

# Eigenvalues Distribution Limit of Covariance Matrices with AR Processes Entries



Zahira Khettab and Tahar Mourid

**Abstract** We consider a class of random matrices  $B_N = X_N T_N X_N^t$ , where  $X_N$  is a matrix ( $N \times n(N)$ ) whose rows are independent, the entries  $X_{ij}$  in each row satisfy an autoregressive relation AR(1), and  $T_N$  is a diagonal matrix independent of  $X_N$ . Under some conditions, we show that if the empirical distribution function of eigenvalues of  $T_N$  converges almost surely to a proper probability distribution as  $N \rightarrow \infty$  and  $\frac{n(N)}{N} \rightarrow c > 0$ , then the empirical distribution function of eigenvalues of  $B_N$  converges almost surely to a non-random limit function given by Marcenko and Pastur. Numerical simulations illustrate the behavior of kernel density estimators and density estimators of Stieltjes transform around the true density and we give a numerical comparison on the base of  $L_1$  error varying different parameters.

**Keywords** Large dimensional random matrix · Empirical distribution function of eigenvalues · Covariance matrix · Autoregressive processes · Stieltjes transform · Kernel density estimators

## 1 Introduction

Theoretical studies on covariance matrices have a long history and appear in many domains in the real world and having links with practical problems (see [1] and [9]). For example, in multivariate statistics, spectral asymptotic results are used in solving the detection problem in signal process [9].

Consider the following random matrix:

$$B_N = X_N T_N X_N^t \quad (1)$$

---

Z. Khettab (✉) · T. Mourid  
Laboratoire de Statistiques et Modélisations Aléatoires, University of Tlemcen,  
13000 Tlemcen, Algeria  
e-mail: khettab.zahira@yahoo.fr

T. Mourid  
e-mail: t\_mourid@mail.univ-tlemcen.dz

where  $X_N = (\frac{1}{\sqrt{N}}X_{ij})$ , ( $i = 1, \dots, N$ ;  $j = 1, \dots, n(N)$ ) is a matrix ( $N \times n(N)$ ) with independent rows with the entries  $X_{ij}$  of each row satisfy an autoregressive relation AR(1) and  $T_N$  is a diagonal matrix ( $n \times n$ ) with real entries independent of  $X_N$  ( $X_N^t$  is the transpose matrix of  $X_N$ ). More precisely, for each  $i \geq 1$  we have

$$X_{ij+1} = \rho X_{ij} + \varepsilon_{ij+1}, \quad j \geq 1 \quad (2)$$

where  $(\varepsilon_{ij}, i, j \geq 1)$  are i.i.d. rv's (random values) with mean 0 and variance  $\sigma^2 > 0$ , such that  $\varepsilon_{ij}$  admits a continuous density function with respect to Lebesgue measure. The parameter  $\rho$  is such that  $|\rho| < 1$  assuring a strictly stationary process. The diagonal matrix  $T_N = \text{diag}(\tau_1, \dots, \tau_n)$  is independent of  $X_N$  and the rv's  $\tau_i$  are real.

The empirical distribution function (*e.d.f.*) of the eigenvalues  $(\lambda_i)$  of the symmetric matrix  $B_N$  is defined by

$$F^{B_N}(x) = \frac{1}{N} \sum_{i=1}^N 1_{(\lambda_i \leq x)},$$

where  $1_A$  denoting the indicator function of the set  $A$ .

A large number of papers have dealt with the problem to identify the limit of the *e.d.f.* of eigenvalues of random matrices  $B_N$  as  $N \rightarrow \infty$  and  $\frac{n(N)}{N} \rightarrow c > 0$ . Marcenko and Pastur [7] originally studied this problem for more general forms of random matrices. They establish, under some conditions on moments, that if the *e.d.f.*  $F^{T_N}$  converges to a proper distribution function  $H$ , then  $F^{B_N}$  converges in probability to a proper distribution function. Their method involves the Stieltjes transform, where they shown that the Stieltjes transform of the limiting distribution function satisfies a first-order partial differential equation, then via the characteristics they shown that this function is a solution of an algebraic equation identifying hence the limit.

Afterward, several authors [4, 5, 8, 11, 12] extended this result giving the almost sure convergence of the *e.d.f.* of eigenvalues under mild conditions on the entries  $X_{ij}$ . Most of the previous papers employ the same transform as [7] and the entries  $X_{ij}$  of the matrices are independent random variables, except the paper [3], where dependent entries are considered.

