Limiting spectral distribution of large-dimensional sample covariance matrices generated by VARMA

Baisuo Jin\textsuperscript{a,}\textsuperscript{*}, Cheng Wang\textsuperscript{a}, Baiqi Miao\textsuperscript{a}, Mong-Na Lo Huang\textsuperscript{b}

\textsuperscript{a} Department of Statistics and Finance, University of Science and Technology of China, Hefei, Anhui 230026, China
\textsuperscript{b} Department of Applied Mathematics, National Sun Yat-sen University, Kaohsiung, 80424, Taiwan

\textbf{ARTICLE INFO}

Article history:
Received 3 February 2008
Available online 3 July 2009

AMS 2008 subject classifications:
60F15
62H99

Keywords:
Large-dimensional random matrices
Limiting spectral distribution
Vector autoregression

\textbf{ABSTRACT}

The existence of a limiting spectral distribution (LSD) for a large-dimensional sample covariance matrix generated by the vector autoregressive moving average (VARMA) model is established. In particular, we obtain explicit forms of the LSDs for random matrices generated by a first-order vector autoregressive (VAR(1)) model and a first-order vector moving average (VMA(1)) model, as well as random coefficients for VAR(1) and VMA(1). The parameters for these explicit forms are also estimated. Finally, simulations demonstrate that the results are effective.

1. Introduction

Let \(X = (x_{i,t})_{N \times T} = (x_1, \ldots, x_T)\) be a random matrix and \(\Sigma = (\varepsilon_{i,t})_{N \times T} = (\varepsilon_1, \ldots, \varepsilon_T)\) be a white noise matrix with a common mean of 0 and variance of 1. Then a stationary and invertible vector autoregressive moving average (VARMA\((p, q)\)) model is of the form:

\[
\phi(B)x_t = \theta(B)\varepsilon_t,
\]

where \(\phi(B) = 1 - b_1B - \cdots - b_pB^p\) and \(\theta(B) = a_0 + a_1B + \cdots + a_qB^q\) are real polynomials in \(B\), which is a backshift operator \(Bx_t = x_{t-j}, j = 0, 1, \ldots\). It is easy to see that the rows of \(X\) are independent and the covariance matrices of the rows of \(X\) are same. The population covariance matrix has a particular sparse structure. For example, the population covariance matrix of VMA\((q)\) model is banded and has small entries far away from the diagonal. Estimation of banding of large-dimensional covariance matrices has been investigated in the literature \cite{1}. Estimation of sparse covariance matrices is of particular practical interest and has recently been studied \cite{2–4}. In the present study, we consider another important topic: the limiting spectral distribution (LSD) of large-dimensional sample covariance matrices and the estimation of parameters of the VARMA model.

Suppose that \(A_n\) is an \(n \times n\) Hermitian matrix with eigenvalues \(\lambda_j, j = 1, 2, \ldots, n\). We can define a one-dimensional distribution function

\[
F_{A_n}(x) = \frac{1}{n} \sum_{j=1}^{n} I(\lambda_j \leq x)
\]
called the empirical spectral distribution (ESD) of the matrix $A_n$, where $I(\cdot)$ denotes the indicator function. Many important statistics in multivariate analysis can be expressed as functions of the ESD (e.g., $\det(A_n) = \exp(n \int \ln x dF_{Ak}(x))$, $\text{tr}(A_n) = n \int x dF_{Ak}(x)$, etc). The limit distribution of $F_{Ak}$ is called the LSD of the sequence $\{A_n\}$. 

A sample covariance matrix is simply defined as $X(X)$, where $X = (X_{ij})_{N \times T}$ is a random matrix. If the dimension $N$ tends to infinity in proportion to the degrees of freedom $T$, namely, $T/N \to y \in (0, \infty)$, we call the sample covariance matrix a large-dimensional sample covariance matrix. Pioneering work (M-P law) [5] derived the LSD of a large-dimensional sample covariance matrix under the assumption that all the variables $X_{ij}$ are independent. When the entries of $X$ are not independent, the case $X = T^{1/2}Y$, where $T$ is a Hermitian matrix and the entries of $Y$ are independent, was considered [6]. Further extensions have been reported in the literature [7–10]. If the dependence structure cannot be expressed as $X = T^{1/2}Y$, other researchers considered the LSD of large-dimensional sample covariance matrices without a column independence structure [11].

Here we establish the existence of LSDs for large-dimensional sample covariance matrices generated by VARMA models. The question arises as to whether the explicit forms of these LSDs can be obtained. We can derive the explicit forms of the LSDs of the sample covariance matrices generated by a first-order vector autoregressive (VAR(1)) model and a first-order vector moving average (VMA(1)) model. Furthermore, if we add the moment condition, we can also obtain the LSDs of the sample covariance matrices generated by random-coefficient VAR(1) and random-coefficient VMA(1) models. It is easy to see that these explicit forms have some parameters that depend on these VARMA and random-coefficient VARMA models. Thus, it is necessary to provide estimates for these parameters. It is of interest to identify these explicit forms because the only known explicit forms of the densities of LSDs of large matrices are the semicircle law [12, 13], the M–P law [5], the LSD of multivariate F matrices [8, 14], the circle law [15], and the LSD of the product of a sample covariance and a Wigner matrix [10].

The remainder of the paper is organized as follows. The next section presents some preliminary results. The existence theorem is given in Section 3. Section 4 contains the explicit forms for the VAR(1) and VMA(1) cases. For the random-coefficient VAR(1) and VMA(1) cases, the explicit forms are derived in Section 5. Estimates and simulations are presented in Section 6.

2. Preliminary results

If the entries of $X$ are not independent, two important situations arise: $X$ has the dependence structure $X = T^{1/2}Y$, and $X$ does not have a column independence structure. Here, we introduce two important theorems for the two cases.

**Theorem 2.1** (Theorem 2.10 of [16]). Suppose that the entries of $\Sigma = (\varepsilon_{i,t})_{N,T}$ are independent complex random variables satisfying for any $\delta > 0$

$$\frac{1}{\delta^2 TN} \sum_{i,t} E(|\varepsilon_{i,t}|^2 I(|\varepsilon_{i,t}| \geq \delta \sqrt{T}) \to 0, \quad (2.3)$$

and assume that $\Gamma (= \Gamma_T)$ is a sequence of $T \times T$ Hermitian matrices independent of $\Sigma$ such that its ESD tends to a non-random and non-degenerate distribution $H$ in probability (or almost surely). Further assume that $T/N \to y \in (0, \infty)$. Then the ESD of the product $S \Gamma$ tends to a non-random limit $F$ in probability (or almost surely), where $S = \frac{1}{\sqrt{N}} \Sigma^{1/2} \Sigma$. 

Moreover, the LSD $F$ of Theorem 2.1 satisfies Eq. (1.2) of [17]:

$$m_{F} = \int \frac{1}{\tau(1-y-yzm)-z} dH(\tau), \quad (2.4)$$

where $m_{F}$ is the Stieltjes transform of $F$ given by:

$$m_{F} = \int \frac{1}{\lambda-z} dF(\lambda), \quad z \in \mathbb{C}^+ \equiv \{z \in \mathbb{C} : \text{Im} z > 0\}. \quad (2.5)$$

Because of the inversion formula

$$F([a, b]) = \lim_{\eta \to 0} \frac{1}{\pi} \int_{\eta}^{b} \text{Im} m_{F}(\xi + i\eta) d\xi, \quad (2.6)$$

where $a < b$ are continuity points of $F$, $F$ is uniquely determined by its Stieltjes transform. Theorem 2.1 also implies that the moments of $S \Gamma$ satisfy the following equation [7]:

$$\beta_{k}(S \Gamma) \equiv \int \lambda^{k} dF(S \Gamma)(\lambda) \overset{a.s.}{\to} \sum_{s=1}^{k} y^{k-s} \sum_{i_{1}+\cdots+i_{u}=s} \sum_{k_{1}+\cdots+k_{u}=k} \frac{k!}{s!} \prod_{m=1}^{i_{1}} \prod_{m=1}^{i_{u}} D_{m}^{i_{m}}, \quad (2.7)$$

where $D_{k} = \lim_{T \to \infty} \frac{1}{T} \text{tr}[F_{T}^{k}]$. 

