The Annals of Applied Probability
2015, Vol. 25, No. 6, 3624–3683
DOI: 10.1214/14-AAP1092

STRONG LIMIT OF THE EXTREME EIGENVALUES OF A SYMMETRIZED AUTO-CROSS COVARIANCE MATRIX

BY CHEN WANG*, BAISUO JIN†,[1], Z. D. BAI[2],‡, K. KRISHNAN NAIR§ AND MATTHEW HARDING¶

*National University of Singapore**, *University of Science and Technology of China*†, *Northeast Normal University*‡, *Stanford University*§ *and Duke University*¶

The auto-cross covariance matrix is defined as

$$\mathbf{M}_n = \frac{1}{2T}\sum_{j=1}^{T}(\mathbf{e}_j\mathbf{e}_{j+\tau}^* + \mathbf{e}_{j+\tau}\mathbf{e}_j^*),$$

where $\mathbf{e}_j$'s are n-dimensional vectors of independent standard complex components with a common mean 0, variance σ^2, and uniformly bounded $2+\eta$th moments and τ is the lag. Jin et al. [*Ann. Appl. Probab.* **24** (2014) 1199–1225] has proved that the LSD of $\mathbf{M}_n$ exists uniquely and nonrandomly, and independent of τ for all $\tau \ge 1$. And in addition they gave an analytic expression of the LSD. As a continuation of Jin et al. [*Ann. Appl. Probab.* **24** (2014) 1199–1225], this paper proved that under the condition of uniformly bounded fourth moments, in any closed interval outside the support of the LSD, with probability 1 there will be no eigenvalues of $\mathbf{M}_n$ for all large n. As a consequence of the main theorem, the limits of the largest and smallest eigenvalue of $\mathbf{M}_n$ are also obtained.

1. Introduction. For a $p \times p$ random Hermitian matrix $\mathbf{A}$ with eigenvalues λ_j, $j = 1, 2, \dots, p$, we define the empirical spectral distribution (ESD) of $\mathbf{A}$ by

$$F^{\mathbf{A}}(x) = \frac{1}{p}\sum_{j=1}^{p} I(\lambda_j \le x).$$

The limit distribution F of $\{F^{\mathbf{A}_n}\}$ for a given sequence of random matrices $\{\mathbf{A}_n\}$ is called the limiting spectral distribution (LSD). Let $\{\varepsilon_{it}\}$ be independent random variables with common mean 0 and variance 1. Define $\mathbf{e}_k = (\varepsilon_{1k}, \dots, \varepsilon_{nk})'$, $\boldsymbol{\gamma}_k =$

Received September 2014; revised December 2014.

[1]Supported by NSF China Young Scientist Grant 11101397.

[2]Supported by NSF China Grant 11171057, Program for Changjiang Scholars and Innovative Research Team in University, and the Fundamental Research Funds for the Central Universities.

MSC2010 subject classifications. Primary 60F15, 15A52, 62H25; secondary 60F05, 60F17.

Key words and phrases. Auto-cross covariance, dynamic factor analysis, Marčenko–Pastur law, limiting spectral distribution, order detection, random matrix theory, strong limit of extreme eigenvalues, Stieltjes transform.

$\frac{1}{\sqrt{2T}}\mathbf{e}_k$ and $\mathbf{M}_n(\tau)=\sum_{k=1}^{T}(\boldsymbol{\gamma}_k\boldsymbol{\gamma}_{k+\tau}^*+\boldsymbol{\gamma}_{k+\tau}\boldsymbol{\gamma}_k^*)$. Here, $\tau\geq 1$ is the number of lags. Under the condition of bounded $2+\eta$th moments, Jin et al. (2014) or under the weaker condition of second moments, Bai and Wang (2015) derived the LSD of $\mathbf{M}_n(\tau)$, namely, $F^{\mathbf{M}_n(\tau)} =: F_n \xrightarrow{w} F_c$ a.s. and F_c has a density function given by

$$\phi_c(x)=\frac{1}{2c\pi}\sqrt{\frac{y_0^2}{1+y_0}-\left(\frac{1-c}{|x|}+\frac{1}{\sqrt{1+y_0}}\right)^2}, \quad -d(c)\leq x\leq d(c). \tag{1.1}$$

Here, $c=\lim_{n\to\infty}c_n:=\lim_{n\to\infty}\frac{n}{T}$ and y_0 is the largest real root of the equation

$$y^3-\frac{(1-c)^2-x^2}{x^2}y^2-\frac{4}{x^2}y-\frac{4}{x^2}=0$$

and

$$d(c)=\begin{cases}\dfrac{(1-c)\sqrt{1+y_1}}{y_1-1}, & c\neq 1,\\ \lim\limits_{c\to 1}\dfrac{(1-c)\sqrt{1+y_1}}{y_1-1}=\lim\limits_{c\to 1}\sqrt{\dfrac{1+y_1}{y_1^3}}\sqrt{1+y_1}=2, & c=1,\end{cases}$$

where y_1 is a real root of the equation:

$$\big((1-c)^2-1\big)y^3+y^2+y-1=0$$

such that $y_1>1$ if $c<1$ and $y_1\in(0,1)$ if $c>1$. Further, if $c>1$, then F_c has a point mass $1-1/c$ at the origin.

The model of consideration comes from a high-dimensional dynamic k-factor model with lag q, that is, $\mathbf{R}_t=\sum_{i=0}^{q}\boldsymbol{\Lambda}_i\mathbf{F}_{t-i}+\mathbf{e}_t$, $t=1,\dots,T$. The factor $\mathbf{F}_{t-\tau}$ captures the structural part of the model at lag τ, while $\mathbf{e}_t$ corresponds to the noise component. Readers are referred to Jin et al. (2014) for more details. An interesting problem to economists is how to estimate k and q. To solve this problem, for $\tau=0,1,\dots,$ define $\Phi_n(\tau)=\frac{1}{2T}\sum_{j=1}^{T}(\mathbf{R}_j\mathbf{R}_{j+\tau}^*+\mathbf{R}_{j+\tau}\mathbf{R}_j^*)$. Note that essentially, $\mathbf{M}_n(\tau)$ and $\Phi_n(\tau)$ are symmetrized auto-cross covariance matrices at lag τ and generalize the standard sample covariance matrices $\mathbf{M}_n(0)$ and $\Phi_n(0)$, respectively. The matrix $\mathbf{M}_n(0)$ has been intensively studied in the literature and it is well known that the LSD has an MP law [Marčenko and Pastur (1967)]. Moreover, when $\tau=0$ and $\mathrm{Cov}(\mathbf{F}_t)=\Sigma_f$, the population covariance matrix of $\mathbf{R}_t$ is a *spiked population model* [Johnstone (2001), Baik and Silverstein (2006), Bai and Yao (2008)]. In fact, under certain conditions, $k(q+1)$ can be estimated by counting the number of eigenvalues of $\Phi(0)$ that are significantly larger than $(1+\sqrt{c})^2$. What remains is to separate the estimates of k and q, which can be achieved using the LSD of $\mathbf{M}_n=\mathbf{M}_n(\tau)$ for general $\tau\geq 1$. A related work has been found in Li, Wang and Yao (2014) in which the number k was detected by a different symmetrized covariance matrix for a factor model without lags. Jin et al. (2014)

has proved that the LSD of $\mathbf{M}_n$ exists uniquely and nonrandomly, and independent of τ for all $\tau \geq 1$, whose Stieltjes transform $m(z)$ satisfies the following equation:

$$\bigl(1 - c^2 m^2(z)\bigr)\bigl(c + czm(z) - 1\bigr)^2 = 1,$$

from which four roots are obtained, with y_0 defined as above:

$$m_1(z) = \frac{((1-c)/z + \sqrt{1+y_0}) + \sqrt{((1-c)/z - 1/\sqrt{1+y_0})^2 - y_0^2/(1+y_0)}}{2c},$$

$$m_2(z) = \frac{((1-c)/z + \sqrt{1+y_0}) - \sqrt{((1-c)/z - 1/\sqrt{1+y_0})^2 - y_0^2/(1+y_0)}}{2c},$$

$$m_3(z) = \frac{((1-c)/z - \sqrt{1+y_0}) + \sqrt{((1-c)/z + 1/\sqrt{1+y_0})^2 - y_0^2/(1+y_0)}}{2c},$$

$$m_4(z) = \frac{((1-c)/z - \sqrt{1+y_0}) - \sqrt{((1-c)/z + 1/\sqrt{1+y_0})^2 - y_0^2/(1+y_0)}}{2c}.$$

Here, as convention, we assume that the square root with a complex number is the one whose imaginary part is positive and the Stieltjes transform for a function of bounded variation G is defined as

$$m_G(z) = \int \frac{1}{x-z}\, dG(x) \qquad \text{for complex } \Im(z) > 0.$$

However, the number of eigenvalues of $\Phi_n(\tau)$ that lie outside the support of the LSD of $\mathbf{M}_n$ at lags $1 \leq \tau \leq q$ is different from that at lags $\tau > q$. Thus, the estimates of k and q can be separated by counting the number of eigenvalues of $\Phi_n(\tau)$ that lie outside the support of the LSD of $\mathbf{M}_n$ from $\tau = 0, 1, 2, \dots, q, q+1, \dots$.

It is worth noting that for the above method to work, one should expect no eigenvalues outside the support of the LSD of $\mathbf{M}_n$ so that if an eigenvalue of $\Phi_n(\tau)$ goes out of the support of the LSD of $\mathbf{M}_n$, it must come from the signal part. As a continuation of Jin et al. (2014), this paper establishes limits of the largest and smallest eigenvalues of $\mathbf{M}_n$, after showing that no eigenvalues exist outside the support of the LSD of $\mathbf{M}_n$, along the similar lines as in Bai and Silverstein (1998).

In Bai and Silverstein (1998), the authors considered the separation problem of the general sample covariance matrices. Later, Paul and Silverstein (2009) extended the result to a more general class of matrices taking the form of $\frac{1}{n}\mathbf{A}_n^{1/2}\mathbf{X}_n\mathbf{B}_n\mathbf{X}_n^*\mathbf{A}_n^{1/2}$ and Bai and Silverstein (2012) established the result for the information-plus-noise matrices.

Compared with Bai and Silverstein (1998), the model we considered here is more complicated and some new techniques are employed. Besides the recursive method to solve a disturbed difference equation as in Jin et al. (2014), a relationship between the convergence rates of polynomial coefficients and those of the roots is established and applied. Moreover, the results in this paper will pave the way for establishing other results such as limit theorems for sample eigenvalues of the spiked model. The main results can now be stated.

THEOREM 1.1. *Assume*:

(a) $\tau \geq 1$ *is a fixed integer.*

(b) $\mathbf{e}_k = (\varepsilon_{1k}, \dots, \varepsilon_{nk})'$, $k = 1, 2, \dots, T+\tau$, *are* n-*vectors of independent standard complex components with* $\sup_{i,t} \mathrm{E}|\varepsilon_{it}|^4 \leq M$ *for some* $M > 0$.

(c) *There exist* $K > 0$ *and a random variable* X *with finite fourth-order moment such that*, *for any* $x > 0$, *for all* n, T

$$\frac{1}{nT} \sum_{i=1}^{n} \sum_{t=1}^{T+\tau} \mathrm{P}\big(|\varepsilon_{it}| > x\big) \leq K \mathrm{P}\big(|X| > x\big). \tag{1.2}$$

(d) $\mathbf{M}_n = \sum_{k=1}^{T} (\boldsymbol{\gamma}_k \boldsymbol{\gamma}_{k+\tau}^* + \boldsymbol{\gamma}_{k+\tau} \boldsymbol{\gamma}_k^*)$, *where* $\boldsymbol{\gamma}_k = \frac{1}{\sqrt{2T}} \mathbf{e}_k$.

(e) $c_n \equiv n/T \to c \in (0,1) \cup (1, \infty)$ *as* $n \to \infty$.

(f) *The interval* $[a, b]$ *lies outside the support of* F_c.

Then P(*no eigenvalues of* $\mathbf{M}_n$ *appear in* $[a, b]$ *for all large* n) $= 1$.

By definition of $\mathbf{e}_k$ and the convergence of the largest eigenvalue of the sample covariance matrix [Yin, Bai and Krishnaiah (1988)], we have, for any $\delta > 0$ and all large n,

$$\begin{aligned} \|\mathbf{M}_n\| &\leq \frac{1}{2T} \big(\|\mathbf{E}\mathbf{E}_\tau^*\| + \|\mathbf{E}_\tau \mathbf{E}^*\| \big) \\ &\leq \frac{1}{T} s_{\max}(\mathbf{E}) s_{\max}(\mathbf{E}_\tau) = s_{\max}\bigg(\frac{\mathbf{E}}{\sqrt{T}}\bigg) s_{\max}\bigg(\frac{\mathbf{E}_\tau}{\sqrt{T}}\bigg) \\ &\leq (1+\sqrt{c})^2 + \delta \qquad \text{a.s.} \end{aligned} \tag{1.3}$$

Here, $\mathbf{E} = (\mathbf{e}_1, \dots, \mathbf{e}_T)$, $\mathbf{E}_\tau = (\mathbf{e}_{1+\tau}, \dots, \mathbf{e}_{T+\tau})$ and $s_{\max}(\mathbf{A})$ denotes the largest singular value of a matrix $\mathbf{A}$. This, together with Theorem 1.1, implies the following result.

THEOREM 1.2. *Assuming conditions* (a)–(e) *in Theorem* 1.1 *hold*, *we have*

$$\lim_{n\to\infty} \lambda_{\min}(\mathbf{M}_n) = -d(c) \qquad \text{a.s.} \quad \text{and} \quad \lim_{n\to\infty} \lambda_{\max}(\mathbf{M}_n) = d(c) \qquad \text{a.s.}$$

Here, $-d(c)$ *and* $d(c)$ *are the left and right boundary points of the support of the LSD of* $\mathbf{M}_n$, *as defined in* (1.1).

PROOF. When $c \in (0,1) \cup (1, \infty)$, let $\varepsilon > 0$ be given and consider the interval $[d(c) + \varepsilon, b]$ with $b > (1+\sqrt{c})^2 + \delta$ for some $\delta > 0$. By (1.3), with probability one, there is no eigenvalue in the interval (b, ∞). This, together with Theorem 1.1, implies that with probability one, there is no eigenvalue in the interval $[d(c) + \varepsilon, \infty)$. Therefore, we have

$$\limsup_{n\to\infty} \lambda_{\max}(\mathbf{M}_n) \le d(c) + \varepsilon \qquad \text{a.s.}$$

Next, we claim that, for all large n, there exists at least one eigenvalue in $[d(c) - \varepsilon, d(c)]$. Otherwise, we have $F_n(d(c)) - F_n(d(c) - \varepsilon) = 0$ for infinitely many n, which contradicts the fact that $F_n \to F_c$, or equivalently that $F_c(d(c)) - F_c(d(c) - \varepsilon) > 0$. Hence, our claim is proved. Therefore, we have

$$\liminf_{n\to\infty} \lambda_{\max}(\mathbf{M}_n) \ge d(c) - \varepsilon \qquad \text{a.s.}$$

Now, let $\varepsilon \to 0$, and we then have $\lim_{n\to\infty} \lambda_{\max}(\mathbf{M}_n) = d(c)$, a.s. By symmetry, $\lim_{n\to\infty} \lambda_{\min}(\mathbf{M}_n) = -d(c)$, a.s. This completes the proof of the theorem. □

One can extend Theorem 1.2 to the case $c = 1$ as follows.

THEOREM 1.3. *When $c = 1$, Theorem 1.2 still holds, that is,*

$$\lim_{n\to\infty} \lambda_{\min}(\mathbf{M}_n) = -d(1) = -2 \qquad a.s.$$

and

$$\lim_{n\to\infty} \lambda_{\max}(\mathbf{M}_n) = d(1) = 2 \qquad a.s.$$

PROOF. To prove this theorem, we need to enlarge the matrix $\mathbf{M}_n$ with a larger dimension. To this end, denote $\mathbf{M}_n = \mathbf{M}_{n,T} = \mathbf{M}_{n,T(n)}$. Fix T, we show that $\lambda_{\max}(\mathbf{M}_{n,T})$ is nondecreasing and $\lambda_{\min}(\mathbf{M}_{n,T})$ is nonincreasing in n, or more precisely, $\lambda_{\max}(\mathbf{M}_{n,T(n)}) \le \lambda_{\max}(\mathbf{M}_{n+1,T(n)})$ and $\lambda_{\min}(\mathbf{M}_{n,T(n)}) \ge \lambda_{\min}(\mathbf{M}_{n+1,T(n)})$.

To prove these relations, we will employ the interlacing theorem (Lemma 2.6) by showing that $\mathbf{M}_{n,T(n)}$ is a major sub-matrix of $\mathbf{M}_{n+1,T(n)}$. Rewrite

$$\mathbf{M}_{n,T(n)} = \sum_{k=1}^{T(n)} (\boldsymbol{\gamma}_k \boldsymbol{\gamma}_{k+\tau}^* + \boldsymbol{\gamma}_{k+\tau} \boldsymbol{\gamma}_k^*) = \sum_{k=1}^{T(n)} (\boldsymbol{\gamma}_{k,n} \boldsymbol{\gamma}_{k+\tau,n}^* + \boldsymbol{\gamma}_{k+\tau,n} \boldsymbol{\gamma}_{k,n}^*).$$

By introducing, $x_{t,n+1} = \frac{1}{\sqrt{2T(n)}} \varepsilon_{(n+1)t}$, we obtain

$$\begin{aligned}
&\mathbf{M}_{n+1,T(n)} \\
&\quad = \sum_{k=1}^{T(n)} (\boldsymbol{\gamma}_{k,n+1} \boldsymbol{\gamma}_{k+\tau,n+1}^* + \boldsymbol{\gamma}_{k+\tau,n+1} \boldsymbol{\gamma}_{k,n+1}^*) \\
&\quad = \sum_{k=1}^{T(n)} \left[\begin{pmatrix} \boldsymbol{\gamma}_{k,n} \\ x_{k,n+1} \end{pmatrix} (\boldsymbol{\gamma}_{k+\tau,n}^*, x_{k+\tau,n+1}^*) + \begin{pmatrix} \boldsymbol{\gamma}_{k+\tau,n} \\ x_{k+\tau,n+1} \end{pmatrix} (\boldsymbol{\gamma}_{k,n}^*, x_{k,n+1}^*) \right]
\end{aligned}$$

$$= \begin{pmatrix} \sum_{k=1}^{T(n)} (\boldsymbol{\gamma}_{k,n}\boldsymbol{\gamma}_{k+\tau,n}^* + \boldsymbol{\gamma}_{k+\tau,n}\boldsymbol{\gamma}_{k,n}^*) & \sum_{k=1}^{T(n)} (\boldsymbol{\gamma}_{k,n}x_{k+\tau,n+1}^* + \boldsymbol{\gamma}_{k+\tau,n}x_{k,n+1}^*) \\ \sum_{k=1}^{T(n)} (x_{k,n+1}\boldsymbol{\gamma}_{k+\tau,n}^* + x_{k+\tau,n+1}\boldsymbol{\gamma}_{k,n}^*) & \sum_{k=1}^{T(n)} (x_{k,n+1}x_{k+\tau,n+1}^* + x_{k+\tau,n+1}x_{k,n+1}^*) \end{pmatrix}$$

$$= \begin{pmatrix} \mathbf{M}_{n,T(n)} & \sum_{k=1}^{T(n)} (\boldsymbol{\gamma}_{k,n}x_{k+\tau,n+1}^* + \boldsymbol{\gamma}_{k+\tau,n}x_{k,n+1}^*) \\ \sum_{k=1}^{T(n)} (x_{k,n+1}\boldsymbol{\gamma}_{k+\tau,n}^* + x_{k+\tau,n+1}\boldsymbol{\gamma}_{k,n}^*) & \sum_{k=1}^{T(n)} (x_{k,n+1}x_{k+\tau,n+1}^* + x_{k+\tau,n+1}x_{k,n+1}^*) \end{pmatrix}.$$

By Lemma 2.6, we have $\lambda_{\max}(\mathbf{M}_{n+1,T(n)}) \geq \lambda_{\max}(\mathbf{M}_{n,T(n)})$. By symmetry, we also have $\lambda_{\min}(\mathbf{M}_{n+1,T(n)}) \leq \lambda_{\min}(\mathbf{M}_{n,T(n)})$. This together with Theorem 1.2 implies that for any $\varepsilon > 0$, we have a.s.

$$\limsup_{\substack{n\to\infty \\ n/T(n)\to 1}} \lambda_{\max}(\mathbf{M}_{n,T(n)}) \leq \lim_{\substack{n\to\infty \\ n/T(n)\to 1}} \lambda_{\max}(\mathbf{M}_{[(1+\varepsilon)n],T(n)}) = d(1+\varepsilon).$$

Note that $d(c)$ is continuous in c. By letting $\varepsilon \to 0$, we have a.s.

$$\limsup_{\substack{n\to\infty \\ n/T(n)\to 1}} \lambda_{\max}(\mathbf{M}_{n,T(n)}) \leq d(1) = 2.$$

Since the LSD of $\mathbf{M}_n$ exists with right support boundary $d(1) = 2$, we have proved that

$$\lim_{\substack{n\to\infty \\ n/T(n)\to 1}} \lambda_{\max}(\mathbf{M}_{n,T(n)}) = 2.$$

By symmetry, we have a.s. $\lim_{n\to\infty, n/T(n)\to 1} \lambda_{\min}(\mathbf{M}_{n,T(n)}) = -d(1) = -2$. The proof of the theorem is complete. □

As an immediate consequence of Theorem 1.3, Corollary 1.1 complements Theorem 1.1 for $c = 1$.

COROLLARY 1.1. *Theorem 1.1 still holds when $c = 1$.*

Figures 1 and 2 display the density functions $\phi_c(x)$ and the distributions of sample eigenvalues with $\tau = 1$, $c = 0.2$ ($n = 200$, $T = 1000$) and $c = 2.5$ ($n = 2500$, $T = 1000$), respectively.

We will now focus on proving Theorem 1.1. As in Jin et al. (2014), we denote the Stieltjes transform of $\mathbf{M}_n$ as $m_n(z) = \frac{1}{n}\operatorname{tr}(\mathbf{M}_n - z\mathbf{I}_n)^{-1}$ where, and throughout the paper, $z = u + iv_n$, $v_n > 0$, and let $m_n^0(z)$ be the Stieltjes transform of ϕ_{c_n} with limiting ratio of $c_n = n/T$. Using the truncation technique employed in Section 3

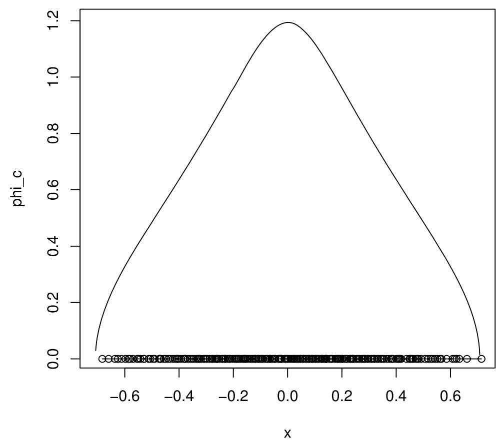

FIG. 1. *Density function $\phi_c(x)$ of F_c and distribution of sample eigenvalues with $\tau = 1, c = 0.2$* ($n = 200, T = 1000$).

of Bai and Silverstein (1998), we further assume that the ε_{ij}'s satisfy the conditions that

$$|\varepsilon_{ij}| \le C, \qquad \mathrm{E}\varepsilon_{ij} = 0, \qquad \mathrm{E}|\varepsilon_{ij}|^2 = 1, \qquad \mathrm{E}|\varepsilon_{ij}|^4 < M \tag{1.4}$$

for some $C, M > 0$. More detailed justifications are provided in the Appendix.

The rest of the paper is structured as follows. Section 2 contains some lemmas of known results. Section 3 provides some technical lemmas. Convergence rates of $\|F_n - F_{c_n}\|$ and $m_n(z) - m_n^0(z)$ are obtained in Sections 4 and 5, respectively. Section 6 concludes the proof of Theorem 1.1. Justifications of variable truncation,

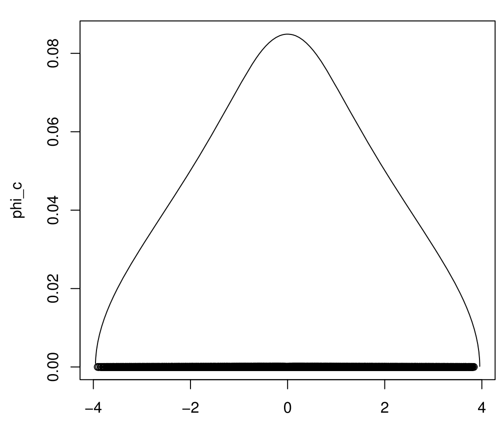

FIG. 2. *Density function $\phi_c(x)$ of F_c and distribution of sample eigenvalues with $\tau = 1, c = 2.5$* ($n = 2500, T = 1000$). *Note that the area under the density function curve is $1/c$.*

centralization and rescaling and proofs of lemmas presented in Section 3 are given in the Appendix.

2. Mathematical tools. In this section, we provide some known results.

LEMMA 2.1 [Burkholder (1973)]. *Let* $\{X_k\}$ *be a complex martingale difference sequence with respect to the increasing* σ*-fields* $\{\mathcal{F}_n\}$*. Then, for* $p \geq 2$, *we have*

$$\mathrm{E}\Big|\sum X_k\Big|^p \leq K_p\Big(\mathrm{E}\Big(\sum \mathrm{E}(|X_k|^2|\mathcal{F}_{k-1})\Big)^{p/2} + \mathrm{E}\sum |X_k|^p\Big).$$

LEMMA 2.2 [Burkholder (1973)]. *Let* $\{X_k\}$ *be as above. Then, for* $p \geq 2$, *we have*

$$\mathrm{E}\Big|\sum X_k\Big|^p \leq K_p \mathrm{E}\Big(\sum |X_k|^2\Big)^{p/2}.$$

LEMMA 2.3 [Theorem A.43 of Bai and Silverstein (2010)]. *Let* $\mathbf{A}$ *and* $\mathbf{B}$ *be two* $n \times n$ *Hermitian matrices. Then*

$$\|F^{\mathbf{A}} - F^{\mathbf{B}}\| \leq \frac{1}{n}\operatorname{rank}(\mathbf{A} - \mathbf{B}),$$

where $F^{\mathbf{A}}$ *is the empirical spectral distribution of* $\mathbf{A}$ *and* $\|f\| = \sup_x |f(x)|$.

LEMMA 2.4 [Bai (1993) or Corollary B.15 of Bai and Silverstein (2010)]. *Let* F *be a distribution function and let* G *be a function of bounded variation satisfying* $\int |F(x) - G(x)|\,dx < \infty$. *Denote their Stieltjes transforms by* $f(z)$ *and* $g(z)$, *respectively. Assume that for some constant* $B > 0$, $F([-B, B]) = 1$ *and* $|G|((-\infty, -B)) = |G|((B, \infty)) = 0$, *where* $|G|((a, b))$ *denotes the total variation of the signed measure* G *on the interval* (a, b). *Then we have*

$$\begin{aligned}\|F - G\| &:= \sup_x \big|F(x) - G(x)\big| \\ &\leq \frac{1}{\pi(1-\kappa)(2\gamma - 1)} \\ &\quad \times \bigg[\int_{-A}^{A} \big|f(z) - g(z)\big|\,du + v^{-1}\sup_x \int_{|y| \leq 2va} \big|G(x+y) - G(x)\big|\,dy\bigg],\end{aligned}$$

where $z = u + iv$, $v > 0$, a *and* γ *are positive constants such that* $\gamma = \frac{1}{\pi}\int_{|u|<a} \frac{1}{u^2+1}\,du > \frac{1}{2}$. A *is a positive constant such that* $A > B$ *and* $\kappa = \frac{4B}{\pi(A-B)(2\gamma-1)} < 1$.

LEMMA 2.5 [Lemma B.26 of Bai and Silverstein (2010)]. *Let* $\mathbf{A} = (a_{ij})$ *be an* $n \times n$ *nonrandom matrix and* $\mathbf{X} = (x_1, \dots, x_n)'$ *be a random vector of independent entries. Assume that* $\mathrm{E}x_i = 0$, $\mathrm{E}|x_i|^2 = 1$, *and* $\mathrm{E}|x_j|^\ell \le v_\ell$. *Then, for any* $p \ge 1$,

$$\mathrm{E}\left|\mathbf{X}^*\mathbf{A}\mathbf{X} - \operatorname{tr}\mathbf{A}\right|^p \le C_p\left(\left(v_4 \operatorname{tr}(\mathbf{A}\mathbf{A}^*)\right)^{p/2} + v_{2p}\operatorname{tr}(\mathbf{A}\mathbf{A}^*)^{p/2}\right),$$

where C_p *is a constant depending on* p *only.*

LEMMA 2.6 [The interlacing theorem, Rao and Rao (1998)]. *If* $\mathbf{C}$ *is an* $(n-1) \times (n-1)$ *major sub-matrix of the* $n \times n$ *Hermitian matrix* $\mathbf{A}$, *then* $\lambda_1(\mathbf{A}) \ge \lambda_1(\mathbf{C}) \ge \lambda_2(\mathbf{A}) \ge \cdots \ge \lambda_{n-1}(\mathbf{C}) \ge \lambda_n(\mathbf{A})$. *Here* $\lambda_i(\mathbf{A})$ *denotes the* i*th largest eigenvalue of the Hermitian matrix* $\mathbf{A}$.

3. Some technical lemmas. Before proceeding, some technical lemmas are presented with proofs postponed in the Appendix. The first three are about the convergence rates of roots of a polynomial.

LEMMA 3.1. *Let* $\{r_n\}$ *be a sequence of positive real numbers converging to* 0 *and* m *be a fixed positive integer, independent of* n. *Let* $B(x_0, r_n)$ *denote the open ball centered at* x_0 *with radius* r_n. *Given* m *points* $x_1, \dots, x_m$ *in* $B(x_0, r_n)$, *one can find* $x \in B(x_0, r_n)$ *and* $d > 0$ *such that* $\min_{i \in \{1,\dots,m\}} |x - x_i| \ge d r_n$.

LEMMA 3.2. *For each* $n \in \mathbb{N}$, *let* $P_n(x) = x^k + a_{n,k-1}x^{k-1} + \cdots + a_{n,1}x + a_{n,0}$ *be a polynomial of degree* k, *with roots* $x_{n1}, \dots, x_{nk}$. *Moreover, for* $i = 0, 1, \dots, k-1$, $\lim_{n\to\infty} a_{n,i} = a_i$. *Let* $P(x) = x^k + a_{k-1}x^{k-1} + \cdots + a_1 x + a_0$. *Suppose* $P(x)$ *has distinct roots* $x_1, \dots, x_m$, *and each* x_j *has multiplicity* ℓ_j *with* $\sum_{j=1}^m \ell_j = k$. *Then for* n *large enough, for each* $j \in \{1, \dots, m\}$, *there are exactly* ℓ_j x_{ni}*'s in* $B(x_j, r_n^{1/\ell_j})$, *where* $r_n = \max_{i \in \{0,1,\dots,k-1\}} |a_{n,i} - a_i|$.

