi^k - \varphi^{2t-k-2}) \\
&+ \frac{1-\varphi}{1+\varphi} a_0 (1 - \varphi^{2t-2})]. \tag{C.14}
\end{aligned}$$

It follows that

$$\begin{aligned}
& \frac{1}{n(T-1)} \sum_{t=2}^T E\dot{\mathbf{x}}_{t,t} \\
&= \frac{\text{tr}[\boldsymbol{\Sigma}^{1/2}(\mathbf{I}-\mathbf{H})\boldsymbol{\Sigma}^{1/2}]}{n} [a_0 + \sum_{k=1}^{T-1} \frac{2(\varphi-1)\varphi^{k-1}(T-k)}{T-1} a_k + 2\frac{1-\varphi}{1+\varphi} \sum_{k=1}^{T-2} \varphi^k a_k \frac{T-k-1}{T-1} \\
&- 2\frac{1}{(1+\varphi)^2} \sum_{k=1}^{T-2} a_k \frac{\varphi^{2+k} - \varphi^{2T-k}}{T-1} + \frac{1-\varphi}{1+\varphi} a_0 - \frac{\varphi^2 - \varphi^{2T}}{(1+\varphi)^2(T-1)} a_0]. \tag{C.15}
\end{aligned}$$

Write $E\dot{\mu}_m$ as

$$E\dot{\mu}_m = \frac{\text{tr}[\boldsymbol{\Sigma}^{1/2}(\mathbf{I}-\mathbf{H})\boldsymbol{\Sigma}^{1/2}]}{n} \sum_{j=0}^{\infty} \kappa_j a_j. \tag{C.16}$$

From (C.13) and (C.15) we conclude that

$$\begin{aligned}
\kappa_0 &= 1 + \frac{1-\varphi}{1+\varphi} - \frac{\varphi^2 - \varphi^{2T}}{(1+\varphi)^2(T-1)} + 2 \sum_{j=1}^m [(\varphi-1)\varphi^{j-1} + \frac{1-\varphi}{1+\varphi} \varphi^j - \frac{1}{(1+\varphi)^2} \frac{\varphi^{2+j} - \varphi^{2T-j}}{T-j-1}] \\
&= \frac{2\varphi^m}{1+\varphi} - \frac{\varphi^2 - \varphi^{2T}}{(1+\varphi)^2(T-1)} - 2 \sum_{j=1}^m \frac{1}{(1+\varphi)^2} \frac{\varphi^{2+j} - \varphi^{2T-j}}{T-1} + 2 \sum_{j=1}^m \frac{1}{(1+\varphi)^2} \frac{j(\varphi^{2+j} - \varphi^{2T-j})}{(T-1)(T-j-1)} \\
&= \frac{2\varphi^m}{1+\varphi} - \frac{\varphi^2 - \varphi^{2T}}{(1+\varphi)^2(T-1)} - 2 \sum_{j=1}^m \frac{1}{(1+\varphi)^2} \frac{\varphi^{2+j} - \varphi^{2T-j}}{T-1} + O\left(\frac{m^2(1-\varphi)}{T}\right), \tag{C.17}
\end{aligned}$$

and

$$1 + 2 \sum_{j=1}^m (\varphi - 1) \varphi^{j-1} < \kappa_0 \leq 1 + \frac{2\varphi}{1+\varphi} \sum_{j=1}^m (\varphi - 1) \varphi^{j-1} + \frac{1-\varphi}{1+\varphi}. \quad (\text{C.18})$$