Our goal in this paper is to study, under some assumptions, the limit of the *e.d.f.*  $F^{B_N}$  of the random matrix  $B_N = X_N T_N X_N^t$ , where the entries  $X_{ij}$  of the matrix  $X_N$  satisfy an autoregressive relation AR(1) for each  $i$ . We follow the approach given in [8] where the authors apply Marcenko and Pastur method to study the limit of Stieltjes transform of the *e.d.f.*  $F^{B_N}$  and then we identify the limit law. We illustrate by numerical simulations the behavior of kernel density estimators and density estimators of Stieltjes transform to identify the true density and give  $L_1$  errors varying different parameters.

The paper is organized as follows. Section 2 provides the main result. Section 3 presents numerical simulations. The proof of the main result will be postponed in Sect. 4.

## 2 Main Result

First, we introduce some random variables and random matrices. We truncate and centralize the entries  $X_{ij}$  of the random matrix  $X_N$  to obtain new corresponding random matrices as follows: for  $i = 1, \dots, N$ ;  $j = 1, \dots, n(N)$ , let

$$\hat{X}_{ij} = X_{ij} 1_{(|X_{ij}| < \sqrt{N})}, \hat{X}_N = \left( \frac{1}{\sqrt{N}} \hat{X}_{ij} \right), \hat{B}_N = \hat{X}_N T_N \hat{X}_N^t, \quad (3)$$

$$\tilde{X}_{ij} = \hat{X}_{ij} - E(\hat{X}_{ij}), \tilde{X}_N = \left( \frac{1}{\sqrt{N}} \tilde{X}_{ij} \right), \tilde{B}_N = \tilde{X}_N T_N \tilde{X}_N^t, \quad (4)$$

and

$$\begin{cases} \bar{X}_{ij} = \tilde{X}_{ij} 1_{(|X_{ij}| \leq \ln N)} - E \tilde{X}_{ij} 1_{(|X_{ij}| \leq \ln N)}, \\ \bar{X}_N = \left( \frac{1}{\sqrt{N}} \bar{X}_{ij} \right), \bar{B}_N = \bar{X}_N T_N \bar{X}_N^t. \end{cases} \quad (5)$$

We pointed out that the problem described above has been often handled by the method of Stieltjes transform. Let  $\mathcal{M}(\mathbb{R})$  be the set of distribution functions on  $\mathbb{R}$ . Recall that the Stieltjes transform of a distribution function  $F \in \mathcal{M}(\mathbb{R})$  is defined by

$$m_F(z) = \int \frac{1}{\lambda - z} dF(\lambda), \quad z \in \mathbb{C}^+ \equiv \{z \in \mathbb{C} : \Im m z > 0\},$$

where  $\Im m$  is the imaginary part. The inversion formula is given by

$$F([a, b]) = \frac{1}{\pi} \lim_{\epsilon \rightarrow 0^+} \int_a^b \Im m m_F(x + i\epsilon) dx,$$

where  $a$  and  $b$  are continuity points of  $F$ . Also, the weak convergence of probability distribution functions is equivalent to the convergence of Stieltjes transforms (Theorem B.9, [1]). From the inversion formula, it follows that for any countable set  $S \subset \mathbb{C}^+$  such that  $\mathbb{R} \subset \bar{S}$  the closure of  $S$ , and a sequence  $(F_N) \in \mathcal{M}(\mathbb{R})$ ,  $F \in \mathcal{M}(\mathbb{R})$ , we have the following equivalence:

$$\lim_{N \rightarrow \infty} m_{F_N}(z) = m_F(z), \quad \forall z \in S \iff F_N \rightarrow F \text{ as } N \rightarrow \infty, \quad (6)$$

where  $F_N \rightarrow F$  is the vague convergence of distributions functions.

Furthermore, we consider the following random matrices. For  $j, l = 1, 2, \dots, n(N)$ , denote by  $\bar{q}_j$  the  $j$ th column of  $\bar{X}_N$  defined by (5) that is

$$\bar{q}_j = \frac{1}{\sqrt{N}} (\bar{X}_{1j}, \dots, \bar{X}_{Nj})^t := \frac{1}{\sqrt{N}} V_j, \quad (7)$$

and by

$$\bar{B}_{(j)} = \bar{B}_{(j)}^N := \bar{B}_N - \tau_j \bar{q}_j \bar{q}_j^t, \quad (8)$$