LEMMA 3.3. *For each* $n \in \mathbb{N}$, *let* $P_n(x) = x^k + a_{n,k-1}x^{k-1} + \cdots + a_{n,1}x + a_{n,0}$ *and* $Q_n(y) = y^k + b_{n,k-1}y^{k-1} + \cdots + b_{n,1}y + b_{n,0}$ *be two polynomials of degree* k, *with roots* $x_{n1}, \dots, x_{nk}$ *and* $y_{n1}, \dots, y_{nk}$, *respectively. Moreover, for* $i = 0, 1, \dots, k-1$, $\lim_{n\to\infty} b_{n,i} = \lim_{n\to\infty} a_{n,i} = a_i$. *Let* $P(x) = x^k + a_{k-1}x^{k-1} + \cdots + a_1 x + a_0$. *Suppose* $P(x)$ *has distinct roots* $x_1, \dots, x_m$, *and each* x_j *has the multiplicity* ℓ_j *with* $\sum_{j=1}^m \ell_j = k$. *Then for* n *large enough, for each* $j \in \{1, \dots, m\}$, *for any* $x_{ni} \in B(x_j, r_n^{1/\ell_j})$, *there exists at least one* y_{nl} *such that* $|x_{ni} - y_{nl}| \le d\widetilde{r}_n^{1/\ell_j}$ *for some* $d > 0$. *Here,* $r_n = \max_{i \in \{0,1,\dots,k-1\}} |a_{n,i} - a_i|$ *and* $\widetilde{r}_n = \max_{i \in \{0,1,\dots,k-1\}} |a_{n,i} - b_{n,i}|$.

To establish the following lemmas, we need some notation: let $z = u + iv_n$, where $u \in [-A, A]$ and $v_n \in [n^{-1/52}, n^{-1/212}]$ and $A > 0$ is a large constant. Define

$$\mathbf{A} = \mathbf{M}_n - z\mathbf{I}_n,$$
$$\mathbf{A}_k = \mathbf{M}_{n,k} - z\mathbf{I}_n = \mathbf{A} - \boldsymbol{\gamma}_k(\boldsymbol{\gamma}_{k-\tau} + \boldsymbol{\gamma}_{k+\tau})^* - (\boldsymbol{\gamma}_{k-\tau} + \boldsymbol{\gamma}_{k+\tau})\boldsymbol{\gamma}_k^*,$$
$$\vdots$$
$$\mathbf{A}_{k,\dots,k+s\tau} = \mathbf{A} - \sum_{t=0}^{s}[\boldsymbol{\gamma}_{k+t\tau}(\boldsymbol{\gamma}_{k+(t-1)\tau} + \boldsymbol{\gamma}_{k+(t+1)\tau})^* + (\boldsymbol{\gamma}_{k+(t-1)\tau} + \boldsymbol{\gamma}_{k+(t+1)\tau})\boldsymbol{\gamma}_{k+t\tau}^*],$$

with the convention that $\boldsymbol{\gamma}_l = 0$ for $l \le 0$ or $l > T + \tau$.

The following lemma will be frequently used.

LEMMA 3.4. *Let r, s be fixed positive integers. For $l \neq k$, we have*

$$\mathrm{E}|\boldsymbol{\gamma}_l^* \mathbf{A}_k^{-s} \boldsymbol{\gamma}_k|^{2r} \le \frac{K}{T^r v_n^{2rs}}$$

for some $K > 0$.

Define $a_n = \frac{c_n \mathrm{E} m_n}{2}$ and let x_{n1}, x_{n0} be two roots of the equation $x^2 = x - a_n^2$ with $|x_{n1}| > |x_{n0}|$. Some properties regarding x_{n1} and x_{n0} are stated in the next lemma.

In the following, if a lemma contains two sets of results simultaneously, then the results labelled by "a" hold for all $z = u + iv_n$, and u lies in a bounded interval $[-A, A] \subseteq \mathbb{R}$, whereas results labelled by "b" hold for all $z = u + iv_n$ with $u \in [a, b]$ and are obtained under the additional condition that $\mathrm{P}(\|F_n - F_{c_n}\| \ge n^{-1/104}) = o(n^{-t})$ for any fixed $t > 0$, where $[a, b]$ is defined in Theorem 1.1. Results "a" will be used to establish a preliminary convergence rate of the ESD of $\mathbf{M}_n$ in Section 4 and the results "b" will be applied to the refinement of the convergence rate when $u \in [a, b]$ in Section 5. If a lemma contains only one set of results, the results will be established for all $u \in [a, b]$ and under the additional assumption that $\mathrm{P}(\|F_n - F_{c_n}\| \ge n^{-1/104}) = o(n^{-t})$.

LEMMA 3.5. *When $u \in [a, b]$, let λ_{kj} denote the jth largest eigenvalue of $\mathbf{M}_n - \boldsymbol{\gamma}_k(\boldsymbol{\gamma}_{k+\tau} + \boldsymbol{\gamma}_{k-\tau})^* - (\boldsymbol{\gamma}_{k+\tau} + \boldsymbol{\gamma}_{k-\tau})\boldsymbol{\gamma}_k^*$, for $\Im(z) \ge n^{-\delta}$ with $\delta = 1/106$, we have, for any $t > 0$*

$$\mathrm{P}\biggl(\frac{1}{2T}\sum \frac{1}{|\lambda_{kj} - z|^2} > K\biggr) = o(n^{-t})$$

for some $K > 0$.

REMARK 3.1. When $u \in [a,b]$, with similar proofs, for $\Im(z) \geq n^{-\delta}$ with $\delta = 1/53$, we have, for any $t > 0$,

$$\mathrm{P}\left(\frac{1}{2T}\left|\operatorname{tr}\mathbf{A}_k^{-1}\right| > K\right) \leq \mathrm{P}\left(\frac{1}{2T}\sum\frac{1}{|\lambda_{kj} - z|} > K\right) = o(n^{-t})$$

and when $\Im(z) \geq n^{-\delta}$ with $\delta = 1/212$,

$$\mathrm{P}\left(\frac{1}{2T}\left|\operatorname{tr}\mathbf{A}_k^{-4}\right| > K\right) \leq \mathrm{P}\left(\frac{1}{2T}\sum\frac{1}{|\lambda_{kj} - z|^4} > K\right) = o(n^{-t})$$

for some $K > 0$.

REMARK 3.2. When $u \in [a,b]$, and λ_{kj}'s are eigenvalues of $\mathbf{M}_{n,k} = \mathbf{M}_n - \boldsymbol{\gamma}_k(\boldsymbol{\gamma}_{k+\tau} + \boldsymbol{\gamma}_{k-\tau})^* - (\boldsymbol{\gamma}_{k+\tau} + \boldsymbol{\gamma}_{k-\tau})\boldsymbol{\gamma}_k^*$, for $\Im(z) \geq n^{-\delta}$ with $\delta = 1/212$, with a similar proof, we have

$$\mathrm{P}\left(\frac{1}{2T}\sum\frac{1}{|\lambda_{kj} - z|^2} > K\right) = o(n^{-t})$$

for some $K > 0$.

LEMMA 3.6. *With x_{n1} and x_{n0} defined as above, for any $v_n \geq n^{-1/52}$, we have*:

(i) *There exists some $\eta > 0$ such that for all large n*:

  (a) $\sup_{u\in[-A,A],\Im(z)=v_n} |\frac{x_{n0}(z)}{x_{n1}(z)}| < 1 - \eta v_n^3$.

  (b) $\sup_{u\in[a,b],\Im(z)=v_n} |\frac{x_{n0}(z)}{x_{n1}(z)}| < 1 - \eta$.

(ii)

  (a) *When $u \in [-A,A]$, we have $|x_{n1}| \geq \frac{1}{2}$ and $|x_{n1}| \leq K v_n^{-1}$ for some constant K.*

  (b) *When $u \in [a,b]$, we have $|x_{n1}| \geq \frac{1}{2}$ and $|x_{n1}| \leq K$ for some constant K.*

(iii)

  (a) *When $u \in [-A,A]$, we have $|x_{n1} - x_{n0}| \geq \eta v_n$ for some constant $\eta > 0$.*

  (b) *When $u \in [a,b]$, we have $|x_{n1} - x_{n0}| \geq \eta$ for some constant $\eta > 0$.*

(iv)

  (a) *When $u \in [-A,A]$, we have $\frac{|x_{n1}|}{|x_{n1}-x_{n0}|} \leq K v_n^{-1}$ for some constant K.*

  (b) *When $u \in [a,b]$, we have $\frac{|x_{n1}|}{|x_{n1}-x_{n0}|} \leq K$ for some constant K.*

(v) *When $u \in [a,b]$, we have $|a_n| < \frac{1}{2} - \eta$ for some constant $\eta > 0$.*

LEMMA 3.7. *For any* $v_n \geq n^{-1/52}$ *and* $t > 0$:

(a) *for any* $u \in [-A, A]$ *and* $k \leq T - v_n^{-4}$, *we have*

$$\mathrm{P}\left(\left|\boldsymbol{\gamma}_{k+\tau}^* \mathbf{A}_k^{-1} \boldsymbol{\gamma}_{k+\tau} - \frac{c_n \mathrm{E} m_n}{2x_{n1}}\right| \geq v_n^6\right) = o(n^{-t})$$

and for any $k \geq v_n^{-4}$,

$$\mathrm{P}\left(\left|\boldsymbol{\gamma}_{k-\tau}^* \mathbf{A}_k^{-1} \boldsymbol{\gamma}_{k-\tau} - \frac{c_n \mathrm{E} m_n}{2x_{n1}}\right| \geq v_n^6\right) = o(n^{-t}),$$

(b1) *for any* $u \in [a, b]$, *there is a constant* $\eta \in (0, \frac{1}{2})$ *such that* $\mathrm{P}(|\boldsymbol{\gamma}_{k+\tau}^* \mathbf{A}_k^{-1} \times \boldsymbol{\gamma}_{k+\tau}| \geq 1 - \eta) = o(n^{-t})$,

(b2) *for any* $u \in [a, b]$, *when* $k \leq T - \log^2 n$, *we have* $|\mathrm{E}\boldsymbol{\gamma}_{k+\tau}^* \mathbf{A}_k^{-1} \boldsymbol{\gamma}_{k+\tau} - \frac{a_n}{x_{n1}}| = o(1/(nv_n))$, *and when* $k \geq \log^2 n$, *we have* $|\mathrm{E}\boldsymbol{\gamma}_{k-\tau}^* \mathbf{A}_k^{-1} \boldsymbol{\gamma}_{k-\tau} - \frac{a_n}{x_{n1}}| = o(1/(nv_n))$,

(b3) *for any* $u \in [a, b]$, *when* $k \leq T - \log^2 n$, *we have* $\mathrm{E}|\boldsymbol{\gamma}_{k+\tau}^* \mathbf{A}_k^{-1} \boldsymbol{\gamma}_{k+\tau} - \frac{a_n}{x_{n1}}|^2 = o(1/(nv_n))$, *and when* $k \geq \log^2 n$, *we have* $\mathrm{E}|\boldsymbol{\gamma}_{k-\tau}^* \mathbf{A}_k^{-1} \boldsymbol{\gamma}_{k-\tau} - \frac{a_n}{x_{n1}}|^2 = o(1/(nv_n))$.

LEMMA 3.8. *For any* $v_n \geq n^{-1/52}$ *and* $t > 0$:

(a) *for any* $u \in [-A, A]$, *we have*

$$\mathrm{P}(|\boldsymbol{\gamma}_{k-\tau}^* \mathbf{A}_k^{-1} \boldsymbol{\gamma}_{k+\tau}| > v_n^6) = o(n^{-t});$$

(b1) *for any* $u \in [a, b]$, *we have* $|\mathrm{E}\boldsymbol{\gamma}_{k-\tau}^* \mathbf{A}_k^{-1} \boldsymbol{\gamma}_{k+\tau}| = o(1/(nv_n))$;

(b2) *for any* $u \in [a, b]$, *we have* $\mathrm{E}|\boldsymbol{\gamma}_{k-\tau}^* \mathbf{A}_k^{-1} \boldsymbol{\gamma}_{k+\tau}|^2 = o(1/(nv_n))$.

LEMMA 3.9. *For any* $v_n \geq n^{-1/212}$, $u \in [a, b]$ *and* $t > 0$, *there exists a constant* $K > 0$ *such that*

$$\mathrm{P}(|\boldsymbol{\gamma}_{k+\tau}^* \mathbf{A}_k^{-1} (\mathbf{A}_k^*)^{-1} \boldsymbol{\gamma}_{k+\tau}| \geq K) = o(n^{-t}).$$

LEMMA 3.10. *For any* $v_n \geq n^{-1/212}$, $u \in [a, b]$ *and* $t > 0$, *we have*

$$\mathrm{P}(|\boldsymbol{\gamma}_{k+\tau}^* \mathbf{A}_k^{-2} (\mathbf{A}_k^*)^{-2} \boldsymbol{\gamma}_{k+\tau}| \geq K) = o(n^{-t})$$

for some $K > 0$.

LEMMA 3.11. *Let* $u \in [a, b]$, *then for any* $v_n \geq n^{-1/212}$, *we have*

$$|\mathrm{E}\,\mathrm{tr}\,\mathbf{A}^{-1} - \mathrm{E}\,\mathrm{tr}\,\mathbf{A}_k^{-1}| = O(1) \quad \textit{and}$$
$$|\mathrm{E}\,\mathrm{tr}\,\mathbf{A}_{k,\dots,k+(s-1)\tau}^{-1} - \mathrm{E}\,\mathrm{tr}\,\mathbf{A}_{k,\dots,k+s\tau}^{-1}| = O(1).$$

4. A convergence rate of the empirical spectral distribution. In this section, we give a convergence rate of $\|F_n - F_{c_n}\|$.

4.1. *A preliminary convergence rate of* $m_n(z) - \mathrm{E}m_n(z)$. Let E_k denote the conditional expectation given $\boldsymbol{\gamma}_{k+1}, \dots, \boldsymbol{\gamma}_{T+\tau}$. With this notation, we have $m_n(z) = \mathrm{E}_0(m_n(z))$ and $\mathrm{E}m_n(z) = \mathrm{E}_T(m_n(z))$. Therefore, we obtain

$$\begin{aligned} m_n(z) - \mathrm{E}m_n(z) &= \sum_{k=1}^{T+\tau} \bigl(\mathrm{E}_{k-1} m_n(z) - \mathrm{E}_k m_n(z)\bigr) \\ &= \sum_{k=1}^{T+\tau} \frac{1}{n} (\mathrm{E}_{k-1} - \mathrm{E}_k)\bigl(\operatorname{tr} A^{-1} - \operatorname{tr} A_k^{-1}\bigr) \\ &\equiv \sum_{k=1}^{T+\tau} \frac{1}{n} (\mathrm{E}_{k-1} - \mathrm{E}_k)\alpha_k. \end{aligned}$$

Write

$$\begin{aligned} \mathbf{M}_n &= \mathbf{M}_{n,k} + (\boldsymbol{\gamma}_{k+\tau}, \boldsymbol{\gamma}_k, \boldsymbol{\gamma}_{k-\tau}) \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} \boldsymbol{\gamma}_{k+\tau}^* \\ \boldsymbol{\gamma}_k^* \\ \boldsymbol{\gamma}_{k-\tau}^* \end{pmatrix} \\ &\equiv \mathbf{M}_{n,k} + \mathbf{C}_k. \end{aligned}$$

Let $\lambda_i(\mathbf{B})$ denote the ith smallest eigenvalue for a Hermitian matrix $\mathbf{B}$. Then, for any $i > 3$, we have

$$\begin{aligned} \lambda_i(\mathbf{M}_n) &= \sup_{\boldsymbol{\alpha}_1, \dots, \boldsymbol{\alpha}_{i-1}} \inf_{\substack{\boldsymbol{\beta} \perp \boldsymbol{\alpha}_1, \dots, \boldsymbol{\alpha}_{i-1} \\ \|\boldsymbol{\beta}\|=1}} \bigl(\boldsymbol{\beta}^* \mathbf{M}_{n,k} \boldsymbol{\beta} + \boldsymbol{\beta}^* \mathbf{C}_k \boldsymbol{\beta}\bigr) \\ &\ge \sup_{\boldsymbol{\alpha}_1, \dots, \boldsymbol{\alpha}_{i-4}} \inf_{\substack{\boldsymbol{\beta} \perp \boldsymbol{\alpha}_1, \dots, \boldsymbol{\alpha}_{i-4}, \boldsymbol{\gamma}_{k+\tau}, \boldsymbol{\gamma}_k, \boldsymbol{\gamma}_{k-\tau} \\ \|\boldsymbol{\beta}\|=1}} \boldsymbol{\beta}^* \mathbf{M}_{n,k} \boldsymbol{\beta} \\ &\ge \sup_{\boldsymbol{\alpha}_1, \dots, \boldsymbol{\alpha}_{i-4}} \inf_{\substack{\boldsymbol{\beta} \perp \boldsymbol{\alpha}_1, \dots, \boldsymbol{\alpha}_{i-4} \\ \|\boldsymbol{\beta}\|=1}} \boldsymbol{\beta}^* \mathbf{M}_{n,k} \boldsymbol{\beta} \\ &= \lambda_{i-3}(\mathbf{M}_{n,k}). \end{aligned} \tag{4.1}$$

Similarly, we have $\lambda_i(\mathbf{M}_n) \le \lambda_{i+3}(\mathbf{M}_{n,k})$. Therefore, with

$$G(x) := \sum_{i=1}^n I_{\{\lambda_i(\mathbf{M}_n) \le x\}} \quad \text{and} \quad G_k(x) := \sum_{i=1}^n I_{\{\lambda_i(\mathbf{M}_{n,k}) \le x\}},$$

we have

$$\begin{aligned} |\alpha_k| &= \bigl|\operatorname{tr} \mathbf{A}^{-1} - \operatorname{tr} \mathbf{A}_k^{-1}\bigr| \\ &= \left| \int \frac{1}{x - z}\, d\bigl(G(x) - G_k(x)\bigr) \right| \\ &\le \int \frac{|G(x) - G_k(x)|}{|x - z|^2}\, dx \end{aligned} \tag{4.2}$$

$$\le 3\int \frac{1}{(x-u)^2+v_n^2}\,dx$$
$$\le \frac{3\pi}{v_n}.$$

Here, the third equality follows from integration by parts. Therefore, by Lemma 2.2,

$$\begin{aligned}
\mathrm{P}\big(|m_n(z)-\mathrm{E}m_n(z)|>v_n\big) &= \mathrm{P}\left(\left|\sum_{k=1}^{T+\tau}(\mathrm{E}_{k-1}-\mathrm{E}_k)\alpha_k\right|>nv_n\right)\\
&\le \mathrm{E}\left(\frac{1}{(nv_n)^p}\left|\sum_{k=1}^{T+\tau}(\mathrm{E}_{k-1}-\mathrm{E}_k)\alpha_k\right|^p\right)\\
&\le \frac{K}{(nv_n)^p}\mathrm{E}\left(\sum_{k=1}^{T+\tau}\left|(\mathrm{E}_{k-1}-\mathrm{E}_k)\alpha_k\right|^2\right)^{p/2}\\
&\le Kn^{-p/2}v_n^{-2p}.
\end{aligned} \tag{4.3}$$

Hence, when $v_n \ge n^{-\alpha}$ for some $0<\alpha<\frac{1}{4}$, we can choose $p>1$ such that $p(\frac{1}{2}-2\alpha)>t$, and thus

$$\mathrm{P}\big(|m_n(z)-\mathrm{E}m_n(z)|>v_n\big)=o\big(n^{-t}\big), \tag{4.4}$$

for any fixed $t>0$. This implies $|m_n(z)-\mathrm{E}m_n(z)|=o(v_n)$, a.s.

4.2. *A preliminary convergence rate of* $\mathrm{E}m_n(z)-m_n^0(z)$. Next, we want to show that when $v_n \ge n^{-1/52}$,

$$\big|\mathrm{E}m_n(z)-m_n^0(z)\big|=o(v_n). \tag{4.5}$$

By

$$\mathbf{A}=\sum_{k=1}^{T}(\boldsymbol{\gamma}_k\boldsymbol{\gamma}_{k+\tau}^*+\boldsymbol{\gamma}_{k+\tau}\boldsymbol{\gamma}_k^*)-z\mathbf{I}_n$$

we have

$$\mathbf{I}_n=\sum_{k=1}^{T}(\mathbf{A}^{-1}\boldsymbol{\gamma}_k\boldsymbol{\gamma}_{k+\tau}^*+\mathbf{A}^{-1}\boldsymbol{\gamma}_{k+\tau}\boldsymbol{\gamma}_k^*)-z\mathbf{A}^{-1}.$$

Taking trace and dividing by n, we obtain

$$1+zm_n(z)=\frac{1}{n}\sum_{k=1}^{T}(\boldsymbol{\gamma}_{k+\tau}^*\mathbf{A}^{-1}\boldsymbol{\gamma}_k+\boldsymbol{\gamma}_k^*\mathbf{A}^{-1}\boldsymbol{\gamma}_{k+\tau}).$$

Taking expectation on both sides, we obtain

$$1 + z\mathrm{E}m_n(z) = \frac{1}{n}\sum_{k=1}^{T}\mathrm{E}\boldsymbol{\gamma}_k^*\mathbf{A}^{-1}(\boldsymbol{\gamma}_{k+\tau}+\boldsymbol{\gamma}_{k-\tau}),$$

or equivalently, by noticing $1-\frac{c_n^2}{2x_{n1}}\mathrm{E}^2m_n(z) = x_{n1}-x_{n0}$,

$$\begin{aligned}
&c_n + c_n z\mathrm{E}m_n(z)\\
&\quad= \frac{1}{T}\sum_{k=1}^{T}\mathrm{E}\boldsymbol{\gamma}_k^*\mathbf{A}^{-1}(\boldsymbol{\gamma}_{k+\tau}+\boldsymbol{\gamma}_{k-\tau})\\
&\quad= \frac{1}{T}\sum_{k=1}^{T}\biggl[1-\mathrm{E}\frac{1}{1+\boldsymbol{\gamma}_k^*\tilde{\mathbf{A}}_k^{-1}(\boldsymbol{\gamma}_{k+\tau}+\boldsymbol{\gamma}_{k-\tau})}\biggr]\\
&\quad= \frac{1}{T}\sum_{k=1}^{T}\biggl[1-\mathrm{E}\biggl(1\Big/\biggl(1+\boldsymbol{\gamma}_k^*\mathbf{A}_k^{-1}(\boldsymbol{\gamma}_{k+\tau}+\boldsymbol{\gamma}_{k-\tau})\\
&\qquad\qquad -\frac{\boldsymbol{\gamma}_k^*\mathbf{A}_k^{-1}\boldsymbol{\gamma}_k(\boldsymbol{\gamma}_{k+\tau}^*+\boldsymbol{\gamma}_{k-\tau}^*)\mathbf{A}_k^{-1}(\boldsymbol{\gamma}_{k+\tau}+\boldsymbol{\gamma}_{k-\tau})}{1+(\boldsymbol{\gamma}_{k+\tau}^*+\boldsymbol{\gamma}_{k-\tau}^*)\mathbf{A}_k^{-1}\boldsymbol{\gamma}_k}\biggr)\biggr)\biggr]\\
&\quad= 1-\frac{1}{1-(c_n^2/(2x_{n1}))\mathrm{E}^2m_n(z)}+\delta_n,
\end{aligned}\tag{4.6}$$

where

$$\begin{aligned}
\tilde{\mathbf{A}}_k &= \mathbf{A}-(\boldsymbol{\gamma}_{k+\tau}+\boldsymbol{\gamma}_{k-\tau})\boldsymbol{\gamma}_k^* = \mathbf{A}_k+\boldsymbol{\gamma}_k(\boldsymbol{\gamma}_{k+\tau}^*+\boldsymbol{\gamma}_{k-\tau}^*),\\
\delta_n &= -\frac{1}{T}\sum_{k=1}^{T}\biggl(\mathrm{E}\biggl(1\Big/\biggl(1+\boldsymbol{\gamma}_k^*\mathbf{A}_k^{-1}(\boldsymbol{\gamma}_{k+\tau}+\boldsymbol{\gamma}_{k-\tau})\\
&\qquad -\frac{\boldsymbol{\gamma}_k^*\mathbf{A}_k^{-1}\boldsymbol{\gamma}_k(\boldsymbol{\gamma}_{k+\tau}^*+\boldsymbol{\gamma}_{k-\tau}^*)\mathbf{A}_k^{-1}(\boldsymbol{\gamma}_{k+\tau}+\boldsymbol{\gamma}_{k-\tau})}{1+(\boldsymbol{\gamma}_{k+\tau}^*+\boldsymbol{\gamma}_{k-\tau}^*)\mathbf{A}_k^{-1}\boldsymbol{\gamma}_k}\biggr)\biggr)\\
&\qquad -\frac{1}{x_{n1}-x_{n0}}\biggr),
\end{aligned}$$

x_{n1}, x_{n0} are the roots of the equation $x^2 = x - a_n^2$ with $|x_{n1}| > |x_{n0}|$, and $a_n = \frac{c_n\mathrm{E}m_n}{2}$, as defined below the statement of Lemma 3.4. Substituting the expression of x_{n1}, we have

$$(1-c_n^2(\mathrm{E}m_n(z))^2)(c_n+c_nz\mathrm{E}m_n(z)-1-\delta_n)^2 = 1.\tag{4.7}$$

Meanwhile, by (3.8) of Jin et al. (2014), we have

$$(1-c^2m^2(z))(c+czm(z)-1)^2 = 1.\tag{4.8}$$

Similarly, $m_n^0(z)$ satisfies

$$(1 - c_n^2 (m_n^0(z))^2)(c_n + c_n z m_n^0(z) - 1)^2 = 1. \tag{4.9}$$

We can regard the three expressions above as polynomials of $\mathrm{E}m_n(u + iv_n)$, $m(u)$ and $m_n^0(u + iv_n)$, respectively. Compared with (4.8), coefficients in (4.7) and (4.9) are different in terms of δ_n and c_n.

4.2.1. *Identification of the solution to equation* (4.8). In this subsection, we show that for $c \neq 1$ and every $A > 0$, there is a constant $\eta > 0$ such that for every z with $\Im(z) \in (0, \eta)$ and $|\Re(z)| \leq A$, equation (4.8)

$$(1 - c^2 m^2(z))(1 - c - czm(z))^2 = 1$$

has only one solution satisfying $\Im(m(z)) > \eta v$ and the other three satisfying $\Im(m(z)) < -\eta v$ when $c < 1$; and one satisfying $\Im(m(z) + \frac{c-1}{cz}) > \eta v$ and the other three satisfying $\Im(m(z) + \frac{c-1}{cz}) < -\eta v$ when $c > 1$.

At first, we claim that the statement is true when $|z| < \delta$ for some small positive δ. In Jin et al. (2014), it has been proved that the four solutions for a z with $\Im(z) > 0$ are

$$m_1(z) = \frac{((1-c)/z + \sqrt{1+y_0}) + \sqrt{((1-c)/z - 1/\sqrt{1+y_0})^2 - y_0^2/(1+y_0)}}{2c},$$

$$m_2(z) = \frac{((1-c)/z + \sqrt{1+y_0}) - \sqrt{((1-c)/z - 1/\sqrt{1+y_0})^2 - y_0^2/(1+y_0)}}{2c},$$

$$m_3(z) = \frac{((1-c)/z - \sqrt{1+y_0}) + \sqrt{((1-c)/z + 1/\sqrt{1+y_0})^2 - y_0^2/(1+y_0)}}{2c},$$

$$m_4(z) = \frac{((1-c)/z - \sqrt{1+y_0}) - \sqrt{((1-c)/z + 1/\sqrt{1+y_0})^2 - y_0^2/(1+y_0)}}{2c},$$

where as convention, we assume that the square root of a complex number is the one with positive imaginary part, and y_0 is the root of the largest absolute value to the equation

$$y^3 - \frac{(1-c)^2 - z^2}{z^2} y^2 - \frac{4}{z^2} y - \frac{4}{z^2} = 0$$

or equivalently

$$z^2 y^3 - ((1-c)^2 - z^2) y^2 - 4y - 4 = 0. \tag{4.10}$$

We first consider the case where $z \to 0$. At first, by Lemma 4.1 of Bai, Miao and Rao (1991), we see that $y_0 \to \infty$ as $z \to 0$. Dividing both sides of (4.10) by y^2, we obtain that $y_0 = \frac{(1-c)^2}{z^2}(1+o(1))$. Writing $y_0 = \frac{(1-c)^2}{z^2} + d$ and substituting it into (4.10), we obtain

$$\begin{aligned}
&\frac{(1-c)^6}{z^4} + 3d\frac{(1-c)^4}{z^2} + 3d^2(1-c)^2 + d^3z^2 \\
&\quad - \left((1-c)^2 - z^2\right)\left(\frac{(1-c)^4}{z^4} + \frac{2d(1-c)^2}{z^2} + d^2\right) - \frac{4(1-c)^2}{z^2} - 4d - 4 \\
&\quad = \frac{d(1-c)^4}{z^2} - \frac{4(1-c)^2 - (1-c)^4}{z^2} + 2\left(d^2+d\right)(1-c)^2 - 4(d+1) \\
&\qquad + \left(d^3+d^2\right)z^2 = 0.
\end{aligned} \tag{4.11}$$

By equation (4.11), we have

$$d = \frac{4}{(1-c)^2} - 1 + O\left(z^2\right).$$

That is,

$$y_0 = \frac{(1-c)^2}{z^2} + \frac{4}{(1-c)^2} - 1 + O\left(z^2\right). \tag{4.12}$$

Therefore, we have

$$\sqrt{1+y_0} = -\frac{|1-c|}{z}\left(1 + \frac{2z^2}{(1-c)^4} + O\left(z^4\right)\right). \tag{4.13}$$

Consequently,

$$\frac{1-c}{z} + \sqrt{1+y_0} = \frac{1-c-|1-c|}{z} - \frac{2z}{|1-c|^3} + O\left(z^3\right), \tag{4.14}$$

$$\frac{1-c}{z} - \sqrt{1+y_0} = \frac{1-c+|1-c|}{z} + \frac{2z}{|1-c|^3} + O\left(z^3\right). \tag{4.15}$$

Because

$$\begin{aligned}
&\left(\frac{1-c}{z} \mp \frac{1}{\sqrt{1+y_0}}\right)^2 - \frac{y_0^2}{1+y_0} = \frac{(1-c)^2}{z^2} \mp 2\frac{1-c}{z\sqrt{1+y_0}} + 1 - y_0 \\
&\quad = -\frac{4}{(1-c)^2} \pm 2\frac{1-c}{|1-c| + O(z^2)} + 2 + O\left(z^2\right) \\
&\quad = -\frac{4}{(1-c)^2} \pm 2\frac{1-c}{|1-c|} + 2 + O\left(z^2\right),
\end{aligned}$$

we obtain

$$\sqrt{\left(\frac{1-c}{z} \mp \frac{1}{\sqrt{1+y_0}}\right)^2 - \frac{y_0^2}{1+y_0}}$$

(4.16)
$$= i\sqrt{\frac{4}{(1-c)^2} \mp 2\frac{1-c}{|1-c|} - 2} + O(z^2).$$

When $c < 1$, from (4.14) and (4.16), as $z \to 0$, we obtain

$$\Im(2cm_1) = \Im\left(O(z) + i\sqrt{\frac{4}{(1-c)^2} - 4}\right) > \frac{\sqrt{c(2-c)}}{(1-c)},$$
$$\Im(2cm_2) = \Im\left(O(z) - i\sqrt{\frac{4}{(1-c)^2} - 4}\right) < -\frac{\sqrt{c(2-c)}}{(1-c)},$$
(4.17)
$$\Im(2cm_3) = \Im\left(\frac{2(1-c)}{z} + i\sqrt{\frac{4}{(1-c)^2}} + O(z)\right) < -\frac{1-c}{|z|^2}v,$$
$$\Im(2cm_4) = \Im\left(\frac{2(1-c)}{z} - i\sqrt{\frac{4}{(1-c)^2}} + O(z)\right) < -\frac{2}{(1-c)}.$$

When $c \in (1, 2]$, as $z \to 0$, we have

$$\Im\left(2c\left(m_1 + \frac{c-1}{cz}\right)\right) = \Im\left(i\sqrt{\frac{4}{(1-c)^2}} + O(z)\right) > \frac{1}{c-1},$$
$$\Im\left(2c\left(m_2 + \frac{c-1}{cz}\right)\right) = \Im\left(-i\sqrt{\frac{4}{(1-c)^2}} + O(z)\right) < -\frac{1}{c-1},$$
$$\Im\left(2c\left(m_3 + \frac{c-1}{cz}\right)\right) = \Im\left(\frac{2(c-1)}{z} + i\sqrt{\frac{4c(2-c)}{(1-c)^2}} + O(z)\right)$$
(4.18)
$$< -\frac{c-1}{|z|^2}v,$$
$$\Im\left(2c\left(m_4 + \frac{c-1}{cz}\right)\right) = \Im\left(\frac{2(c-1)}{z} - i\sqrt{\frac{4c(2-c)}{(1-c)^2}} + O(z)\right)$$
$$< -\frac{c-1}{|z|^2}v.$$

When $c > 2$, as $z \to 0$, we have

$$\Im\left(2c\left(m_1 + \frac{c-1}{cz}\right)\right) = \Im\left(i\sqrt{\frac{4}{(1-c)^2}} + O(z)\right) > \frac{1}{c-1},$$
$$\Im\left(2c\left(m_2 + \frac{c-1}{cz}\right)\right) = \Im\left(-i\sqrt{\frac{4}{(1-c)^2}} + O(z)\right) < -\frac{1}{c-1},$$
$$\Im\left(2c\left(m_3 + \frac{c-1}{cz}\right)\right) = \Im\left(\frac{2(c-1)}{z} - \sqrt{\frac{4c(c-2)}{(1-c)^2}} + O(z)\right)$$

$$< -\frac{c-1}{|z|^2}v, \tag{4.19}$$

$$\Im\left(2c\left(m_4 + \frac{c-1}{cz}\right)\right) = \Im\left(\frac{2(c-1)}{z} + \sqrt{\frac{4c(c-2)}{(1-c)^2}} + O(z)\right)$$
$$< -\frac{c-1}{|z|^2}v.$$

This proves the result when $|z| < \delta$ for some $\delta > 0$.