It follows that

$$|1 - \kappa_0| = O(m(1 - \varphi)). \quad (\text{C.19})$$

Similarly, for $1 \leq f \leq m$,

$$\begin{aligned} \kappa_f/2 &= (\varphi - 1) \varphi^{f-1} \frac{T-f}{T-1} + \frac{1-\varphi}{1+\varphi} \varphi^f \frac{T-f-1}{T-1} - \frac{\varphi^{2+f} - \varphi^{2T-f}}{(1+\varphi)^2(T-1)} \\ &+ 1 + \frac{1-\varphi}{1+\varphi} - \frac{\varphi^2 - \varphi^{2T-2f}}{(1+\varphi)^2(T-f-1)} \\ &+ \sum_{j=f+1}^m \left((\varphi - 1) \varphi^{j-1-f} + \frac{1-\varphi}{1+\varphi} \varphi^{j-f} - \frac{1}{(1+\varphi)^2} \frac{\varphi^{2+j-f} - \varphi^{2T-j-f}}{T-j-1} \right) \\ &+ \sum_{j=1}^{f-1} \left((\varphi - 1) \varphi^{f-1-j} \frac{T-f}{T-1-j} + \frac{1-\varphi}{1+\varphi} \varphi^{f-j} \frac{T-f-1}{T-1-j} - \frac{1}{(1+\varphi)^2} \frac{\varphi^{2+f-j} - \varphi^{2T-j-f}}{T-j-1} \right) \\ &+ \sum_{j=1}^m \left((\varphi - 1) \varphi^{f-1+j} \frac{T-f-j}{T-1-j} + \frac{1-\varphi}{1+\varphi} \varphi^{f+j} \frac{T-f-j-1}{T-1-j} - \frac{1}{(1+\varphi)^2} \frac{\varphi^{2+f+j} - \varphi^{2T-j-f}}{T-j-1} \right) \\ &- \sum_{j=1}^m \frac{1}{(1+\varphi)^2} \frac{\varphi^{2+|f-j|} - \varphi^{2T-j-f}}{T-1} + O\left(\frac{m^2(1-\varphi)}{T}\right) \\ &= \kappa_0 + \frac{\varphi^{m-f} + \varphi^{m+f} - 2\varphi^m}{1+\varphi} - \frac{\varphi^{2+f} - \varphi^2 - \varphi^{2T-f} + \varphi^{2T}}{(1+\varphi)^2(T-1)} \\ &- \sum_{j=1}^m \frac{1}{(1+\varphi)^2} \frac{\varphi^{2+j+f} - \varphi^{2+j} - \varphi^{2T-j-f} + \varphi^{2T-j}}{T-1} \\ &- \sum_{j=1}^m \frac{1}{(1+\varphi)^2} \frac{\varphi^{2+|f-j|} - \varphi^{2+j} - \varphi^{2T-j-f} + \varphi^{2T-j}}{T-1} + O\left(\frac{m^2(1-\varphi)}{T}\right) \\ &= \kappa_0 + O\left(\frac{mf(1-\varphi)}{T}\right) + O\left(\frac{m^2(1-\varphi)}{T}\right) \\ &= \kappa_0 + o(m(1-\varphi)). \end{aligned} \quad (\text{C.20})$$

As for $f \geq m_1$, we also have

$$|\kappa_j/2| = O(m(1 - \varphi)) \quad (\text{C.21})$$

and

$$\sum_{j=m+1}^{\infty} |\kappa_j| |a_j| \leq \frac{\max_{j>m} \{|\kappa_j|\}}{m+1} \sum_{j=m+1}^{\infty} j |a_j| = O(1 - \varphi). \quad (\text{C.22})$$

Then

$$E\hat{\mu}_m = \frac{\text{tr}(\boldsymbol{\Sigma}^{1/2}(\mathbf{I} - \mathbf{H})\boldsymbol{\Sigma}^{1/2})}{n} \kappa_0 (a_0 + 2 \sum_{j=1}^{\infty} a_j) + o(m(1 - \varphi)) + O(1 - \varphi). \quad (\text{C.23})$$

Similarly, one can verify that

$$\text{Var}(\mu_m) = O\left(\sum_{i=2}^T \frac{\mathbf{y}_i^\top \boldsymbol{\Sigma}^{1/2}(\mathbf{I} - \mathbf{H})\boldsymbol{\Sigma}^{1/2} \mathbf{y}_i}{n(T-1)} + 2 \sum_{j=1}^m \sum_{i=2}^{T-j} \frac{\mathbf{y}_i^\top \boldsymbol{\Sigma}^{1/2}(\mathbf{I} - \mathbf{H})\boldsymbol{\Sigma}^{1/2} \mathbf{y}_{i+j}}{n(T-j-1)} \right). \quad (\text{C.24})$$