Theorem 2.2 (Theorem 1.1 of [11]). Let \( X = (x_{i,t})_{N,T} \) and \( X' = (X_1, \ldots, X_N) \). As \( T \to \infty \), assume:

1. The following moment conditions for \( X \) hold: For all \( k \), \( E_{x_{i,t}} x_{i,t} = \gamma_{kl} \), and for any non-random \( p \times p \) matrix \( B = (b_{ij}) \) with bounded spectral norm,

\[
E|X_i B X_i - \text{tr}(B^T)|^2 = o(n^2),
\]

where \( \Gamma = (\gamma_{ij}) \).

2. \( T/N \to y \in (0, \infty) \).

3. The spectral norm of the matrix \( \Gamma = \Gamma_T \) is uniformly bounded and \( F^{\Gamma} \) tends to a non-random probability distribution \( H \).

Then, with probability 1, \( F^{\frac{1}{N}X'X} \) tends to a probability distribution, whose Stieltjes transform \( m = m(z) (z \in \mathbb{C}^+) \) satisfies (2.4).

Remark 2.1. If the conditions of Theorem 2.2 are satisfied, then we know the LSD of \( \frac{1}{N}X'X \) is the same as the LSD of \( S \Gamma \), where \( \Gamma \) is defined by Theorem 2.2 and \( S \) is defined by Theorem 2.1. Furthermore, we have:

\[
\beta_k \left( \frac{1}{N}X'X \right) \xrightarrow{a.s.} \sum_{s=1}^{k} y^{k-s} \sum_{i_1 + \cdots + i_q = k+1} \frac{k!}{s!} \prod_{m=1}^{k} D_{m}^{\Gamma} \tag{2.9}
\]

where \( D_k = \lim_{T \to \infty} \frac{1}{T} \text{tr}(\Gamma_T^k) \).

3. Existence theorem

Let \( \widetilde{X} = (\varepsilon_1^{1-q}, \ldots, \varepsilon_0) \) and \( \widetilde{X} = (x_{1-p}, \ldots, x_0) \) be two auxiliary matrices. Then a matrix expression of model (1.1) can be obtained whereby

\[
H_0 \Sigma' + L_0 \Sigma' = H_{\omega} \widetilde{X}' + L_{\omega} X',
\]

where

\[
H_0 = \begin{pmatrix}
\begin{array}{cccc}
a_q & a_{q-1} & \cdots & a_2 & a_1 \\
0 & a_q & \cdots & a_3 & a_2 \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & \cdots & 0 & a_q \\
0 & 0 & \cdots & 0 & 0 \\
\end{array}
\end{pmatrix}
\quad \text{and} \quad
H_{\omega} = \begin{pmatrix}
\begin{array}{cccc}
-b_p & -b_{p-1} & \cdots & -b_2 & -b_1 \\
0 & -b_p & \cdots & -b_3 & -b_2 \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & \cdots & 0 & -b_p \\
0 & 0 & \cdots & 0 & 0 \\
\end{array}
\end{pmatrix}.
\]

\[
L_0 = \begin{pmatrix}
1 & 0 & \cdots & \cdots & 0 \\
a_1 & 1 & 0 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
a_q & \cdots & a_1 & 1 & 0 \\
0 & \cdots & a_q & \cdots & a_1 \\
\end{pmatrix} \quad \text{and} \quad
L_{\omega} = \begin{pmatrix}
1 & 0 & \cdots & \cdots & 0 \\
-b_1 & 1 & 0 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
-b_p & \cdots & -b_1 & 1 & 0 \\
0 & \cdots & -b_p & \cdots & -b_1 \\
\end{pmatrix}.
\]

Then:

\[
X' = A \Sigma' + H,
\]

where \( A = L_{\omega}^{-1} L_0 \) and \( H = L_{\omega}^{-1} (H_0 \Sigma' - H_{\omega} \widetilde{X}) \).

Noting that \( \text{rank}(H) \leq \text{rank}(H_0) + \text{rank}(H_{\omega}) \leq q + p \), using Theorem 11.43 of [18], we can obtain:

\[
\| F^{\frac{1}{N}X'X} - F^{\frac{1}{N}A \Sigma' \Sigma' A'} \| \leq \frac{1}{T} \text{rank}(X' - A \Sigma') \leq \frac{p + q}{T},
\]

where \( \| f \| = \sup_{x \in \mathbb{R}} |f(x)| \).

Therefore, \( \frac{1}{N}X'X \) and \( \frac{1}{N}A \Sigma' \Sigma' A' \) have the same LSD. Let \( S = \Sigma' \Sigma' / N \). Then the ESD of \( \frac{1}{N}A \Sigma' \Sigma' A' \) is the same as that of \( A' \Sigma A \). Thus, if it can be proved that \( F^{\frac{1}{N}X'X} \) tends to a non-random limit, the following theorem can be obtained by Theorem 2.1.

Theorem 3.1. Let \( X \) be generated by a causal VARMA\((p,q)\) model (1.1) and \( \lim_{T \to \infty} T/N = y \in (0, \infty) \). Then, with probability 1, \( F^{\frac{1}{N}X'X} \) tends to a non-random probability distribution.

The proof is contained in the Appendix.

4. Explicit forms I

After the existence of the LSD of \( \frac{1}{N}X'X \) is established, (2.4) is satisfied. Then, if an analytic solution to Eq. (2.4) for \( m \) can be obtained, we obtain the explicit form of the density function of the LSD by (2.6). Therefore, we can only obtain the explicit
form of the density function of the LSD of $\frac{1}{N}X'X$ generated from VMA(1) and VAR(1), as the Stieltjes transforms of other VARMA models require solution of equations for $m$ with order greater than four, which do not have an analytic solution. In the following theorems, the function $R = R(p, q, r)$ is used extensively and is defined as follows:

$$
R = \frac{-2p}{3} + \frac{(2p^3 + 27q^2 - 72pr - \sqrt{4(-p^2 - 12r)^3 + (2p^3 + 27q^2 - 72pr)^2})^{1/3}}{3 \times 2^{1/3}} \\
+ \frac{(2p^3 + 27q^2 - 72pr + \sqrt{4(-p^2 - 12r)^3 + (2p^3 + 27q^2 - 72pr)^2})^{1/3}}{3 \times 2^{1/3}},
$$

where $p, q, r$ are real numbers. It is easy to see $R$ satisfies the equation $R^3 + 2pR^2 + (p^2 - 4r)R - q^2 = 0$. Throughout the paper, we define

$$
(t_1 \pm \sqrt{t_2})^{1/3} = \begin{cases} 
\text{sign}(t_1 \pm \sqrt{t_2})|t_1|^{1/3}, & \text{if } t_2 > 0, \\
(t_1^2 - t_2)^{1/6} \left(\cos \phi \frac{2}{3} \pm \sin \phi \frac{2}{3}\right), & \text{otherwise},
\end{cases}
$$

where $t_1, t_2$ are real numbers and $\cos \phi = \frac{(t_1^2 - t_2)^{1/6}}{\sqrt{t_1^2 - t_2}}, \; \phi \in [0, \pi].$

**Theorem 4.1.** Under the assumptions of Theorem 3.1, taking $p = 0$ and $q = 1$ in model (1.1), then the LSD of $\frac{1}{N}X'X$ has a density function

$$
f(x) = \begin{cases} 
\frac{1}{2\pi} \frac{R_1 + 2p_1 + 2\sqrt{q_1^2 R_1}}{R_1}, & \text{if } x \in \Omega_1, \\
0, & \text{otherwise}
\end{cases}
$$

and has a point mass $1 - 1/y$ at the origin if $y > 1$, where $c_1 = \max((a_0 - a_1)^2, (a_0 + a_1)^2), c_2 = \min((a_0 - a_1)^2, (a_0 + a_1)^2), R_1 = R(p_1, q_1, r_1) \in (0, -2p_1), p_1 = -\frac{((c_1 + c_2)x - 2c_1c_2(1 - y))^2 + 2(c_1 - c_2)^2x^2}{8c_1^2c_2^2x^2y^2}, q_1 = \frac{(c_1 - c_2)^2((c_1 + c_2)x - 2c_1c_2(1 - y))}{8c_1^2c_2^2xy^2}, r_1 = \frac{((c_1 + c_2)x - 2c_1c_2(1 - y))^4 - 4(c_1 + c_2)x - 2c_1c_2(1 - y))2(c_1 - c_2)^2x^2 - 256c_1^2c_2^2x^2y^2}{256c_1^2c_2^2x^2y^4},$

and $\Omega_1 = \{x : c_2(1 - \sqrt{y})^2 \leq x \leq c_1(1 + \sqrt{y})^2, (2p_1^3 + 27q_1^2 - 72p_1r_1)^2 > 4(p_1^2 + 12r_1)^3\}.$