For $|z| \geq \delta$, we first consider the case where $c < 1$. Suppose that $m(z)$ is one of the four continuous branches of the solutions of the equation (4.8). If the conclusion is incorrect for $m(z)$, then there exist a sequence of constants $\zeta_n \downarrow 0$ and a sequence of complex numbers $z_n = u_n + iv_n$ satisfying $|z_n| \geq \delta$, $|u_n| \leq A$, $v_n \in (0, \eta)$ with $\eta = \delta^2/2$ and $|\Im(m(z_n))| \leq \zeta_n v_n$. Then there is a subsequence $\{n'\}$ such that $z_{n'} \to z_0 = u_0 + iv_0$ with $u_{n'} \to u_0 \in [-A, A]$ and $v_{n'} \to v_0 \in [0, \eta]$.

Write $m(z_n) = m_1(z_n) + im_2(z_n)$, where $m_1(z_n)$ and $m_2(z_n)$ are real. Since $m(z_n)$ satisfies the equation (4.8), we have

$$\left(1 - c^2m^2(z_n)\right)\left(1 - c - cz_nm(z_n)\right)^2 = 1. \tag{4.20}$$

Comparing the imaginary parts of both sides of (4.20), we obtain

$$\begin{aligned} &c^2m_1(z_n)m_2(z_n) \\ &\quad \times \left[\left(1 - c - cu_nm_1(z_n) + cv_nm_2(z_n)\right)^2 - \left(cu_nm_2(z_n) + cv_nm_1(z_n)\right)^2\right] \\ &\quad + \left(1 - c^2m_1^2(z_n) + c^2m_2^2(z_n)\right)\left(cu_nm_2(z_n) + cv_nm_1(z_n)\right) \\ &\quad \times \left(1 - c - cu_nm_1(z_n) + cv_nm_2(z_n)\right) = 0. \end{aligned}$$

Dividing by v_n both sides of the equation above, we obtain

$$\left(1 - c^2m_1^2(z_0)\right)\left(cm_1(z_0)\right)\left(1 - c - cu_0m_1(z_0)\right) = 0. \tag{4.21}$$

By the condition that $|\Im(m(z_n))| \leq \zeta_n v_n \to 0$, we have that $m(z_0) = m_1(z_0)$ is real. The solutions $\pm 1/c$ and 0 of the equation (4.21) for $m(z_0)$ do not satisfy equation (4.8). Therefore, we have $1 - c - cu_0m(z_0) = 0$, and hence by (4.8)

$$-\left(1 - c^2m^2(z_0)\right)c^2v_0^2m^2(z_0) = 1. \tag{4.22}$$

Note that $v_0 = 0$ contradicts to the equation above. Thus, we have $v_0 \in (0, \delta^2/2]$. By (4.22) and the fact that $1 - c - cu_0m(z_0) = 0$, we obtain

$$\frac{(1-c)^2}{u_0^2} = \frac{v_0^2 + \sqrt{v_0^4 + 4v_0^2}}{2v_0^2} \quad \text{or} \quad u_0^2 = \frac{2v_0^2(1-c)^2}{v_0^2 + \sqrt{v_0^4 + 4v_0^2}}.$$

The expression of u_0^2 implies that $u_0^2 < v_0 < \delta^2/2$. On the other hand, by the assumption that $|z_0| > \delta$, we have $u_0^2 + v_0^2 > \delta^2$ and $v_0^2 < v_0 < \delta^2/2$ which implies that $u_0^2 > \delta^2/2$, the contradiction proves our assertion.

Now, we consider the case $c > 1$. Let $\underline{m}(z) = cm(z) + \frac{c-1}{z}$. Then equation (4.8) becomes

$$z^2\underline{m}^2(z)\left(1 - \left(\frac{1-c}{z} + \underline{m}(z)\right)^2\right) = 1. \tag{4.23}$$

If the conclusion is untrue, similar to the case where $c < 1$, there exist sequences $\zeta_n \downarrow 0$ and $z_n = u_n + iv_n \to z_0 = u_0 + i0$ such that $|\Im(\underline{m}(z_n))| \le \zeta_n v_n$, and $|u_n| \le A$. By the continuity of the solution $\underline{m}(z)$ for $|z| \ge \delta$, we may assume the inequality above is an equality, for otherwise, one may shift $\Re(z_n) = u_n$ toward the origin. Write $\underline{m}(z_n) = \underline{m}_1(z_n) + i\underline{m}_2(z_n)$, where $\underline{m}_1(z_n)$ and $\underline{m}_2(z_n)$ are both real. By the equality of imaginary parts of (4.23), we have

$$\begin{aligned}
&\underline{m}_1(z_n)\underline{m}_2(z_n)\\
&\quad\times \big(u_n^2 - v_n^2 - \big(1-c+u_n\underline{m}_1(z_n) - v_0\underline{m}_2(z_n)\big)^2\\
&\qquad\qquad + \big(u_n\underline{m}_2(z_n) + v_n\underline{m}_1(z_n)\big)^2\big)\\
&\quad - \big(\underline{m}_1^2(z_n) - \underline{m}_2^2(z_n)\big)\\
&\qquad \times \big(u_nv_n - \big(1-c+u_n\underline{m}_1(z_n) - v_n\underline{m}_2(z_n)\big)\big(u_n\underline{m}_2(z_n) + v_n\underline{m}_1(z_n)\big)\big)\\
&= 0
\end{aligned} \tag{4.24}$$

Dividing both sides by v_n and making $n \to \infty$ on both sides of the equation above, by assumption, we obtain

$$\underline{m}_1^2(z_0)\big(u_0 - \big(1-c+u_0\underline{m}_1(z_0)\big)\underline{m}_1(z_0)\big) = 0. \tag{4.25}$$

This implies that

$$u_0 = \frac{(1-c)\underline{m}_1(z_0)}{(1-\underline{m}_1^2(z_0))}. \tag{4.26}$$

Similarly, we have $\underline{m}(u_0) = \underline{m}_1(u_0)$ which is real. By the real part of (4.23), we have

$$\underline{m}^2(u_0)\big(u_0^2 - \big(1-c+u_0\underline{m}(u_0)\big)^2\big) = 1.$$

The solution to the equation above in u_0 is

$$u_0 = \frac{\underline{m}^3(u_0)(1-c) \pm \sqrt{\underline{m}^2(u_0) - c(2-c)\underline{m}^4(u_0)}}{\underline{m}^2(u_0)(1-\underline{m}^2(u_0))}. \tag{4.27}$$

If $\underline{m}^2(u_0) \ne \frac{1}{c(2-c)}$, then (4.27) contradicts (4.26).

Now, we consider the case where $c \in (1,2)$ and $\underline{m}^2(u_0) = \frac{1}{c(2-c)}$. By differentiating (4.23) with respect to z, we obtain

$$\begin{aligned}
\frac{d\underline{m}(z)}{dz} &= -\frac{\underline{m}(z - \underline{m}(1-c+z\underline{m}))}{z^2 - (1-c+z\underline{m})^2 - z\underline{m}(1-c+z\underline{m})}\\
&= -\frac{\underline{m}(z - \underline{m}(1-c+z\underline{m}))}{z^2 - (1-c)^2 - z(1-c)\underline{m}}.
\end{aligned}$$

Because

$$\Im\big(z_n - \underline{m}\big(1-c+z_n\underline{m}(z_n)\big)\big) = v_n\big[\big(1-\underline{m}_1^2(u_0)\big)+o(1)\big],$$
$$\begin{aligned}
&\Re\big(z_n - \underline{m}\big(1-c+z_n\underline{m}(z_n)\big)\big)\\
&\quad= \big[u_n - \underline{m}_1(z_n)\big(1-c+u_n\underline{m}_1(z_n)\big)\big] + O\big(\underline{m}_2(z_n)\big)\\
&\quad= \big[u_n\big(1-\underline{m}_1^2(z_n)\big) - (1-c)\underline{m}_1(z_n)\big] + O\big(\underline{m}_2(z_n)\big) \qquad \text{(by (4.24))}\\
&\quad= -\frac{\underline{m}_2(z_n)}{v_n\underline{m}_1(z_n)}\big[u_n^2 - (1-c)^2 - u_n(1-c)\underline{m}_1(z_n) + o(1)\big]\\
&\quad\simeq \zeta_n \frac{(1-c)^2[1-2\underline{m}^2(u_0)]}{\underline{m}(u_0)(1-\underline{m}(u_0)^2)^2},
\end{aligned}$$
$$\frac{z_n^2 - (1-c)^2 - z(1-c)\underline{m}(z_n)}{\underline{m}(z_n)} \simeq \frac{(1-c)^2[2\underline{m}^2(u_0)-1]}{\underline{m}(u_0)(1-\underline{m}^2(u_0))^2}.$$

Therefore,

$$\frac{\partial \underline{m}_2(z_n)}{\partial u} \simeq v_n \frac{\underline{m}(u_0)(1-\underline{m}^2(u_0))^3}{(1-c)^2(2\underline{m}^2(u_0)-1)},$$

and

$$\frac{\partial \underline{m}_1(z_n)}{\partial u} \simeq \zeta_n.$$

Hence,

$$\begin{aligned}
G_n &= \underline{m}_1^2(z_n)\big(u_n - \big(1-c+u_n\underline{m}_1(z_n)\big)\underline{m}_1(z_n)\big)\\
&\quad - \underline{m}_1^2(z_0)\big(u_0 - \big(1-c+u_0\underline{m}_1(z_0)\big)\underline{m}_1(z_0)\big) \qquad (4.28)\\
&= (u_n-u_0)\big(\underline{m}_1^2(z_n^*)\big(1-\underline{m}_1^2(z_0)\big)+O(\zeta_n)\big).
\end{aligned}$$

On the other hand, we have

$$\begin{aligned}
\zeta_n v_n &= \underline{m}_2(z_n) - \underline{m}_2(z_0)\\
&= (u_n-u_0)\frac{\partial \underline{m}_2(z_n^*)}{\partial u} \qquad (4.29)\\
&\simeq (u_n-u_0)\frac{v_n\underline{m}(z_0)(1-\underline{m}^2(z_0))^3}{(1-c)^2(2\underline{m}^2(z_0)-1)}.
\end{aligned}$$

Therefore,

$$G_n \simeq \zeta_n \frac{(1-c)^2\underline{m}(u_0)(2\underline{m}^2(z_0)-1)}{(1-\underline{m}^2(z_0))^2}. \qquad (4.30)$$

Substituting the above into (4.24) and dividing $\underline{m}_2(z_n) = \zeta_n v_n$ on both sides and letting $n \to \infty$, we obtain

$$\begin{aligned} 0 &= \underline{m}(u_0)\big(u_0^2 - \big(1-c+u_0\underline{m}(u_0)\big)^2\big) + \underline{m}^2(u_0)\big(1-c+u_0\underline{m}(u_0)\big)u_0 \\ &\quad + \frac{(1-c)^2\underline{m}(u_0)(2\underline{m}^2(u_0)-1)}{(1-\underline{m}^2(u_0))^2} \\ &= \underline{m}(u_0)\big(u_0^2 - (1-c)^2 - u_0(1-c)\underline{m}(u_0)\big) \\ &\quad + \frac{(1-c)^2\underline{m}(u_0)(2\underline{m}^2(u_0)-1)}{(1-\underline{m}^2(u_0))^2}. \end{aligned} \tag{4.31}$$

By substitution of (4.26), the equation above becomes

$$\frac{2(1-c)^2\underline{m}(u_0)(2\underline{m}^2(u_0)-1)}{(1-\underline{m}^2(u_0))^2} = 0$$

which also implies that $\underline{m}^2(u_0) = \frac{1}{2}$. This contradicts to the assumption that $\underline{m}^2(u_0) = \frac{1}{c(2-c)}$ and the assertion is finally proved.

Consequently, under the condition that $|\delta_n| \le K v_n^\eta$ with $\eta > 1$, we have $\max_{j=2,3,4,z=u+iv_n} |m_j(z) - \mathrm{E}m_n(z)| \ge \eta v_n$ and thus $\max_{z=u+iv_n} |m_1(z) - \mathrm{E}m_n(z)| \le K v_n^\eta$ when $c < 1$. Similarly for $\underline{m}(z)$ when $c > 1$.

Hence, to prove (4.5), it remains to show

$$|\delta_n| \le K v_n^\eta \tag{4.32}$$

for some $K > 0$, and $\eta > 1$.

4.2.2. *Convergence rate of* δ_n. Let $v_n \ge n^{-1/52}$. By (4.6), we have

$$\delta_n = c_n + c_n z\mathrm{E}m_n(z) - 1 + \frac{1}{x_{n1} - x_{n0}} =: \frac{1}{T}\sum_{k=1}^{T} \mathrm{E}\eta_k,$$

where

$$\eta_k = \boldsymbol{\gamma}_k^* \mathbf{A}^{-1}(\boldsymbol{\gamma}_{k+\tau} + \boldsymbol{\gamma}_{k-\tau}) - 1 + \frac{1}{x_{n1} - x_{n0}}.$$

When $k \le v_n^{-4}$ or $\ge T - v_n^{-4}$, by (iii)(a) of Lemma 3.6, we have

$$\begin{aligned} |\mathrm{E}\eta_k| &\le v_n^{-1}\sqrt{\mathrm{E}|\boldsymbol{\gamma}_k|^2\big(\mathrm{E}|\boldsymbol{\gamma}_{k-\tau}|^2 + \mathrm{E}|\boldsymbol{\gamma}_{k+\tau}|^2\big)} + 1 + \frac{1}{|x_{n1} - x_{n0}|} \\ &\le K v_n^{-1}. \end{aligned}$$

Therefore, for all large n,

$$\frac{1}{T}\left(\sum_{k=1}^{[v_n^{-4}]} + \sum_{k=[T-v_n^{-4}]}^{T}\right)|\mathrm{E}\eta_k| \le \frac{K}{T v_n^5} \le K v_n^{47}. \tag{4.33}$$

When $k \in ([v_n^{-4}], [T - v_n^{-4}])$, denote

$$
\begin{aligned}
&\varepsilon_1 = (\boldsymbol{\gamma}_{k+\tau}^* + \boldsymbol{\gamma}_{k-\tau}^*)\mathbf{A}_k^{-1}\boldsymbol{\gamma}_k,\\
&\varepsilon_2 = \boldsymbol{\gamma}_k^*\mathbf{A}_k^{-1}(\boldsymbol{\gamma}_{k+\tau} + \boldsymbol{\gamma}_{k-\tau}),\\
&\varepsilon_3 = \boldsymbol{\gamma}_k^*\mathbf{A}_k^{-1}\boldsymbol{\gamma}_k - \frac{1}{2T}\operatorname{tr}\mathbf{A}_k^{-1},\\
&\varepsilon_4 = \frac{1}{2T}\operatorname{tr}\mathbf{A}_k^{-1} - \frac{c_n}{2}\mathrm{E}m_n(z),\\
&\varepsilon_5 = (\boldsymbol{\gamma}_{k+\tau}^* + \boldsymbol{\gamma}_{k-\tau}^*)\mathbf{A}_k^{-1}(\boldsymbol{\gamma}_{k+\tau} + \boldsymbol{\gamma}_{k-\tau}) - \frac{c_n}{x_{n1}}\mathrm{E}m_n(z).
\end{aligned}
\tag{4.34}
$$

Then, by the fact that $x_{n1} - x_{n0} = 1 - 2a_n^2/x_{n1}$, we have

$$
\begin{aligned}
-\mathrm{E}\eta_k &= \mathrm{E}\biggl(1\Big/\biggl(1 + \boldsymbol{\gamma}_k^*\mathbf{A}_k^{-1}(\boldsymbol{\gamma}_{k+\tau} + \boldsymbol{\gamma}_{k-\tau})\\
&\qquad - \frac{\boldsymbol{\gamma}_k^*\mathbf{A}_k^{-1}\boldsymbol{\gamma}_k(\boldsymbol{\gamma}_{k+\tau}^* + \boldsymbol{\gamma}_{k-\tau}^*)\mathbf{A}_k^{-1}(\boldsymbol{\gamma}_{k+\tau} + \boldsymbol{\gamma}_{k-\tau})}{1 + (\boldsymbol{\gamma}_{k+\tau}^* + \boldsymbol{\gamma}_{k-\tau}^*)\mathbf{A}_k^{-1}\boldsymbol{\gamma}_k}\biggr)\biggr) - \frac{1}{x_{n1} - x_{n0}}\\
&= \frac{1}{x_{n1} - x_{n0}}\mathrm{E}\beta_k\biggl(-2\varepsilon_1\frac{a_n^2}{x_{n1}} - \varepsilon_2 - \varepsilon_1\varepsilon_2\\
&\qquad + (\boldsymbol{\gamma}_{k+\tau}^* + \boldsymbol{\gamma}_{k-\tau}^*)\mathbf{A}_k^{-1}(\boldsymbol{\gamma}_{k+\tau} + \boldsymbol{\gamma}_{k-\tau})(\varepsilon_3 + \varepsilon_4) + a_n\varepsilon_5\biggr),
\end{aligned}
$$

where

$$
\begin{aligned}
\beta_k &= \frac{1}{1 + \varepsilon_1 + \varepsilon_2 + \varepsilon_1\varepsilon_2 - \boldsymbol{\gamma}_k^*\mathbf{A}_k^{-1}\boldsymbol{\gamma}_k(\boldsymbol{\gamma}_{k+\tau}^* + \boldsymbol{\gamma}_{k-\tau}^*)\mathbf{A}_k^{-1}(\boldsymbol{\gamma}_{k+\tau} + \boldsymbol{\gamma}_{k-\tau})}\\
&= \frac{1}{1 + \varepsilon_1 + \varepsilon_2 + \varepsilon_1\varepsilon_2 - (a_n + \varepsilon_3 + \varepsilon_4)(2a_n/x_{n1} + \varepsilon_5)}.
\end{aligned}
$$

Define a random set $\mathcal{E}_n = \{|\varepsilon_i| \le v_n^6, i = 1, 2, 3, 4, 5\}$. When $\mathcal{E}_n$ happens, by the facts $|a_n| \le K v_n^{-1}$, $|\frac{2a_n}{x_{n1}}| \le 2$ and Lemma 3.6(iii)(a), we have

$$
\begin{aligned}
|\beta_k| &\le \frac{1}{|1 - 2a_n^2/x_{n1} - 9v_n^6 - K v_n^5|}\\
&= \frac{1}{|1 - 2x_{n0} - 9v_n^6 - K v_n^5|}\\
&= \frac{1}{|x_{n1} - x_{n0} - 9v_n^6 - K v_n^5|}\\
&\le K v_n^{-1}.
\end{aligned}
$$

Together with Lemma 3.6(ii)(a) and (iii)(a), we obtain that

$$
\begin{aligned}
|\eta_k| &\le \frac{1}{|x_{n1}-x_{n0}|} \\
&\quad \times K v_n^{-1}\big(v_n^6(2|x_{n0}|)+v_n^6+v_n^{12}+v_n^{-1}\|\boldsymbol{\gamma}_{k+\tau}+\boldsymbol{\gamma}_{k-\tau}\|^2(2v_n^6)+Kv_n^5\big) \\
&\le K v_n^3.
\end{aligned}
$$

Therefore, by Lemmas 3.4, 3.7(a) and 3.8(a), when $v_n \ge n^{-1/52}$, we have

$$
\begin{aligned}
\mathrm{E}|\eta_k| &\le K v_n^3 + K v_n^{-1}\left(\sum_{i=1}^5 \mathrm{P}\big(|\varepsilon_i| \ge v_n^6\big)\right) \\
&\le K v_n^3.
\end{aligned}
\tag{4.35}
$$

Then the conclusion (4.32) follows from (4.33) and (4.35).

4.3. *Convergence rate of* $\|F_n - F_{c_n}\|$. Choose $v_n = n^{-1/52}$. Let F_n be the empirical distribution function of $\mathbf{M}_n$ and F_{c_n} be the LSD with the ratio parameter $c_n = n/T$ whose Stieltjes transform is denoted by m_n^0. By (1.3), let $B = (1+\sqrt{c})^2 + \delta$, and we have $F_{c_n}([-B, B]) = 1$. By Lemma 2.4 we have, for some $A > B$ and $a > 0$,

$$
\begin{aligned}
&\mathrm{P}\big(\|F_n - F_{c_n}\| > c'\sqrt{v_n}\big) \\
&\quad \le \mathrm{P}\Big(\sup_{u\in[-A,A]} \big|m_n(z) - m_n^0(z)\big| > K_0\sqrt{v_n}\Big) \\
&\qquad + \mathrm{P}\Big(\sup_x \int_{|y|\le 2v_n a} \big|F_{c_n}(x+y) - F_{c_n}(x)\big|\,dy > K_0\big(c'-1\big)v_n^{3/2}\Big) \\
&\quad \le \mathrm{P}\Big(\sup_{u\in[-A,A]} \big|m_n(z) - \mathrm{E}m_n(z)\big| > \frac{K_0\sqrt{v_n}}{2}\Big) \\
&\qquad + \mathrm{P}\Big(\sup_{u\in[-A,A]} \big|\mathrm{E}m_n(z) - m_n^0(z)\big| > \frac{K_0\sqrt{v_n}}{2}\Big) \\
&\qquad + \mathrm{P}\Big(\sup_x \int_{|y|\le 2v_n a} \big|F_{c_n}(x+y) - F_{c_n}(x)\big|\,dy > K_0\big(c'-1\big)v_n^{3/2}\Big),
\end{aligned}
$$

where $K_0 = \pi(1-\kappa)(2\gamma - 1)$, and a is a constant defined in Lemma 2.4. By $|\mathrm{E}m_n(z) - m_n^0(z)| = o(v_n)$, the second probability is 0 for all large n.

By the analysis of Section 3 of Jin et al. (2014), we see that $\phi_{c_n}(x) := \frac{d}{dx}F_{c_n}(x) \le K|x|^{-1/2}$, which implies that F_{c_n} satisfies the Lipschitz condition with index $\frac{1}{2}$. Hence, for some large c', we have

$$
\begin{aligned}
&\sup_x \int_{|y|\le 2v_n a} \big|F_{c_n}(x+y) - F_{c_n}(x)\big|\,dy \\
&\qquad \le K\int_{|y|\le 2v_n a} |y|^{1/2}\,dy = 4Ka^2 v_n^{3/2} < K_0\big(c'-1\big)v_n^{3/2}.
\end{aligned}
$$

Therefore, the third probability is 0.

For the first probability, let $\mathcal{S}_n$ be the set containing n^2 points that are equally spaced between $-n$ and n and note that $[-A, A] \subseteq [-n, n]$ for all large n. When $|u_1 - u_2| \le \frac{2}{n}$, we have

$$\big|m_n(u_1 + iv_n) - m_n(u_2 + iv_n)\big| \le |u_1 - u_2| v_n^{-2} < \frac{K_0\sqrt{v_n}}{2},$$

$$\big|m_n^0(u_1 + iv_n) - m_n^0(u_2 + iv_n)\big| \le |u_1 - u_2| v_n^{-2} < \frac{K_0\sqrt{v_n}}{2}.$$

Therefore, by (4.3), for any $t > 0$, we have

$$\begin{aligned}
&\mathrm{P}\Big(\sup_{u\in[-A,A]} \big|m_n(z) - \mathrm{E}m_n(z)\big| > \frac{K_0\sqrt{v_n}}{2}\Big) \\
&\quad = \mathrm{P}\Big(\sup_{u\in\mathcal{S}_n} \big|m_n(z) - \mathrm{E}m_n(z)\big| > \frac{K_0\sqrt{v_n}}{2}\Big) \\
&\quad \le n^2 \mathrm{P}\Big(\big|m_n(z) - \mathrm{E}m_n(z)\big| > \frac{K_0\sqrt{v_n}}{2}\Big) \\
&\quad \le K n^{2-p/2} v_n^{-p} \\
&\quad = o(n^{-t})
\end{aligned}$$

by selecting p large enough. Thus, we have proved, for any fixed $t > 0$

$$\mathrm{P}\big(\|F_n - F_{c_n}\| > c' n^{-1/104}\big) = o(n^{-t}). \tag{4.36}$$

Next, let $a' = a - \underline{\varepsilon}$ and $b' = b + \underline{\varepsilon}$ for some $\underline{\varepsilon} > 0$ such that $(a', b') \supseteq [a, b]$ is an open interval outside the support of F_{c_n} for all n large enough. By $|d(c_n) - d(c)| \to 0$, and hence $[a', b']$ is also outside the support of F_{c_n}. We conclude that $F_{c_n}(b') - F_{c_n}(a') = 0$ for all large n. Hence, we have

$$\begin{aligned}
F_n\{[a', b']\} &= F_n(b') - F_n(a') - \big(F_{c_n}(b') - F_{c_n}(a')\big) \\
&\le 2\|F_n - F_{c_n}\|.
\end{aligned}$$

Therefore,

$$\begin{aligned}
&\mathrm{P}\Big(\max_{k\le n} \mathrm{E}_k\big(F_n\{[a', b']\}\big) \ge 4c' n^{-1/104}\Big) \\
&\quad \le \mathrm{P}\Big(\max_{k\le n} \mathrm{E}_k\big(F_n\{[a', b']\} I_{\{\|F_n - F_{c_n}\| < c' n^{-1/104}\}}\big) \ge 2c' n^{-1/104}\Big) \\
&\qquad + \mathrm{P}\Big(\max_{k\le n} \mathrm{E}_k\big(F_n\{[a', b']\} I_{\{\|F_n - F_{c_n}\| \ge c' n^{-1/104}\}}\big) \ge 2c' n^{-1/104}\Big) \\
&\quad \le 0 + \mathrm{P}\Big(\max_{k\le n} \mathrm{E}_k I_{\{\|F_n - F_{c_n}\| \ge c' n^{-1/104}\}} \ne 0\Big) \\
&\quad \le n\mathrm{P}\big(\|F_n - F_{c_n}\| \ge c' n^{-1/104}\big) = o(n^{-t})
\end{aligned} \tag{4.37}$$

for any $t > 0$.

5. A refined convergence rate of Stieltjes transform when $u \in [a, b]$. In this section, we are to prove that for $v_n = n^{-1/212}$,

$$m_n - m_n^0 = o\big(1/(nv_n)\big) \qquad \text{a.s.} \tag{5.1}$$

by refining the convergence rates obtained in the last section.