This, together with (2.66) in [31], implies

$$\text{Var}(\mu_m) = o(n^{-1}). \quad (\text{C.25})$$

Write

$$\begin{aligned} & \kappa(\theta_{\varphi,1}) - \kappa_0 \\ &= \varphi^m - m \frac{\varphi^3 - \varphi^{2T-m}}{2(T-1)} - \frac{2\varphi^m}{1+\varphi} + \frac{\varphi^2 - \varphi^{2T}}{(1+\varphi)^2(T-1)} \\ &+ 2 \sum_{j=1}^m \frac{1}{(1+\varphi)^2} \frac{\varphi^{2+j} - \varphi^{2T-j}}{T-1} + O\left(\frac{m^2(1-\varphi)}{T}\right). \end{aligned} \quad (\text{C.26})$$

Note that

$$\varphi^m - \frac{2\varphi^m}{1+\varphi} = \varphi^m \frac{1-\varphi}{1+\varphi} = O(1-\varphi), \quad (\text{C.27})$$

$$\frac{\varphi^2 - \varphi^{2T}}{(1+\varphi)^2(T-1)} = O(1-\varphi), \quad (\text{C.28})$$

$$m \frac{\varphi^3 - \varphi^{2T-m}}{2(T-1)} - 2 \sum_{j=1}^m \frac{1}{(1+\varphi)^2} \frac{\varphi^{2+j} - \varphi^{2T-j}}{T-1} = O\left(\frac{m^2(1-\varphi)}{T}\right) \quad (\text{C.29})$$

and

$$\frac{m^2(1-\varphi)}{T} = o(m(1-\varphi)). \quad (\text{C.30})$$

(C.26)-(C.30) imply (A.1). (C.23)-(C.25) and (A.1) imply (A.2). \square

C.2 The proof of Lemma 4

proof of Lemma 4. Now we define

$$\check{\mathbf{x}}_{t,s} = (\mathbf{x}_t - \mathbf{x}_{t-1})^\top (\mathbf{x}_s - \mathbf{x}_{s-1}).$$

We also replace $\hat{\mathbf{x}}_{t,s}$ by $\check{\mathbf{x}}_{t,s}$ in $\hat{S}_{\sigma^2,0}$ and $\hat{S}_{\sigma^2,m}$ and get $S_{\sigma^2,0}$ and $S_{\sigma^2,m}$. In fact the differences between $\hat{S}_{\sigma^2,0}$ and $S_{\sigma^2,0}$ are the ones between $\Sigma^{1/2}(\mathbf{I} - \mathbf{H})\Sigma^{1/2}$ and Σ . So we just need to study $S_{\sigma^2,0}$ and $S_{\sigma^2,m}$ then replace Σ by $\Sigma^{1/2}(\mathbf{I} - \mathbf{H})\Sigma^{1/2}$.

. We rewrite $(\mathbf{x}_t - \mathbf{x}_{t-1})$ as follows

$$\mathbf{x}_t - \mathbf{x}_{t-1} = \Sigma^{1/2} \left(\mathbf{y}_t + (\varphi - 1) \sum_{k=0}^{t-2} \varphi^k \mathbf{y}_{t-k-1} \right) \quad (\text{C.31})$$

$$\begin{aligned} &= \Sigma^{1/2} \left(\sum_{s=0}^{\infty} b_s Z_{t-s} + (\varphi - 1) \sum_{k=0}^{t-2} \varphi^k \sum_{s=0}^{\infty} b_s Z_{t-1-k-s} \right) \\ &= \Sigma^{1/2} \left(\sum_{s=-\infty}^t b_{t-s} Z_s + (\varphi - 1) \sum_{k=0}^{t-2} \varphi^k \sum_{s=-\infty}^{t-1-k} b_{t-1-k-s} Z_s \right) \\ &= \Sigma^{1/2} \left(b_0 Z_t + \sum_{s=1}^{t-1} [b_{t-s} + (\varphi - 1) \sum_{k=0}^{t-1-s} \varphi^k b_{t-1-k-s}] Z_s \right. \\ &\quad \left. + \sum_{s=-\infty}^0 [b_{t-s} + (\varphi - 1) \sum_{k=0}^{t-2} \varphi^k b_{t-1-k-s}] Z_s \right). \end{aligned} \quad (\text{C.32})$$