where  $\tau_j$  are the elements of  $T_N$ , and define

$$x = x_N := \frac{1}{N} \sum_{j=1}^n \frac{\tau_j}{1 + \tau_j m_{F^{\bar{B}_N}}(z)}, \quad x_{(j)} = x_{(j)}^N := \frac{1}{N} \sum_{l=1}^n \frac{\tau_l}{1 + \tau_l m_{F^{\bar{B}_{(j)}}}(z)}, \quad (9)$$

where  $m_{F^{\bar{B}_N}}$  and  $m_{F^{\bar{B}_{(j)}}}$  are Stieltjes transform of the matrices  $\bar{B}_N$  and  $\bar{B}_{(j)}$ , respectively. Finally, set

$$C_{(j)}^1 := (\bar{B}_{(j)} - zI)^{-1}, \quad C_{(j)}^2 := (x_{(j)} - z)^{-1} (\bar{B}_{(j)} - zI)^{-1} \quad (10)$$

where  $I$  is the identity matrix.

Now, we state the main result of this paper giving the almost sure limit of the *e.d.f.* of the eigenvalues of the random matrix  $B_N$  ( $tr$  is the trace of the matrix).

**Theorem 1** *Assume*

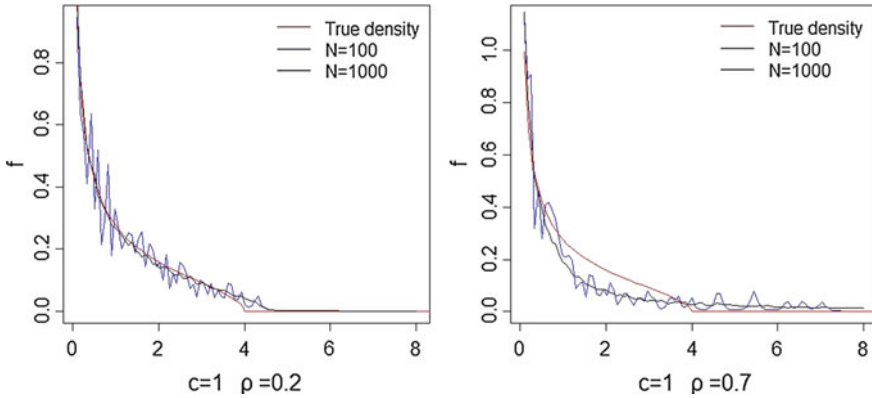
- (a) For  $N = 1, 2, \dots$ , let  $X_N = \left(\frac{1}{\sqrt{N}} X_{ij}\right)$  be a matrix ( $N \times n(N)$ ) with independent rows and an AR(1) autoregressive relation (2) in each row. The entries  $X_{ij}$ ,  $i, j \geq 1$ , have all their moments finite and  $\frac{n(N)}{N} \rightarrow c > 0$  as  $N \rightarrow \infty$ .
- (b)  $T_N = \text{diag}(\tau_1, \dots, \tau_n)$ ,  $\tau_i \in \mathbb{R}$ , and the *e.d.f.* of  $T_N$  converges almost surely to a distribution function  $H$  as  $N \rightarrow \infty$ .
- (c) The matrices  $X_N$  and  $T_N$  are independent.
- (d) For  $k = 1, 2$  and  $j = 1, 2, \dots, n(N)$ , the matrices  $C_{(j)}^k$  defined in (10) satisfy

$$E \left| V_j^t C_{(j)}^k V_j - \text{tr} C_{(j)}^k \right|^6 \leq K N^3, \text{ where } V_j \text{ given by (7) and } K > 0.$$

Then, the *e.d.f.*  $F^{B_N}$  of the random matrix  $B_N = X_N T_N X_N^t$  converges vaguely almost surely to a distribution function  $F$ , as  $N \rightarrow \infty$ , whose Stieltjes transform  $m_F(z)$  satisfies the following functional relation:

$$m_F(z) = - \left( z - c \int \frac{\tau dH(\tau)}{1 + \tau m_F(z)} \right)^{-1}; \quad z \in \mathbb{C}^+. \quad (11)$$

*Remark* Assumption (a) is fulfilled in part if the white noise ( $\varepsilon_{ij}$ ) has all moments (Gaussian white noise). Assumptions (b), (c) are standard and analogous of that



**Fig. 1** Densities of the limit law and STE with  $T_N$  the identity matrix

existing in [8]. Assumption (d) requires a control of sixth moment of quadratic form of a matrix and its trace by a third power of size  $N$ . In the case of *i.i.d.* entries  $X_{ij}, i, j \geq 1$ , assumption (d) is fulfilled (cf. Lemma 3.1 in [8]).