5.1. *A refined convergence rate of $m_n - \mathrm{E}m_n$.* In this subsection, we want to show that

$$\sup_{u\in[a,b]} \big|m_n(z) - \mathrm{E}m_n(z)\big| = o\big(1/(nv_n)\big), \qquad \text{a.s.} \tag{5.2}$$

First, by recalling that $\tilde{\mathbf{A}}_k = \mathbf{A} - (\boldsymbol{\gamma}_{k+\tau} + \boldsymbol{\gamma}_{k-\tau})\boldsymbol{\gamma}_k^*$ and $\mathbf{A}_k = \tilde{\mathbf{A}}_k - \boldsymbol{\gamma}_k(\boldsymbol{\gamma}_{k+\tau} + \boldsymbol{\gamma}_{k-\tau})^*$, we have

$$\begin{aligned}
&m_n(z) - \mathrm{E}m_n(z)\\
&\quad= \sum_{k=1}^{T}\big(\mathrm{E}_{k-1}m_n(z) - \mathrm{E}_k m_n(z)\big)\\
&\quad= \sum_{k=1}^{T}\frac{1}{n}(\mathrm{E}_k - \mathrm{E}_{k-1})\big(\big(\operatorname{tr}\mathbf{A}_k^{-1} - \operatorname{tr}\tilde{\mathbf{A}}_k^{-1}\big) + \big(\operatorname{tr}\tilde{\mathbf{A}}_k^{-1} - \operatorname{tr}\mathbf{A}^{-1}\big)\big)\\
&\quad= \sum_{k=1}^{T}\frac{1}{n}(\mathrm{E}_k - \mathrm{E}_{k-1})\\
&\qquad\quad\times\left(\frac{(\boldsymbol{\gamma}_{k+\tau} + \boldsymbol{\gamma}_{k-\tau})^*\mathbf{A}_k^{-2}\boldsymbol{\gamma}_k}{1 + (\boldsymbol{\gamma}_{k+\tau} + \boldsymbol{\gamma}_{k-\tau})^*\mathbf{A}_k^{-1}\boldsymbol{\gamma}_k} + \frac{\boldsymbol{\gamma}_k^*\tilde{\mathbf{A}}_k^{-2}(\boldsymbol{\gamma}_{k+\tau} + \boldsymbol{\gamma}_{k-\tau})}{1 + \boldsymbol{\gamma}_k^*\tilde{\mathbf{A}}_k^{-1}(\boldsymbol{\gamma}_{k+\tau} + \boldsymbol{\gamma}_{k-\tau})}\right)\\
&\quad= \sum_{k=1}^{T}\frac{1}{n}(\mathrm{E}_k - \mathrm{E}_{k-1})\frac{d}{dz}\big(\log\big(1 + (\boldsymbol{\gamma}_{k+\tau} + \boldsymbol{\gamma}_{k-\tau})^*\mathbf{A}_k^{-1}\boldsymbol{\gamma}_k\big)\\
&\qquad\qquad + \log\big(1 + \boldsymbol{\gamma}_k^*\tilde{\mathbf{A}}_k^{-1}(\boldsymbol{\gamma}_{k+\tau} + \boldsymbol{\gamma}_{k-\tau})\big)\big)\\
&\quad= \sum_{k=1}^{T}\frac{1}{n}(\mathrm{E}_k - \mathrm{E}_{k-1})\frac{d}{dz}\\
&\qquad\quad\times\big(\log\big(\big(1 + (\boldsymbol{\gamma}_{k+\tau} + \boldsymbol{\gamma}_{k-\tau})^*\mathbf{A}_k^{-1}\boldsymbol{\gamma}_k\big)\big(1 + \boldsymbol{\gamma}_k^*\mathbf{A}_k^{-1}(\boldsymbol{\gamma}_{k+\tau} + \boldsymbol{\gamma}_{k-\tau})\big)\\
&\qquad\qquad - \boldsymbol{\gamma}_k^*\mathbf{A}_k^{-1}\boldsymbol{\gamma}_k\big(\boldsymbol{\gamma}_{k+\tau}^* + \boldsymbol{\gamma}_{k-\tau}^*\big)\mathbf{A}_k^{-1}(\boldsymbol{\gamma}_{k+\tau} + \boldsymbol{\gamma}_{k-\tau})\big)\\
&\qquad\qquad - \log(x_{n1} - x_{n0})\big)\\
&\quad= \sum_{k=1}^{T}\frac{1}{n}(\mathrm{E}_k - \mathrm{E}_{k-1})
\end{aligned}$$

$$
\begin{aligned}
&\times \left(\frac{d}{dz}\log\biggl(1+\frac{\varepsilon_1}{x_{n1}-x_{n0}}+\frac{\varepsilon_2}{x_{n1}-x_{n0}}\right.\\
&\qquad\qquad -\frac{\varepsilon_3(\boldsymbol{\gamma}_{k+\tau}^*+\boldsymbol{\gamma}_{k-\tau}^*)\mathbf{A}_k^{-1}(\boldsymbol{\gamma}_{k+\tau}+\boldsymbol{\gamma}_{k-\tau})}{x_{n1}-x_{n0}}\\
&\qquad\qquad \left.+\frac{\varepsilon_1\varepsilon_2-\varepsilon_4(\boldsymbol{\gamma}_{k+\tau}^*+\boldsymbol{\gamma}_{k-\tau}^*)\mathbf{A}_k^{-1}(\boldsymbol{\gamma}_{k+\tau}+\boldsymbol{\gamma}_{k-\tau})-a_n\varepsilon_5}{x_{n1}-x_{n0}}\biggr)\right)\\
&:=\sum_{k=1}^{T}\frac{1}{n}(\mathrm{E}_k-\mathrm{E}_{k-1})\frac{d}{dz}\log\bigl(1+\alpha_{k1}(z)+\alpha_{k2}(z)+\alpha_{k3}(z)+r_k(z)\bigr)\\
&:=\sum_{k=1}^{T}\frac{1}{n}(\mathrm{E}_k-\mathrm{E}_{k-1})\frac{d}{dz}f_k(z),
\end{aligned}
$$

where ε_i's, $i=1,\dots,5$, are defined in (4.34).

Let $\alpha_{k4}(z):=f_k(z)-\alpha_{k1}(z)-\alpha_{k2}(z)-\alpha_{k3}(z)-r_k(z)$. It is easy to derive that

$$
\begin{aligned}
\frac{d}{dz}\alpha_{k1}(z)&=\frac{1}{x_{n1}-x_{n0}}(\boldsymbol{\gamma}_{k+\tau}^*+\boldsymbol{\gamma}_{k-\tau}^*)\mathbf{A}_k^{-2}\boldsymbol{\gamma}_k\\
&\qquad -\frac{x'_{n1}-x'_{n0}}{(x_{n1}-x_{n0})^2}(\boldsymbol{\gamma}_{k+\tau}^*+\boldsymbol{\gamma}_{k-\tau}^*)\mathbf{A}_k^{-1}\boldsymbol{\gamma}_k,
\end{aligned}\tag{5.3}
$$

$$
\begin{aligned}
\frac{d}{dz}\alpha_{k2}(z)&=\frac{1}{x_{n1}-x_{n0}}\boldsymbol{\gamma}_k^*\mathbf{A}_k^{-2}(\boldsymbol{\gamma}_{k+\tau}+\boldsymbol{\gamma}_{k-\tau})\\
&\qquad -\frac{x'_{n1}-x'_{n0}}{(x_{n1}-x_{n0})^2}\boldsymbol{\gamma}_{k+}^*\mathbf{A}_k^{-1}(\boldsymbol{\gamma}_{k+\tau}+\boldsymbol{\gamma}_{k-\tau})
\end{aligned}\tag{5.4}
$$

and

$$
\begin{aligned}
&\frac{d}{dz}\alpha_{k3}(z)\\
&\quad=\frac{1}{x_{n1}-x_{n0}}\\
&\qquad\times\biggl(\Bigl(\boldsymbol{\gamma}_k^*\mathbf{A}_k^{-2}\boldsymbol{\gamma}_k-\frac{1}{2T}\operatorname{tr}\mathbf{A}_k^{-2}\Bigr)(\boldsymbol{\gamma}_{k+\tau}^*+\boldsymbol{\gamma}_{k-\tau}^*)\mathbf{A}_k^{-1}(\boldsymbol{\gamma}_{k+\tau}+\boldsymbol{\gamma}_{k-\tau})\\
&\qquad\quad+\Bigl(\boldsymbol{\gamma}_k^*\mathbf{A}_k^{-1}\boldsymbol{\gamma}_k-\frac{1}{2T}\operatorname{tr}\mathbf{A}_k^{-1}\Bigr)(\boldsymbol{\gamma}_{k+\tau}^*+\boldsymbol{\gamma}_{k-\tau}^*)\mathbf{A}_k^{-2}(\boldsymbol{\gamma}_{k+\tau}+\boldsymbol{\gamma}_{k-\tau})\biggr)\\
&\qquad+\frac{x'_{n1}-x'_{n0}}{(x_{n1}-x_{n0})^2}\\
&\qquad\quad\times\Bigl(\boldsymbol{\gamma}_k^*\mathbf{A}_k^{-1}\boldsymbol{\gamma}_k-\frac{1}{2T}\operatorname{tr}\mathbf{A}_k^{-1}\Bigr)(\boldsymbol{\gamma}_{k+\tau}^*+\boldsymbol{\gamma}_{k-\tau}^*)\mathbf{A}_k^{-1}(\boldsymbol{\gamma}_{k+\tau}+\boldsymbol{\gamma}_{k-\tau}).
\end{aligned}\tag{5.5}
$$

Note that by (iii)(b) of Lemma 3.6, we have $\frac{1}{|x_{n1}-x_{n0}|} \leq K$. Also, by Remarks 3.1 and 3.2, we have $|x'_{n1} - x'_{n0}| = |-\frac{4a_n a'_n}{x_{n1}-x_{n0}}| \leq K$. Together with Cauchy's formula and the fact that $|\ln(1+x) - x| \leq |x|^2$ for any complex x with absolute value smaller than $\frac{1}{2}$, we have

$$\begin{aligned}
&\left|\frac{d}{dz}\alpha_{k4}(z)\right| \\
&\quad = \left|\frac{d}{dz}\big(\log\big(1+\alpha_{k1}(z)+\alpha_{k2}(z)+\alpha_{k3}(z)+r_k(z)\big)\right. \\
&\qquad\qquad \left. - \alpha_{k1}(z) - \alpha_{k2}(z) - \alpha_{k3}(z) - r_k(z)\big)\right| \\
&\quad = \left|\frac{1}{2\pi i}\oint_{|\xi - z| = v_n/2}\big(\big(\log\big(1+\alpha_{k1}(\xi)+\alpha_{k2}(\xi)+\alpha_{k3}(z)+r_k(\xi)\big)\right. \\
&\qquad\qquad - \alpha_{k1}(\xi) - \alpha_{k2}(\xi) - \alpha_{k3}(\xi) - r_k(\xi)\big) \\
&\qquad\qquad \left. /(\xi - z)^2\big)\,d\xi\right|.
\end{aligned} \tag{5.6}$$

Therefore, for each $u \in [a,b]$, $\ell \geq 1$, we have

$$\begin{aligned}
&\mathrm{E}\big|nv_n\big(m_n(z) - \mathrm{E}m_n(z)\big)\big|^{2\ell} \\
&\quad = \mathrm{E}\left|v_n\sum_{k=1}^{T}(\mathrm{E}_k - \mathrm{E}_{k-1})\frac{d}{dz}f_k(z)\right|^{2\ell} \\
&\quad \leq K\sum_{i=1}^{4}\mathrm{E}\left|v_n\sum_{k=1}^{T}(\mathrm{E}_k - \mathrm{E}_{k-1})\frac{d}{dz}\alpha_{ki}\right|^{2\ell} + K\mathrm{E}\left|v_n\sum_{k=1}^{T}(\mathrm{E}_k - \mathrm{E}_{k-1})\frac{d}{dz}r_k\right|^{2\ell}.
\end{aligned} \tag{5.7}$$

By Lemma 2.1, for $i = 1,2,3,4$, we have

$$\begin{aligned}
&\mathrm{E}\left|v_n\sum_{k=1}^{T}(\mathrm{E}_k - \mathrm{E}_{k-1})\frac{d}{dz}\alpha_{ki}\right|^{2\ell} \\
&\quad \leq K_\ell v_n^{2\ell}\left[\mathrm{E}\left(\sum_{k=1}^{T}\mathrm{E}_{k-1}\left|(\mathrm{E}_k - \mathrm{E}_{k-1})\frac{d}{dz}\alpha_{ki}\right|^2\right)^{\ell}\right. \\
&\qquad\qquad \left. + \sum_{k=1}^{T}\mathrm{E}\left|(\mathrm{E}_k - \mathrm{E}_{k-1})\frac{d}{dz}\alpha_{ki}\right|^{2\ell}\right] \\
&\quad \leq K'_\ell v_n^{2\ell}\left[\mathrm{E}\left(\sum_{k=1}^{T}\mathrm{E}_{k-1}\left|\frac{d}{dz}\alpha_{ki}\right|^2\right)^{\ell} + \sum_{k=1}^{T}\mathrm{E}\left|\frac{d}{dz}\alpha_{ki}\right|^{2\ell}\right].
\end{aligned}$$

Now we are ready to estimate the terms above. By elementary calculation, we have

$$\begin{aligned}\mathrm{E}_k|\boldsymbol{\gamma}_{k+\tau}^*\mathbf{A}_k^{-1}\boldsymbol{\gamma}_k|^2 &= \frac{1}{2T}\mathrm{E}_k\boldsymbol{\gamma}_{k+\tau}^*\mathbf{A}_k^{-1}(\mathbf{A}_k^*)^{-1}\boldsymbol{\gamma}_{k+\tau}\\ &\le \frac{K}{T}+\frac{1}{2Tv_n^2}\mathrm{E}_k I\big(|\boldsymbol{\gamma}_{k+\tau}^*\mathbf{A}_k^{-1}(\mathbf{A}_k^*)^{-1}\boldsymbol{\gamma}_{k+\tau}|\ge K\big)\end{aligned}\tag{5.8}$$

and

$$\begin{aligned}\mathrm{E}_k|\boldsymbol{\gamma}_{k+\tau}^*\mathbf{A}_k^{-2}\boldsymbol{\gamma}_k|^2 &= \frac{1}{2T}\mathrm{E}_k\boldsymbol{\gamma}_{k+\tau}^*\mathbf{A}_k^{-2}(\mathbf{A}_k^*)^{-2}\boldsymbol{\gamma}_{k+\tau}\\ &\le \frac{K}{T}+\frac{1}{2Tv_n^4}\mathrm{E}_k I\big(|\boldsymbol{\gamma}_{k+\tau}^*\mathbf{A}_k^{-2}(\mathbf{A}_k^*)^{-2}\boldsymbol{\gamma}_{k+\tau}|\ge K\big),\end{aligned}\tag{5.9}$$

for the constant $K>0$ such that Lemmas 3.9 and 3.10 hold.

Come back to the expressions of (5.3), (5.4) and (5.5). By definition of x_{ni} one can verify that $x'_{n1}-x'_{n0}=-\frac{4a_na'_n}{x_{n1}-x_{n0}}$ which is bounded. By Remarks 3.1, 3.2, Lemma 3.4 and estimates (5.8), (5.9), we have

$$\begin{aligned}&v_n^{2\ell}\Bigg[\mathrm{E}\Bigg(\sum_{k=1}^{T}\mathrm{E}_k\bigg|\frac{d}{dz}\alpha_{k1}\bigg|^2\Bigg)^\ell+\sum_{k=1}^{T}\mathrm{E}\bigg|\frac{d}{dz}\alpha_{k1}\bigg|^{2\ell}\Bigg]\\
&\quad\le Kv_n^{2\ell}\Bigg[\mathrm{E}\Bigg(\sum_{k=1}^{T}\mathrm{E}_k|(\boldsymbol{\gamma}_{k+\tau}^*+\boldsymbol{\gamma}_{k-\tau}^*)\mathbf{A}_k^{-2}\boldsymbol{\gamma}_k|^2\\
&\qquad\qquad+\sum_{k=1}^{T}\mathrm{E}_k|(\boldsymbol{\gamma}_{k+\tau}^*+\boldsymbol{\gamma}_{k-\tau}^*)\mathbf{A}_k^{-1}\boldsymbol{\gamma}_k|^2\Bigg)^\ell\\
&\qquad\qquad+\sum_{k=1}^{T}\mathrm{E}|(\boldsymbol{\gamma}_{k+\tau}^*+\boldsymbol{\gamma}_{k-\tau}^*)\mathbf{A}_k^{-2}\boldsymbol{\gamma}_k|^{2\ell}\\
&\qquad\qquad+\sum_{k=1}^{T}\mathrm{E}|(\boldsymbol{\gamma}_{k+\tau}^*+\boldsymbol{\gamma}_{k-\tau}^*)\mathbf{A}_k^{-1}\boldsymbol{\gamma}_k|^{2\ell}\Bigg]\\
&\quad\le Kv_n^{2\ell}\\
&\qquad+Kv_n^{-2\ell}\mathrm{E}\Big(\max_k\mathrm{E}_k I\big(|(\boldsymbol{\gamma}_{k+\tau}+\boldsymbol{\gamma}_{k-\tau})^*\mathbf{A}_k^{-2}(\mathbf{A}_k^*)^{-2}(\boldsymbol{\gamma}_{k+\tau}+\boldsymbol{\gamma}_{k-\tau})|\ge K\big)\Big)^\ell\\
&\qquad+Kv_n^{2\ell}\\
&\qquad+K\mathrm{E}\Big(\max_k\mathrm{E}_k I\big(|(\boldsymbol{\gamma}_{k+\tau}+\boldsymbol{\gamma}_{k-\tau})^*\mathbf{A}_k^{-1}(\mathbf{A}_k^*)^{-1}(\boldsymbol{\gamma}_{k+\tau}+\boldsymbol{\gamma}_{k-\tau})|\ge K\big)\Big)^\ell\\
&\qquad+Kv_n^{2\ell}(T^{1-\ell}v_n^{-4\ell}+T^{1-\ell}v_n^{-2\ell})\\
&\quad\le Kv_n^{2\ell}\end{aligned}$$

$$
\begin{aligned}
&+ K v_n^{-2\ell} \sum_{k=1}^{T} \mathrm{E}\big(\mathrm{E}_k I\big(\big|(\boldsymbol{\gamma}_{k+\tau} + \boldsymbol{\gamma}_{k-\tau})^* \mathbf{A}_k^{-2} (\mathbf{A}_k^*)^{-2} (\boldsymbol{\gamma}_{k+\tau} + \boldsymbol{\gamma}_{k-\tau})\big| \geq K\big)\big) \\
&+ K \sum_{k=1}^{T} \mathrm{E}\big(\mathrm{E}_k I\big(\big|(\boldsymbol{\gamma}_{k+\tau} + \boldsymbol{\gamma}_{k-\tau})^* \mathbf{A}_k^{-1} (\mathbf{A}_k^*)^{-1} (\boldsymbol{\gamma}_{k+\tau} + \boldsymbol{\gamma}_{k-\tau})\big| \geq K\big)\big) \\
&\leq K v_n^{2\ell},
\end{aligned}
$$

where Lemmas 3.9 and 3.10 are used in the last estimation. By similar arguments, one can show that

$$
v_n^{2\ell} \left[\mathrm{E}\left(\sum_{k=1}^{T} \mathrm{E}_k \left| \frac{d}{dz} \alpha_{k2} \right|^2 \right)^{\ell} + \sum_{k=1}^{T} \mathrm{E} \left| \frac{d}{dz} \alpha_{k2} \right|^{2\ell} \right] \leq K v_n^{2\ell}.
$$

By Remarks 3.1, 3.2, (5.8), (5.9) and Lemmas 2.5 and 3.5 we have

$$
\begin{aligned}
&v_n^{2\ell} \left[\mathrm{E}\left(\sum_{k=1}^{T} \mathrm{E}_k \left| \frac{d}{dz} \alpha_{k3} \right|^2 \right)^{\ell} + \sum_{k=1}^{T} \mathrm{E} \left| \frac{d}{dz} \alpha_{k3} \right|^{2\ell} \right] \\
&\quad \leq K v_n^{2\ell} \Bigg[\mathrm{E}\Bigg(\sum_{k=1}^{T} \mathrm{E}_k \left| \boldsymbol{\gamma}_k^* \mathbf{A}_k^{-2} \boldsymbol{\gamma}_k - \frac{1}{2T} \operatorname{tr} \mathbf{A}_k^{-2} \right|^2 \\
&\qquad\qquad \times \big|(\boldsymbol{\gamma}_{k+\tau} + \boldsymbol{\gamma}_{k-\tau})^* \mathbf{A}_k^{-1} (\boldsymbol{\gamma}_{k+\tau} + \boldsymbol{\gamma}_{k-\tau})\big|^2 \\
&\qquad\qquad + \sum_{k=1}^{T} \mathrm{E}_k \left| \boldsymbol{\gamma}_k^* \mathbf{A}_k^{-1} \boldsymbol{\gamma}_k - \frac{1}{2T} \operatorname{tr} \mathbf{A}_k^{-1} \right|^2 \\
&\qquad\qquad\qquad \times \big|(\boldsymbol{\gamma}_{k+\tau} + \boldsymbol{\gamma}_{k-\tau})^* \mathbf{A}_k^{-2} (\boldsymbol{\gamma}_{k+\tau} + \boldsymbol{\gamma}_{k-\tau})\big|^2 \Bigg)^{\ell} \\
&\qquad + \sum_{k=1}^{T} \mathrm{E} \left| \boldsymbol{\gamma}_k^* \mathbf{A}_k^{-2} \boldsymbol{\gamma}_k - \frac{1}{2T} \operatorname{tr} \mathbf{A}_k^{-2} \right|^{2\ell} \\
&\qquad\qquad \times \big|(\boldsymbol{\gamma}_{k+\tau} + \boldsymbol{\gamma}_{k-\tau})^* \mathbf{A}_k^{-1} (\boldsymbol{\gamma}_{k+\tau} + \boldsymbol{\gamma}_{k-\tau})\big|^{2\ell} \\
&\qquad + \sum_{k=1}^{T} \mathrm{E} \left| \boldsymbol{\gamma}_k^* \mathbf{A}_k^{-1} \boldsymbol{\gamma}_k - \frac{1}{2T} \operatorname{tr} \mathbf{A}_k^{-1} \right|^{2\ell} \\
&\qquad\qquad\qquad \times \big|(\boldsymbol{\gamma}_{k+\tau} + \boldsymbol{\gamma}_{k-\tau})^* \mathbf{A}_k^{-2} (\boldsymbol{\gamma}_{k+\tau} + \boldsymbol{\gamma}_{k-\tau})\big|^{2\ell} \Bigg] \\
&\quad \leq K v_n^{2\ell} \mathrm{E}\left(\left(\sum_{k=1}^{T} \frac{1}{4T^2} \mathrm{E}_k \operatorname{tr} \mathbf{A}_k^{-2} \bar{\mathbf{A}}_k^{-2} \big|(\boldsymbol{\gamma}_{k+\tau} + \boldsymbol{\gamma}_{k-\tau})^* \mathbf{A}_k^{-1} (\boldsymbol{\gamma}_{k+\tau} + \boldsymbol{\gamma}_{k-\tau})\big|^2 \right)^{\ell} \right.
\end{aligned}
$$

$$
\begin{aligned}
&+\Biggl(\sum_{k=1}^{T}\frac{1}{4T^2}\mathrm{E}_k\operatorname{tr}\mathbf{A}_k^{-1}\bar{\mathbf{A}}_k^{-1} \\
&\qquad\times\bigl|(\boldsymbol{\gamma}_{k+\tau}+\boldsymbol{\gamma}_{k-\tau})^*\mathbf{A}_k^{-2}(\boldsymbol{\gamma}_{k+\tau}+\boldsymbol{\gamma}_{k-\tau})\bigr|^2\Biggr)^{\ell} \\
&+\sum_{k=1}^{T}\mathrm{E}\biggl(\frac{1}{4T^2}\operatorname{tr}\mathbf{A}_k^{-2}\bar{\mathbf{A}}_k^{-2} \\
&\qquad\times\bigl|(\boldsymbol{\gamma}_{k+\tau}+\boldsymbol{\gamma}_{k-\tau})^*\mathbf{A}_k^{-1}(\boldsymbol{\gamma}_{k+\tau}+\boldsymbol{\gamma}_{k-\tau})\bigr|^2\biggr)^{\ell} \\
&+\sum_{k=1}^{T}\mathrm{E}\biggl(\frac{1}{4T^2}\operatorname{tr}\mathbf{A}_k^{-1}\bar{\mathbf{A}}_k^{-1} \\
&\qquad\qquad\times\bigl|(\boldsymbol{\gamma}_{k+\tau}+\boldsymbol{\gamma}_{k-\tau})^*\mathbf{A}_k^{-2}(\boldsymbol{\gamma}_{k+\tau}+\boldsymbol{\gamma}_{k-\tau})\bigr|^2\biggr)^{\ell}\Biggr) \\
&\le K v_n^{2\ell}.
\end{aligned}
$$

By (5.6) and similar arguments, we have

$$
\begin{aligned}
&v_n^{2\ell}\Biggl[\mathrm{E}\Biggl(\sum_{k=1}^{T}\mathrm{E}_k\biggl|\frac{d}{dz}\alpha_{k4}\biggr|^2\Biggr)^{\ell}+\sum_{k=1}^{T}\mathrm{E}\biggl|\frac{d}{dz}\alpha_{k4}\biggr|^{2\ell}\Biggr] \\
&\quad\le K v_n^{2\ell}\Biggl[\mathrm{E}\Biggl(\frac{1}{v_n^2}\sup_{|\xi-z|=v_n/2}\sum_{k=1}^{T}\mathrm{E}_k\bigl(\bigl|\alpha_{k1}(\xi)\bigr|^4+\bigl|\alpha_{k2}(\xi)\bigr|^4 \\
&\qquad\qquad+\bigl|\alpha_{k3}(\xi)\bigr|^4+\bigl|r_k(\xi)\bigr|^4\bigr)\Biggr)^{\ell} \\
&\qquad+\frac{1}{v_n^{2\ell}}\sup_{|\xi-z|=v_n/2}\sum_{k=1}^{T}\mathrm{E}\bigl(\bigl|\alpha_{k1}(\xi)\bigr|^{4\ell}+\bigl|\alpha_{k2}(\xi)\bigr|^{4\ell} \\
&\qquad\qquad+\bigl|\alpha_{k3}(\xi)\bigr|^{4\ell}+\bigl|r_k(\xi)\bigr|^{4\ell}\bigr)\Biggr] \\
&\quad\le K T^{-\ell}v_n^{-4\ell}.
\end{aligned}
$$

Finally, by measurable properties of some terms of r_k, we have

$$
(\mathrm{E}_{k-1}-\mathrm{E}_k)r_k=(\mathrm{E}_{k-1}-\mathrm{E}_k)\frac{\varepsilon_1\varepsilon_2}{x_{n1}-x_{n0}}
$$

from which and similar argument for α_{k1} and α_{k2}, we conclude that

$$v_n^{2\ell}\mathrm{E}\left|\frac{d}{dz}\sum_{k=1}^{T}(\mathrm{E}_{k-1}-\mathrm{E}_k)r_k\right|^{2\ell}=KT^{-\ell}v_n^{-4\ell}.$$

Substituting the five upper-bounds into (5.7), we have

$$\begin{aligned}
&\mathrm{P}\Bigl(\max_{u\in S_n}\bigl|nv_n\bigl(m_n(z)-\mathrm{E}m_n(z)\bigr)\bigr|>\varepsilon\Bigr)\\
&\qquad=Kn^2\mathrm{E}\bigl|nv_n\bigl(m_n(z)-\mathrm{E}m_n(z)\bigr)\bigr|^{2\ell}\\
&\qquad\le Kn^2\bigl(v_n^{2\ell}+v_n^{-4\ell}T^{-\ell}\bigr)
\end{aligned}$$

which is summable when $\ell>318$ and $v_n\ge n^{-\alpha}$ for $\alpha=1/212$. Therefore, we have proved that $\max_{u\in[a,b]}|m_n(z)-\mathrm{E}m_n(z)|=o(\frac{1}{nv_n})$ a.s.

5.2. *A refined convergence rate of* $\mathrm{E}m_n(z)-m_n^0(z)$. To show

$$\sup_{u\in[a,b]}\bigl|\mathrm{E}m_n(z)-m_n^0(z)\bigr|=o\biggl(\frac{1}{nv_n}\biggr),$$

we follow the notation and expressions in Section 4.2. Recall

$$\begin{aligned}
&c_n+c_nz\mathrm{E}m_n(z)\\
&\quad=\frac{1}{T}\sum_{k=1}^{T}\biggl[1-\mathrm{E}\frac{1}{1+\boldsymbol{\gamma}_k^*\widetilde{\mathbf{A}}_k^{-1}(\boldsymbol{\gamma}_{k+\tau}+\boldsymbol{\gamma}_{k-\tau})}\biggr]\\
&\quad=\frac{1}{T}\sum_{k=1}^{T}\biggl[1-\mathrm{E}\biggl(1\Big/\biggl(1+\boldsymbol{\gamma}_k^*\mathbf{A}_k^{-1}(\boldsymbol{\gamma}_{k+\tau}+\boldsymbol{\gamma}_{k-\tau})\\
&\qquad\qquad-\frac{\boldsymbol{\gamma}_k^*\mathbf{A}_k^{-1}\boldsymbol{\gamma}_k(\boldsymbol{\gamma}_{k+\tau}^*+\boldsymbol{\gamma}_{k-\tau}^*)\mathbf{A}_k^{-1}(\boldsymbol{\gamma}_{k+\tau}+\boldsymbol{\gamma}_{k-\tau})}{1+(\boldsymbol{\gamma}_{k+\tau}^*+\boldsymbol{\gamma}_{k-\tau}^*)\mathbf{A}_k^{-1}\boldsymbol{\gamma}_k}\biggr)\biggr)\biggr]\\
&\quad=1-\frac{1}{x_{n1}-x_{n0}}+\delta_n,
\end{aligned}\tag{5.10}$$

where

$$\delta_n=\frac{1}{T}\sum_{k=1}^{T}\mathrm{E}\eta_k$$

with

$$\begin{aligned}
\eta_k=-\biggl(&1\Big/\biggl(1+\boldsymbol{\gamma}_k^*\mathbf{A}_k^{-1}(\boldsymbol{\gamma}_{k+\tau}+\boldsymbol{\gamma}_{k-\tau})\\
&-\frac{\boldsymbol{\gamma}_k^*\mathbf{A}_k^{-1}\boldsymbol{\gamma}_k(\boldsymbol{\gamma}_{k+\tau}^*+\boldsymbol{\gamma}_{k-\tau}^*)\mathbf{A}_k^{-1}(\boldsymbol{\gamma}_{k+\tau}+\boldsymbol{\gamma}_{k-\tau})}{1+(\boldsymbol{\gamma}_{k+\tau}^*+\boldsymbol{\gamma}_{k-\tau}^*)\mathbf{A}_k^{-1}\boldsymbol{\gamma}_k}\biggr)-\frac{1}{x_{n1}-x_{n0}}\biggr).
\end{aligned}$$

Consider expressions of (4.7) and (4.8). To apply Lemma 3.2, we only need to show $|\delta_n| = o(\frac{1}{nv_n})$, which can be reduced to showing $|\mathrm{E}\eta_k| = o(\frac{1}{nv_n})$ for $\log^2 n < k < T - \log^2 n$ and $|\mathrm{E}\eta_k| = O(1)$ for $k \le \log^2 n$ or $\ge T - \log^2 n$.