Define a new variable $\check{b}_{t,s}$ as

$$\check{b}_{t,s} = \begin{cases} 0 & s > t, \\ b_0 & s = t, \\ b_{t-s} + (\varphi - 1) \sum_{k=0}^{t-1-s} \varphi^k b_{t-1-k-s} & 1 \leq s \leq t-1, \\ b_{t-s} + (\varphi - 1) \sum_{k=0}^{t-2} \varphi^k b_{t-1-k-s} & s \leq 0, \end{cases} \quad (\text{C.33})$$

so that

$$\mathbf{x}_t - \mathbf{x}_{t-1} = \Sigma^{1/2} \sum_{s=-\infty}^t \check{b}_{t,s} Z_s = \Sigma^{1/2} \sum_{s=-\infty}^{\infty} \check{b}_{t,s} Z_s. \quad (\text{C.34})$$

Then

$$\begin{aligned} \check{\mathbf{x}}_{f,g}^2 &= \left(\sum_{i=1}^n \sum_{j=1}^n \Sigma_{ij} \sum_{h_1=-\infty}^{\infty} \sum_{h_2=-\infty}^{\infty} \check{b}_{f,h_1} \check{b}_{g,h_2} Z_{h_1,i} Z_{h_2,j} \right)^2 \\ &= \sum_{i_1=1}^n \sum_{j_1=1}^n \sum_{i_2=1}^n \sum_{j_2=1}^n \Sigma_{i_1 j_1} \Sigma_{i_2 j_2} \sum_{h_1=-\infty}^{\infty} \sum_{h_2=-\infty}^{\infty} \sum_{h_3=-\infty}^{\infty} \sum_{h_4=-\infty}^{\infty} \\ &\quad \check{b}_{f,h_1} \check{b}_{g,h_2} \check{b}_{f,h_3} \check{b}_{g,h_4} Z_{h_1,i_1} Z_{h_2,j_1} Z_{h_3,i_2} Z_{h_4,j_2}. \end{aligned} \quad (\text{C.35})$$

Write $S_{\sigma^2,0} = S_{\sigma^2,0,h} + S_{\sigma^2,0,l}$ where

$$\begin{aligned} S_{\sigma^2,0,h} &= \frac{1}{(T - \frac{3}{2}\lfloor T/2 \rfloor)(\lfloor T/2 \rfloor - 1)} \sum_{f=2}^{\lfloor T/2 \rfloor} \sum_{g=f+\lfloor T/2 \rfloor}^T \left(2 \sum_{i=1}^n \sum_{j=1}^n \Sigma_{ii} \Sigma_{ij} \sum_{h_1=-\infty}^{\infty} \sum_{h_2=-\infty}^{\infty} \right. \\ &\quad \left. Z_{h_1,i}^3 Z_{h_2,j} (\check{b}_{f,h_1}^2 \check{b}_{g,h_1} \check{b}_{g,h_2} + \check{b}_{f,h_1} \check{b}_{f,h_2} \check{b}_{g,h_1}^2) - 3 \sum_{i=1}^n \Sigma_{ii}^2 \sum_{h=-\infty}^{\infty} Z_{h,i}^4 \check{b}_{f,h}^2 \check{b}_{g,h}^2 \right). \end{aligned} \quad (\text{C.36})$$