**Theorem 4.2.** Under the assumptions of Theorem 3.1, taking $p = 1$ and $q = 0$ in model (1.1), then the LSD of $\frac{1}{N}X'X$ has a density function

$$
f(x) = \begin{cases} 
\frac{1}{2\pi x} \frac{R_2 + 2p_2 + 2\sqrt{q_2^2 R_2}}{R_2}, & \text{if } x \in \Omega_2, \\
0, & \text{otherwise}
\end{cases}
$$

and has a point mass $1 - 1/y$ at the origin if $y > 1$, where $c = \max(\frac{(1+b_1)^2}{a_0}, \frac{(1-b_1)^2}{a_0}), d = \min(\frac{(1+b_1)^2}{a_0}, \frac{(1-b_1)^2}{a_0}), R_2 = R(p_2, q_2, r_2) \in (0, -2p_2), p_2 = -1 - \frac{3(1-cx)^2 + 3(1-dx)^2 - 2(1-cx)(1-dx)}{8y^2}, q_2 = \frac{(2 - (c + d)x)(4y^2 + (c - d)^2x^2) - 2}{y}, r_2 = -\frac{(2 - (c + d)x)^2(3(c - d)^2x^2) - 4(1-cx)(1-dx) + 16y^2}{256y^4} - \frac{(c + d)x}{2y^2},$

and $\Omega_2 = \{x : (1 - \sqrt{y})^2/c \leq x \leq (1 + \sqrt{y})^2/d, (2p_2^3 + 27q_2^2 - 72p_2r_2)^2 > 4(p_2^2 + 12r_2)^3\}.$
The proofs of Theorems 4.1 and 4.2 are given in the Appendix. As an example, we plot the density function curves in Fig. 1, where the VMA(1) model is $x_t = 0.5\epsilon_{t1} + 0.2\epsilon_{t-1}$, the VAR(1) model is $x_t = 0.2x_{t-1} + 0.5\epsilon_t$, and $y = 0.4$.

**Remark 4.1.** In fact, if we have another root $R$ of the equation $R^3 + 2p_1R^2 + (p_1^2 - 4r_1)R - q_1^2 = 0$ (satisfying $R \in [0, -2p_1]$), then Theorem 4.1 (Theorem 4.2) also holds. The LSD of $\frac{1}{N}X'X$ has a point mass $1 - 1/y$ at the origin if $y > 1$, because $\frac{1}{N}X'X$ has $T - N$ zero eigenvalues.

It is conceivable that the M-P law can be derived from Theorem 4.1 or Theorem 4.2, because the M-P law is a special case of Theorem 4.1 or Theorem 4.2.

**Remark 4.2.** Taking $a_1 = 0$ in Theorem 4.1, we have

$$c_1 = c_2 = a_0^2, \quad p_1 = \frac{(x - a_0^2(1 - y))^2}{2a_0^4xy^2}, \quad q_1 = 0, \quad r_1 = \frac{(x - a_0^2(1 - y))^4 - 16a_0^4xy^2}{16a_0^3x^4y^4}.$$ 

By $(2p_1^3 - 72p_1r_1)^2 - 4(p_1^2 + 12r_1)^3 = -432(p_1^2 - 4r_1)^2r_1$, we get

$$\Omega_1 = \{x : a_0^2(1 - \sqrt{\gamma})^2 \leq x \leq a_0^2(1 + \sqrt{\gamma})^2, r_1 < 0\}$$ 

$$= \{x : a_0^2(1 - \sqrt{\gamma})^2 < x < a_0^2(1 + \sqrt{\gamma})^2\}.$$ 

For $R_1 = 0 \in [0, -2p_1)$ is a root of the equation $R^3 + 2p_1R^2 + (p_1^2 - 4r_1)R = 0$ and $\lim_{q_1 \to 0} q_1^2/R_1 = \lim_{q_1 \to 0} R_1^2 + 2p_1R_1 + (p_1^2 - 4r_1) = \frac{4}{a_0^2xy}$, we have

$$f(x) = \frac{\sqrt{R_1 + 2p_1 + 2\sqrt{\frac{q_1^2}{r_1}}}}{2\pi} = \sqrt{\frac{(x - a_0^2(1 - y))^2 + \frac{4}{a_0^2xy}}{2\pi}}$$
\[
= \frac{\sqrt{(a_0^2(1 + \sqrt{y})^2 - x)(x - a_0^2(1 - \sqrt{y})^2)}}{2\pi a_0^2xy}, \quad a_0^2(1 - \sqrt{y})^2 < x < a_0^2(1 + \sqrt{y})^2.
\]

**Remark 4.3.** Taking \(b_1 = 0\) in **Theorem 4.2**, we have \(c = d = 1/a_0^2\), \(p_2 = -1 - \frac{(1-x/a_0^2)^2}{2y^2}\), \(q_2 = -1 + x/a_0^2/y\), \(r_2 = -\frac{(1-x/a_0^2)^2 - (1-x/a_0^2)^2 + 4y^2}{16y^4} - x/a_0^2/y^2\).

By
\[
(2p_2^3 + 27q_2^2 - 72p_2r_2)^2 - 4(p_2^3 + 12r_2^3) = -432x^2((1 + \sqrt{y})^2 - x/a_0^2)\times((1 - \sqrt{y})^2 - x/a_0^2)((1 - x/a_0^2)^2 - 2y(1 - x/a_0^2) + y^2 + 4y) \frac{a_0^2y^8}{a_0^2y^8},
\]
we obtain \(\Omega_2 = \{x : a_0^2(1 - \sqrt{y})^2 < x < a_0^2(1 + \sqrt{y})^2\}\). Because \(R_2 = 1 \in [0, -2p_2)\) is a root of the equation \(R^3 + 2p_2R^2 + (p_2^3 + 4r_2)^2 - q_2^2 = 0\), we have
\[
f(x) = \sqrt{R_2 + 2p_2 + 2R_2^2} = \frac{1}{2\pi} \sqrt{1 - 2 - \frac{(1-x/a_0^2)^2}{y^2} + \frac{2+2x/a_0^2}{y}}.
\]

5. **Explicit forms II**

The coefficients of the models of **Theorems 4.1** and **4.2** are constant. In this section, we derive the explicit forms of the density functions of the LSDs of large-dimensional sample covariance matrices generated by random-coefficient VAR(1) and random-coefficient VMA(1) models. According to **Remark 2.1**, we only need to show that the conditions of **Theorem 2.2** are satisfied, and then the explicit forms can be easily obtained following similar steps as in **Theorems 4.1** and **4.2**.

**Theorem 5.1.** Let \(T/N \rightarrow y \in (0, \infty)\) and
\[
x_{i,t} = a_{i,t}e_{i,t} + b_{i,t}e_{i,t-1}, \quad i = 1, \ldots, N, \ t = 1, \ldots, T,
\]
where \(\{e_{i,t}\}\) is independent with common first moment 0, second moment 1, finite third moment \(Ee^3\) and finite fourth moment \(Ee^4\), \(\{a_{i,t}\}\) is independent of the four common finite moments \(Ea\), \(Ea^2\), \(Ea^3\) and \(Ea^4\), and \(\{b_{i,t}\}\), \(\{a_{i,t}\}\) and \(\{e_{i,t}\}\) are independent. Then the density function of the LSD of \(\frac{1}{N}X'X\) is given by:
\[
f(x) = \begin{cases} 
\frac{1}{2\pi} \sqrt{R_1 + 2p_1 + 2R_1^2} & \text{if } x \in \Omega_1, \\
0 & \text{otherwise},
\end{cases}
\]
where \(c_1 = \max(Ea^2 + Eb^2 + 2EaEb, Ea^2 + Eb^2 - 2EaEb), c_2 = \min(Ea^2 + Eb^2 + 2EaEb, Ea^2 + Eb^2 - 2EaEb)\) and \(R_1, \Omega_1, p_1, q_1, r_1\) are defined by **Theorem 4.1**.