When $\log^2 n < k < T - \log^2 n$, rewrite η_k as

$$\begin{aligned}
-\eta_k &= 1\Big/\Big(1 + \boldsymbol{\gamma}_k^* \mathbf{A}_k^{-1}(\boldsymbol{\gamma}_{k+\tau} + \boldsymbol{\gamma}_{k-\tau}) \\
&\qquad - \frac{\boldsymbol{\gamma}_k^* \mathbf{A}_k^{-1}\boldsymbol{\gamma}_k(\boldsymbol{\gamma}_{k+\tau}^* + \boldsymbol{\gamma}_{k-\tau}^*)\mathbf{A}_k^{-1}(\boldsymbol{\gamma}_{k+\tau} + \boldsymbol{\gamma}_{k-\tau})}{1 + (\boldsymbol{\gamma}_{k+\tau}^* + \boldsymbol{\gamma}_{k-\tau}^*)\mathbf{A}_k^{-1}\boldsymbol{\gamma}_k}\Big) - \frac{1}{1 - (2a_n^2/x_{n1})} \\
&= \big(1 + (\boldsymbol{\gamma}_{k+\tau}^* + \boldsymbol{\gamma}_{k-\tau}^*)\mathbf{A}_k^{-1}\boldsymbol{\gamma}_k\big) \\
&\quad /\big(\big(1 + \boldsymbol{\gamma}_k^* \mathbf{A}_k^{-1}(\boldsymbol{\gamma}_{k+\tau} + \boldsymbol{\gamma}_{k-\tau})\big)\big(1 + (\boldsymbol{\gamma}_{k+\tau}^* + \boldsymbol{\gamma}_{k-\tau}^*)\mathbf{A}_k^{-1}\boldsymbol{\gamma}_k\big) \\
&\qquad - \boldsymbol{\gamma}_k^* \mathbf{A}_k^{-1}\boldsymbol{\gamma}_k\big(\boldsymbol{\gamma}_{k+\tau}^* + \boldsymbol{\gamma}_{k-\tau}^*)\mathbf{A}_k^{-1}(\boldsymbol{\gamma}_{k+\tau} + \boldsymbol{\gamma}_{k-\tau})\big) \\
&\quad - \frac{1}{1 - (2a_n^2/x_{n1})} \\
&= (1 + \varepsilon_1) \\
&\quad \Big/\Big(1 + \varepsilon_1 + \varepsilon_2 + \varepsilon_1\varepsilon_2 \\
&\qquad - (\boldsymbol{\gamma}_{k+\tau}^* + \boldsymbol{\gamma}_{k-\tau}^*)\mathbf{A}_k^{-1}(\boldsymbol{\gamma}_{k+\tau} + \boldsymbol{\gamma}_{k-\tau})(\varepsilon_3 + \varepsilon_4) - a_n\varepsilon_5 - \frac{2a_n^2}{x_{n1}}\Big) \\
&\quad - \frac{1}{1 - (2a_n^2/x_{n1})} \\
&= \frac{1}{1 - (2a_n^2/x_{n1})} \\
&\quad \times \Big(-\varepsilon_1 \frac{2a_n^2}{x_{n1}} - \varepsilon_2 - \varepsilon_1\varepsilon_2 \\
&\qquad + (\boldsymbol{\gamma}_{k+\tau}^* + \boldsymbol{\gamma}_{k-\tau}^*)\mathbf{A}_k^{-1}(\boldsymbol{\gamma}_{k+\tau} + \boldsymbol{\gamma}_{k-\tau})(\varepsilon_3 + \varepsilon_4) + a_n\varepsilon_5\Big) \\
&\qquad \Big/\Big(1 + \varepsilon_1 + \varepsilon_2 + \varepsilon_1\varepsilon_2 \\
&\qquad\quad - (\boldsymbol{\gamma}_{k+\tau}^* + \boldsymbol{\gamma}_{k-\tau}^*)\mathbf{A}_k^{-1}(\boldsymbol{\gamma}_{k+\tau} + \boldsymbol{\gamma}_{k-\tau})(\varepsilon_3 + \varepsilon_4) - a_n\varepsilon_5 - \frac{2a_n^2}{x_{n1}}\Big),
\end{aligned}$$

where ε_i's are defined as in Section 4.2.

For simplicity, denote $\tilde{\varepsilon} = \varepsilon_2 + \varepsilon_1\varepsilon_2 - (\boldsymbol{\gamma}_{k+\tau}^* + \boldsymbol{\gamma}_{k-\tau}^*)\mathbf{A}_k^{-1}(\boldsymbol{\gamma}_{k+\tau} + \boldsymbol{\gamma}_{k-\tau})(\varepsilon_3 + \varepsilon_4) - a_n\varepsilon_5$. Applying the identity $\frac{x}{1+x+y} = \frac{x}{1+y} - \frac{x^2}{(1+x+y)(1+y)}$ repeatedly, we have

$$\begin{aligned}
-\eta_k &= \frac{1}{1-(2a_n^2/x_{n1})} \times \frac{-\varepsilon_1(2a_n^2/x_{n1}) - \tilde{\varepsilon}}{1+\varepsilon_1+\tilde{\varepsilon}-(2a_n^2/x_{n1})} \\
&= -\frac{2a_n^2/x_{n1}}{1-(2a_n^2/x_{n1})} \times \frac{\varepsilon_1+\tilde{\varepsilon}}{1+\varepsilon_1+\tilde{\varepsilon}-(2a_n^2/x_{n1})} - \frac{\tilde{\varepsilon}}{1+\varepsilon_1+\tilde{\varepsilon}-(2a_n^2/x_{n1})} \\
&= -\frac{2a_n^2/x_{n1}}{1-(2a_n^2/x_{n1})} \\
&\quad \times \left(\frac{\varepsilon_1+\tilde{\varepsilon}}{1-(2a_n^2/x_{n1})} - \frac{(\varepsilon_1+\tilde{\varepsilon})^2}{(1-(2a_n^2/x_{n1}))(1+\varepsilon_1+\tilde{\varepsilon}-(2a_n^2/x_{n1}))}\right) \\
&\quad - \left(\frac{\tilde{\varepsilon}}{1+\varepsilon_1-(2a_n^2/x_{n1})}\right. \\
&\qquad \left. - \frac{\tilde{\varepsilon}^2}{(1+\varepsilon_1-(2a_n^2/x_{n1}))(1+\varepsilon_1+\tilde{\varepsilon}-(2a_n^2/x_{n1}))}\right) \\
&= -\frac{2a_n^2/x_{n1}}{1-(2a_n^2/x_{n1})} \\
&\quad \times \left(\frac{\varepsilon_1+\tilde{\varepsilon}}{1-(2a_n^2/x_{n1})} - \frac{(\varepsilon_1+\tilde{\varepsilon})^2}{(1-(2a_n^2/x_{n1}))(1+\varepsilon_1+\tilde{\varepsilon}-(2a_n^2/x_{n1}))}\right) \\
&\quad - \left(\frac{\tilde{\varepsilon}}{1-(2a_n^2/x_{n1})} - \frac{\tilde{\varepsilon}\varepsilon_1}{(1+\varepsilon_1-(2a_n^2/x_{n1}))(1-(2a_n^2/x_{n1}))}\right) \\
&\quad + \frac{\tilde{\varepsilon}^2}{(1+\varepsilon_1-(2a_n^2/x_{n1}))(1+\varepsilon_1+\tilde{\varepsilon}-(2a_n^2/x_{n1}))}.
\end{aligned}$$

Therefore, by Lemma 3.6(iv)(b), we have $|-\frac{2a_n^2/x_{n1}}{1-(2a_n^2/x_{n1})}| = |\frac{2x_{n0}}{x_{n1}-x_{n0}}| \le |\frac{2x_{n1}}{x_{n1}-x_{n0}}|$ is bounded. Together with the fact that all the denominators being bounded below and the Cauchy–Schwarz inequality, to show $|\mathrm{E}\eta_k| = o(\frac{1}{nv_n})$, it suffices to show $|\mathrm{E}\varepsilon_1|$, $|\mathrm{E}\tilde{\varepsilon}|$, $|\mathrm{E}\varepsilon_1^2|$, $|\mathrm{E}\tilde{\varepsilon}^2|$ are of $o(\frac{1}{nv_n})$. As $|\mathrm{E}\varepsilon_i| = 0$ for $i = 1, 2, 3$, it is clear that the above convergence rates achieve $o(\frac{1}{nv_n})$ provided that so do $\mathrm{E}|\varepsilon_i|^2$, $i = 1, 2, 3, 4, 5$, $|\mathrm{E}\varepsilon_4|$ and $|\mathrm{E}\varepsilon_5|$ for $\log^2 n < k < T - \log^2 n$.

When $\log^2 n < k < T - \log^2 n$, for $i = 1$, by Lemma 3.9, we have, for any $t > 0$,

$$\begin{aligned}
\mathrm{E}|(\boldsymbol{\gamma}_{k+\tau} + \boldsymbol{\gamma}_{k-\tau})^*\mathbf{A}_k^{-1}\boldsymbol{\gamma}_k|^2 &= \frac{1}{2T}\mathrm{E}(\boldsymbol{\gamma}_{k+\tau} + \boldsymbol{\gamma}_{k-\tau})^*\mathbf{A}_k^{-1}(\mathbf{A}_k^*)^{-1}(\boldsymbol{\gamma}_{k+\tau} + \boldsymbol{\gamma}_{k-\tau}) \\
&= \frac{K}{T} + v_n^{-2}o(n^{-t}) = O(1/n) = o\left(\frac{1}{nv_n}\right).
\end{aligned}$$

Similarly, for $i=2$, $\mathrm{E}|\varepsilon_2|^2 = O(1/n) = o(\frac{1}{nv_n})$.

For $i=3$, by Lemmas 2.5 and 3.5, we have

$$\begin{aligned}
\mathrm{E}|\varepsilon_3|^2 &= \mathrm{E}\left|\boldsymbol{\gamma}_k^*\mathbf{A}_k^{-1}\boldsymbol{\gamma}_k - \frac{1}{2T}\operatorname{tr}\mathbf{A}_k^{-1}\right|^2 \le \frac{K}{4T^2}\mathrm{E}\left|\operatorname{tr}\mathbf{A}_k^{-1}(\mathbf{A}_k^*)^{-1}\right| \\
&= \frac{K}{4T^2}\mathrm{E}\sum\frac{1}{|\lambda_{kj}-z|^2} \\
&\le \frac{K}{2T} + \frac{K}{Tv_n^2}F_n([a',b']) \le \frac{K}{T} + o(T^{-1}) = O(1/n) = o\left(\frac{1}{nv_n}\right).
\end{aligned}$$

For $|\mathrm{E}\varepsilon_4|$, by Lemma 3.11 we have

$$|\mathrm{E}\varepsilon_4| = \left|\frac{1}{2T}\mathrm{E}\operatorname{tr}\mathbf{A}_k^{-1} - a_n\right| = \frac{1}{2T}\left|\mathrm{E}(\operatorname{tr}\mathbf{A}_k^{-1} - \operatorname{tr}\mathbf{A}^{-1})\right| = O(T^{-1}) = o\left(\frac{1}{nv_n}\right).$$

For $\mathrm{E}|\varepsilon_4|^2$, by (4.2) and the convergence rate obtained in Section 5.1, we have

$$\begin{aligned}
&\mathrm{E}\left|\frac{1}{2T}\operatorname{tr}\mathbf{A}_k^{-1} - a_n\right|^2 \\
&\quad\le 2\mathrm{E}\left|\frac{1}{2T}\operatorname{tr}\mathbf{A}_k^{-1} - \mathrm{E}\frac{1}{2T}\operatorname{tr}\mathbf{A}_k^{-1}\right|^2 + 2\left|\frac{1}{2T}\mathrm{E}\operatorname{tr}\mathbf{A}_k^{-1} - a_n\right|^2 \\
&\quad\le \frac{K}{n^2v_n^2} + O(n^{-1}) = o\left(\frac{1}{nv_n}\right).
\end{aligned}$$

Bounds of $|\mathrm{E}\varepsilon_5|$ and $\mathrm{E}|\varepsilon_5|^2$ will follow Lemmas 3.7(b2), (b3) and 3.8(b1), (b2).

To show $|\mathrm{E}\eta_k| = O(1)$ when $k \le \log^2 n$ or $\ge T - \log^2 n$, we just prove the case for $k \ge T - \log^2 n$, as the case for $k \le \log^2 n$ follows by symmetry.

When $k \ge T - \log^2 n$, by Lemma 3.7(b1), we have $\mathrm{P}(|\boldsymbol{\gamma}_{k+\tau}^*\mathbf{A}_k^{-1}\boldsymbol{\gamma}_{k+\tau}| \ge 1-\eta) = o(n^{-t})$. By Lemma 3.7(a), we have $\mathrm{P}(|\boldsymbol{\gamma}_{k-\tau}^*\mathbf{A}_k^{-1}\boldsymbol{\gamma}_{k-\tau} - \frac{c_n\mathrm{E}m_n}{2x_{n1}}| \ge v_n^6) = o(n^{-t})$, by Lemma 3.4, $\mathrm{P}(|\boldsymbol{\gamma}_k^*\mathbf{A}_k^{-1}\boldsymbol{\gamma}_{k\pm\tau}| \ge v_n^3) = o(n^{-t})$, and by Lemmas 2.5 and inequalities (4.2) and (4.3), $\mathrm{P}(|\boldsymbol{\gamma}_k^*\mathbf{A}_k^{-1}\boldsymbol{\gamma}_k - a_n| \ge v_n^3) = o(n^{-t})$. By Lemma 3.8(a), $\mathrm{P}(|\boldsymbol{\gamma}_{k\pm\tau}^*\mathbf{A}_k^{-1}\boldsymbol{\gamma}_{k\mp\tau}| \ge v_n^6) = o(n^{-t})$. By Lemma 3.6(ii)(b) and (iv)(b), we have $|\frac{1}{x_{n1}-x_{n0}}| \le K$ and $|\mathrm{E}\eta_k| \le Kv_n^{-1}$. Substitute the above results into the definition of η_k, and we finally have

$$\begin{aligned}
|\mathrm{E}\eta_k| \le{}& \Bigg|\mathrm{E}\Bigg(1\Big/\Bigg(1 + \boldsymbol{\gamma}_k^*\mathbf{A}_k^{-1}(\boldsymbol{\gamma}_{k+\tau} + \boldsymbol{\gamma}_{k-\tau}) \\
&\qquad - \frac{\boldsymbol{\gamma}_k^*\mathbf{A}_k^{-1}\boldsymbol{\gamma}_k(\boldsymbol{\gamma}_{k+\tau}^* + \boldsymbol{\gamma}_{k-\tau}^*)\mathbf{A}_k^{-1}(\boldsymbol{\gamma}_{k+\tau} + \boldsymbol{\gamma}_{k-\tau})}{1 + (\boldsymbol{\gamma}_{k+\tau}^* + \boldsymbol{\gamma}_{k-\tau}^*)\mathbf{A}_k^{-1}\boldsymbol{\gamma}_k}\Bigg)\Bigg)\Bigg| \\
&+ \left|\frac{1}{x_{n1} - x_{n0}}\right|
\end{aligned}$$

$$\le \left|\frac{1+v_n^3}{(1-2v_n^3)-(1/2-\eta+v_n^3)(1-\eta+3v_n^3+|a_n|/|x_{n1}|)}\right| + K + Kv_n^{-1}o(n^{-t}) = O(1).$$

6. Completing the proof. In this section, we follow the idea of Bai and Silverstein (1998) and give the main steps here. From what has been obtained in the last two sections, we have, with $v_n = n^{-1/212}$,

$$\sup_{u\in[a,b]}\left|m_n(z)-m_n^0(z)\right| = o\left(\frac{1}{nv_n}\right) \qquad \text{a.s.} \tag{6.1}$$

It is clear from the last two sections that (6.1) is true when $\Im(z)$ is replaced by a constant multiple of v_n. In fact, we have

$$\max_{k\in\{1,2,\dots,106\}}\sup_{u\in[a,b]}\left|m_n(u+i\sqrt{k}v_n)-m_n^0(u+i\sqrt{k}v_n)\right| = o(v_n^{211}) \qquad \text{a.s.}$$

Taking the imaginary part, we get

$$\max_{k\in\{1,2,\dots,106\}}\sup_{u\in[a,b]}\left|\int\frac{d(F_n(\lambda)-F_n^0(\lambda))}{(u-\lambda)^2+kv_n^2}\right| = o(v_n^{210}) \qquad \text{a.s.}$$

After taking difference, we obtain

$$\max_{k_1\neq k_2}\sup_{u\in[a,b]}\left|\int\frac{v_n^2 d(F_n(\lambda)-F_n^0(\lambda))}{((u-\lambda)^2+k_1v_n^2)((u-\lambda)^2+k_2v_n^2)}\right| = o(v_n^{210}) \qquad \text{a.s.}$$

$$\vdots$$

$$\sup_{u\in[a,b]}\left|\int\frac{(v_n^2)^{105}d(F_n(\lambda)-F_n^0(\lambda))}{((u-\lambda)^2+v_n^2)((u-\lambda)^2+2v_n^2)\cdots((u-\lambda)^2+106v_n^2)}\right| = o(v_n^{210}) \qquad \text{a.s.}$$

Therefore,

$$\sup_{u\in[a,b]}\left|\int\frac{d(F_n(\lambda)-F_n^0(\lambda))}{((u-\lambda)^2+v_n^2)((u-\lambda)^2+2v_n^2)\cdots((u-\lambda)^2+106v_n^2)}\right| = o(1) \qquad \text{a.s.}$$

After splitting the integral, we get

$$\sup_{u\in[a,b]}\left|\int\frac{I_{[a',b']^c}(\lambda)d(F_n(\lambda)-F_n^0(\lambda))}{((u-\lambda)^2+v_n^2)((u-\lambda)^2+2v_n^2)\cdots((u-\lambda)^2+106v_n^2)}\right.$$
$$\left.+\sum_{\lambda_j\in[a',b']}\frac{v_n^{212}}{((u-\lambda_j)^2+v_n^2)((u-\lambda_j)^2+2v_n^2)\cdots((u-\lambda_j)^2+106v_n^2)}\right|$$
$$= o(1) \qquad \text{a.s.}$$

Note that the first term tends to 0 by dominated convergence theorem. Now, if there is at least one eigenvalue contained in $[a,b]$, then the second sum will be away from zero when u takes one of such eigenvalues. This contradicts the right-hand side. Therefore, with probability 1, there are no eigenvalues of $\mathbf{M}_n$ in $[a,b]$ for all n large and the proof is complete.

APPENDIX A: JUSTIFICATION OF TRUNCATION, CENTRALIZATION AND RESCALING

Here, we give some justifications of (1.4), which will be divided into two parts.

A.1. Truncation and centralization. Fix some $C > 0$, define $\hat{\varepsilon}_{it} = \varepsilon_{it} I_{\{|x_{it}| \le C\}} - \mathrm{E}\varepsilon_{it} I_{\{|x_{it}| \le C\}}$, $\hat{\boldsymbol{\gamma}}_k = \frac{1}{\sqrt{2T}}(\hat{\varepsilon}_{1k}, \dots, \hat{\varepsilon}_{nk})' \equiv \frac{1}{\sqrt{2T}}\hat{\mathbf{e}}_k$, $\hat{\mathbf{E}} = (\hat{\mathbf{e}}_1, \dots, \hat{\mathbf{e}}_T)$, $\hat{\mathbf{E}}_\tau = (\hat{\mathbf{e}}_{1+\tau}, \dots, \hat{\mathbf{e}}_{T+\tau})$ and $\hat{\mathbf{M}}_n = \sum_{k=1}^T (\hat{\boldsymbol{\gamma}}_k \hat{\boldsymbol{\gamma}}_{k+\tau}^* + \hat{\boldsymbol{\gamma}}_{k+\tau}\hat{\boldsymbol{\gamma}}_k^*) = \frac{1}{2T}(\hat{\mathbf{E}}\hat{\mathbf{E}}_\tau^* + \hat{\mathbf{E}}_\tau \hat{\mathbf{E}}^*)$. By Theorem A.46 of Bai and Silverstein (2010),

$$\begin{aligned}
&\max_k \bigl|\lambda_k(\hat{\mathbf{M}}_n) - \lambda_k(\mathbf{M}_n)\bigr| \\
&\quad \le \|\hat{\mathbf{M}}_n - \mathbf{M}_n\| \\
&\quad = \frac{1}{2T}\bigl\|(\mathbf{E} - \hat{\mathbf{E}})\hat{\mathbf{E}}_\tau^* + \hat{\mathbf{E}}_\tau(\mathbf{E} - \hat{\mathbf{E}})^* + \mathbf{E}(\mathbf{E}_\tau - \hat{\mathbf{E}}_\tau)^* + (\mathbf{E}_\tau - \hat{\mathbf{E}}_\tau)\mathbf{E}^*\bigr\| \\
&\quad \le \frac{1}{T}\bigl(\|\mathbf{E} - \hat{\mathbf{E}}\|\|\hat{\mathbf{E}}_\tau\| + \|\mathbf{E} - \hat{\mathbf{E}}\|\|\mathbf{E}\|\bigr).
\end{aligned}$$

By a similar approach as in Yin, Bai and Krishnaiah (1988), one can show that almost surely

$$\limsup_n \frac{1}{\sqrt{T}}\|\mathbf{E}\| \le (1+\sqrt{c})^2,$$

$$\limsup_n \frac{1}{\sqrt{T}}\|\hat{\mathbf{E}}_\tau\| \le (1+\sqrt{c})^2$$

and

$$\begin{aligned}
&\limsup_n \frac{1}{\sqrt{T}}\|\mathbf{E} - \hat{\mathbf{E}}\| \\
&\quad \le (1+\sqrt{c})^2 \max_{i,t} \operatorname{var}(\varepsilon_{it} - \hat{\varepsilon}_{it}) \\
&\quad = (1+\sqrt{c})^2 \max_{i,t} \operatorname{var}(\varepsilon_{it} I_{\{|x_{it}| \ge C\}}) \\
&\quad \le (1+\sqrt{c})^2 \max_{i,t} \mathrm{E}(\varepsilon_{it} I_{\{|x_{it}| \ge C\}})^2 \\
&\quad \le \frac{(1+\sqrt{c})^2}{C^2} \max_{i,t} \mathrm{E}\varepsilon_{it}^4 \\
&\quad \le \frac{(1+\sqrt{c})^2 M}{C^2},
\end{aligned}$$

which can be arbitrarily small by choosing C large enough. This verifies the truncation at a fixed point and centralization.

A.2. Rescaling. Define $\sigma_{it}^2 = \mathrm{E}|\hat{\varepsilon}_{it}|^2$, $\check{\varepsilon}_{it} = \hat{\varepsilon}_{it}/\sigma_{it}$, $\check{\boldsymbol{\gamma}}_k = \frac{1}{\sqrt{2T}}(\check{\varepsilon}_{1k}, \dots, \check{\varepsilon}_{nk})' \equiv \frac{1}{\sqrt{2T}}\check{\mathbf{e}}_k$, $\check{\mathbf{E}} = (\check{\mathbf{e}}_1, \dots, \check{\mathbf{e}}_T)$, $\check{\mathbf{E}}_\tau = (\check{\mathbf{e}}_{1+\tau}, \dots, \check{\mathbf{e}}_{T+\tau})$, $\mathbf{D} = (\sigma_{it}^{-1})_{n\times T}$, $\mathbf{D}_\tau = (\sigma_{i(t+\tau)}^{-1})_{n\times T}$ and $\check{\mathbf{M}}_n = \sum_{k=1}^T (\check{\boldsymbol{\gamma}}_k \check{\boldsymbol{\gamma}}_{k+\tau}^* + \check{\boldsymbol{\gamma}}_{k+\tau}\check{\boldsymbol{\gamma}}_k^*) = \frac{1}{2T}(\check{\mathbf{E}}\check{\mathbf{E}}_\tau^* + \check{\mathbf{E}}_\tau\check{\mathbf{E}}^*)$. By Theorem A.46 and Corollary A.21 of Bai and Silverstein (2010),

$$\begin{aligned}
&\max_k \bigl|\lambda_k(\check{\mathbf{M}}_\tau) - \lambda_k(\hat{\mathbf{M}}_\tau)\bigr| \\
&\quad \le \|\check{\mathbf{M}}_\tau - \hat{\mathbf{M}}_\tau\| \\
&\quad \le \frac{1}{T}\|\hat{\mathbf{E}}\circ(\mathbf{D}-\mathbf{J})\|\,\|\hat{\mathbf{E}}_\tau\circ(\mathbf{D}_\tau-\mathbf{J})\| \\
&\quad \le \frac{1}{T}\|\hat{\mathbf{E}}\|\,\|\hat{\mathbf{E}}_\tau\| \max_{i,t}\bigl(\sigma_{it}^{-1}-1\bigr)^2.
\end{aligned}$$

Here, $\circ$ denotes the Hadamard product and $\mathbf{J}$ is the $n\times T$ matrix of all entries 1.

From Yin, Bai and Krishnaiah (1988), we have, with probability 1 that $\limsup_n \frac{1}{T}\|\hat{\mathbf{E}}\|\,\|\hat{\mathbf{E}}_\tau\| \le (1+\sqrt{c})^4$.

Also, we have

$$\begin{aligned}
\max_{i,t}\bigl|1-\sigma_{it}^2\bigr| &\le \max_{i,t}\bigl(\mathrm{E}|\varepsilon_{it}|^2 I\bigl(|\varepsilon_{it}|>C\bigr) + \bigl(\mathrm{E}|\varepsilon_{it}|I\bigl(|\varepsilon_{it}|>C\bigr)\bigr)^2\bigr) \\
&\le \max_{i,t}\frac{2}{C^2}\mathrm{E}|\varepsilon_{it}|^4 \le \frac{2M}{C^2} \to 0 \qquad \text{as } C\to\infty.
\end{aligned}$$

Since $\min_{i,t}\sigma_{it} \to 1$ as $n\to\infty$ and thus $\sigma_{it}(1+\sigma_{it}) \ge 1$ for all large n. Therefore, we have

$$\sigma_{it}^{-1} - 1 = \frac{1-\sigma_{it}^2}{\sigma_{it}(1+\sigma_{it})} \le 1-\sigma_{it}^2,$$

which implies $\max_k |\lambda_k(\check{\mathbf{M}}_\tau) - \lambda_k(\hat{\mathbf{M}}_\tau)| \to 0$ as $n\to\infty$.

APPENDIX B: PROOFS OF LEMMAS IN SECTION 3

B.1. Proofs of Lemmas 3.1, 3.2 and 3.3. To show Lemma 3.1, take $d = \sqrt{\frac{1}{2m}}$ and denote S the total area covered by the m balls $B(x_i, dr_n)$, $i = 1, \dots, m$. Then we have $S \le m\pi(dr_n)^2 < \pi r_n^2$, which is the total area of $B(x_0, r_n)$. Therefore, such x must exist.

For Lemma 3.2, write $P_n(x) = \prod_{j=1}^k (x - x_{nj})$ and $P(x) = \prod_{j=1}^m (x-x_j)^{\ell_j}$. Let

$$\delta = \frac{1}{3}\min_{\substack{i,j\in\{1,\dots,m\}\\ i\ne j}} |x_i - x_j| > 0.$$

First, we claim that for any $i \in \{1, \dots, k\}$, there exists $j \in \{1, \dots, m\}$ such that $x_{ni} \in B(x_j, \delta)$. Suppose not, that is, there is some x_{ni} with $|x_{ni} - x_j| \geq \delta$ for any $j \in \{1, \dots, m\}$. Then it follows that $|P(x_{ni})| = \prod_{j=1}^m |x_{ni} - x_j|^{\ell_j} \geq \delta^k$. On the other hand, as $P_n(x_{ni}) = 0$, we have $Lr_n \geq |P_n(x_{ni}) - P(x_{ni})| = |P(x_{ni})|$. This is a contradiction.

Also, by our construction of δ, it follows that all the $B(x_j, \delta)$'s are disjoint.

Suppose the lemma is not true, then as the sum of ℓ_j's is fixed, there is at least one j such that, there are ℓ_0 x_{ni}'s in $B(x_j, r_n^{1/\ell_j})$, with $0 \leq \ell_0 < \ell_j$. WLOG, we can assume $j = 1$ and denote these ℓ_0 x_{ni}'s by $x_{n1}^1, \dots, x_{n\ell_0}^1$. By Lemma 3.1, we can choose $x^* \in B(x_1, r_n^{1/\ell_1})$ such that $\min_{i \in \{1,\dots,\ell_0\}} |x^* - x_{ni}^1| \geq dr_n^{1/\ell_1}$ for some $d > 0$. By the construction of δ, we have $|x^* - x| > \delta$ for all $x \in B(x_j, r_n^{1/\ell_j})$, $j = 2, \dots, m$. Therefore, we have $|P(x^*)| = \prod_{j=1}^m |x^* - x_j|^{\ell_j} = |x^* - x_1|^{\ell_1} \prod_{j=2}^m |x^* - x_j|^{\ell_j} = O(r_n)$. On the other hand, we have $|P_n(x^*)| = \prod_{j=1}^k |x^* - x_{nj}| = \prod_{i=1}^{\ell_0} |x^* - x_{ni}^1| \prod_{x_{nj} \notin B(x_1, r_n^{1/\ell_1})} |x^* - x_{nj}| > \delta^{k-\ell_0} r_n^{\ell_0/\ell_1}$, contradicting $|P(x^*) - P(x_n^*)| = O(r_n)$. Therefore, the lemma is proved.

For Lemma 3.3, write $P_n(x) = \prod_{j=1}^k (x - x_{nj})$, $Q_n(y) = \prod_{j=1}^k (y - y_{nj})$ and $P(x) = \prod_{j=1}^m (x - x_j)^{\ell_j}$. Let $\delta = \frac{1}{3} \min_{i,j \in \{1,\dots,m\}, i \neq j} |x_i - x_j| > 0$. By the definition of $\widetilde{r}_n$, there exists some $L > 0$ such that $L\widetilde{r}_n \geq |P_n(x_{ni}) - Q_n(x_{ni})|$ for all x_{ni}. Let $j \in \{1, \dots, m\}$ be given, and let $d := (\frac{L}{\delta^{k-\ell_j}})^{1/\ell_j} > 0$. By Lemma 3.2, we have exactly ℓ_j x_{ni}'s and exactly ℓ_j y_{ni}'s in $B(x_j, r_n^{1/\ell_j})$. Let $x_{ni} \in B(x_j, r_n^{1/\ell_j})$ be fixed. By our construction in the proof of Lemma 3.2, if $y_{nl} \notin B(x_j, r_n^{1/\ell_j})$, one has $d(x_{ni}, y_{nl}) > \delta$. Therefore, for the lemma to be true, we only need to look at those $y_{nl} \in B(x_j, r_n^{1/\ell_j})$ and show that at least one such y_{nl} satisfies the desired distance. Suppose not, that is, for this $x_{ni} \in B(x_j, r_n^{1/\ell_j})$, for any $y_{nl} \in B(x_j, r_n^{1/\ell_j})$, one has $d(x_{ni}, y_{nl}) > \widetilde{r}_n^{1/\ell_j}$. Note that when $y_{nl} \notin B(x_j, r_n^{1/\ell_j})$, we have $d(x_{ni}, y_{nl}) > \delta$. Hence, we have $|Q_n(x_{ni})| = \prod_{l=1}^k |x_{ni} - y_{nl}| > \delta^{k-\ell_j} (d\widetilde{r}_n^{1/\ell_j})^{\ell_j} = L\widetilde{r}_n$. However, we also have $L\widetilde{r}_n \geq |Q_n(x_{ni}) - P_n(x_{ni})| = |Q_n(x_{ni})|$, which is a contradiction.