When $t > 2$

$$\begin{aligned} &\sum_{f=2}^{\lfloor T/2 \rfloor} \sum_{g=f+\lfloor T/2 \rfloor}^T \sum_{i=1}^n \sum_{j=1}^n |\Sigma_{ii} \Sigma_{ij}| \sum_{h_1=-\infty}^{\infty} \sum_{h_2=-\infty}^{\infty} \check{b}_{f,h_1}^2 |\check{b}_{g,h_1}| |\check{b}_{g,h_2}| \\ &= \sum_{i=1}^n \sum_{j=1}^n |\Sigma_{ii} \Sigma_{ij}| \sum_{f=2}^{\lfloor T/2 \rfloor} \sum_{h_1=-\infty}^f \check{b}_{f,h_1}^2 \sum_{g=f+\lfloor T/2 \rfloor}^T |\check{b}_{g,h_1}| \sum_{h_2=-\infty}^g |\check{b}_{g,h_2}|. \end{aligned} \quad (\text{C.37})$$

When $h_2 \leq g$

$$\sum_{h_2=-\infty}^g |\check{b}_{g,h_2}| \leq 2 \sum_{s=0}^{\infty} |b_s| < \infty. \quad (\text{C.38})$$

When $h_1 \leq f \leq g - \lfloor T/2 \rfloor$

$$\sum_{g=f+\lfloor T/2 \rfloor}^T |\check{b}_{g,h_1}| \leq \sum_{s=\lfloor T/2 \rfloor}^{\infty} |b_s| + \sum_{s=0}^{\infty} |b_s| T(1-\varphi) = o(T^{-1}) + O(T(1-\varphi)) < \infty \quad (\text{C.39})$$

and

$$\sum_{h_1=-\infty}^f \check{b}_{f,h_1}^2 \leq \left(\sum_{h_1=-\infty}^f |\check{b}_{f,h_1}| \right)^2 < \infty. \quad (\text{C.40})$$

Moreover,

$$\sum_{i=1}^n \sum_{j=1}^n |\Sigma_{ii} \Sigma_{ij}| \leq \max_{1 \leq i \leq p} \{|\Sigma_{ii}|\} \left(\sum_{i=1}^n \sum_{j=1}^n \Sigma_{ij}^2 n^2 \right)^{1/2} = \max_{1 \leq i \leq p} \{|\Sigma_{ii}|\} n (\text{tr}(\Sigma^2))^{1/2} = O(n^{3/2}). \quad (\text{C.41})$$

From (C.38)-(C.41) we conclude that

$$\sum_{f=2}^{\lfloor T/2 \rfloor} \sum_{g=f+\lfloor T/2 \rfloor}^T \sum_{i=1}^n \sum_{j=1}^n |\Sigma_{ii} \Sigma_{ij}| \sum_{h_1=-\infty}^{\infty} \sum_{h_2=-\infty}^{\infty} \check{b}_{f,h_1}^2 |\check{b}_{g,h_1}| |\check{b}_{g,h_2}| = O(n^{3/2}T).$$

Similarly, we can obtain the order

$$\sum_{f=2}^{\lfloor T/2 \rfloor} \sum_{g=f+\lfloor T/2 \rfloor}^T \sum_{i=1}^n \sum_{j=1}^n |\Sigma_{ii} \Sigma_{ij}| \sum_{h_1=-\infty}^{\infty} \sum_{h_2=-\infty}^{\infty} \check{b}_{g,h_1}^2 |\check{b}_{f,h_1}| |\check{b}_{f,h_2}| = O(n^{3/2}T)$$

and

$$\sum_{i=1}^n \Sigma_{ii}^2 \sum_{h=-\infty}^{\infty} \check{b}_{f,h}^2 \check{b}_{g,h}^2 = O(n).$$