**Theorem 5.2.** Let \(T/N \rightarrow y \in (0, \infty)\) and
\[
x_{i,t} = b_{i,t}x_{i,t-1} + a_{i,t}e_{i,t-1}, \quad i = 1, \ldots, N, \ t = 1, \ldots, T.
\]
In addition to the assumptions of \(\{e_{i,t}\}, \{a_{i,t}\}\) and \(\{b_{i,t}\}\) in **Theorem 5.1**, we assume that \(\{b_{i,t}\}\) and \(\{x_{i,t}\}\) are independent and there is a constant \(0 < M < 1\) such that \(|b_{i,t}| \leq M\) for any \(i, t\). Then the density function of the LSD of \(\frac{1}{N}X'X\) is given by:
\[
f(x) = \begin{cases} 
\frac{1}{2\pi} \sqrt{R_2 + 2p_2 + 2R_2^2} & \text{if } x \in \Omega_2, \\
0 & \text{otherwise},
\end{cases}
\]
where \(c = \max\left(\frac{(1-	ext{EP}^2)(1+\text{EB})}{\text{EP}^2(1-\text{EB})}, \frac{(1-	ext{EP}^2)(1-\text{EB})}{\text{EP}^2(1+\text{EB})}\right), d = \min\left(\frac{(1-	ext{EB}^2)(1+\text{EB})}{\text{EB}^2(1-\text{EB})}, \frac{(1-	ext{EB}^2)(1-\text{EB})}{\text{EB}^2(1+\text{EB})}\right)\) and \(R_2, \Omega_2, p_2, q_2, r_2\) are defined by **Theorem 4.2**.
The proof of Theorems 5.1 and 5.2 is contained in the Appendix.

6. Estimations and simulation

All these explicit forms have two parameters $c_1$, $c_2$ or $c$, $d$. In this section we consider how to estimate these parameters. Because a constant coefficient is a special case of a random coefficient, we first discuss random coefficient cases. By (2.9), we have:

$$\beta_1 \equiv \beta_1 \left( \frac{1}{N} X' X \right) = \frac{1}{T} \text{tr} \left( \frac{1}{N} X' X \right) \xrightarrow{a.s} D_1,$$

$$\beta_2 \equiv \beta_2 \left( \frac{1}{N} X' X \right) = \frac{1}{T} \text{tr} \left( \frac{1}{N} X' X \right)^2 \xrightarrow{a.s} yD_1^2 + D_2,$$

where $D_1 = Ea^2 + Eb^2$, $D_2 = D_1^2 + 2(EaEb)^2$ in Theorem 5.1 and

$$D_1 = \frac{Ea^2}{1 - Eb^2}, \quad D_2 = \frac{(Ea^2)^2 (1 + (Eb)^2)}{(1 - Eb^2)^2 (1 - (Eb)^2)}$$

in Theorem 5.2. Let $\hat{D}_1 = \beta_1$ and $y\hat{D}_1^2 + \hat{D}_2 = \beta_2$. Then we have the following theorem.

**Theorem 6.1.** Under the conditions of Theorem 5.1, we have:

$$\hat{c}_1 = \beta_1 + \sqrt{2(\beta_2 - (y + 1)\beta_1^2)} \xrightarrow{a.s} c_1, \quad \hat{c}_2 = \beta_1 - \sqrt{2(\beta_2 - (y + 1)\beta_1^2)} \xrightarrow{a.s} c_2.$$ 

Under the conditions of Theorem 5.2, we have:

$$\hat{c} = \frac{\beta_2 - y\beta_1^2 + \sqrt{(\beta_2 - y\beta_1^2)^2 - \beta_1^4}}{\beta_1^2} \xrightarrow{a.s} c, \quad \hat{d} = \frac{\beta_2 - y\beta_1^2 - \sqrt{(\beta_2 - y\beta_1^2)^2 - \beta_1^4}}{\beta_1^3} \xrightarrow{a.s} d.$$ 

Furthermore, as a special case of Theorem 6.1, we have the corollary for the case of a constant coefficient.

**Corollary 1.** Under the conditions of Theorem 4.1, if $\{ε_{i,t}\}$ has a finite fourth moment and $a_0 > a_1 > 0$, we have:

$$\hat{a}_0 = \frac{\beta_1 + \sqrt{(3 + 2y)\beta_1^2 - 2\beta_2}}{2} \xrightarrow{a.s} a_0, \quad \hat{a}_1 = \frac{\beta_1 - \sqrt{(3 + 2y)\beta_1^2 - 2\beta_2}}{2} \xrightarrow{a.s} a_1.$$ 

Under the conditions of Theorem 4.2, if $\{ε_{i,t}\}$ has a finite fourth moment and $0 \leq b_1 < 1$, $a_1 > 0$, we have:

$$\hat{b}_1 = \sqrt{(\beta_2 - (y + 1)\beta_1^2) / (\beta_2 + (1 - y)\beta_1^2)} \xrightarrow{a.s} b_1, \quad \hat{a}_0 = \frac{2\beta_1^2 / (\beta_2 + (1 - y)\beta_1^2)}{a.s} a_0.$$ 

To assess the results, we carried out some simulations. First, we generated data $X_{N,T}$ from different VARMA models and random-coefficient VARMA models for which the noise $ε_{i,t}, i = 1, \ldots, N, t = 1, \ldots, T$ are iid N(0,1). Then, by Theorem 6.1, we can obtain an estimate of $c_1$, $c_2$ or $c$, $d$. Applying these estimates, the density function of the LSD in Theorem 5.1 (or Theorem 4.1) or Theorem 5.2 (or Theorem 4.2) is obtained. To examine this LSD, we used the Kolmogorov–Smirnov test, for which the $P$-value tends to 1 when the ESD of $\frac{1}{N} X' X$ tends to the LSD. Tables 1–3 list the average estimates and $P$-values over 100 replicates, with the mean standard error or sample variance reported in parentheses or brackets, respectively.

We note that the estimates for VAR(1) are quite accurate, with $P$-values close to 1. The estimates for VMA(1) are also quite good, except in the case of $a_1 = 1.5$ in Table 1. The average estimates of $a_0$ and $a_1$ are complex numbers for $a_1 = 1.5$ in the VMA(1) model. This is because our moment method can produce some complex numbers in the 100 estimates when $a_1$ is close to $a_0$. Comparison of the results for $T = 100$ and $T = 200$ reveals that the proportion of complex numbers in the 100 estimates decreases and that estimates of $c_1$ and $c_2$ are increasingly accurate with $P$-values approaching 1 as $T$ and $N$ increase. For the random coefficient model, four simple random-coefficient VMA(1) and VAR(1) models for different $N$ and $T$ are shown in Table 2. All the simulation results demonstrate that the estimates are good.

If $X_{N,T}$ is not generated by VAR(1) or VMA(1), the question arises as to whether we could use the method presented here to estimate the LSD of $\frac{1}{N} X' X$. Table 3 shows that the VAR(1) model performs better than the VMA(1) model, since the LSD of $\frac{1}{N} X' X$ can be estimated by the LSD of the large-sample random matrix generated by VAR(1) in some cases, such as for small $b_2$ in VAR(2) and small $a_1$ in VMA(2) and VARMA(1,1).
7. Conclusion

Regardless of whether the VAR(1) and VMA(1) coefficients are random or not, we can use the same moment estimation method to obtain parametric estimates of the density function of the LSD for large N and T. Then we can obtain an accurate estimate of the explicit form of the density function even if we do not know if the VAR(1) or VMA(1) coefficients are random or constant for observed data. Although we have only provided estimates of VAR(1) and VMA(1), the LSDs of sample covariance matrices generated by other VAMA models can be estimated from those generated by the VAR(1) model according to simulation experiments. Therefore, the LSD of sample covariance matrices generated by a VAR(1) model or a random-coefficient VAR(1) model is a robust limiting distribution.