### 3 Numerical Simulations

As a practical impact of the main result, we illustrated in [6] the behavior of the empirical density estimator of *e.d.f.* of eigenvalues  $(\lambda_i, i = 1, \dots, N)$  of large random matrices  $B_N$ , and identify the density function of the limit law by numerical simulations. First, we recall the formulas giving density of limit law and the empirical Stieltjes transform estimator. From [10], we have for all  $x \in \mathbb{R} - \{0\}$ , and  $z = x + iy, y > 0$ , the distribution function  $F$  (limit of the *e.d.f.*  $F^{B_N}$ ) has a continuous derivative  $f$  defined by  $f(x) = (1/\pi) \Im m m_0(x)$ , where  $m_0(x)$  is given by Stieltjes transform  $m_F(z)$  as  $\lim_{z \rightarrow x} m_F(z) := m_0(x)$  (Figs. 1, 2 and Table 1).

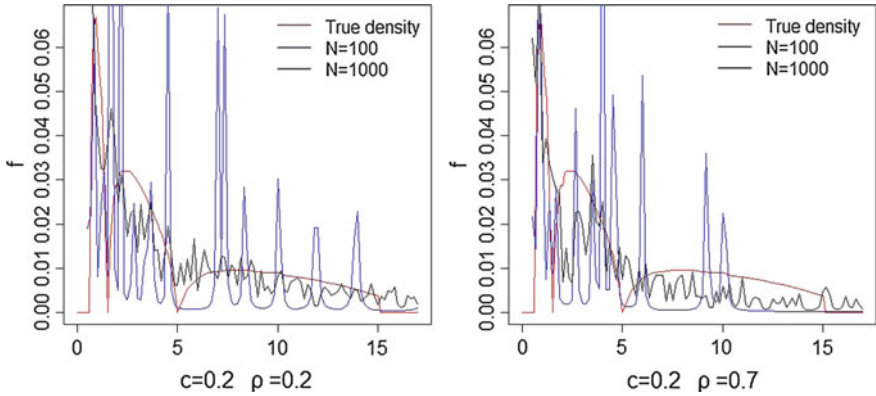
The Stieltjes Transform Estimator (STE) is defined by

$$f_N(x) = (1/\pi) \Im m m_{F^{B_N}}(z),$$

where  $m_{F^{B_N}}(z) = \frac{1}{N} \text{tr} (B_N - zI)^{-1} = \frac{1}{N} \sum_{i=1}^N (\lambda_i - z)^{-1}$ .

Now, we apply Gaussian Kernel Estimators (GKE) defined by

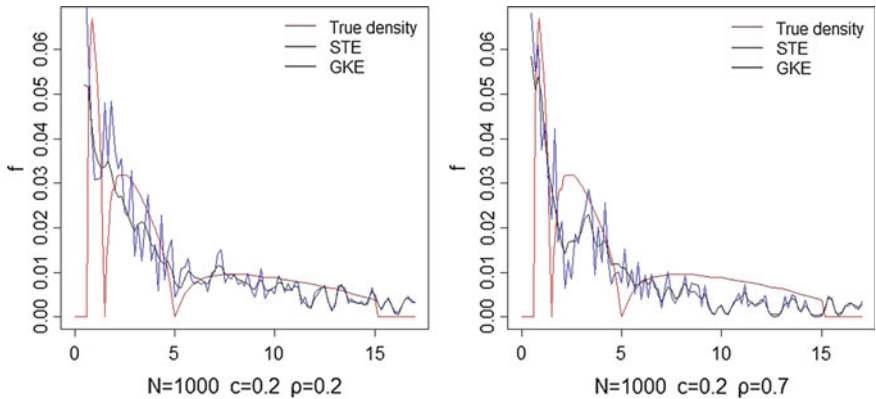
$$\hat{f}_N(\lambda) = \frac{1}{Nh_N} \sum_{i=1}^N K\left(\frac{\lambda - \lambda_i}{h_N}\right); \lambda \in \mathbb{R}$$



**Fig. 2** Densities of the limit law and STE with  $T_N$  having three eigenvalues: 1, 3, 10

**Table 1**  $L_1$ -errors of STE in Case 1:  $T_N$  identity matrix with  $c = 1$  and Case 2:  $T_N$  diagonal matrix having three eigenvalues: 1, 3, 10 with  $c = 0.2$ . For weak and strong dependence and different sample size values  $N$

	$\rho = 0.2$			$\rho = 0.7$		
$N$	100	500	1000	100	500	1000
Case 1	0.0297	0.0245	0.0200	0.0390	0.0327	0.0298
Case 2	0.0206	0.0020	0.0009	0.0151	0.0020	0.0009



**Fig. 3** Behavior of STE and GKE with matrix  $T_N$  having three eigenvalues: 1, 3, 10

where  $h_N$  is the bandwidth converging to 0 and  $Nh_N \rightarrow \infty$ , and  $K$  is a Gaussian kernel :  $K(u) = \frac{1}{\sqrt{2\pi}} \exp(-\frac{1}{2}u^2)$ .