Summarizing the above we have

$$E|S_{\sigma^2,0,h}| = O(n^{3/2}T^{-1}) = o(n). \quad (\text{C.42})$$

Similarly write

$$\begin{aligned} ES_{\sigma^2,0,l} &= \frac{1}{(T - \frac{3}{2}\lfloor T/2 \rfloor)(\lfloor T/2 \rfloor - 1)} \sum_{f=2}^{\lfloor T/2 \rfloor} \sum_{g=f+\lfloor T/2 \rfloor}^T \\ &\left(\sum_{i=1}^n \sum_{j=1}^n \Sigma_{ij}^2 \sum_{h_1=-\infty}^{\infty} \sum_{h_2=-\infty}^{\infty} \check{b}_{f,h_1}^2 \check{b}_{g,h_2}^2 \right. \\ &+ \sum_{i=1}^n \sum_{j=1}^n \Sigma_{ij}^2 \sum_{h_1=-\infty}^{\infty} \sum_{h_2=-\infty}^{\infty} \check{b}_{f,h_1} \check{b}_{g,h_2} \check{b}_{f,h_2} \check{b}_{g,h_1} \\ &+ \sum_{i=1}^n \sum_{j=1}^n \Sigma_{ii} \Sigma_{jj} \sum_{h_1=-\infty}^{\infty} \sum_{h_2=-\infty}^{\infty} \check{b}_{f,h_1} \check{b}_{g,h_1} \check{b}_{f,h_2} \check{b}_{g,h_2} \\ &\left. - 3 \sum_{i=1}^n \Sigma_{ii}^2 \sum_{h=-\infty}^{\infty} \check{b}_{f,h}^2 \check{b}_{g,h}^2 \right). \end{aligned} \quad (\text{C.43})$$

Note that

$$\left| \sum_{h=-\infty}^{\infty} \check{b}_{f,h} \check{b}_{g,h} - a_{g-f} \right| = \left| \sum_{h=-\infty}^{\infty} \check{b}_{f,h} \check{b}_{g,h} - \sum_{s=0}^{\infty} b_s b_{g-f+s} \right| = O(1 - \varphi) \quad (\text{C.44})$$

and

$$\left| \sum_{h=-\infty}^{\infty} \check{b}_{f,h}^2 - a_0 \right| = \left| \sum_{h=-\infty}^{\infty} \check{b}_{f,h}^2 - \sum_{s=0}^{\infty} b_s^2 \right| = O(1 - \varphi). \quad (\text{C.45})$$

We then conclude that

$$\frac{\sum_{f=2}^{\lfloor T/2 \rfloor} \sum_{g=f+\lfloor T/2 \rfloor}^T \sum_{i=1}^n \sum_{j=1}^n \Sigma_{ij}^2 \sum_{h_1=-\infty}^{\infty} \sum_{h_2=-\infty}^{\infty} \check{b}_{f,h_1}^2 \check{b}_{g,h_2}^2}{(T - \frac{3}{2}\lfloor T/2 \rfloor)(\lfloor T/2 \rfloor - 1)} = a_0^2 \text{tr}(\Sigma^2) + O((1 - \varphi)n), \quad (\text{C.46})$$

$$\begin{aligned}
& \frac{\sum_{f=2}^{\lfloor T/2 \rfloor} \sum_{g=f+\lfloor T/2 \rfloor}^T \sum_{i=1}^n \sum_{j=1}^n \boldsymbol{\Sigma}_{ij}^2 \sum_{h_1=-\infty}^{\infty} \sum_{h_2=-\infty}^{\infty} \check{b}_{f,h_1} \check{b}_{g,h_2} \check{b}_{f,h_2} \check{b}_{g,h_1}}{(T - \frac{3}{2}\lfloor T/2 \rfloor)(\lfloor T/2 \rfloor - 1)} \\
&= \frac{\text{tr}(\boldsymbol{\Sigma}^2)}{(T - \frac{3}{2}\lfloor T/2 \rfloor)(\lfloor T/2 \rfloor - 1)} \sum_{f=2}^{\lfloor T/2 \rfloor} \sum_{g=f+\lfloor T/2 \rfloor}^T (a_{g-f}^2 + O(a_{g-f}(1-\varphi)) + O((1-\varphi)^2)) \\
&= O(nT^{-3} + n(1-\varphi)^2) \\
&= o(T^{-1} + T^2(1-\varphi)^2) = o(1), \tag{C.47}
\end{aligned}$$