Acknowledgments

The authors would like to thank Professor Z.D. Bai for his valuable suggestions, which have improved the paper significantly.
Table 3
Other model.

<table>
<thead>
<tr>
<th></th>
<th>Estimated by VAR(1)</th>
<th>K-S test</th>
<th>Estimated by VMA(1)</th>
<th>K-S test</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( \hat{c} )</td>
<td>( d )</td>
<td>( \hat{c}_1 )</td>
<td>( \hat{c}_2 )</td>
</tr>
<tr>
<td>( a_2 )</td>
<td>VMA(2), ( x_t = 2x_{t-1} + \varepsilon_{t-1} + \varepsilon_t ), ( N = 150, T = 100. )</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.4</td>
<td>0.50 [1e-04]</td>
<td>0.08 [6e-06]</td>
<td>1.00 (2e-05)</td>
<td>10.20 [0.055]</td>
</tr>
<tr>
<td>0.8</td>
<td>0.53 [2e-04]</td>
<td>0.06 [4e-06]</td>
<td>0.99 (5e-04)</td>
<td>12.09 [0.064]</td>
</tr>
<tr>
<td>1.2</td>
<td>0.51 [1e-04]</td>
<td>0.05 [2e-06]</td>
<td>0.60 (0.170)</td>
<td>14.37 [0.091]</td>
</tr>
<tr>
<td>( b_2 )</td>
<td>VAR(2), ( x_t = 0.2x_{t-1} + b_2x_{t-2} + 2\varepsilon_t ), ( N = 150, T = 100. )</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.1</td>
<td>0.41 [2e-04]</td>
<td>0.14 [3e-05]</td>
<td>1.00 (3e-09)</td>
<td>6.602 [0.037]</td>
</tr>
<tr>
<td>0.3</td>
<td>0.56 [3e-04]</td>
<td>0.08 [1e-05]</td>
<td>0.99 (3e-04)</td>
<td>9.683 [0.069]</td>
</tr>
<tr>
<td>0.5</td>
<td>0.85 [7e-04]</td>
<td>0.03 [2e-06]</td>
<td>0.40 (0.373)</td>
<td>17.96 [0.295]</td>
</tr>
<tr>
<td>( a_1 )</td>
<td>VARMA(1,1), ( x_t = 0.2x_{t-1} + 0.5x_{t-2} + 2\varepsilon_t + \varepsilon_{t-1} ), ( N = 150, T = 100. )</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.2</td>
<td>0.42 [2e-04]</td>
<td>0.13 [2e-05]</td>
<td>1.00 (6e-08)</td>
<td>7.026 [0.037]</td>
</tr>
<tr>
<td>0.6</td>
<td>0.48 [1e-04]</td>
<td>0.08 [7e-06]</td>
<td>0.99 (2e-04)</td>
<td>9.557 [0.050]</td>
</tr>
<tr>
<td>1.0</td>
<td>0.47 [1e-04]</td>
<td>0.06 [3e-06]</td>
<td>0.48 (0.279)</td>
<td>12.57 [0.063]</td>
</tr>
</tbody>
</table>

Appendix

Proof of Theorem 3.1

Let

\[
\Gamma = E \begin{pmatrix}
X_{1,1} \\
X_{1,2} \\
\vdots \\
X_{1,T}
\end{pmatrix} =
A \begin{pmatrix}
\varepsilon_{1,1} \\
\varepsilon_{1,2} \\
\vdots \\
\varepsilon_{1,T}
\end{pmatrix} + L_{\phi}^{-1}H_0 \begin{pmatrix}
\varepsilon_{1,1-q} \\
\varepsilon_{1,2-q} \\
\vdots \\
\varepsilon_{1,0}
\end{pmatrix} - L_{\phi}^{-1}H_0 \begin{pmatrix}
X_{1-1,p} \\
X_{1-2,p} \\
\vdots \\
X_{0}
\end{pmatrix}.
\]

Using Theorem 11.43 of [18], it is easy to know that \( \Gamma \) and \( \Gamma' \) have the same LSD.

From the properties of ACVF, we have \( \gamma(k) = \sum_{i=1}^{p} \alpha_i z_i^{-|k|} \) for \( k > q \) where \( \alpha_1, \ldots, \alpha_p \) are constants and \( z_1, \ldots, z_p \) are the roots of equation \( 1 - b_1z_1 - \cdots - b_pz_p = 0. \) For ensuring the stationarity and invertibility, \( |z_j| > 1 \) for all \( j. \) Therefore, \( \sum_{i=1}^{p} \gamma(i) \) is finite.

Next, we use mathematical induction to prove for arbitrary \( k \geq 1, \beta_{T,k} = \int x^k dF_T \) converges to a finite limit \( \beta_k \) as \( T \to \infty \) and \( \beta_{T,k} \leq \gamma(0) + 2 \sum_{i=1}^{\infty} |\gamma(i)|^k. \) Let \( \Gamma^k = (\gamma_{k,i,j}). \) Then \( \beta_{T,k} = \frac{1}{T} \sum_{i=1}^{T-k} \gamma_{k,i,i+j}. \)

Suppose \( \frac{1}{T} \sum_{i=1}^{T-k} \gamma_{k,i,i+j} \) converges to a finite limit and

\[
\frac{1}{T} \sum_{i=1}^{T-k} |\gamma_{k,i,i+j}| \leq (\gamma(0) + 2 \sum_{i=1}^{\infty} |\gamma(i)|^k), \quad l = 0, \ldots, T - 1
\]

(the results are right when \( k = 1 \)).

Then we have

\[
\frac{1}{T} \sum_{i=1}^{T-k} \gamma_{k+1,i,i+j} = (\gamma(0))^l \sum_{i=1}^{T-k} \gamma_{k,i,i+j} + \sum_{j=1}^{T-k} \gamma(j) \left( \frac{1}{T} \sum_{i=1}^{T-j} \gamma_{k,i,i+j} + \frac{1}{T} \sum_{l=1}^{T-I} \gamma_{k,i,l+j} \right). \]

It is easy to see that \( \frac{1}{T} \sum_{i=1}^{T-k} \gamma_{k+1,i,i+j} \) converges to a finite limit and

\[
\frac{1}{T} \sum_{i=1}^{T-k} |\gamma_{k+1,i,i+j}| \leq \left( (\gamma(0) + 2 \sum_{i=1}^{\infty} |\gamma(i)|^k \right)^{k+1}, \quad l = 0, \ldots, T - 1.
\]
Therefore, for arbitrary $k \geq 1$, $\beta_{T,k} = \int x^4 dF_T$ converges to a finite limit $\beta_k$ and Carleman condition is also satisfied:
\[
\sum_{k=1}^{\infty} \beta_{2k}^{-\frac{1}{2k}} \geq \sum_{k=1}^{\infty} \left( \gamma(0) + 2 \sum_{i=1}^{\infty} |\gamma(i)| \right)^{-1} = \infty.
\]
By Moment Convergence Theorem, the spectral distribution of $A^TA$ tends to a non random probability distribution $H$. The proof of the theorem is complete.

Before proving Theorems 4.1 and 4.2, we need the following lemmas.

**Lemma A.1.** Under the definition of (4.11), suppose that $-2p^3 + 8pr - q^2 > 0$, $p^2 - 4r > 0$ and $4(-p^2 - 12r)^3 + (2p^3 + 27q^2 - 72pr)^2 > 0$, we have $0 \leq R < -2p$.