We compare the performance of Stieltjes Transform Estimators (STE) and Gaussian Kernel Estimators (GKE) on the base of  $L_1$ -errors (Fig. 3 and Table 2).

**Table 2**  $L_1$ -errors of the STE and GKE, with  $T_N$  having three eigenvalues and  $c = 0.2$  for different sample size values  $N$

$N$	$\rho = 0.2$			$\rho = 0.7$		
	100	500	1000	100	500	1000
STE	0.0160	0.0102	0.0092	0.0126	0.0102	0.0093
GKE	0.0109	0.0094	0.0088	0.0106	0.0093	0.0091

**Conclusion**

From more numerical simulations, we may observe that the performance of estimators strongly depends on the choices of the dimension  $c$ , AR parameter  $\rho$ , and sample size  $N$ . The variability of parameters has a direct impact on the stabilization and convergence rate of the estimators. Particular choices of parameters confirm a good performance of the estimators and lead to indicate optimal values for these parameters. We also observe an effect of the dimension  $c$  on density estimator convergence rate. For  $c > 1$ , for both weak and strong dependencies ( $\rho = 0.2, \rho = 0.7$ ), the estimators perform well from on  $N = 100$ . However, for small values of  $c$  ( $c < 0.2$ ), there is an influence of parameter values on the convergence rate. For weak dependence ( $\rho = 0.2$ ) STE perform quite well for moderate value  $N = 100$ , whereas for strong dependence ( $\rho = 0.7$ ) the estimator accurate enough well only for large  $N$  (when  $N > 1000$ ). The number of eigenvalues of  $T_N$  has an effect on the behavior of the estimators as well as on their performance. Both estimators perform well and give a good representation of the true density with a small advantage of GKE.

**4 Proof of the Main Result**

Recall these well-known facts. For each  $i$ , the process  $(X_{ij}, j \in \mathbb{Z})$  satisfying relation (2) is a stationary AR(1) process, then it satisfies the geometric strong mixing property (G.S.M) with strong mixing coefficient  $\alpha_k = \alpha_k(\mathcal{F}_0^m, \mathcal{F}_{m+k}^\infty) = O(\rho^k)$ , where  $0 < \rho < 1$  and  $\mathcal{F}_a^b = \mathcal{F}_{a,i}^b = \sigma(X_{ij}, a \leq j \leq b)$ , whenever  $\varepsilon_{ij}$  has a strictly positive continuous density (see [2] p. 58).

The covariance between two real-valued rv's is bounded as follows: if  $\eta \in L^p$  and  $\xi \in L^q$  are  $\mathcal{F}_0^m$  and  $\mathcal{F}_{m+k}^\infty$ -measurable, respectively, then we have

$$|E(\eta\xi) - E(\eta)E(\xi)| \leq 12 \|\eta\|_p \|\xi\|_q \alpha_k^{\frac{1}{r}} \tag{12}$$

for all  $1 \leq p, q, r \leq \infty$  with  $\frac{1}{p} + \frac{1}{q} + \frac{1}{r} = 1$ , and the norm  $\|\cdot\|_p = E^{\frac{1}{p}}|\cdot|^p$ .

On the other hand, there exists a distance  $D(\cdot, \cdot)$  on the space  $\mathcal{M}(\mathbb{R})$ , such that for two sequences  $(F_N), (G_N) \in \mathcal{M}(\mathbb{R})$ , we have (see [8]).

$$\lim_{N \rightarrow \infty} \|F_N - G_N\| = 0 \implies \lim_{N \rightarrow \infty} D(F_N, G_N) = 0, \tag{13}$$

where  $\|\cdot\|$  denotes the sup-norm of bounded functions from  $\mathbb{R}$  to  $\mathbb{R}$ .