B.2. Proof of Lemma 3.4. Let $\boldsymbol{\gamma}_l^* \mathbf{A}_k^{-s} = \mathbf{b} = (b_1, \dots, b_n)$. Noting $|\varepsilon_{it}| < C$, we have

$$
\begin{aligned}
&\mathrm{E}\left|\boldsymbol{\gamma}_l^* \mathbf{A}_k^{-s} \boldsymbol{\gamma}_k\right|^{2r} \\
&\quad = \frac{1}{2^r T^r} \mathrm{E}\left(\left|\sum_{i=1}^n \varepsilon_{ki} b_i\right|^{2r}\right) \\
&\quad = \frac{1}{2^r T^r} \mathrm{E} \sum_{\substack{i_1+\cdots+i_n=r \\ j_1+\cdots+j_n=r}} \frac{(r!)^2}{i_1! j_1! \cdots i_n! j_n!} (\varepsilon_{k1} b_1)^{i_1} (\bar{\varepsilon}_{k1} \bar{b}_1)^{j_1} \cdots (\varepsilon_{kn} b_n)^{i_n} (\bar{\varepsilon}_{kn} \bar{b}_n)^{j_n}
\end{aligned}
$$

$$= \frac{1}{2^r T^r} \mathrm{E} \sum_{\substack{i_1+\cdots+i_n=r \\ j_1+\cdots+j_n=r \\ i_1+j_1\neq 1}} \frac{(r!)^2}{i_1! j_1! \cdots i_n! j_n!} (\varepsilon_{k1} b_1)^{i_1} (\bar{\varepsilon}_{k1} \bar{b}_1)^{j_1} \cdots (\varepsilon_{kn} b_n)^{i_n} (\bar{\varepsilon}_{kn} \bar{b}_n)^{j_n}.$$

Let l denote the number $k \le n$ such that $i_k + j_k \ge 2$. By the fact that $\frac{(r!)^2}{(2r)!} \le \frac{r}{2r} \frac{r-1}{2r-1} \cdots \frac{1}{r+1} \le \frac{1}{2^r}$, we have

$$\begin{aligned}
&\mathrm{E}\left|\boldsymbol{\gamma}_l^* \mathbf{A}_k^{-s} \boldsymbol{\gamma}_k\right|^{2r} \\
&\quad \le \frac{1}{2^{2r} T^r} \sum_{l=1}^{r} \sum_{1\le j_1<\cdots<j_l\le n} \sum_{\substack{i_1+\cdots+i_l=2r \\ i_1\ge 2,\dots,i_l\ge 2}} \frac{(2r)!}{i_1!\cdots i_l! l!} \mathrm{E}\left|\varepsilon_{kj_1}^{i_1} b_{j_1}^{i_1} \cdots \varepsilon_{kj_l}^{i_l} b_{j_l}^{i_l}\right| \\
&\quad \le \frac{1}{2^{2r} T^r} \mathrm{E} \sum_{l=1}^{r} C^{2r} \sum_{1\le j_1<\cdots<j_l\le n} \sum_{\substack{i_1+\cdots+i_l=2r \\ i_1\ge 2,\dots,i_l\ge 2}} \frac{(2r)!}{i_1!\cdots i_l! l!} |b_{j_1}|^{i_1} \cdots |b_{j_l}|^{i_l} \\
&\quad \le \frac{K_r}{T^r} \sum_{l=1}^{r} \sum_{i_1+\cdots+i_l=2r} \mathrm{E} \prod_{t=1}^{l} \left(\sum_{j=1}^{n} |b_j|^{i_t} \right) \\
&\quad \le \frac{K_r}{T^r} \mathrm{E} \left(\sum_{j=1}^{n} |b_j^2| \right)^r \\
&\quad \le \frac{K_r}{T^r} \mathrm{E}\left(\boldsymbol{\gamma}_l^* \mathbf{A}_k^{-s} (\mathbf{A}_k^*)^{-s} \boldsymbol{\gamma}_l\right)^r.
\end{aligned}$$

Note that $\|\boldsymbol{\gamma}_l\| \le K$ and $\|\mathbf{A}_k^{-1}\| \le v_n^{-1}$, we finally obtain that

$$\mathrm{E}\left|\boldsymbol{\gamma}_l^* \mathbf{A}_k^{-s} \boldsymbol{\gamma}_k\right|^{2r} \le \frac{K}{T^r v_n^{2rs}}$$

for some $K > 0$. The proof of the lemma is complete.

B.3. Proof of Lemma 3.5. Recall that $a' = a - \underline{\varepsilon}$ and $b' = b + \underline{\varepsilon}$, as defined at the end of Section 4. Therefore, we have

$$\begin{aligned}
&\mathrm{P}\left(\frac{1}{2T} \sum \frac{1}{|\lambda_{kj} - z|^2} > K \right) \\
&\qquad \le \mathrm{P}\left(\sum_{\lambda_{kj} \notin [a',b']} \frac{1}{|\lambda_{kj} - u|^2 + v_n^2} > TK \right) \\
&\qquad\quad + \mathrm{P}\left(\sum_{\lambda_{kj} \in [a',b']} \frac{1}{|\lambda_{kj} - u|^2 + v_n^2} > TK \right)
\end{aligned}$$

$$\leq \mathrm{P}\big(n\underline{\varepsilon}^{-2} > TK\big) + \mathrm{P}\big(nv_n^{-2}F_{nk}([a', b']) > TK\big)$$
$$\leq 0 + \mathrm{P}\left(\|F_n - F_{c_n}\| \geq \frac{K}{2c}n^{-1/53}\right) = o\big(n^{-t}\big).$$

Here, we pick $K > c\underline{\varepsilon}^{-2}$ so that the first probability is 0. The second probability follows (4.36). The proof is complete.

B.4. Proof of Lemma 3.6, part (a). For (i)(a), by definition of x_{nj}, $j = 0, 1$, we have

$$x_{n0,1} = \tfrac{1}{2}\Big(1 \pm \sqrt{1 - 4a_n^2}\Big) := \tfrac{1}{2}\big(1 \pm (\tilde{\alpha} + i\tilde{\beta})\big).$$

Therefore,

$$\left|\frac{x_{n0}}{x_{n1}}\right| = \begin{cases} \sqrt{\dfrac{(1-\tilde{\alpha})^2 + \tilde{\beta}^2}{(1+\tilde{\alpha})^2 + \tilde{\beta}^2}} < 1 - \dfrac{2\tilde{\alpha}}{(1+\tilde{\alpha})^2 + \tilde{\beta}^2}, & \text{if } \tilde{\alpha} > 0, \\ \sqrt{\dfrac{(1+\tilde{\alpha})^2 + \tilde{\beta}^2}{(1-\tilde{\alpha})^2 + \tilde{\beta}^2}} < 1 - \dfrac{2|\tilde{\alpha}|}{(1-\tilde{\alpha})^2 + \tilde{\beta}^2}, & \text{if } \tilde{\alpha} < 0 \end{cases}$$
$$= 1 - \frac{|\tilde{\alpha}|}{2|x_{n1}^2|} < 1 - \eta_1 v_n^2 |\tilde{\alpha}|,$$

where the last inequality follows from the fact that $x_{n1}^2 = x_{n1} - a_n^2 = O(v_n^{-2})$.

Thus, to complete the proof of (i)(a), it suffices to show that there is a constant $\eta_2 > 0$ such that $|\tilde{\alpha}| > \eta_2 v_n$.

Write $c_n \mathrm{E} m_n(z) = 2a_n = \alpha + i\beta$ where α and β are real. Then, by the formula of square root of complex numbers [see (2.3.2) of Bai and Silverstein (2010)] we have

$$\sqrt{1 - 4a_n^2} = \tilde{\alpha} + i\tilde{\beta},$$

where

$$\tilde{\alpha} = \frac{-\sqrt{2}\alpha\beta}{\sqrt{\sqrt{(1 - \alpha^2 + \beta^2)^2 + 4\alpha^2\beta^2} - (1 - \alpha^2 + \beta^2)}}.$$

Obviously, when $1 - \alpha^2 + \beta^2 > 0$, by $\sqrt{(1 - \alpha^2 + \beta^2)^2 + 4\alpha^2\beta^2} - (1 - \alpha^2 + \beta^2) < 2|\alpha|\beta$ we have

$$|\tilde{\alpha}| > 1/\sqrt{|\alpha|\beta} > 1/\big|c_n \mathrm{E} m_n(z)\big| > \eta_2 v_n,$$

for all large n such that $c_n\eta_2 < 1$, where $\eta_2 \in (0, c^{-1})$.

On the other hand, if $1-\alpha^2+\beta^2<0$, by $\alpha^2>1+\beta^2$ we have

$$|\tilde{\alpha}|>\frac{|\alpha|\beta}{\sqrt[4]{(1-\alpha^2+\beta^2)^2+4\alpha^2\beta^2}}=\frac{|\alpha|\beta}{\sqrt[4]{(1-\alpha^2-\beta^2)^2+4\beta^2}}>\beta/\sqrt{2}.$$

Then the assertion that $|\tilde{\alpha}|>\eta_2 v_n$ is proved if one can show that $\beta>\eta_3 v_n$ for some $\eta_3>0$. This is trivial if one notices

$$\beta=v\int\frac{1}{(x-u)^2+v^2}\,d\mathrm{E}F_n(x)>v_n\big(4A^2+1\big)^{-1}\mathrm{E}F_n\big([-A,A]\big),$$

when $|z|<A$ and $v\in(v_n,1)$. The conclusion (i) is proved.

For (ii)(a), by $x_{n1}+x_{n0}=1$ and $|x_{n1}|>|x_{n0}|$, we conclude that $|x_{n1}|\ge\frac{1}{2}$. Since $x_{n1}=\frac{1}{2}(1\pm\sqrt{1-4a_n^2})$, we conclude that

$$|x_{n1}|\le\frac{1}{2}\Big(1+\Big|\sqrt{1-4a_n^2}\Big|\Big)\le Kv_n^{-1}.$$

For (iii)(a), by noting that

$$|x_{n1}-x_{n0}|^2=\big(1-\alpha^2+\beta^2\big)^2+4\alpha^2\beta^2=\big(1-\alpha^2-\beta^2\big)^2+4\beta^2.$$

Then the conclusion (iii)(a) follows from the fact $|\beta|>\eta_3 v_n$ that is shown in the proof of part (i)(a) of the lemma.

The conclusion (iv)(a) follows from

$$\frac{|x_{n0}|}{|x_{n1}-x_{n0}|}\le\frac{1}{2}\bigg(\frac{1}{|\sqrt{1-4a_n^2}|}+1\bigg)\le Kv_n^{-1},$$

where the last inequality follows from conclusion (iii)(a).

The proof of the lemma is complete.

B.5. Proof of Lemma 3.7(a). Recall that $a_n=\frac{c_n\mathrm{E}m_n}{2}$. Write $W_k=\boldsymbol{\gamma}_{k+\tau}^*\times\mathbf{A}_k^{-1}\boldsymbol{\gamma}_{k+\tau}$ and $W_{k,k+\tau,\dots,k+s\tau}=\boldsymbol{\gamma}_{k+(s+1)\tau}^*\mathbf{A}_{k,k+\tau,\dots,k+s\tau}^{-1}\boldsymbol{\gamma}_{k+(s+1)\tau}$. Denote $\widetilde{\mathbf{A}}_{k,\dots,k+(s-1)\tau}=\mathbf{A}_{k,\dots,k+s\tau}+\boldsymbol{\gamma}_{k+(s+1)\tau}\boldsymbol{\gamma}_{k+s\tau}^*$. Apply the identity

$$\big(\mathbf{B}+\boldsymbol{\alpha}\boldsymbol{\gamma}^*\big)^{-1}=\mathbf{B}^{-1}-\frac{\mathbf{B}^{-1}\boldsymbol{\alpha}\boldsymbol{\gamma}^*\mathbf{B}^{-1}}{1+\boldsymbol{\gamma}^*\mathbf{B}^{-1}\boldsymbol{\alpha}},$$

we have

$$\begin{aligned}\mathbf{A}_{k,\dots,k+(s-1)\tau}^{-1}&=\big(\widetilde{\mathbf{A}}_{k,\dots,k+(s-1)\tau}+\boldsymbol{\gamma}_{k+s\tau}\boldsymbol{\gamma}_{k+(s+1)\tau}^*\big)^{-1}\\&=\widetilde{\mathbf{A}}_{k,\dots,k+(s-1)\tau}^{-1}-\frac{\widetilde{\mathbf{A}}_{k,\dots,k+(s-1)\tau}^{-1}\boldsymbol{\gamma}_{k+(s+1)\tau}\boldsymbol{\gamma}_{k+s\tau}^*\widetilde{\mathbf{A}}_{k,\dots,k+(s-1)\tau}^{-1}}{1+\boldsymbol{\gamma}_{k+s\tau}^*\widetilde{\mathbf{A}}_{k,\dots,k+(s-1)\tau}^{-1}\boldsymbol{\gamma}_{k+(s+1)\tau}},\\\widetilde{\mathbf{A}}_{k,\dots,k+(s-1)\tau}&=\big(\mathbf{A}_{k,\dots,k+s\tau}+\boldsymbol{\gamma}_{k+(s+1)\tau}\boldsymbol{\gamma}_{k+s\tau}^*\big)^{-1}\\&=\mathbf{A}_{k,\dots,k+s\tau}^{-1}-\frac{\mathbf{A}_{k,\dots,k+s\tau}^{-1}\boldsymbol{\gamma}_{k+(s+1)\tau}\boldsymbol{\gamma}_{k+s\tau}^*\mathbf{A}_{k,\dots,k+s\tau}^{-1}}{1+\boldsymbol{\gamma}_{k+s\tau}^*\mathbf{A}_{k,\dots,k+s\tau}^{-1}\boldsymbol{\gamma}_{k+(s+1)\tau}}.\end{aligned}$$

Therefore, we have

$$\begin{aligned}
&\boldsymbol{\gamma}_{k+s\tau}^{*}\mathbf{A}_{k,\dots,k+(s-1)\tau}^{-1}\\
&\quad=\boldsymbol{\gamma}_{k+s\tau}^{*}\widetilde{\mathbf{A}}_{k,\dots,k+(s-1)\tau}^{-1}\\
&\qquad-\frac{\boldsymbol{\gamma}_{k+s\tau}^{*}\widetilde{\mathbf{A}}_{k,\dots,k+(s-1)\tau}^{-1}\boldsymbol{\gamma}_{k+(s+1)\tau}\boldsymbol{\gamma}_{k+s\tau}^{*}\widetilde{\mathbf{A}}_{k,\dots,k+(s-1)\tau}^{-1}}{1+\boldsymbol{\gamma}_{k+s\tau}^{*}\widetilde{\mathbf{A}}_{k,\dots,k+(s-1)\tau}^{-1}\boldsymbol{\gamma}_{k+(s+1)\tau}}\\
&\quad=\frac{\boldsymbol{\gamma}_{k+s\tau}^{*}\widetilde{\mathbf{A}}_{k,\dots,k+(s-1)\tau}^{-1}}{1+\boldsymbol{\gamma}_{k+s\tau}^{*}\widetilde{\mathbf{A}}_{k,\dots,k+(s-1)\tau}^{-1}\boldsymbol{\gamma}_{k+(s+1)\tau}}
\end{aligned}$$

and

$$\begin{aligned}
&\boldsymbol{\gamma}_{k+s\tau}^{*}\mathbf{A}_{k,\dots,k+(s-1)\tau}^{-1}\boldsymbol{\gamma}_{k+s\tau}\\
&\quad=\frac{\boldsymbol{\gamma}_{k+s\tau}^{*}\widetilde{\mathbf{A}}_{k,\dots,k+(s-1)\tau}^{-1}\boldsymbol{\gamma}_{k+s\tau}}{1+\boldsymbol{\gamma}_{k+(s+1)\tau}^{*}\widetilde{\mathbf{A}}_{k,\dots,k+(s-1)\tau}^{-1}\boldsymbol{\gamma}_{k+s\tau}}\\
&\quad=\Bigg(\boldsymbol{\gamma}_{k+s\tau}^{*}\mathbf{A}_{k,\dots,k+s\tau}^{-1}\boldsymbol{\gamma}_{k+s\tau}\\
&\qquad\quad-\frac{\boldsymbol{\gamma}_{k+s\tau}^{*}\mathbf{A}_{k,\dots,k+s\tau}^{-1}\boldsymbol{\gamma}_{k+(s+1)\tau}\boldsymbol{\gamma}_{k+s\tau}^{*}\mathbf{A}_{k,\dots,k+s\tau}^{-1}\boldsymbol{\gamma}_{k+s\tau}}{1+\boldsymbol{\gamma}_{k+s\tau}^{*}\mathbf{A}_{k,\dots,k+s\tau}^{-1}\boldsymbol{\gamma}_{k+(s+1)\tau}}\Bigg)\\
&\qquad\Bigg/\Bigg(1+\boldsymbol{\gamma}_{k+(s+1)\tau}^{*}\mathbf{A}_{k,\dots,k+s\tau}^{-1}\boldsymbol{\gamma}_{k+s\tau}\\
&\qquad\quad-\frac{\boldsymbol{\gamma}_{k+(s+1)\tau}^{*}\mathbf{A}_{k,\dots,k+s\tau}^{-1}\boldsymbol{\gamma}_{k+(s+1)\tau}\boldsymbol{\gamma}_{k+s\tau}^{*}\mathbf{A}_{k,\dots,k+s\tau}^{-1}\boldsymbol{\gamma}_{k+s\tau}}{1+\boldsymbol{\gamma}_{k+s\tau}^{*}\mathbf{A}_{k,\dots,k+s\tau}^{-1}\boldsymbol{\gamma}_{k+(s+1)\tau}}\Bigg)\\
&\quad=\frac{(c_n/2)\mathrm{E}m_n(z)+r_1(k+s\tau)}{1-(c_n/2)\mathrm{E}m_n(z)\boldsymbol{\gamma}_{k+(s+1)\tau}^{*}\mathbf{A}_{k,\dots,k+s\tau}^{-1}\boldsymbol{\gamma}_{k+(s+1)\tau}+r_2(k+s\tau)},
\end{aligned}\tag{B.1}$$

that is,

$$W_{k,\dots,k+(s-1)\tau}=\frac{a_n+r_1(k+s\tau)}{1-a_nW_{k,\dots,k+s\tau}+r_2(k+s\tau)},\tag{B.2}$$

where

$$\begin{aligned}
r_1(k+s\tau)&=\boldsymbol{\gamma}_{k+s\tau}^{*}\mathbf{A}_{k,\dots,k+s\tau}^{-1}\boldsymbol{\gamma}_{k+s\tau}-a_n,\\
r_2(k+s\tau)&=-(\boldsymbol{\gamma}_{k+s\tau}^{*}\mathbf{A}_{k,\dots,k+s\tau}^{-1}\boldsymbol{\gamma}_{k+s\tau}-a_n)\boldsymbol{\gamma}_{k+(s+1)\tau}^{*}\mathbf{A}_{k,\dots,k+s\tau}^{-1}\boldsymbol{\gamma}_{k+(s+1)\tau}\\
&\quad+\boldsymbol{\gamma}_{k+(s+1)\tau}^{*}\mathbf{A}_{k,\dots,k+s\tau}^{-1}\boldsymbol{\gamma}_{k+s\tau}+\boldsymbol{\gamma}_{k+s\tau}^{*}\mathbf{A}_{k,\dots,k+s\tau}^{-1}\boldsymbol{\gamma}_{k+(s+1)\tau}\\
&\quad+\boldsymbol{\gamma}_{k+(s+1)\tau}^{*}\mathbf{A}_{k,\dots,k+s\tau}^{-1}\boldsymbol{\gamma}_{k+s\tau}\boldsymbol{\gamma}_{k+s\tau}^{*}\mathbf{A}_{k,\dots,k+s\tau}^{-1}\boldsymbol{\gamma}_{k+(s+1)\tau}.
\end{aligned}$$

When $k \le T - v_n^{-4}$, applying this relation ℓ times ($\ell = [v_n^{-4}]$), we may express W_k in the following form:

$$W_k = \frac{(a_n + r_1(k+\tau))(\alpha_{k+\tau,\ell} - a_n \gamma_{k+\tau,\ell} W_{k,k+\tau,\dots,k+(\ell+1)\tau})}{\alpha_{k,\ell} - a_n \gamma_{k,\ell} W_{k,k+\tau,\dots,k+(\ell+1)\tau}},$$

where the coefficients satisfy the recursive relation

$$\begin{aligned}
\alpha_{k+s\tau,\ell} &= \big(1 + r_2(k+s\tau)\big)\alpha_{k+(s+1)\tau,\ell} \\
&\qquad - a_n\big(a_n + r_1(k+s\tau)\big)\alpha_{k+(s+2)\tau,\ell}, \\
\alpha_{k+\ell\tau,\ell} &= 1 + r_2(k+\ell\tau), \qquad \alpha_{k+(\ell+1)\tau,\ell} = 1, \\
\gamma_{k+s\tau,\ell} &= \big(1 + r_2(k+s\tau)\big)\gamma_{k+(s+1)\tau,\ell} \\
&\qquad - a_n\big(a_n + r_1(k+s\tau)\big)\gamma_{k+(s+2)\tau,\ell}, \\
\gamma_{k+\ell\tau,\ell} &= 1, \qquad \gamma_{k+(\ell+1)\tau,\ell} = 0.
\end{aligned} \tag{B.3}$$

Notice that $v_n = n^{-1/52}$. Employing Lemma 2.5 and an estimation similar to (4.3), for any fixed t, one has

$$\mathrm{P}\big(|r_i(k+\ell\tau)| \ge v_n^{12}\big) = o\big(n^{-t}\big) \qquad \text{for } i = 1, 2. \tag{B.4}$$

As in the proof of Lemma B.3 of Jin et al. (2014), by letting $\ell = [v_n^{-4}]$, it follows by induction that

$$\alpha_{k+l\tau,\ell} = (1-\alpha) \prod_{\mu=1}^{\ell-l+1} \nu_{\mu,1} + \alpha \prod_{\mu=1}^{\ell-l+1} \nu_{\mu,0}, \tag{B.5}$$

where $\nu_{1,i}$, $i = 1, 0$ (with $|\nu_{1,1}| > |\nu_{1,0}|$) are defined by the two roots of the quadratic equation

$$x^2 = \big(1 + r_2(k+\ell\tau)\big)x - a_n\big(a_n + r_1(k+\ell\tau)\big)$$

and α is such that

$$(1-\alpha)\nu_{1,1} + \alpha\nu_{1,0} = 1 + r_2(k+\ell\tau) = \alpha_{k+\ell\tau,\ell}.$$

Recall that x_{ni}, $i = 1, 0$ (with $|x_{n1}| > |x_{n0}|$) are two roots of the quadratic equation

$$x^2 = x - a_n^2.$$

Applying Lemmas 3.1–3.3 to the above two quadratic equations and using (B.4), we have

$$\begin{aligned}
&\mathrm{P}\big(|\nu_{1,i} - x_{ni}| \ge 2v_n^6\big) \\
&\qquad \le \mathrm{P}\big(|r_1(k+\ell\tau)| \ge v_n^{12}\big) + \mathrm{P}\big(|r_2(k+\ell\tau)| \ge v_n^{12}\big) = o\big(n^{-t}\big),
\end{aligned} \tag{B.6}$$

$$\begin{aligned}
&\mathrm{P}\left(\left|\alpha - \frac{x_{n0}}{x_{n0} - x_{n1}}\right| \ge 3v_n^6\right) \\
&\qquad \le \mathrm{P}\big(|\nu_{1,0} - x_{n0}| \ge v_n^6\big) + \mathrm{P}\big(|\nu_{1,1} - x_{n1}| \ge v_n^6\big) + \mathrm{P}\big(|r_2(k+\ell\tau)| \ge v_n^6\big) \\
&\qquad = o\big(n^{-t}\big).
\end{aligned} \tag{B.7}$$

By induction, one has for $\mu \in [1, \ell]$

$$\nu_{\mu+1,i} = 1 + r_2\big(k + (\ell - \mu)\tau\big) - \frac{a_n(a_n + r_1(k + (\ell - \mu)\tau))}{\nu_{\mu,i}}$$

and can similarly verify that

$$\mathrm{P}\big(|\nu_{\mu,i} - x_{ni}| \geq 2\mu v_n^6\big) \leq \sum_{l=1}^{\mu} \sum_{j=1}^{2} \mathrm{P}\big(\big|r_j(k + l\tau)\big| \geq v_n^{12}\big) = o\big(n^{-t}\big).$$

Therefore, we have

$$\begin{aligned}
\mathrm{P}\big(\big|\alpha_{k+\tau,\ell} - \big((1-\alpha)x_{n1}^{\ell} + \alpha x_{n0}^{\ell}\big)\big| \geq v_n^6\big) &\leq \sum_{\mu=1}^{\ell} \sum_{i=0}^{1} \mathrm{P}\big(|\nu_{\mu,i} - x_{ni}| \geq 2\mu v_n^6\big) \\
&= o\big(n^{-t}\big), \\
\mathrm{P}\big(\big|\alpha_{k,\ell} - \big((1-\alpha)x_{n1}^{\ell+1} + \alpha x_{n0}^{\ell+1}\big)\big| \geq v_n^6\big) &\leq \sum_{\mu=1}^{\ell+1} \sum_{i=0}^{1} \mathrm{P}\big(|\nu_{\mu,i} - x_{ni}| \geq 2\mu v_n^6\big) \\
&= o\big(n^{-t}\big),
\end{aligned}$$

and

$$\begin{aligned}
&\mathrm{P}\left(\left|\frac{\alpha_{k+\tau,\ell}}{\alpha_{k,\ell}} - \frac{1}{x_{n1}}\right| \geq v_n^6\right) \\
&\quad \leq \mathrm{P}\big(\big|\alpha_{k+\tau,\ell} - \big((1-\alpha)x_{n1}^{\ell} + \alpha x_{n0}^{\ell}\big)\big| \geq v_n^6\big) \\
&\qquad + \mathrm{P}\big(\big|\alpha_{k,\ell} - \big((1-\alpha)x_{n1}^{\ell+1} + \alpha x_{n0}^{\ell+1}\big)\big| \geq v_n^6\big) \\
&\qquad + \mathrm{P}\big(|\nu_{\ell+1,1} - x_{n1}| \geq 2(\ell+1)v_n^6\big) \\
&\quad = o\big(n^{-t}\big).
\end{aligned}$$

Similarly, we have

$$\gamma_{k+l\tau,\ell} = (1 - \tilde{\alpha}) \prod_{\mu=1}^{\ell-l+1} \tilde{\nu}_{\mu,1} + \tilde{\alpha} \prod_{\mu=1}^{\ell-l+1} \tilde{\nu}_{\mu,0},$$

where $\tilde{\nu}_{\mu,i}$, $i = 1, 0$, are the two roots of the quadratic equation

$$x^2 = \big(1 + r_2\big(k + (\ell - 1)\tau\big)\big)x - a_n\big(a_n + r_1\big(k + (\ell - 1)\tau\big)\big),$$

and $\tilde{\alpha}$ satisfies

$$(1 - \tilde{\alpha})\tilde{\nu}_{1,1} + \tilde{\alpha}\tilde{\nu}_{1,0} = 1 + r_2\big(k + (\ell - 1)\tau\big) = \gamma_{k+(\ell-1)\tau,\ell}.$$

One can similarly prove that $\tilde{\nu}_{\mu,i}$, $i = 0, 1$, satisfy

$$\mathrm{P}\big(|\tilde{\nu}_{\mu,i} - x_{ni}| \geq 2\mu v_n^6\big) \leq \sum_{l=0}^{\mu} \sum_{j=1}^{2} \mathrm{P}\big(\big|r_j(k + l\tau)\big| \geq v_n^{12}\big) = o\big(n^{-t}\big),$$

and

$$\mathrm{P}\left(\left|\tilde{\alpha} - \frac{x_{n0}}{x_{n0} - x_{n1}}\right| \geq 3v_n^6\right) = o(n^{-t}).$$

Therefore, we have

$$\begin{aligned}
\mathrm{P}\big(\big|\gamma_{k+\tau,\ell} - \big((1-\tilde{\alpha})x_{n1}^{\ell} + \tilde{\alpha}x_{n0}^{\ell}\big)\big| \geq v_n^6\big) &\leq \sum_{\mu=1}^{\ell}\sum_{i=0}^{1}\mathrm{P}\big(|\tilde{v}_{\mu,i} - x_{ni}| \geq 2\mu v_n^6\big) \\
&= o(n^{-t}), \\
\mathrm{P}\big(\big|\gamma_{k,\ell} - \big((1-\tilde{\alpha})x_{n1}^{\ell+1} + \tilde{\alpha}x_{n0}^{\ell+1}\big)\big| \geq v_n^6\big) &\leq \sum_{\mu=1}^{\ell+1}\sum_{i=0}^{1}\mathrm{P}\big(|\tilde{v}_{\mu,i} - x_{ni}| \geq 2\mu v_n^6\big) \\
&= o(n^{-t}),
\end{aligned}$$

and

$$\begin{aligned}
&\mathrm{P}\left(\left|\frac{\gamma_{k+\tau,\ell}}{\gamma_{k,\ell}} - \frac{1}{x_{n1}}\right| \geq v_n^6\right) \\
&\quad\leq \mathrm{P}\big(\big|\gamma_{k+\tau,\ell} - \big((1-\tilde{\alpha})x_{n1}^{\ell} + \tilde{\alpha}x_{n0}^{\ell}\big)\big| \geq v_n^6\big) \\
&\qquad + \mathrm{P}\big(\big|\gamma_{k,\ell} - \big((1-\tilde{\alpha})x_{n1}^{\ell+1} + \tilde{\alpha}x_{n0}^{\ell+1}\big)\big| \geq v_n^6\big) \\
&\qquad + \mathrm{P}\big(|\tilde{v}_{\ell+1,1} - x_{n1}| \geq 2(\ell+1)v_n^6\big) \\
&\quad= o(n^{-t}).
\end{aligned}$$

Substituting back to the recursive expression of W_k, we thus have

$$\mathrm{P}\left(\left|W_k - \frac{a_n}{x_{n1}}\right| \geq v_n^6\right) = o(n^{-t}). \tag{B.8}$$

The proof of this lemma is complete.

B.6. Proof of Lemma 3.8(a). When $\tau < k \leq 2\tau$, the lemma is obviously true because $\boldsymbol{\gamma}_{k-\tau}$ is independent of $\mathbf{A}_k$. Similarly, the lemma is true when $T - \tau < k \leq T$.