**Proof.** $-2p^3 + 8pr - q^2 > 0$ and $p^2 - 4r > 0$ imply $p < 0$. Let
\[
k_0 = (2p^3 + 27q^2 - 72pr - \sqrt{4(-p^2 - 12r)^3 + (2p^3 + 27q^2 - 72pr)^2})^{1/3}
\]
\[+(2p^3 + 27q^2 - 72pr + \sqrt{4(-p^2 - 12r)^3 + (2p^3 + 27q^2 - 72pr)^2})^{1/3}.
\]
Then we have $k_0 - 3 \times 4^{1/3}(p^2 + 12r)k_0 - 2(2p^3 + 27q^2 - 72pr) = 0$. Now let
\[
f(k) = k^3 - 3 \times 4^{1/3}(p^2 + 12r)k - 2(2p^3 + 27q^2 - 72pr).
\]
Thus we have $f(k_0) = 0$, $f(2^{4/3}p) = -54q^2 \leq 0$, and $f(-2^{7/3}p) = 54(-2p^3 + 8pr - q^2) > 0$.

By (A.2), $f'(k) = 3(k^2 - 4^{1/3}(p^2 + 12r))$. If $p^2 + 12r < 0$, then $f(k)$ is increasing. Thus $2^{4/3}p \leq k_0 < -2^{7/3}p$ which implies $0 \leq R < -2p$.

If $p^2 + 12r < 0$, noticing $2^{4/3}p > 2^{1/3}(p^2 + 12r)^{1/3}, -2^{7/3}p < -2^{1/3}(p^2 + 12r)^{1/3}, f(-2^{7/3}p) > 0$ and $f(2^{4/3}p) \leq 0$, we know that the real root of $f(k) = 0$ lies in the interval $[2^{4/3}p, -2^{7/3}p]$. Therefore $2^{4/3}p \leq k_0 < -2^{7/3}p$ implies $0 \leq R < -2p$.

**Lemma A.2.** Assume $0 \leq R < -2p$, then we have $R + 2p + 2\sqrt{\frac{q^2}{R}} > 0$ if and only if $-4(p^2 + 12r)^3 + (2p^3 + 27q^2 - 72pr)^2 > 0$.

**Proof.** Let $t_1 = 2p^3 + 27q^2 - 72pr, t_2 = -4(p^2 + 12r)^3 + (2p^3 + 27q^2 - 72pr)^2$, we have
\[
R = -\frac{2p}{3} + \frac{1}{32\pi}((t_1 - \sqrt{t_2})^{\frac{3}{2}} + (t_1 + \sqrt{t_2})^{\frac{3}{2}}).
\]
By $0 \leq R < -2p$ and $R^3 + 2pR^2 + (p^2 - 4r)R - q^2 = 0$, we can get: $R + 2p + 2\sqrt{\frac{q^2}{R}} > 0$ if and only if $R + 2p + 2\sqrt{R^2 + 2pR + (p^2 - 4r)} > 0, i.e. (R + \frac{2p}{3})^2 > \frac{4}{9}(p^2 + 12r)$. Thus $R + 2p + 2\sqrt{\frac{q^2}{R}} > 0$ if and only if $(t_1 - \sqrt{t_2})^{\frac{3}{2}} + (t_1 + \sqrt{t_2})^{\frac{3}{2}} > 4(t_1^2 - t_2)^{\frac{1}{2}}, i.e. t_2 > 0$ (by Definition (4.12)).

**Proof of Theorem 4.1**

From (3.10), we have
\[
AA' = \begin{pmatrix}
  a_0^2 & a_0a_1 & 0 & \cdots & 0 & 0 \\
  a_0a_1 & a_0^2 + a_1^2 & a_0a_1 & \cdots & 0 & 0 \\
  0 & a_0a_1 & a_0^2 + a_1^2 & \cdots & 0 & 0 \\
  \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
  0 & 0 & 0 & \cdots & a_0a_1 & a_0^2 + a_1^2 \\
\end{pmatrix} \equiv C.
\]
Define
\[
\tilde{C} = \begin{pmatrix}
  a_0^2 + a_1^2 & a_0a_1 & 0 & \cdots & 0 & 0 \\
  a_0a_1 & a_0^2 + a_1^2 & a_0a_1 & \cdots & 0 & 0 \\
  0 & a_0a_1 & a_0^2 + a_1^2 & \cdots & 0 & 0 \\
  \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
  0 & 0 & 0 & \cdots & a_0a_1 & a_0^2 + a_1^2 \\
\end{pmatrix},
\]
by Lemma 2.2 of [2], then the LSD of $\tilde{C}$ is the same as that of $C$. The eigenvalues of $\tilde{C}$ are
\[
\lambda_k = a_0^2 + a_1^2 + 2a_0a_1 \cos \left( \frac{k\pi}{T+1} \right), \quad k = 1, 2, \ldots, T.
\]
Let
\[ m = -(1 - y)/x + ym(x), \]
and the LSD of \( \hat{C} \) is \( H \). Then by (2.4), we have
\[ x = -\frac{1}{m} + y \int_0^1 \frac{t}{1 + tm} \, dH(t) \]
(A.5)

From (A.3), it can be rewritten that
\[ x = -\frac{1}{m} + y \int_0^1 \frac{(a_2^2 + a_1^2 + 2a_0a_1\cos(\pi t))\, dt}{1 + (a_0^2 + a_1^2 + 2a_0a_1\cos(\pi t))m} \]
\[ = \frac{y - 1}{m} - \frac{y}{2\pi im} \int_{|z|=1} \frac{d\zeta}{(1 + (a_0^2 + a_1^2 + a_0a_1(\zeta + \zeta^{-1}))m)} \]
\[ = \frac{y - 1}{m} - \frac{a_0a_1y}{\sqrt{(1 + m(a_0 + a_1)^2)(1 + m(a_0 - a_1)^2)}} \]

And it yields that
\[ c_1c_2x^2y^2m^4 + xy((c_1 + c_2)x - 2c_1c_2(1 - y))m^3 + (x - c_1(1 - y))(x - c_2(1 - y))m^2 - 1 = 0, \]
where \( c_1, c_2 \) are defined in Theorem 4.1. Letting
\[ m = Z - \frac{(c_1 + c_2)x - 2c_1c_2(1 - y)}{4c_1c_2xy}, \]
we get \( Z^4 + p_1Z^2 + q_1Z + r_1 = 0 \), where \( p_1, q_1, r_1 \) are defined in Theorem 4.1. Let \( R_1 \) satisfy the equation \( R_1^2 + 2p_1R_1^2 + (p_1^2 - 4r_1)R_1 - q_1^2 = 0 \). Through simple computation, it may be expressed that
\[ Z^4 + p_1Z^2 + q_1Z + r_1 = \left(Z^2 + \frac{p_1 + R_1}{2}\right)^2 - \left(R_1Z^2 - q_1Z + \frac{(p_1 + R_1)^2}{4} - r_1\right) \]
\[ = \left(Z^2 + \frac{p_1 + R_1}{2}\right)^2 - R_1\left(Z_1 - \frac{q_1}{2R_1}\right)^2 = 0. \]

Therefore the four roots of the equation about \( Z \) can be solved:
\[ Z_{1,2} = 1/2 \left(\sqrt{R_1} \pm \sqrt{-R_1 - 2p_1 - 2q_1/\sqrt{R_1}}\right), \]
\[ Z_{3,4} = 1/2 \left(-\sqrt{R_1} \pm \sqrt{-R_1 - 2p_1 + 2q_1/\sqrt{R_1}}\right). \]

In the function \( f(R) = R^3 + 2p_1R^2 + (p_1^2 - 4r_1)R - q_1^2, f(0) = -q_1^2 \leq 0, \) and \( f(-2p_1) = -2p_1^3 + 8r_1p_1 - q_1^2 \). Let \( v_1 = (c_1 + c_2)x - 2c_1c_2(1 - y), v_2 = (c_1 - c_2)x \). From
\[ -2p_1^3 + 8p_1r_1 - q_1^2 = \frac{v_1^6 + v_2^4v_1^2 + 2v_1^2v_2^4 + 32c^3d^3x^3y^3v_1^2 + 64c^3d^3x^3y^3v_2^2}{32c^6d^6x^6y^6} > 0, \]
(A.7)
we have \( f(-2p_1) > 0 \). Thus we can take the root \( R_1 \) of the equation \( R^3 + 2p_1R^2 + (p_1^2 - 4r_1)R - q_1^2 = 0 \) satisfying \( 0 \leq R_1 < -2p_1 \).