To lighten the writing, the dependency of most of variables on  $N$  will occasionally be dropped from the notation. Now, we replace  $T$  by a suitable matrix for further analysis: for  $\theta \geq 0$  define,  $T_\theta = \text{diag}(\tau_1 1_{(|\tau_1| \leq \theta)}, \dots, \tau_n 1_{(|\tau_n| \leq \theta)})$ , and let  $Q$  be any  $(N \times n)$  matrix. If  $\theta$  and  $(-\theta)$  are continuity points of  $H$ , then by Lemma 2.5 of [8] and assumption (b) of the Theorem 1, as  $N \rightarrow \infty$  and  $\frac{n}{N} \rightarrow c > 0$ , we have

$$\left\| F^{QTQ'} - F^{QT_\theta Q'} \right\| \leq \frac{1}{N} \text{rg}(T - T_\theta) = \frac{1}{N} \sum_{j=1}^n 1_{(|\tau_j| > \theta)} \rightarrow cH\{[-\theta, \theta]^c\} \text{ a.s.}$$

It follows that if  $\theta = \theta_N \rightarrow \infty$ , then

$$\left\| F^{QTQ'} - F^{QT_\theta Q'} \right\| \rightarrow 0 \text{ a.s.} \quad (14)$$

Choose  $\theta$  such that

$$\theta^4 \left[ E^{\frac{2}{3}} |X_{11}|^2 1_{(|X_{11}| \geq \ln N)} + \frac{1}{N} \right] \rightarrow 0, \quad (15)$$

and

$$\sum_{N=1}^{\infty} \theta^8 \left[ \frac{1}{N^{7/6}} E^{1/6} |X_{11}|^4 1_{(\ln N \leq |X_{11}| < \sqrt{N})} + \frac{1}{N^2} \right] < \infty. \quad (16)$$

For continue, we need the following result.

**Lemma 1** *Let the  $(N \times n)$  matrices  $X = \left(\frac{1}{\sqrt{N}} X_{ij}\right)$  verifying assumption (a) of Theorem 1, and  $\hat{X} = \left(\frac{1}{\sqrt{N}} \hat{X}_{ij}\right)$  where  $\hat{X}_{ij} = X_{ij} 1_{(|X_{ij}| < \sqrt{N})}$ . For  $\theta \geq 0$  set  $T_\theta = \text{diag}(\tau_1 1_{(|\tau_1| \leq \theta)}, \dots, \tau_n 1_{(|\tau_n| \leq \theta)})$ ,  $\tau_i \in \mathbb{R}$ . We have*

$$D\left(F^{XT_\theta X^t}, F^{\hat{X}T_\theta \hat{X}^t}\right) \rightarrow 0 \text{ a.s.}$$

*Proof* From Corollary A.42 of [1], we find

$$D^2\left(F^{XT_\theta X^t}, F^{\hat{X}T_\theta \hat{X}^t}\right) \leq \left[ \frac{2}{N} \text{tr}(XX^t - \hat{X}\hat{X}^t) + \frac{4}{N} \text{tr}\hat{X}\hat{X}^t \right] \left[ \frac{\theta^2}{N} \text{tr}(XX^t - \hat{X}\hat{X}^t) \right].$$

In order that this distance tends almost surely to 0, we can show by Borel–Cantelli lemma that  $\left[ \frac{\theta^2}{N} \text{tr}(XX^t - \hat{X}\hat{X}^t) \right]$  tends to 0 and  $\left[ \frac{4}{N} \text{tr}\hat{X}\hat{X}^t \right]$  is bounded almost surely. So the result.



### ***Proof of the Theorem 1***

For  $(N \times n)$  matrix,  $X = \left(\frac{1}{\sqrt{N}} X_{ij}\right)$  verifying assumption (a) of Theorem 1. With help to inequality (12) and the fact that  $(X_{ij})$  satisfies the G.S.M. property, we obtain

$$M_1 \leq E \left( \text{tr} (X X^t)^2 \right) \leq M_2 \quad (17)$$

where

$$M_1 = \frac{K}{N} \{nE |X_{11}|^4 + n(N+n-2) E^2 |X_{11}|^2 - E^{\frac{2}{3}} |X_{11}|^6\},$$

$$M_2 = \frac{K}{N} \{nE |X_{11}|^4 + n(N+n-2) E^2 |X_{11}|^2 + E^{\frac{2}{3}} |X_{11}|^6 + (N-1) E^{\frac{4}{3}} |X_{11}|^3\}.$$