$$\begin{aligned}
& \frac{\sum_{f=2}^{\lfloor T/2 \rfloor} \sum_{g=f+\lfloor T/2 \rfloor}^T \sum_{i=1}^n \sum_{j=1}^n \boldsymbol{\Sigma}_{ii} \boldsymbol{\Sigma}_{jj} \sum_{h_1=-\infty}^{\infty} \sum_{h_2=-\infty}^{\infty} \check{b}_{f,h_1} \check{b}_{g,h_1} \check{b}_{f,h_2} \check{b}_{g,h_2}}{(T - \frac{3}{2}\lfloor T/2 \rfloor)(\lfloor T/2 \rfloor - 1)} \\
&= \frac{\text{tr}(\boldsymbol{\Sigma})^2}{(T - \frac{3}{2}\lfloor T/2 \rfloor)(\lfloor T/2 \rfloor - 1)} \sum_{f=2}^{\lfloor T/2 \rfloor} \sum_{g=f+\lfloor T/2 \rfloor}^T (a_{g-f}^2 + O(a_{g-f}(1-\varphi)) + O((1-\varphi)^2)) \\
&= O(n^2T^{-3} + n^2(1-\varphi)^2) = o(n^{1/2} + nT^2(1-\varphi)^2) = o(n), \tag{C.48}
\end{aligned}$$

and

$$\begin{aligned}
& \frac{\sum_{f=2}^{\lfloor T/2 \rfloor} \sum_{g=f+\lfloor T/2 \rfloor}^T \sum_{i=1}^n \boldsymbol{\Sigma}_{ii}^2 \sum_{h=-\infty}^{\infty} \check{b}_{f,h}^2 \check{b}_{g,h}^2}{(T - \frac{3}{2}\lfloor T/2 \rfloor)(\lfloor T/2 \rfloor - 1)} \\
&\leq \frac{\sum_{i=1}^n \boldsymbol{\Sigma}_{ii}^2}{(T - \frac{3}{2}\lfloor T/2 \rfloor)(\lfloor T/2 \rfloor - 1)} \sum_{f=2}^{\lfloor T/2 \rfloor} \sum_{g=f+\lfloor T/2 \rfloor}^T \left(\sum_{h=-\infty}^f \check{b}_{f,h}^2 \right) \left(\sum_{h=-\infty}^f \check{b}_{g,h}^2 \right) \\
&= o(nT^{-1}) = o(n^{1/2}). \tag{C.49}
\end{aligned}$$

These ensure that

$$ES_{\sigma^2,0,l} = a_0^2 \text{tr}(\boldsymbol{\Sigma}^2) + o(n). \tag{C.50}$$

One can similarly obtain the order of $\text{Var}(S_{\sigma^2,0,l})$ as follows

$$\text{Var}(S_{\sigma^2,0,l}) = o(n^3T^{-2} + n) = o(n^2). \tag{C.51}$$

From (C.42), (C.50) and (C.51), we conclude that

$$\frac{S_{\sigma^2,0}}{n} = a_0^2 \frac{\text{tr}(\boldsymbol{\Sigma}^2)}{n} + o_p(1) \tag{C.52}$$

and

$$\frac{\dot{S}_{\sigma^2,0}}{n} = a_0^2 \frac{\text{tr}[\boldsymbol{\Sigma}^{1/2}(\mathbf{I} - \mathbf{H})\boldsymbol{\Sigma}(\mathbf{I} - \mathbf{H})\boldsymbol{\Sigma}^{1/2}]}{n} + o_p(1). \tag{C.53}$$

This, together with Lemma 3, implies that

$$\dot{S}_{\sigma^2,m} = \frac{|\dot{\mu}_m|(2\frac{\dot{S}_{\sigma^2,0}}{n})^{1/2}}{\sum_{i=2}^T \frac{\dot{\mathbf{x}}_{i,i}}{n(T-1)}} = (a_0 + 2 \sum_{i=1}^{\infty} a_i) \left(\frac{2}{n} \text{tr}[\boldsymbol{\Sigma}^{1/2}(\mathbf{I} - \mathbf{H})\boldsymbol{\Sigma}(\mathbf{I} - \mathbf{H})\boldsymbol{\Sigma}^{1/2}] \right)^{1/2} + o_p(1). \tag{C.54}$$

□