When $2\tau < k \leq T/2$, similar to (B.1), we have

$$\begin{aligned}
&\widetilde{W}_{k,\dots,k+s\tau} \\
&\quad:= \boldsymbol{\gamma}_{k-\tau}^{*}\mathbf{A}_{k,k+\tau,\dots,k+(s-1)\tau}^{-1}\boldsymbol{\gamma}_{k+s\tau} \\
&\quad= \frac{\boldsymbol{\gamma}_{k-\tau}^{*}(\mathbf{A}_{k,k+\tau,\dots,k+s\tau} + \boldsymbol{\gamma}_{k+(s+1)\tau}\boldsymbol{\gamma}_{k+s\tau}^{*})^{-1}\boldsymbol{\gamma}_{k+s\tau}}{1 + \boldsymbol{\gamma}_{k+(s+1)\tau}^{*}(\mathbf{A}_{k,k+\tau,\dots,k+s\tau} + \boldsymbol{\gamma}_{k+(s+1)\tau}\boldsymbol{\gamma}_{k+s\tau}^{*})^{-1}\boldsymbol{\gamma}_{k+s\tau}} \\
&\quad= \Big(\boldsymbol{\gamma}_{k-\tau}^{*}\mathbf{A}_{k,k+\tau,\dots,k+s\tau}^{-1}\boldsymbol{\gamma}_{k+s\tau}
\end{aligned}$$

$$- \frac{\boldsymbol{\gamma}_{k-\tau}^* \mathbf{A}_{k,k+\tau,\dots,k+s\tau}^{-1} \boldsymbol{\gamma}_{k+(s+1)\tau} \boldsymbol{\gamma}_{k+s\tau}^* \mathbf{A}_{k,k+\tau,\dots,k+s\tau}^{-1} \boldsymbol{\gamma}_{k+s\tau}}{1 + \boldsymbol{\gamma}_{k+s\tau}^* \mathbf{A}_{k,k+\tau,\dots,k+s\tau}^{-1} \boldsymbol{\gamma}_{k+(s+1)\tau}} \Bigg)$$

$$\Bigg/ \Bigg(1 + \boldsymbol{\gamma}_{k+(s+1)\tau}^* \mathbf{A}_{k,k+\tau,\dots,k+s\tau}^{-1} \boldsymbol{\gamma}_{k+s\tau} - \frac{\boldsymbol{\gamma}_{k+(s+1)\tau}^* \mathbf{A}_{k,k+\tau,\dots,k+s\tau}^{-1} \boldsymbol{\gamma}_{k+(s+1)\tau} \boldsymbol{\gamma}_{k+s\tau}^* \mathbf{A}_{k,k+\tau,\dots,k+s\tau}^{-1} \boldsymbol{\gamma}_{k+s\tau}}{1 + \boldsymbol{\gamma}_{k+s\tau}^* \mathbf{A}_{k,k+\tau,\dots,k+s\tau}^{-1} \boldsymbol{\gamma}_{k+(s+1)\tau}} \Bigg)$$

$$= \frac{\widetilde{r}_1(k+s\tau) - \widetilde{W}_{k,\dots,k+(s+1)\tau} a_n}{1 + r_2(k+s\tau) - a_n W_{k,\dots,k+s\tau}},$$

where

$$\widetilde{r}_1(k+s\tau) = \boldsymbol{\gamma}_{k-\tau}^* \mathbf{A}_{k,\dots,k+s\tau}^{-1} \boldsymbol{\gamma}_{k+s\tau} \big(1 + \boldsymbol{\gamma}_{k+s\tau}^* \mathbf{A}_{k,\dots,k+s\tau}^{-1} \boldsymbol{\gamma}_{k+(s+1)\tau}\big) - \widetilde{W}_{k,\dots,k+(s+1)\tau} \big(\boldsymbol{\gamma}_{k+s\tau}^* \mathbf{A}_{k,\dots,k+s\tau}^{-1} \boldsymbol{\gamma}_{k+s\tau} - a_n\big).$$

Similarly, one can show that

$$\mathrm{P}\big(\big|\widetilde{r}_1(k+s\tau)\big| \ge v_n^{12}\big) = o\big(n^{-t}\big).$$

When $|\widetilde{r}_1(t+s\tau)| \le v_n^{12}$, $|r_2(k+s\tau)| \le v_n^{12}$, and $|W_{k,\dots,k+s\tau} - \frac{a_n}{x_{n1}}| \le v_n^6$, we have

$$|\widetilde{W}_{k,\dots,k+s\tau}| \le \frac{v_n^{12}}{|x_{n1}| - v_n^5} + |\widetilde{W}_{k,\dots,k+(s+1)\tau}| \left| \frac{|a_n|}{|x_{n1}|} + v_n^5 \right| \le 3v_n^{12} + |\widetilde{W}_{k,\dots,k+(s+1)\tau}| \left(1 - \frac{1}{2}\eta v_n^3 + v_n^4\right),$$

where the second term follows from the fact that

$$\frac{|a_n|}{|x_{n1}|} = \sqrt{\frac{|x_{n0}|}{|x_{n1}|}} \le 1 - \frac{1}{2}\eta v_n^3.$$

Therefore, when $v_n^{-4} < \ell < v_n^{-5}$,

$$|\widetilde{W}_k| \le 3\ell v_n^{12} + |\widetilde{W}_{k,\dots,k+\ell\tau}| \left|1 - \tfrac{1}{2}\eta v_n^3 + v_n^4\right|^\ell \le v_n^6.$$

The lemma then follows by the fact that

$$\mathrm{P}\big(|\widetilde{W}_k| \ge v_n^6\big) \le \sum_{s=1}^{\ell} \left(\mathrm{P}\big(\big|\widetilde{r}_1(k+s\tau)\big| \ge v_n^{12}\big) + \mathrm{P}\big(\big|r_2(k+s\tau)\big| \ge v_n^{12}\big) + \mathrm{P}\left(\left|W_{k,\dots,k+s\tau} - \frac{a_n}{x_{n1}}\right| \ge v_n^6\right)\right) = o\big(n^{-t}\big).$$

The proof of the lemma is complete.

B.7. Proof of Lemma 3.6, part (b). Let x_1 and x_0 be the two roots of the quadratic equation

$$x^2 = x - \breve{a}^2,$$

where $\breve{a} = \breve{a}(z) = cm(z)/2$ and $m(z)$ satisfies (4.8). We claim that

$$\sup_{u\in[a,b]} \frac{|x_0(z)|}{|x_1(z)|} \le 1 - \eta \tag{B.9}$$

for some $\eta \in (0, 1)$. Otherwise, there will be a sequence $\{z_k\}$ with $\Re(z_k) \in [a, b]$ and

$$\frac{|x_0(z_k)|}{|x_1(z_k)|} \to 1.$$

Then we can select a convergent subsequence $\{z_{k'}\} \to z_0$. If $z_0 = \infty$, then $\breve{a}(z_0) = 0$ and hence $x_1 = 1$ and $x_0 = 0$. It contradicts the fact that

$$\frac{|x_0(z_0)|}{|x_1(z_0)|} = 1.$$

The only case to make the equality above true is that $\breve{a}(z_0)$ is real and its absolute value is $\ge\frac{1}{2}$. That is, z_0 is real and $|\breve{a}(z_0)| \ge \frac{1}{2}$. Since $\breve{a}(\infty) = 0$, there is a real number z' between z_0 and $\operatorname{sgn}(z_0)\infty$ such that $|\breve{a}(z')| = \frac{1}{2}$ which contradicts the equation (4.8). Therefore, (B.9) is proved.

Since $m_n^0(z) \to m(z)$ uniformly for all $\Re(z) \in [a, b]$, we conclude that there is a constant $\eta \in (0, 1)$ such that

$$\sup_{\Re(z)\in[a,b]} \frac{|\widetilde{x}_{n0}|}{|\widetilde{x}_{n1}|} < 1 - \eta,$$

where $\widetilde{x}_{n1}$ and $\widetilde{x}_{n0}$ are the two roots of the equation

$$x^2 = x - \tfrac{1}{4}c_n^2\bigl(m_n^0(z)\bigr)^2.$$

By what has been proved in Section 4, we have $\sup_{1>\Im(z)\ge n^{-1/52}} |\mathrm{E}m_n(z) - m_n^0(z)| \to 0$. Thus,

$$\sup_{\substack{\Re(z)\in[a,b]\\ 1>\Im(z)\ge n^{-1/52}}} \frac{|x_{n0}|}{|x_{n1}|} \le 1 - \eta.$$

The conclusion (i)(b) follows.

We then prove the conclusion (v). In the proof of (i)(b), we actually proved that there is a constant $\eta \in (0, \frac{1}{2})$ such that for all $u \in [a, b]$,

$$\bigl|\breve{a}(u)\bigr| < \tfrac{1}{2} - \eta.$$

By the uniform continuity of $\breve{a}(z)$ for all $\Re(z) \in [a, b]$. we have

$$\sum_{u\in[a,b],v\in(0,\delta_n)} \left|\breve{a}(u+iv) - \breve{a}(u)\right| \to 0 \qquad \text{as } \delta_n \to 0.$$

Then conclusion (v) follows from the fact that $\sup_{1>\Im(z)\geq n^{-1/52}} |\mathrm{E}m_n(z) - m_n^0(z)| \to 0$.

The first conclusion of (ii)(b) is the same as (ii)(a) and the second follows easily from the fact that $|a_n(z)| \leq \frac{1}{2}$ and the argument that $|x_{n1}| \leq \frac{1}{2}(1+\sqrt{1+4|a_n^2|}) \leq \frac{3}{2}$.

The conclusion (iii)(b) follows from the fact that $|x_{n1} - x_{n0}| = |\sqrt{1-4a_n^2}| \geq \sqrt{4\eta(1-\eta)}$. The conclusion (iv)(b) follows from conclusions (ii)(b) and (iii)(b). The goal of this section is reached.

B.8. Proof of Lemma 3.7(b1). When $k \leq T - \log^2 n$, noticing $|x_{n0}|/|x_{n1}| \leq 1-\eta$ established in part (b) of Lemma 3.6, so (B.8) remains true, hence in turn implies the lemma. When $k > T - \log^2 n$, we shall recursively show the lemma by proving

$$\mathrm{P}\big(|W_{k,\dots,k+s\tau}| > 1-\eta\big) = o\big(n^{-t}\big), \tag{B.10}$$

for some $\eta \in (0, \frac{1}{2})$. In fact, when $k + s\tau \geq T > k + (s-1)\tau$, (B.10) follows easily by the fact that $\boldsymbol{\gamma}_{k+(s+1)\tau}$ is independent of $\mathbf{A}_{k,\dots,k+s\tau}^{-1}$, and hence $\mathrm{P}(|W_{k,\dots,k+s\tau} - a_n| \geq v_n^3) = o(n^{-t})$ and $|a_n| \leq 1/2 - \eta$.

By induction, assume that (B.10) is true for some $s \geq 1$. By (B.2) and Lemma 3.6(v), when $|r_1(k+s\tau)| \leq v_n^3$ and $|r_2(k+s\tau)| \leq v_n^3$, we have

$$|W_{k,\dots,k+(s-1)\tau}| \leq \frac{1/2 - \eta + v_n^3}{1 - (1/2-\eta)(1-\eta) - v_n^3} \leq 1-\eta \qquad \text{for all large } n.$$

Thus,

$$\begin{aligned}
&\mathrm{P}\big(|W_{k,\dots,k+(s-1)\tau}| > 1-\eta\big) \\
&\quad \leq \mathrm{P}\big(|W_{k,\dots,k+s\tau}| > 1-\eta\big) + \mathrm{P}\big(\left|r_1(k+s\tau)\right| \geq v_n^3\big) + \mathrm{P}\big(\left|r_2(k+s\tau)\right| \geq v_n^3\big) \\
&\quad = o\big(n^{-t}\big).
\end{aligned}$$

The assertion (B.10) is proved, and thus the proof of the lemma is complete.

B.9. Proof of Lemma 3.9. Define $\widetilde{\mathbf{A}}_k = \mathbf{A}_{k,k+\tau} + \boldsymbol{\gamma}_{k+\tau}\boldsymbol{\gamma}_{k+2\tau}^*$. Recall $\mathbf{A}_k = \mathbf{A}_{k,k+\tau} + \boldsymbol{\gamma}_{k+\tau}\boldsymbol{\gamma}_{k+2\tau}^* + \boldsymbol{\gamma}_{k+2\tau}\boldsymbol{\gamma}_{k+\tau}^*$, so we have

$$\mathbf{A}_k^{-1} = \big(\widetilde{\mathbf{A}}_k + \boldsymbol{\gamma}_{k+2\tau}\boldsymbol{\gamma}_{k+\tau}^*\big)^{-1} = \widetilde{\mathbf{A}}_k^{-1} - \frac{\widetilde{\mathbf{A}}_k^{-1}\boldsymbol{\gamma}_{k+2\tau}\boldsymbol{\gamma}_{k+\tau}^*\widetilde{\mathbf{A}}_k^{-1}}{1 + \boldsymbol{\gamma}_{k+\tau}^*\widetilde{\mathbf{A}}_k^{-1}\boldsymbol{\gamma}_{k+2\tau}}.$$

Hence, we have

$$\boldsymbol{\gamma}_{k+\tau}^* \mathbf{A}_k^{-1} = \boldsymbol{\gamma}_{k+\tau}^* \widetilde{\mathbf{A}}_k^{-1} - \frac{\boldsymbol{\gamma}_{k+\tau}^* \widetilde{\mathbf{A}}_k^{-1} \boldsymbol{\gamma}_{k+2\tau} \boldsymbol{\gamma}_{k+\tau}^* \widetilde{\mathbf{A}}_k^{-1}}{1 + \boldsymbol{\gamma}_{k+\tau}^* \widetilde{\mathbf{A}}_k^{-1} \boldsymbol{\gamma}_{k+2\tau}} = \frac{\boldsymbol{\gamma}_{k+\tau}^* \widetilde{\mathbf{A}}_k^{-1}}{1 + \boldsymbol{\gamma}_{k+\tau}^* \widetilde{\mathbf{A}}_k^{-1} \boldsymbol{\gamma}_{k+2\tau}}.$$

Next, we have

$$\begin{aligned} \boldsymbol{\gamma}_{k+\tau}^* \widetilde{\mathbf{A}}_k^{-1} &= \boldsymbol{\gamma}_{k+\tau}^* \mathbf{A}_{k,k+\tau}^{-1} - \frac{\boldsymbol{\gamma}_{k+\tau}^* \mathbf{A}_{k,k+\tau}^{-1} \boldsymbol{\gamma}_{k+\tau} \boldsymbol{\gamma}_{k+2\tau}^* \mathbf{A}_{k,k+\tau}^{-1}}{1 + \boldsymbol{\gamma}_{k+2\tau}^* \mathbf{A}_{k,k+\tau}^{-1} \boldsymbol{\gamma}_{k+\tau}} \\ &= \boldsymbol{\gamma}_{k+\tau}^* \mathbf{A}_{k,k+\tau}^{-1} - a_n \boldsymbol{\gamma}_{k+2\tau}^* \mathbf{A}_{k,k+\tau}^{-1} + R_{k1}, \end{aligned}$$

where

$$\begin{aligned} R_{k1} &= a_n \boldsymbol{\gamma}_{k+2\tau}^* \mathbf{A}_{k,k+\tau}^{-1} - \frac{\boldsymbol{\gamma}_{k+\tau}^* \mathbf{A}_{k,k+\tau}^{-1} \boldsymbol{\gamma}_{k+\tau} \boldsymbol{\gamma}_{k+2\tau}^* \mathbf{A}_{k,k+\tau}^{-1}}{1 + \boldsymbol{\gamma}_{k+2\tau}^* \mathbf{A}_{k,k+\tau}^{-1} \boldsymbol{\gamma}_{k+\tau}} \\ &= \left(\frac{a_n - \boldsymbol{\gamma}_{k+\tau}^* \mathbf{A}_{k,k+\tau}^{-1} \boldsymbol{\gamma}_{k+\tau} + a_n \boldsymbol{\gamma}_{k+2\tau}^* \mathbf{A}_{k,k+\tau}^{-1} \boldsymbol{\gamma}_{k+\tau}}{1 + \boldsymbol{\gamma}_{k+2\tau}^* \mathbf{A}_{k,k+\tau}^{-1} \boldsymbol{\gamma}_{k+\tau}} \right) \boldsymbol{\gamma}_{k+2\tau}^* \mathbf{A}_{k,k+\tau}^{-1}. \end{aligned}$$

Substituting back, we obtain

$$\begin{aligned} \boldsymbol{\gamma}_{k+\tau}^* \mathbf{A}_k^{-1} &= \frac{\boldsymbol{\gamma}_{k+\tau}^* \mathbf{A}_{k,k+\tau}^{-1} - a_n \boldsymbol{\gamma}_{k+2\tau}^* \mathbf{A}_{k,k+\tau}^{-1} + R_{k1}}{1 + \boldsymbol{\gamma}_{k+\tau}^* \mathbf{A}_{k,k+\tau}^{-1} \boldsymbol{\gamma}_{k+2\tau} - a_n \boldsymbol{\gamma}_{k+2\tau}^* \mathbf{A}_{k,k+\tau}^{-1} \boldsymbol{\gamma}_{k+2\tau} + R_{k1} \boldsymbol{\gamma}_{k+2\tau}} \\ &= (\boldsymbol{\gamma}_{k+\tau}^* \mathbf{A}_{k,k+\tau}^{-1} - a_n \boldsymbol{\gamma}_{k+2\tau}^* \mathbf{A}_{k,k+\tau}^{-1} + R_{k1}) \\ &\quad / \big(x_{n1} + \boldsymbol{\gamma}_{k+\tau}^* \mathbf{A}_{k,k+\tau}^{-1} \boldsymbol{\gamma}_{k+2\tau} \\ &\qquad - a_n (\boldsymbol{\gamma}_{k+2\tau}^* \mathbf{A}_{k,k+\tau}^{-1} \boldsymbol{\gamma}_{k+2\tau} - a_n / x_{n1}) + R_{k1} \boldsymbol{\gamma}_{k+2\tau}\big). \end{aligned} \tag{B.11}$$

When $|\boldsymbol{\gamma}_{k+2\tau}^* \mathbf{A}_{k,k+\tau}^{-1} \boldsymbol{\gamma}_{k+\tau}| \le v_n^3$, $|a_n - \boldsymbol{\gamma}_{k+\tau}^* \mathbf{A}_{k,k+\tau}^{-1} \boldsymbol{\gamma}_{k+\tau}| \le v_n^3$, we have

$$\|R_{k1}\| \le K v_n^2.$$

Using similar approach of the proof of Lemma 3.7(a), one can prove that when $k \le T - \log^2 n$, $|\boldsymbol{\gamma}_{k+l\tau}^* \mathbf{A}_{k,\dots,k+l\tau}^{-1} \boldsymbol{\gamma}_{k+(l+1)\tau}| \le v_n^3$, $|\boldsymbol{\gamma}_{k+(l+1)\tau}^* \mathbf{A}_{k,\dots,k+l\tau}^{-1} \boldsymbol{\gamma}_{k+l\tau}| \le v_n^3$, and $|\boldsymbol{\gamma}_{k+l\tau}^* \mathbf{A}_{k,\dots,k+l\tau}^{-1} \boldsymbol{\gamma}_{k+l\tau} - a_n| \le v_n^3$, for $l = 1, \dots, [\log^2 n]$, we have

$$\mathrm{P}(|\boldsymbol{\gamma}_{k+2\tau}^* \mathbf{A}_{k,k+\tau}^{-1} \boldsymbol{\gamma}_{k+2\tau} - a_n / x_{n1}| \ge v_n^3) = o(n^{-t}).$$

Therefore, by (B.11), we have

$$\|\boldsymbol{\gamma}_{k+\tau}^* \mathbf{A}_k^{-1}\| \le 2 \|\boldsymbol{\gamma}_{k+\tau}^* \mathbf{A}_{k,k+\tau}^{-1}\| + (1 - \eta') \|\boldsymbol{\gamma}_{k+2\tau}^* \mathbf{A}_{k,k+\tau}^{-1}\| + K v_n. \tag{B.12}$$

Similarly, one can prove that

$$\begin{aligned} &\|\boldsymbol{\gamma}_{k+2\tau}^* \mathbf{A}_{k,k+\tau}^{-1}\| \\ &\quad \le 2 \|\boldsymbol{\gamma}_{k+2\tau}^* \mathbf{A}_{k,k+\tau,k+2\tau}^{-1}\| + (1 - \eta') \|\boldsymbol{\gamma}_{k+3\tau}^* \mathbf{A}_{k,k+\tau,k+2\tau}^{-1}\| + K v_n. \end{aligned} \tag{B.13}$$

By induction, for any $k \le T - [\log^2 n]$ and $\ell \le [\log^2 n]$, one obtains

$$\begin{aligned}
&\|\boldsymbol{\gamma}_{k+\tau}\mathbf{A}_k^{-1}\| \\
&\quad \le 2\sum_{l=1}^{\ell}(1-\eta')^{l-1}\|\boldsymbol{\gamma}^*_{k+l\tau}\mathbf{A}^{-1}_{k,\dots,k+l\tau}\| \\
&\qquad + (1-\eta')^{\ell}\|\boldsymbol{\gamma}^*_{k+(\ell+1)\tau}\mathbf{A}^{-1}_{k,\dots,k+\ell\tau}\| + K\ell v_n,
\end{aligned} \tag{B.14}$$

where $\eta' \in (0, \eta)$ is a constant. Since

$$\|\boldsymbol{\gamma}^*_{k+l\tau}\mathbf{A}^{-1}_{k,\dots,k+l\tau}\|^2 \to \frac{c}{2}\int \frac{1}{(x-u)^2}\,dF_c(x) =: K_1$$

uniformly for $k \le T + \tau - [\log^2 n]$ and $l \le [\log^2 n]$, then for any $K > \frac{2\sqrt{K_1+\varepsilon}}{\eta'}$, when n is large, we have

$$\begin{aligned}
&\mathrm{P}\big(\|\boldsymbol{\gamma}^*_{k+\tau}\mathbf{A}_k^{-1}\| \ge K\big) \\
&\quad \le \sum_{l=1}^{[\log^2 n]}\big[\mathrm{P}\big(|\boldsymbol{\gamma}^*_{k+(l+1)\tau}\mathbf{A}^{-1}_{k,\dots,k+l\tau}\boldsymbol{\gamma}_{k+l\tau}| \ge v_n^3\big) \\
&\qquad + \mathrm{P}\big(|\boldsymbol{\gamma}^*_{k+l\tau}\mathbf{A}^{-1}_{k,\dots,k+l\tau}\boldsymbol{\gamma}_{k+(l+1)\tau}| \ge v_n^3\big) \\
&\qquad + \mathrm{P}\big(|\boldsymbol{\gamma}^*_{k+l\tau}\mathbf{A}^{-1}_{k,\dots,k+l\tau}\boldsymbol{\gamma}_{k+l\tau} - a_n| \ge v_n^3\big)\big] \\
&\quad = o\big(n^{-t}\big).
\end{aligned} \tag{B.15}$$

This proves the lemma for $k \le T + \tau - [\log^2 n]$.

When $k > T + \tau - [\log^2 n]$, by the first equality of (B.11) and Lemma 3.6(v), when $|\boldsymbol{\gamma}^*_{k+2\tau}\mathbf{A}^{-1}_{k,k+\tau}\boldsymbol{\gamma}_{k+2\tau}| \le 1$ [which, by (B.10), occurs with probability $1 - o(n^{-t})$], we have

$$\begin{aligned}
&\big|1 + \boldsymbol{\gamma}^*_{k+\tau}\mathbf{A}^{-1}_{k,k+\tau}\boldsymbol{\gamma}_{k+2\tau} - a_n\boldsymbol{\gamma}^*_{k+2\tau}\mathbf{A}^{-1}_{k,k+\tau}\boldsymbol{\gamma}_{k+2\tau} + R_{k1}\boldsymbol{\gamma}_{k+2\tau}\big| \\
&\quad \ge 1 - v_n^3 - \big(\tfrac{1}{2} - \eta\big) - K v_n^2 \ge \tfrac{1}{2} + \eta',
\end{aligned}$$

for some constant $\eta' > 0$. Therefore,

$$\|\boldsymbol{\gamma}^*_{k+\tau}\mathbf{A}_k^{-1}\| \le 2\|\boldsymbol{\gamma}^*_{k+\tau}\mathbf{A}^{-1}_{k,k+\tau}\| + (1-\eta')\|\boldsymbol{\gamma}^*_{k+2\tau}\mathbf{A}^{-1}_{k,k+\tau}\| + K v_n.$$

Again, by using induction, the lemma can be proved for the case where $k > T - \log^2 n$.

Therefore, the proof of the lemma is complete.

B.10. Proof of Lemma 3.10. As in last subsection, we first consider the case $k \le T + \tau - [\log^2 n]$. Note that

$$\mathbf{A}_k^{-1} = (\widetilde{\mathbf{A}}_k + \boldsymbol{\gamma}_{k+2\tau}\boldsymbol{\gamma}_{k+\tau}^*)^{-1} = \widetilde{\mathbf{A}}_k^{-1} - \frac{\widetilde{\mathbf{A}}_k^{-1}\boldsymbol{\gamma}_{k+2\tau}\boldsymbol{\gamma}_{k+\tau}^*\widetilde{\mathbf{A}}_k^{-1}}{1+\boldsymbol{\gamma}_{k+\tau}^*\widetilde{\mathbf{A}}_k^{-1}\boldsymbol{\gamma}_{k+2\tau}},$$

$$\widetilde{\mathbf{A}}_k^{-1} = \mathbf{A}_{k,k+\tau}^{-1} - \frac{\mathbf{A}_{k,k+\tau}^{-1}\boldsymbol{\gamma}_{k+\tau}\boldsymbol{\gamma}_{k+2\tau}^*\mathbf{A}_{k,k+\tau}^{-1}}{1+\boldsymbol{\gamma}_{k+2\tau}^*\mathbf{A}_{k,k+\tau}^{-1}\boldsymbol{\gamma}_{k+\tau}},$$

and

$$\boldsymbol{\gamma}_{k+\tau}^*\mathbf{A}_k^{-1} = \boldsymbol{\gamma}_{k+\tau}^*\widetilde{\mathbf{A}}_k^{-1} - \frac{\boldsymbol{\gamma}_{k+\tau}^*\widetilde{\mathbf{A}}_k^{-1}\boldsymbol{\gamma}_{k+2\tau}\boldsymbol{\gamma}_{k+\tau}^*\widetilde{\mathbf{A}}_k^{-1}}{1+\boldsymbol{\gamma}_{k+\tau}^*\widetilde{\mathbf{A}}_k^{-1}\boldsymbol{\gamma}_{k+2\tau}} = \frac{\boldsymbol{\gamma}_{k+\tau}^*\widetilde{\mathbf{A}}_k^{-1}}{1+\boldsymbol{\gamma}_{k+\tau}^*\widetilde{\mathbf{A}}_k^{-1}\boldsymbol{\gamma}_{k+2\tau}}.$$

By similar approach to prove Lemmas 3.7 and 3.9, we have

$$\begin{aligned}
|\boldsymbol{\gamma}_{k+\tau}^*\mathbf{A}_{k,k+\tau}^{-1}\boldsymbol{\gamma}_{k+2\tau}| &\le v_n^3 \qquad \text{with probability } 1-o(n^{-t}),\\
|\boldsymbol{\gamma}_{k+\tau}^*\mathbf{A}_{k,k+\tau}^{-2}\boldsymbol{\gamma}_{k+2\tau}| &\le v_n^3 \qquad \text{with probability } 1-o(n^{-t}),\\
|\boldsymbol{\gamma}_{k+2\tau}^*\mathbf{A}_{k,k+\tau}^{-1}\boldsymbol{\gamma}_{k+2\tau} - a_n/x_{n1}| &\le v_n^3 \qquad \text{with probability } 1-o(n^{-t}),\\
|\boldsymbol{\gamma}_{k+\tau}^*\mathbf{A}_{k,k+\tau}^{-1}\boldsymbol{\gamma}_{k+\tau} - a_n| &\le v_n^3 \qquad \text{with probability } 1-o(n^{-t}).
\end{aligned}$$

By Remark 3.2,

$$\boldsymbol{\gamma}_{k+\tau}^*\mathbf{A}_{k,k+\tau}^{-2}\boldsymbol{\gamma}_{k+\tau} = \frac{1}{2T}\operatorname{tr}\mathbf{A}^{-2} + o(v_n^3) \le K \qquad \text{with probability } 1-o(n^{-t}).$$

By Lemma 3.9,

$$\|\boldsymbol{\gamma}_{k+2\tau}^*\mathbf{A}_{k,k+\tau}^{-1}\|^2 = |\boldsymbol{\gamma}_{k+2\tau}^*\mathbf{A}_{k,k+\tau}^{-1}(\mathbf{A}_{k,k+\tau}^*)^{-1}\boldsymbol{\gamma}_{k+2\tau}| \le K \qquad \text{with probability } 1-o(n^{-t}),$$

$$|\boldsymbol{\gamma}_{k+2\tau}^*\mathbf{A}_{k,k+\tau}^{-2}\boldsymbol{\gamma}_{k+2\tau}| \le |\boldsymbol{\gamma}_{k+2\tau}^*\mathbf{A}_{k,k+\tau}^{-1}(\mathbf{A}_{k,k+\tau}^*)^{-1}\boldsymbol{\gamma}_{k+2\tau}| \le K \qquad \text{with probability } 1-o(n^{-t}).$$

By Lemma 3.5,

$$\|\boldsymbol{\gamma}_{k+\tau}^*\mathbf{A}_{k,k+\tau}^{-1}\|^2 = \frac{1}{2T}\operatorname{tr}\mathbf{A}_{k,k+\tau}^{-1}(\mathbf{A}_{k,k+\tau}^*)^{-1} + o(v_n^3) \le K \qquad \text{with probability } 1-o(n^{-t}).$$

Also, we have

$$\begin{aligned}
\boldsymbol{\gamma}_{k+\tau}^*\widetilde{\mathbf{A}}_k^{-1}\boldsymbol{\gamma}_{k+2\tau} &= \boldsymbol{\gamma}_{k+\tau}^*\mathbf{A}_{k,k+\tau}^{-1}\boldsymbol{\gamma}_{k+2\tau} - \frac{\boldsymbol{\gamma}_{k+\tau}^*\mathbf{A}_{k,k+\tau}^{-1}\boldsymbol{\gamma}_{k+\tau}\boldsymbol{\gamma}_{k+2\tau}^*\mathbf{A}_{k,k+\tau}^{-1}\boldsymbol{\gamma}_{k+2\tau}}{1+\boldsymbol{\gamma}_{k+2\tau}^*\mathbf{A}_{k,k+\tau}^{-1}\boldsymbol{\gamma}_{k+\tau}}\\
&= -x_{n0} + o(v_n^3) \qquad \text{with probability } 1-o(n^{-t}).
\end{aligned}$$