With the same arguments, we may deduce a bound of the variance

$$\text{var} \left( \text{tr} (X X^t)^2 \right) \leq \frac{K}{N^4} \{N^4 E |X_{11}|^4 E^2 |X_{11}|^2 \quad (18)$$

$$+ N^3 [E^{\frac{4}{3}} |X_{11}|^6 + E^{\frac{1}{2}} |X_{11}|^4 E^{\frac{1}{3}} |X_{11}|^6 E^{\frac{1}{3}} |X_{11}|^{12} + E |X_{11}|^2 E^{\frac{2}{3}} |X_{11}|^9] \\ + N^2 [E^{\frac{2}{3}} |X_{11}|^6 E^{\frac{1}{6}} |X_{11}|^{24} + E^{\frac{1}{3}} |X_{11}|^{15} E |X_{11}|^3 + E^{\frac{2}{3}} |X_{11}|^{12} + E^{\frac{5}{12}} |X_{11}|^{12} E^{\frac{1}{6}} |X_{11}|^{18}].$$

Using (14) and (13), we may write

$$D \left( F^{X^T X^t}, F^{X^T \hat{\theta} X^t} \right) \rightarrow 0 \quad \text{and} \quad D \left( F^{\hat{X}^T \hat{\theta} \hat{X}^t}, F^{\hat{X}^T \hat{X}^t} \right) \rightarrow 0 \quad a.s.$$

Furthermore, by Lemma 1, we get

$$D \left( F^{X^T X^t}, F^{\hat{X}^T \hat{X}^t} \right) \rightarrow 0 \quad a.s. \quad (19)$$

For  $\hat{B}_N$  and  $\tilde{B}_N$  defined by relations (3) and (4), we have from Lemma 2.5 of [8],

$$\left\| F^{\hat{B}_N} - F^{\tilde{B}_N} \right\| \rightarrow 0. \quad (20)$$

Let  $\bar{\bar{X}}_{ij} = \tilde{X}_{ij} - \bar{X}_{ij}$ . Hence,

$$\bar{\bar{X}}_{ij} = \tilde{X}_{ij} 1_{(|X_{ij}| \geq \ln N)} + E \tilde{X}_{ij} 1_{(|X_{ij}| < \ln N)}, \quad \bar{\bar{X}} = \left( \frac{1}{\sqrt{N}} \bar{\bar{X}}_{ij} \right)$$

where  $\tilde{X}_{ij}$  and  $\bar{X}_{ij}$  are defined by the relations (4), (5), respectively. Then, from Cauchy–Schwartz inequality, we can show that the squared distance

$$D^2 \left( F^{\tilde{X}T_\theta \tilde{X}'} , F^{\bar{X}T_\theta \bar{X}'} \right)$$

is bounded by

$$\frac{1}{N} \left\{ \theta^2 \text{tr} \left( \bar{X} \bar{X}' \right)^2 + 4 \left[ \theta^4 \text{tr} \left( \bar{X} \bar{X}' \right)^2 \text{tr} \left( \bar{X} \bar{X}' \right)^2 \right]^{\frac{1}{2}} + 4 \left[ \left[ \theta^4 \text{tr} \left( \bar{X} \bar{X}' \right)^2 \text{tr} \left( \bar{X} \bar{X}' \right)^2 \right]^{\frac{1}{2}} \theta^2 \text{tr} \left( \bar{X} \bar{X}' \right)^2 \right]^{\frac{1}{2}} \right\}.$$

Therefore, in order to show that almost surely

$$D \left( F^{\tilde{X}T_\theta \tilde{X}'} , F^{\bar{X}T_\theta \bar{X}'} \right) \rightarrow 0, \quad (21)$$

it suffices to verify that

$$\theta^4 \frac{1}{N} \text{tr} \left( \bar{X} \bar{X}' \right)^2 \rightarrow 0, \quad \frac{1}{N} \text{tr} \left( \bar{X} \bar{X}' \right)^2 = O(1) \text{ a.s.} \quad (22)$$

Since  $E \left( \bar{X}_{11} \right) = 0$  and  $\bar{X}_{ij} = \tilde{X}_{ij} 1_{(|X_{ij}| \geq \ln N)} + E \tilde{X}_{ij} 1_{(|X_{ij}| < \ln N)}$ , we have

$$E \left| \bar{X}_{11} \right|^2 \leq K E |X_{11}|^2 1_{(|X_{11}| \geq \ln N)} \rightarrow 0. \quad (23)$$

For  $p \geq 4$ ,

$$E \left| \bar{X}_{11} \right|^p \leq K \left( N^{\frac{p-4}{2}} E |X_{11}|^4 1_{(\ln N \leq |X_{11}| < \sqrt{N})} + 1 \right). \quad (24)$$