Note that the density function \( f(x) = \lim_{y \to 0} \frac{I_m(m(x + iy))}{\pi} > 0 \), the imaginary part of \( m(x) \) is positive. Thus we have
\[ f(x) = \lim_{y \to 0} \frac{I_m(m(x + iy))}{\pi} = \sqrt{R_1 + 2p_1 + 2\sqrt{\frac{v_1^2}{R_1}}} \frac{\sqrt{R_1 + 2p_1 + 2\sqrt{\frac{v_1^2}{R_1}}}}{2\pi}. \]

Moreover \( c_2(1 - \sqrt{y})^2 = \lambda_{\min}(A)\lambda_{\min}(B) \leq \lambda(AB) \leq \lambda_{\max}(A)\lambda_{\max}(B) = c_1(1 + \sqrt{y})^2, \) domain of \( f(x) \) is \( \Omega_1 = \{ x : c_2(1 - \sqrt{y})^2 \leq x \leq c_1(1 + \sqrt{y})^2, R_1 + 2p_1 + 2\sqrt{\frac{v_1^2}{R_1}} > 0 \} \). By Lemma A.2, \( \Omega_1 = \{ x : c_2(1 - \sqrt{y})^2 \leq x \leq c_1(1 + \sqrt{y})^2, (2p_1^2 + 27q_1^2 - 72p_1r_1)^2 > 4(p_1^2 + 12r_1)^3 \} \).

For \( x \in \Omega_1 \),
\[ p_1^2 - 4r_1 = \frac{2v_1^2v_2^2 + v_1^4 + 64c^3d^3x^3y^3}{16c^4d^4x^4y^4} > 0 \]
and (A.7), by Lemma A.1, we take \( R_1 = R(p_1, q_1, r_1) \) in (4.11), then \( 0 \leq R_1 < -2p_1 \). Thus the theorem is proved.
Proof of Theorem 4.2

From (3.10), we have

$$AA' = a^2_0 \begin{pmatrix} 1 & -b_1 & 0 & \cdots & 0 & 0 \\ -b_1 & 1 + b_1^2 & -b_1 & \cdots & 0 & 0 \\ 0 & -b_1 & 1 + b_1^2 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & -b_1 & 1 + b_1^2 \end{pmatrix}^{-1} = C.$$  

Let

$$\tilde{C}^{-1} = a^{-2}_0 \begin{pmatrix} 1 + b_1^2 & -b_1 & 0 & \cdots & 0 & 0 \\ -b_1 & 1 + b_1^2 & -b_1 & \cdots & 0 & 0 \\ 0 & -b_1 & 1 + b_1^2 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & -b_1 & 1 + b_1^2 \end{pmatrix}$$

by Lemma 2.2 of [2], then the LSD of \( \tilde{C}^{-1} \) is the same as that of \( C^{-1} \). The eigenvalues of \( \tilde{C}^{-1} \) are \( \lambda_k = a^{-2}_0 (1 + b_1^2 + 2b_1 \cos(\frac{k\pi}{r+1})), \) \( k = 1, 2, \ldots, T \). Hence by (A.5), we have

$$x = -\frac{1}{m} + y \int_0^1 \frac{dt}{m + a^{-2}_0(1 + b_1^2 + 2b_1 \cos(\pi t))} = -\frac{1}{m} + \frac{1}{2\pi i} \oint_{|\zeta|=1} \frac{d\zeta}{\zeta(m + a^{-2}_0(1 + b_1^2 + b_1(\zeta + \zeta^{-1})))} \equiv x_0.$$  

Letting \( z = -x_0m(x) - 1 \), we have

$$z^4 + \frac{2 - (c + d)x}{y} z^2 + \left( \frac{1 - cx}{y^2} - 1 \right) z^2 - \frac{2}{y} z - \frac{1}{y^2} = 0,$$  

(A.8)

where \( c, d \) are defined in Theorem 4.2. Let \( Z = z + \frac{2 - (c + d)x}{4y} \), it yields from (A.8)

$$Z^4 + p_2 Z^2 + q_2Z + r_2 = 0,$$  

(A.9)

where \( p_2, q_2, r_2 \) are defined in Theorem 4.2. Similar as in proof of Theorem 4.1, letting \( v_1 = 1 - cx, \) \( v_2 = 1 - dx, \)

$$-2p_2^3 + 8p_2^2r_2 - q_2^2 = 2 + \frac{(v_1 - v_2)^2(v_1^2 + v_2^2)^2}{8y^6} + \frac{(v_1 - v_2)^2 + (v_1^2 + v_2^2)((v_1 - v_2)^2 + (v_1 - 1)^2 + (v_2 - 1)^2 + 1)/2}{y^4} + \frac{(v_1 - v_2)^2 + 4(v_1 - 1/2)^2 + 4(v_2 - 1/2)^2 + 6}{2y^2} > 0,$$  

(A.10)

we can take \( R_2 \in [0, -2p_2) \) satisfying the equation \( R_2^2 + 2p_2R_2^2 + (p_2^2 - 4r_2)R_2 - q_2^2 = 0 \). Then the four roots of the Eq. (A.9) about \( Z \) can be found as:

$$Z_{1,2} = 1/2(\sqrt{R_2} \pm \sqrt{-R_2 - 2p_2 - 2q_2/\sqrt{R_2}}),$$

$$Z_{3,4} = 1/2(-\sqrt{R_2} \pm \sqrt{-R_2 - 2p_2 + 2q_2/\sqrt{R_2}}).$$

Just as the discussion of the proof of Theorem 4.1, we have:

$$f(x) = \begin{cases} \sqrt{R_2 + 2p_2 + 2\sqrt{\frac{q_2}{R_2}}}, & \text{if } x \in \Omega_2, \\ \frac{2\pi x}{2\pi x}, & \text{otherwise} \end{cases}$$  

(A.11)

where \( \Omega_2 = \{x : (1 - \sqrt{y})^2/c \leq x \leq (1 + \sqrt{y})^2/d, (2p_2^3 + 27q_2^2 - 72p_2r_2)^2 > 4(p_2^3 + 12r_2)^3\}. \)
For \( x \in \Omega_2 \),
\[
p_2^2 - 4r_2 = \frac{(v_1 - v_2)^4 + 2(v_1^2 - v_2^2)^2 + 16((v_1 - 1)^2 + (v_2 - 1)^2 + 2)^2y^2 + 16y^4}{16y^4} > 0,
\]
and (A.10), by Lemma A.1 we take \( R_2 = R(p_2, q_2, r_2) \) in (4.11), then \( 0 \leq R_2 < -2p_2 \). Thus the theorem is proved.

**Proof of Theorem 5.1**

From (5.15), we have \( Ex_{t,i}^2 = Ea + Eb \) and
\[
\gamma_{i,l} = Ex_{j,i}x_{t,l} = (Ea^2 + Eb^2)I(i = l) + EdEbI(j = l = 1).
\]
Let \( v_1 \geq v_2 \geq v_3 \geq v_4 \) be the order of \( j_1, j_2, l_1, l_2 \).

If \( v_2 > v_3 + 1, Ex_{v_1,t}x_{v_2,t}x_{v_3,t}x_{v_4,t} = y_{v_1,v_2}y_{v_3,v_4} \) and \( Ex_{v_1,v_3}y_{v_2,v_4} = 0 \), then \( Ex_{v_1,t}x_{v_2,t}x_{v_3,t}x_{v_4,t} - y_{j_1,l_1}y_{j_2,l_2} = y_{j_1,l_2}y_{j_2,l_1} + y_{j_1,l_1}y_{j_2,l_2} \).

If \( v_2 \leq v_3 + 1 \) and \( v_1 - v_{i+1} \) does not equal to 1 and 0 where \( i = 1, 2, 3 \), then \( Ex_{v_1,t}x_{v_2,t}x_{v_3,t}x_{v_4,t} - y_{j_1,l_1}y_{j_2,l_2} - y_{j_1,l_2}y_{j_2,l_1} - y_{j_1,l_1}y_{j_2,l_2} = 0 \).