Therefore, with probability $1-o(n^{-t})$, we have

$$\begin{aligned}
&\|\boldsymbol{\gamma}_{k+\tau}^{*}\mathbf{A}_{k}^{-1}\widetilde{\mathbf{A}}_{k}^{-1}\boldsymbol{\gamma}_{k+2\tau}\boldsymbol{\gamma}_{k+\tau}^{*}\widetilde{\mathbf{A}}_{k}^{-1}\| \\
&\quad=\bigg\|\frac{\boldsymbol{\gamma}_{k+\tau}^{*}\widetilde{\mathbf{A}}_{k}^{-1}}{1+\boldsymbol{\gamma}_{k+\tau}^{*}\widetilde{\mathbf{A}}_{k}^{-1}\boldsymbol{\gamma}_{k+2\tau}} \\
&\qquad\times\bigg(\mathbf{A}_{k,k+\tau}^{-1}-\frac{\mathbf{A}_{k,k+\tau}^{-1}\boldsymbol{\gamma}_{k+\tau}\boldsymbol{\gamma}_{k+2\tau}^{*}\mathbf{A}_{k,k+\tau}^{-1}}{1+\boldsymbol{\gamma}_{k+2\tau}^{*}\mathbf{A}_{k,k+\tau}^{-1}\boldsymbol{\gamma}_{k+\tau}}\bigg)\boldsymbol{\gamma}_{k+2\tau}\boldsymbol{\gamma}_{k+\tau}^{*}\widetilde{\mathbf{A}}_{k}^{-1}\bigg\| \\
&\quad=\bigg|\frac{1}{1+\boldsymbol{\gamma}_{k+\tau}^{*}\widetilde{\mathbf{A}}_{k}^{-1}\boldsymbol{\gamma}_{k+2\tau}}\bigg| \\
&\qquad\times\bigg|\boldsymbol{\gamma}_{k+\tau}^{*}\bigg(\mathbf{A}_{k,k+\tau}^{-1}-\frac{\mathbf{A}_{k,k+\tau}^{-1}\boldsymbol{\gamma}_{k+\tau}\boldsymbol{\gamma}_{k+2\tau}^{*}\mathbf{A}_{k,k+\tau}^{-1}}{1+\boldsymbol{\gamma}_{k+2\tau}^{*}\mathbf{A}_{k,k+\tau}^{-1}\boldsymbol{\gamma}_{k+\tau}}\bigg)^{2}\boldsymbol{\gamma}_{k+2\tau}\bigg| \\
&\qquad\times\bigg\|\boldsymbol{\gamma}_{k+\tau}^{*}\bigg(\mathbf{A}_{k,k+\tau}^{-1}-\frac{\mathbf{A}_{k,k+\tau}^{-1}\boldsymbol{\gamma}_{k+\tau}\boldsymbol{\gamma}_{k+2\tau}^{*}\mathbf{A}_{k,k+\tau}^{-1}}{1+\boldsymbol{\gamma}_{k+2\tau}^{*}\mathbf{A}_{k,k+\tau}^{-1}\boldsymbol{\gamma}_{k+\tau}}\bigg)\bigg\| \\
&\quad\le M_1
\end{aligned}$$

for some $M_1>0$. By Remark 3.1,

$$\|\boldsymbol{\gamma}_{k+\tau}^{*}\mathbf{A}_{k,k+\tau}^{-2}\|^{2}=\frac{1}{2T}\operatorname{tr}\mathbf{A}^{-2}(\mathbf{A}^{*})^{-2}+o(v_n^{3})\le K \quad \text{with probability } 1-o(n^{-t}).$$

This implies, with probability $1-o(n^{-t})$

$$\begin{aligned}
&\|\boldsymbol{\gamma}_{k+\tau}^{*}\mathbf{A}_{k}^{-1}\widetilde{\mathbf{A}}_{k}^{-1}\| \\
&\quad=\bigg\|\frac{\boldsymbol{\gamma}_{k+\tau}^{*}}{1+\boldsymbol{\gamma}_{k+\tau}^{*}\widetilde{\mathbf{A}}_{k}^{-1}\boldsymbol{\gamma}_{k+2\tau}}\bigg(\mathbf{A}_{k,k+\tau}^{-1}-\frac{\mathbf{A}_{k,k+\tau}^{-1}\boldsymbol{\gamma}_{k+\tau}\boldsymbol{\gamma}_{k+2\tau}^{*}\mathbf{A}_{k,k+\tau}^{-1}}{1+\boldsymbol{\gamma}_{k+2\tau}^{*}\mathbf{A}_{k,k+\tau}^{-1}\boldsymbol{\gamma}_{k+\tau}}\bigg)^{2}\bigg\| \\
&\quad\le M_2+|b_n|\|\boldsymbol{\gamma}_{k+2\tau}^{*}\mathbf{A}_{k,k+\tau}^{-2}\|
\end{aligned}$$

for some $M_2>0$ and

$$b_n=-\frac{c_n\mathrm{E}m_n/2}{1-(c_n\mathrm{E}m_n/2)(c_n\mathrm{E}m_n/2x_{n1})}=-\frac{a_n}{x_{n1}}$$

with

$$\bigg|\frac{a_n}{x_{n1}}\bigg|\le\sqrt{|x_{n0}|/|x_{n1}|}\le\sqrt{1-\eta}.$$

Therefore, we have

$$\begin{aligned}
&\|\boldsymbol{\gamma}_{k+\tau}^*\mathbf{A}_k^{-2}\|\\
&\quad=\left\|\boldsymbol{\gamma}_{k+\tau}^*\mathbf{A}_k^{-1}\left(\widetilde{\mathbf{A}}_k^{-1}-\frac{\widetilde{\mathbf{A}}_k^{-1}\boldsymbol{\gamma}_{k+2\tau}\boldsymbol{\gamma}_{k+\tau}^*\widetilde{\mathbf{A}}_k^{-1}}{1+\boldsymbol{\gamma}_{k+\tau}^*\widetilde{\mathbf{A}}_k^{-1}\boldsymbol{\gamma}_{k+2\tau}}\right)\right\|\\
&\quad\le\|\boldsymbol{\gamma}_{k+\tau}^*\mathbf{A}_k^{-1}\widetilde{\mathbf{A}}_k^{-1}\|+\left|\frac{1}{1+\boldsymbol{\gamma}_{k+\tau}^*\widetilde{\mathbf{A}}_k^{-1}\boldsymbol{\gamma}_{k+2\tau}}\right|\|\boldsymbol{\gamma}_{k+\tau}^*\mathbf{A}_k^{-1}\widetilde{\mathbf{A}}_k^{-1}\boldsymbol{\gamma}_{k+2\tau}\boldsymbol{\gamma}_{k+\tau}^*\widetilde{\mathbf{A}}_k^{-1}\|\\
&\quad\le(2+\varepsilon)M_1+M_2+\sqrt{1-\eta}\|\boldsymbol{\gamma}_{k+2\tau}^*\mathbf{A}_{k,k+\tau}^{-2}\|,
\end{aligned}$$

where $\varepsilon>0$ is a constant. Then similar to the proof of Lemma 3.9, using the recursion above we have

$$\mathrm{P}\big(|\boldsymbol{\gamma}_{k+\tau}^*\mathbf{A}_k^{-2}(\mathbf{A}_k^*)^{-2}\boldsymbol{\gamma}_{k+\tau}|\ge K\big)=o(n^{-t})$$

for some $K>0$. When $k>T-\log^2 n$, one can similarly prove the inequality above. The proof of the lemma is complete.

B.11. Proof of Lemma 3.11. We first consider the case where $\log^2 n<k<T-\log^2 n$. Note that $\mathbf{A}=\mathbf{A}_k+\boldsymbol{\gamma}_k\boldsymbol{\beta}_k^*+\boldsymbol{\beta}_k\boldsymbol{\gamma}_k^*$, where $\boldsymbol{\beta}_k=\boldsymbol{\gamma}_{k-\tau}+\boldsymbol{\gamma}_{k+\tau}$. We have

$$\begin{aligned}
&\operatorname{tr}\mathbf{A}_k^{-1}-\operatorname{tr}\mathbf{A}^{-1}\\
&\quad=\frac{d}{dz}\log\big((1+\varepsilon_1)(1+\varepsilon_2)-\boldsymbol{\gamma}_k^*\mathbf{A}_k^{-1}\boldsymbol{\gamma}_k\boldsymbol{\beta}_k^*\mathbf{A}_k^{-1}\boldsymbol{\beta}_k\big)\\
&\quad=\frac{d}{dz}\log\left((1+\varepsilon_1)(1+\varepsilon_2)-(\varepsilon_3+\varepsilon_4+a_n)\left(\varepsilon_5+\frac{2a_n}{x_{n1}}\right)\right)\\
&\quad=\frac{d}{dz}\log\left(x_{n1}-x_{n0}+\varepsilon_1+\varepsilon_2+\varepsilon_1\varepsilon_2-a_n\varepsilon_5-\left(\frac{2a_n}{x_{n1}}+\varepsilon_5\right)(\varepsilon_3+\varepsilon_4)\right),
\end{aligned}\tag{B.16}$$

where ε_i's are defined in (4.34). Note that

$$\mathrm{E}(\varepsilon_i|\boldsymbol{\gamma}_j,j\ne k)=0\qquad\text{for } i=1,2,3.$$

Therefore, by Taylor's expansion, Cauchy integral and Lemma 3.6 part (b), we have

$$\begin{aligned}
&\left|\mathrm{E}\big(\operatorname{tr}\mathbf{A}_k^{-1}-\operatorname{tr}\mathbf{A}^{-1}\big)-\frac{d}{dz}\log(x_{n1}-x_{n0})\right|\\
&\quad\le\left|\frac{d}{dz}\mathrm{E}\left[\log\left(1+\frac{\varepsilon_1+\varepsilon_2+\varepsilon_1\varepsilon_2-a_n\varepsilon_5-((2a_n/x_{n1})+\varepsilon_5)(\varepsilon_3+\varepsilon_4)}{x_{n1}-x_{n0}}\right)\right.\right.\\
&\qquad\qquad\left.\left.-\frac{\varepsilon_1+\varepsilon_2}{x_{n1}-x_{n0}}-\frac{2\varepsilon_3a_n}{x_{n1}(x_{n1}-x_{n0})}\right]\right|\\
&\quad\le Kv_n^{-1}\sup_{|\xi-z|=v_n/2}\left[\sum_{i=1}^5\big(\mathrm{E}|\varepsilon_i^2(\xi)|\big)+|\mathrm{E}\varepsilon_4(\xi)|+|\mathrm{E}\varepsilon_5(\xi)|\right].
\end{aligned}\tag{B.17}$$

By applying Lemmas 3.9 and 3.10, one can easily verify that

$$\mathrm{E}\big|\varepsilon_i^2(\xi)\big| = O\big(n^{-1}\big) \qquad \text{for } i = 1, 2, 3. \tag{B.18}$$

Also, by (4.2),

$$\big|\mathrm{E}\varepsilon_4(\xi)\big| = \left|\frac{1}{2T}\mathrm{E}\big(\operatorname{tr}\mathbf{A}_k^{-1}(\xi) - \operatorname{tr}\mathbf{A}^{-1}(\xi)\big)\right| \le \frac{K}{T v_n}, \tag{B.19}$$

and similar to the proof of (4.4)

$$\big|\mathrm{E}\varepsilon_4^2(\xi)\big| \le \frac{1}{4T^2}\mathrm{E}\big|\operatorname{tr}\mathbf{A}_k^{-1}(\xi) - \mathrm{E}\operatorname{tr}\mathbf{A}_k^{-1}(\xi)\big|^2 + \big|\mathrm{E}\varepsilon_4(\xi)\big|^2 = O\left(\frac{1}{n}\right). \tag{B.20}$$

By the proof of Lemma 3.7(a) with noticing $|x_{n0}/x_{n1}| \le 1-\eta$, when $\log^2 n \le k \le T - \log^2 n$, for $i = 1, 2$, one can prove that

$$\begin{aligned} &\mathrm{E}\left|\boldsymbol{\gamma}_{k+\tau}^*\mathbf{A}_k^{-1}\boldsymbol{\gamma}_{k+\tau} - \frac{a_n}{x_{n1}}\right|^i = o(1), \\ &\mathrm{E}\left|\boldsymbol{\gamma}_{k-\tau}^*\mathbf{A}_k^{-1}\boldsymbol{\gamma}_{k-\tau} - \frac{a_n}{x_{n1}}\right|^i = o(1), \end{aligned} \tag{B.21}$$

and by the proof of Lemma 3.8(a),

$$\mathrm{E}\big|\boldsymbol{\gamma}_{k-\tau}^*\mathbf{A}_k^{-1}\boldsymbol{\gamma}_k\big|^i = o(1), \qquad \big|\mathrm{E}\boldsymbol{\gamma}_{k+\tau}^*\mathbf{A}_k^{-1}\boldsymbol{\gamma}_{k-\tau}\big|^i = o(1). \tag{B.22}$$

inequalities (B.21) and (B.22) imply that

$$\mathrm{E}\big|\varepsilon_5(\xi)\big|^i = o(1). \tag{B.23}$$

Combining (B.17), (B.18), (B.19), (B.20) and (B.23), the first conclusion of Lemma 3.11 is proved when $\log^2 n \le k \le T - \log^2 n$. If $k > T - \log^2 n$, by Lemmas 3.7(b1) and 3.8(a), one may modify the right-hand sides of (B.21)–(B.22) as $O(1)$. This also proves the lemma. The conclusion for $k < \log^2 n$ can be proved similarly.

The second conclusion of the lemma can be proved similarly. The proof of the lemma is complete.

B.12. Proof of Lemma 3.7(b2). We assume that $k < T - \log^2 n$ and prove the first statement only, as the second follows by symmetry. As in the proof of Lemma 3.7(a), write $W_k = \boldsymbol{\gamma}_{k+\tau}^*\mathbf{A}_k^{-1}\boldsymbol{\gamma}_{k+\tau}$ and $W_{k,k+\tau,\dots,k+s\tau} = \boldsymbol{\gamma}_{k+(s+1)\tau}^* \times \mathbf{A}_{k,k+\tau,\dots,k+s\tau}^{-1}\boldsymbol{\gamma}_{k+(s+1)\tau}$. Then by (B.2), we have

$$W_{k,\dots,k+(s-1)\tau} = \frac{a_n + r_1(k+s\tau)}{1 - a_n W_{k,\dots,k+s\tau} + r_2(k+s\tau)},$$

where

$$r_1(k+s\tau) = \boldsymbol{\gamma}_{k+s\tau}^* \mathbf{A}_{k,\dots,k+s\tau}^{-1} \boldsymbol{\gamma}_{k+s\tau} - a_n,$$

$$\begin{aligned} r_2(k+s\tau) = &-(\boldsymbol{\gamma}_{k+s\tau}^* \mathbf{A}_{k,\dots,k+s\tau}^{-1} \boldsymbol{\gamma}_{k+s\tau} - a_n)\boldsymbol{\gamma}_{k+(s+1)\tau}^* \mathbf{A}_{k,\dots,k+s\tau}^{-1} \boldsymbol{\gamma}_{k+(s+1)\tau} \\ &+ \boldsymbol{\gamma}_{k+(s+1)\tau}^* \mathbf{A}_{k,\dots,k+s\tau}^{-1} \boldsymbol{\gamma}_{k+s\tau} + \boldsymbol{\gamma}_{k+s\tau}^* \mathbf{A}_{k,\dots,k+(s+1)\tau}^{-1} \boldsymbol{\gamma}_{k+s\tau} \\ &+ \boldsymbol{\gamma}_{k+(s+1)\tau}^* \mathbf{A}_{k,\dots,k+s\tau}^{-1} \boldsymbol{\gamma}_{k+s\tau} \boldsymbol{\gamma}_{k+s\tau}^* \mathbf{A}_{k,\dots,k+(s+1)\tau}^{-1} \boldsymbol{\gamma}_{k+s\tau}. \end{aligned}$$

Therefore, we have

$$\begin{aligned} W_k - \frac{a_n}{x_{n1}} &= \frac{a_n + r_1(k+\tau)}{1 - a_n W_{k,k+\tau} + r_2(k+\tau)} - \frac{a_n}{x_{n1}} \\ &= \frac{r_1(k+\tau)}{1 - a_n W_{k,k+\tau} + r_2(k+\tau)} - \frac{a_n r_2(k+\tau)}{x_{n1}(1 - a_n W_{k,k+\tau} + r_2(k+\tau))} \\ &\quad + \frac{a_n^2 (W_{k,k+\tau} - (a_n/x_{n1}))}{x_{n1}(1 - a_n W_{k,k+\tau} + r_2(k+\tau))}. \end{aligned} \tag{B.24}$$

By Lemma 3.11, when $k + s\tau \le T$,

$$|\mathrm{E} r_1(k+s\tau)| = \left| \frac{1}{2T} \mathrm{E}\, \mathrm{tr}\, \mathbf{A}_{k,\dots,k+s\tau}^{-1} - a_n \right| = O\left(\frac{s}{n}\right) = O\left(\frac{\log^2 n}{n}\right).$$

Using this estimate together with Lemmas 3.4 and 3.9, one can prove that

$$\begin{aligned} &\mathrm{E}(|r_1(k+s\tau)|^p) \\ &\quad \le K(|\mathrm{E} r_1(k+s\tau)|^p + \mathrm{E}|r_1(k+s\tau) - \mathrm{E} r_1(k+s\tau)|^p) \\ &\quad \le K(n^{-p} \log^{2p} n + n^{-p} \mathrm{E}(\mathrm{tr}\, \mathbf{A}_{k,\dots,k+s\tau}^{-1} (\mathbf{A}_{k,\dots,k+s\tau}^*)^{-1})^{p/2}) \\ &\quad \le K n^{-p/2}, \end{aligned} \tag{B.25}$$

which implies that for any fixed $\delta > 0$,

$$\mathrm{P}(|r_1(k+s\tau)| \ge n^{-0.5+\delta}) = o(n^{-t}). \tag{B.26}$$

By this and Lemmas 3.7(b1) and 3.4, one can prove that

$$\mathrm{P}(|r_2(k+s\tau)| \ge n^{-0.5+\delta}) = o(n^{-t}). \tag{B.27}$$

In Section 4, we have proved that with probability $1 - o(n^{-t})$, $|W_{k,k+\tau} - \frac{a_n}{x_{n1}}| \le \upsilon_n^6$. Also by Lemma 3.6(ii)(b), we have $|x_{n1}| \ge \frac{1}{2}$ which implies that $|\frac{1}{1 - a_n W_{k,k+\tau} + r_2(k+\tau)}|$ is bounded by 3 with probability $1 - o(n^{-t})$.

Moreover, by the fact that $|\frac{a_n}{x_{n1}}| = \sqrt{|\frac{x_{n0}}{x_{n1}}|} \leq \sqrt{1-\eta} < 1 - \frac{1}{2}\eta$, we have, with probability $1 - o(n^{-t})$,

$$\begin{aligned}\left|\frac{a_n}{1 - a_n W_{k,k+\tau} + r_2(k+\tau)}\right| &\leq \frac{|a_n|}{|x_{n1}| - v_n^4} \leq \frac{(1-(1/2)\eta)|x_{n1}|}{|x_{n1}| - v_n^4} \\ &\leq \frac{1-(1/2)\eta}{1 - 2v_n^4} \leq 1 - \eta',\end{aligned}$$

for some $0 < \eta' < \frac{1}{2}\eta$. In (B.24), split the first term as

$$\begin{aligned}&\frac{r_1(k+\tau)}{1 - a_n W_{k,k+\tau} + r_2(k+\tau)} \\ &\quad = \frac{r_1(k+\tau)}{1 - a_n W_{k,k+\tau}} - \frac{r_1(k+\tau) r_2(k+\tau)}{(1 - a_n W_{k,k+\tau})(1 - a_n W_{k,k+\tau} + r_2(k+\tau))}\end{aligned}$$

and the second term as

$$\begin{aligned}&\frac{a_n r_2(k+\tau)}{x_{n1}(1 - a_n W_{k,k+\tau} + r_2(k+\tau))} \\ &\quad = \frac{a_n r_2(k+\tau)}{x_{n1}(1 - a_n W_{k,k+\tau})} - \frac{a_n r_2^2(k+\tau)}{x_{n1}(1 - a_n W_{k,k+\tau})(1 - a_n W_{k,k+\tau} + r_2(k+\tau))}.\end{aligned}$$

Noting that $|W_k| \leq K v_n^{-1}$, we have

$$\begin{aligned}&\left|\mathrm{E}W_k - \frac{a_n}{x_{n1}}\right| \\ &\quad \leq K n^{-1+2\delta} + K\left|\mathrm{E}r_1(k+\tau)\right| + K\left|\mathrm{E}r_2(k+\tau)\right| \\ &\qquad + (1-\eta')^2\left|\mathrm{E}W_{k,k+\tau} - \frac{a_n}{x_{n1}}\right| \\ &\quad \vdots \\ &\quad \leq K\ell n^{-1+2\delta} + K\sum_{s=1}^{\ell}\left|\mathrm{E}r_1(k+s\tau)\right| + K\sum_{s=1}^{\ell}\left|\mathrm{E}r_2(k+s\tau)\right| \\ &\qquad + (1-\eta')^{2\ell}\left|\mathrm{E}W_{k,\dots,k+\ell\tau} - \frac{a_n}{x_{n1}}\right|.\end{aligned} \tag{B.28}$$

By choosing $\ell = [\log^2 n]$ and $\delta < 1/106$, we can show that $\sum_{s=1}^{\ell} |\mathrm{E}r_i(k+s\tau)| = o(1/(nv_n))$, $i = 1, 2$ and that $(1-\eta')^{2\ell}|\mathrm{E}W_{k,\dots,k+\ell\tau} - \frac{a_n}{x_{n1}}| = o(1/(nv_n))$. Substituting all the above into (B.28), we have $|\mathrm{E}W_k - \frac{a_n}{x_{n1}}| = o(1/(nv_n))$.

B.13. Proof of Lemma 3.7(b3). Again, we assume that $k < T - \log^2 n$ and prove the first statement only, as the second follows by symmetry. As in the proof of Lemma 3.7(b2), we have

$$\begin{aligned}
\mathrm{E}\left|W_k - \frac{a_n}{x_{n1}}\right|^2 & \\
&\le K\mathrm{E}\bigl|r_1(k+\tau)\bigr|^2 + K\mathrm{E}\bigl|r_2(k+\tau)\bigr|^2 + \bigl(1-\eta'\bigr)^4\mathrm{E}\left|W_{k,k+\tau} - \frac{a_n}{x_{n1}}\right|^2 \\
&\vdots \\
&\le K\sum_{s=1}^{\ell}\mathrm{E}\bigl|r_1(k+s\tau)\bigr|^2 + K\sum_{s=1}^{\ell}\mathrm{E}\bigl|r_2(k+s\tau)\bigr|^2 \\
&\quad + \bigl(1-\eta'\bigr)^{4\ell}\mathrm{E}\left|W_{k,\dots,k+\ell\tau} - \frac{a_n}{x_{n1}}\right|^2 \\
&\le K\ell n^{-1+2\delta} = o\bigl(1/(nv_n)\bigr).
\end{aligned} \tag{B.29}$$

The proof of the lemma is complete.

B.14. Proof of Lemma 3.8(b1). By symmetry, we only consider the case $k \le T/2$. As in the proof of Lemma 3.8(a), write

$$\widetilde{W}_{k,\dots,k+s\tau} := \boldsymbol{\gamma}_{k-\tau}^* \mathbf{A}_{k,k+\tau,\dots,k+(s-1)\tau}^{-1} \boldsymbol{\gamma}_{k+s\tau}.$$

Then we have

$$\widetilde{W}_{k,\dots,k+s\tau} = \frac{\widetilde{r}_1(k+s\tau) - \widetilde{W}_{k,\dots,k+(s+1)\tau}(a_n + \widetilde{r}_2(k+s\tau))}{1 + r_2(k+s\tau) - a_n W_{k,\dots,k+s\tau}}, \tag{B.30}$$

where

$$\widetilde{r}_1(k+s\tau) = \boldsymbol{\gamma}_{k-\tau}^* \mathbf{A}_{k,\dots,k+s\tau}^{-1} \boldsymbol{\gamma}_{k+s\tau} \bigl(1 + \boldsymbol{\gamma}_{k+s\tau}^* \mathbf{A}_{k,\dots,k+s\tau}^{-1} \boldsymbol{\gamma}_{k+(s+1)\tau}\bigr),$$

$$\widetilde{r}_2(k+s\tau) = \boldsymbol{\gamma}_{k+s\tau}^* \mathbf{A}_{k,\dots,k+s\tau}^{-1} \boldsymbol{\gamma}_{k+s\tau} - a_n.$$

Similar to the proof of (B.27), one has

$$\mathrm{P}\bigl(\bigl|\widetilde{r}_i(k+\tau)\bigr| \ge n^{-0.5+\delta}\bigr) = o\bigl(n^{-t}\bigr), \qquad i = 1, 2. \tag{B.31}$$

Similar to the proof of (B.28), one can prove that for some $\eta' > 0$,

$$|\mathrm{E}\widetilde{W}_{k,\dots,k+s\tau}| \le Kn^{-1+2\delta} + K\bigl|\mathrm{E}\widetilde{r}_1(k+s\tau)\bigr| + \bigl(1-\eta'\bigr)|\mathrm{E}\widetilde{W}_{k,\dots,k+(s+1)\tau}|.$$

Therefore, when $k \le T/2$,

$$\begin{aligned}
|\mathrm{E}\widetilde{W}_k| &\le K\ell n^{-1+2\delta} + K\sum_{s=1}^{\ell}\bigl|\mathrm{E}\widetilde{r}_1(k+s\tau)\bigr| + \bigl(1-\eta'\bigr)^{\ell}|\mathrm{E}\widetilde{W}_{k,\dots,k+\ell\tau}| \\
&= o\bigl(1/(nv_n)\bigr).
\end{aligned}$$

The proof of the lemma is complete.

B.15. Proof of Lemma 3.8(b2). Using the notation of Lemma 3.8(b1), by triangle inequality, we have

$$\big(\mathrm{E}|\widetilde{W}_{k+s\tau}|^2\big)^{1/2} \le K\big(\mathrm{E}\big|\widetilde{r}_1(k+s\tau)\big|^2\big)^{1/2} + \big((1-\eta')\mathrm{E}|\widetilde{W}_{k,\dots,k+(s+1)\tau}|^2\big)^{1/2}.$$

Therefore, when $k \le T/2$ and $\ell = [\log^2 n]$,

$$\begin{aligned}\big(\mathrm{E}|\widetilde{W}_k|^2\big)^{1/2} &\le K\sum_{s=1}^{\ell}\big(\mathrm{E}\big|\widetilde{r}_1(k+s\tau)\big|^2\big)^{1/2} + \big(1-\eta'\big)^{\ell/2}\big(\mathrm{E}|\widetilde{W}_{k,\dots,k+\ell\tau}|^2\big)^{1/2}\\ &\le K\log^2 nn^{-1/2+\delta}.\end{aligned}$$

Therefore, when $2\delta < 1/212$,

$$\mathrm{E}|\widetilde{W}_k|^2 \le K\log^4 nn^{-1+2\delta} = o\big(1/(nv_n)\big)$$

and the proof of the lemma is complete.

Acknowledgements. The authors would like to thank the referees for their careful reading and invaluable comments which greatly improved the quality of the paper.

REFERENCES

BAI, Z. D. (1993). Convergence rate of expected spectral distributions of large random matrices. I. Wigner matrices. *Ann. Probab.* **21** 625–648. MR1217559

BAI, Z. D., MIAO, B. Q. and RAO, C. R. (1991). Estimation of directions of arrival of signals: Asymptotic results. In *Advances in Spectrum Analysis and Array Processing*, *Vol. I* (S. Haykin, ed.) 327–347. Prentice Hall, West Nyack, NY.

BAI, Z. D. and SILVERSTEIN, J. W. (1998). No eigenvalues outside the support of the limiting spectral distribution of large-dimensional sample covariance matrices. *Ann. Probab.* **26** 316–345. MR1617051

BAI, Z. D. and SILVERSTEIN, J. W. (2010). *Spectral Analysis of Large Dimensional Random Matrices*, 2nd ed. Springer, New York. MR2567175

BAI, Z. D. and SILVERSTEIN, J. W. (2012). No eigenvalues outside the support of the limiting spectral distribution of information-plus-noise type matrices. *Random Matrices Theory Appl.* **1** 1150004, 44. MR2930382

BAI, Z. D. and WANG, C. (2015). A note on the limiting spectral distribution of a symmetrized auto-cross covariance matrix. *Statist. Probab. Lett.* **96** 333–340. MR3281785

BAI, Z. D. and YAO, J.-F. (2008). Central limit theorems for eigenvalues in a spiked population model. *Ann. Inst. Henri Poincaré Probab. Stat.* **44** 447–474. MR2451053

BAIK, J. and SILVERSTEIN, J. W. (2006). Eigenvalues of large sample covariance matrices of spiked population models. *J. Multivariate Anal.* **97** 1382–1408. MR2279680

BURKHOLDER, D. L. (1973). Distribution function inequalities for martingales. *Ann. Probab.* **1** 19–42. MR0365692

JIN, B., WANG, C., BAI, Z. D., NAIR, K. K. and HARDING, M. (2014). Limiting spectral distribution of a symmetrized auto-cross covariance matrix. *Ann. Appl. Probab.* **24** 1199–1225. MR3199984

JOHNSTONE, I. M. (2001). On the distribution of the largest eigenvalue in principal components analysis. *Ann. Statist.* **29** 295–327. MR1863961

LI, Z., WANG, Q. and YAO, J. F. (2014). Identifying the number of factors from singular values of a large sample auto-covariance matrix. Preprint. Available at arXiv:1410.3687v2.

MARČENKO, V. A. and PASTUR, L. A. (1967). Distribution of eigenvalues in certain sets of random matrices. *Mat. Sb.* **72 (114)** 507–536. MR0208649

PAUL, D. and SILVERSTEIN, J. W. (2009). No eigenvalues outside the support of the limiting empirical spectral distribution of a separable covariance matrix. *J. Multivariate Anal.* **100** 37–57. MR2460475

RAO, C. R. and RAO, M. B. (1998). *Matrix Algebra and Its Applications to Statistics and Econometrics*. World Scientific, River Edge, NJ. MR1660868

YIN, Y. Q., BAI, Z. D. and KRISHNAIAH, P. R. (1988). On the limit of the largest eigenvalue of the large-dimensional sample covariance matrix. *Probab. Theory Related Fields* **78** 509–521. MR0950344

C. WANG
DEPARTMENT OF STATISTICS
AND APPLIED PROBABILITY
NATIONAL UNIVERSITY OF SINGAPORE
SINGAPORE 117546
E-MAIL: stawc@nus.edu.sg

B. JIN
DEPARTMENT OF STATISTICS
AND FINANCE
UNIVERSITY OF SCIENCE
AND TECHNOLOGY OF CHINA
96, JINZHAI ROAD
HEFEI 23 0026
P. R. CHINA
E-MAIL: jbs@ustc.edu.cn

Z. D. BAI
KLASMOE AND SCHOOL OF MATH. AND STAT.
NORTHEAST NORMAL UNIVERSITY
5268 RENMIN STREET, CHANGCHUN
JILIN PROVINCE 130024
P. R. CHINA
E-MAIL: baizd@nenu.edu.cn
URL: http://www.nenu.edu.cn/professor/pro/show.php?flag=1&id=123

K. K. NAIR
DEPARTMENT OF CIVIL AND
ENVIRONMENTAL ENGINEERING
STANFORD UNIVERSITY
439 PANAMA MALL
STANFORD, CALIFORNIA 94305
USA
E-MAIL: kknair@stanford.edu

M. HARDING
SANFORD SCHOOL OF PUBLIC POLICY
DUKE UNIVERSITY
DURHAM NORTH CALIFORNIA 27708
USA
E-MAIL: matthew.harding@duke.edu