Letting \( B \) be a non-random matrix with bounded spectral norm, we have
\[
\sum_{j_1,l_1,j_2,l_2} b_{j_1,l_1}b_{j_2,l_2} (\gamma_{j_1,l_1}\gamma_{j_2,l_2} + \gamma_{j_1,l_1}\gamma_{j_2,l_2}) = tr(\Gamma'\Gamma) + tr(\Gamma'\Gamma) = O(T)
\]
where \( \Gamma = (\gamma_{i,j}) \). Because the entries of \( B \) are bounded, we have
\[
E|X|^2 - tr(\Gamma')^2 = Ex_{t,i}^2(\gamma_{i,j} - \gamma_{i,j})^2 = \left( \sum_{j_1,l_1} b_{j_1,l_1} (X_{j_1,l_1} - \gamma_{j_1,l_1}) \right)^2 = \sum_{j_1,l_1} b_{j_1,l_1}b_{j_2,l_2} (Ex_{j_1,l_1}x_{j_2,l_2} - \gamma_{j_1,l_1}y_{j_1,l_2} - \gamma_{j_1,l_2}y_{j_1,l_2})
\]
\[
+ \sum_{j_1,l_1,j_2,l_2} b_{j_1,l_1}b_{j_2,l_2} (\gamma_{j_1,l_1}\gamma_{j_2,l_2} + \gamma_{j_1,l_2}\gamma_{j_2,l_2}) = O(T).
\]
Thus, the moment condition of Theorem 2.2 is satisfied.

By Remark 2.1, similar as in Theorem 4.1, the explicit form of the density function of the LSD may be obtained and the differences are \( c_1 = \max(Ea^2 + Eb^2 + 2EaEb, Ea^2 + Eb^2 - 2EaEb) \) and \( c_2 = \min(Ea^2 + Eb^2 + 2EaEb, Ea^2 + Eb^2 - 2EaEb) \).

**Proof of Theorem 5.2**

Under these assumptions, \( \{x_{t,i} \} \) have four common and finite moments \( Ex, Ex^2, Ex^3 \) and \( Ex^4 \). From (5.17), we have \( Ex = 0 \), and \( Ex^2 = \frac{\varepsilon^2}{1 - Ed} \). Let \( \sum_{x=b+1}^{\infty} f(x + 1) = 1 \) and \( \sum_{x=b}^{\infty} f(x) = 0 \). If \( v_1 \geq v_2 \), we have
\[
x_{v_1,t} = \sum_{v_1-t+1}^{v_1} b_{v_1-t,x_{v_1-t}} + \sum_{i=0}^{v_1-t} d_{v_1-t,i}x_{v_1-t} \prod_{v_1-t-i+1}^{v_1} b_{v_1-t}.
\]
Let \( v_1 \geq v_2 \) be the order of \( j, l \). Then we get
\[
\gamma_{i,l} = Ex_{v_1,t}x_{v_2,t} = (Eb)v_1+v_2 Ex^2.
\]
Let \( v_1 \geq v_2 \geq v_3 \geq v_4 \) be the order of \( j_1, j_2, l_1, l_2 \). From (A.12), we have
\[
Ex_{v_1,t}x_{v_2,t}x_{v_3,t}x_{v_4,t} = (Eb)^{v_1+v_2} Ex^2 x_{v_1,t}x_{v_2,t}x_{v_3,t}x_{v_4,t}
\]
\[
= (Eb)^{v_1+v_2} \left( Ex^3 x_{v_1,t}x_{v_2,t}x_{v_3,t}x_{v_4,t} + Eq^2 Ex^2 (1 - (Eb)^{v_2-v_3}) \right)
\]
\[
= (Eb)^{v_1+v_2} (Eb)^{v_2-v_3} Ex^3 x_{v_1,t}x_{v_2,t}x_{v_3,t}x_{v_4,t} + (Eb)^{v_1+v_2} (Eb)^{v_3-v_4} (Ex^2)^2 (1 - (Eb)^{v_2-v_3})
\]
\[
= (Eb)^{v_1+v_2} (Eb)^{v_2-v_3} Ex^3 x_{v_1,t}x_{v_2,t}x_{v_3,t}x_{v_4,t} - (Eb)^{v_1+v_2} (Eb)^{v_3-v_4} (Ex^2)^2 (Eb)^{v_2-v_3}
\]
\[
- 2(Ex^2)^2 (Eb)^{v_1+v_2} (Eb)^{v_2-v_3} (Eb)^{v_3-v_4} + y_{v_1,v_2}y_{v_3,v_4} + y_{v_1,v_3}y_{v_2,v_4} + y_{v_1,v_3}y_{v_2,v_4}
\]
where
\[
Ex^3 x_{v_1,t}x_{v_2,t}x_{v_3,t}x_{v_4,t} = (Eb)^{v_3-v_4} Ex^4 + 3 \sum_{i=0}^{v_3-v_4-1} Ed^2 Ex^2 (Eb)^{v_3-v_4-i} (Eb)^{v_3-i} Ex^2
\]
\[
= (Eb)^{v_3-v_4} Ex^4 + \frac{3Ed^2 Ex^2 Eb((Eb)^{v_3-v_4} - (Eb)^{v_3-v_4})}{Eb - Eb^3}.
\]
Let B be a non-random matrix with bounded spectral norm. Because the entries of B are bounded and
\[ \sum_{j_1,j_2,i_1,i_2} b_{j_1}^* b_{j_2}^{*\top} x_{i_1} x_{i_2} = O(T) \]
where \(|b_i| < 1, i = 1, \ldots, 3\), we have
\[ \sum_{j_1,j_2,i_1,i_2} b_{j_1} b_{j_2} b_{i_1} b_{i_2} ((Eb)^{v_1-v_2} (Eb)^{v_3-v_4} X_{i_1} x_{i_2,t} - (Eb)^{v_1-v_2} (Eb)^{v_3-v_4} (Ex_t^2)^2 (Eb)^{v_2-v_3} \]
\[ - 2 (Ex_t^2)^2 (Eb)^{v_1-v_2} (Eb)^{v_3-v_4} (Eb)^{v_2-v_3} = O(T) \]
and
\[ \sum_{j_1,j_2,i_1,i_2} b_{j_1} b_{j_2} (y_{j_1,j_2} y_{i_1,i_2} + y_{j_1,i_2} y_{j_2,i_1}) = \text{trace}(B \Gamma B^\top) + \text{trace}(B \Gamma^2 \Gamma^\top) = O(T) \]
where \( \Gamma = (y_{j_1,j_1}) \). Then we have
\[ E|\mathbf{x}^\top B \mathbf{x} - \text{trace}(B \Gamma)|^2 = E \left( \sum_{j_1,j_2} b_{j_1}(x_{j_1} x_{j_2} - y_{j_1,j_2}) \right)^2 \]
\[ = \sum_{j_1,j_2} b_{j_1} b_{j_2} (Ex_{j_1} x_{j_2} x_{j_2,t} x_{j_2,t} - y_{j_1,j_2} y_{j_1,j_2}) = o(T^2). \]
Thus, the moment condition of Theorem 2.2 is satisfied. For
\[ r^{-1} = \frac{1 - Eb^2}{Ea^2 (1 - Eb^2)} \begin{pmatrix} 1 & -Eb & 0 & \cdots & 0 & 0 \\ -Eb & 1 + (Eb)^2 & -Eb & \cdots & 0 & 0 \\ 0 & -Eb & 1 + (Eb)^2 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 + (Eb)^2 & -Eb \\ 0 & 0 & 0 & \cdots & -Eb & 1 \end{pmatrix}, \]
similarly as in Theorem 4.2, the explicit form of the density function of the LSD is obtained and the differences are
\[ c = \max \left( \frac{(1 - Eb^2)(1 + Eb)}{Ea^2 (1 - Eb)}, \frac{(1 - Eb^2)(1 - Eb)}{Ea^2 (1 + Eb)} \right), \]
\[ d = \min \left( \frac{(1 - Eb^2)(1 + Eb)}{Ea^2 (1 - Eb)}, \frac{(1 - Eb^2)(1 - Eb)}{Ea^2 (1 + Eb)} \right). \]
References