By dominated convergence theorem, we get

$$E \left| \bar{X}_{11} \right|^2 \rightarrow E |X_{11}|^2 = \gamma. \quad (25)$$

For  $p \geq 4$  and definition of rv's  $\bar{X}_{11}$ , we have

$$E \left| \bar{X}_{11} \right|^p \leq K (\ln N)^{p-2}. \quad (26)$$

From (15), (23), (24),  $E(|X_{11}|^4 1_{(\ln N \leq |X_{11}| < \sqrt{N})}) \leq N E |X_{11}|^2$  and relation (17), we may write

$$E \left[ \frac{1}{N} \theta^4 \text{tr} \left( \bar{\bar{X}} \bar{\bar{X}}^t \right)^2 \right] \leq K \theta^4 \left[ E^{\frac{2}{3}} |X_{11}|^2 \mathbf{1}_{(|X_{11}| \geq \ln N)} + \frac{1}{N} \right] \rightarrow 0.$$

Also (18) gives

$$\text{var} \left( \frac{1}{N} \theta^4 \text{tr} \left( \bar{\bar{X}} \bar{\bar{X}}^t \right)^2 \right) \leq K \theta^8 \left[ \frac{1}{N^{7/6}} E^{1/6} |X_{11}|^4 \mathbf{1}_{(\ln N \leq |X_{11}| < \sqrt{N})} + \frac{1}{N^2} \right],$$

where the latter bound is summable by (16).

Hence, we obtain  $\frac{1}{N} \theta^4 \text{tr} \left( \bar{\bar{X}} \bar{\bar{X}}^t \right)^2 \rightarrow 0$  a.s.

Now it remains to show that  $\frac{1}{N} \text{tr} \left( \bar{X} \bar{X}^t \right)^2 = O(1)$  a.s. Using (17), (25) and (26), we find

$$\begin{aligned} K \left\{ -\frac{(\ln N)^{\frac{8}{3}}}{N^2} \right\} &\leq E \left[ \frac{1}{N} \text{tr} \left( \bar{X} \bar{X}^t \right)^2 \right] - \frac{n}{N} \left( \frac{n}{N} + 1 - \frac{2}{N} \right) E^2 |\bar{X}_{11}|^2 \\ &\leq K \left\{ \frac{n}{N} \frac{(\ln N)^2}{N} + \frac{(\ln N)^{\frac{8}{3}}}{N^2} + \frac{(\ln N)^{\frac{4}{3}}}{N} \right\}. \end{aligned}$$

Consequently,  $E \left[ \frac{1}{N} \text{tr} \left( \bar{X} \bar{X}^t \right)^2 \right] - \frac{n}{N} \left( \frac{n}{N} + 1 - \frac{2}{N} \right) E^2 |\bar{X}_{11}|^2 \rightarrow 0$ , and,

$$E \left[ \frac{1}{N} \text{tr} \left( \bar{X} \bar{X}^t \right)^2 \right] \rightarrow \gamma^2 [c(c+1)].$$

Concerning the variance, by (18), (25) and (26), we may obtain

$$\text{var} \left( \frac{1}{N} \text{tr} \left( \bar{X} \bar{X}^t \right)^2 \right) \leq K \frac{(\ln N)^{17/3}}{N^3},$$

which is summable. Then, (22) is verified from which (21) follows. This result with (14) allow us to write,  $D \left( F^{\bar{X}T\bar{X}^t}, F^{\bar{X}T\bar{X}^t} \right) \rightarrow 0$  a.s.

From (19) and (20), in order to prove  $D \left( F^{XTX^t}, F \right) \rightarrow 0$  a.s, it suffices to verify that,  $D \left( F^{\bar{X}T\bar{X}^t}, F \right) \rightarrow 0$  a.s. For this aim, we shall show that for any  $z \in \mathbb{C}^+$ ,  $m_{F^{\bar{X}T\bar{X}^t}}(z) \rightarrow m_F(z)$  a.s.

Let  $z \in \mathbb{C}^+$  and  $\bar{B}_N = \bar{X}T\bar{X}^t$ , the sequence  $\{F^{\bar{B}_N}\}$  satisfies the assumptions of Lemma 2.8 of [8]. So  $\exists m > 0$  such that

$$\inf_N F^{\bar{B}_N} [-m, m] > 0, \quad \delta = \inf_N \Im m \left( m_{F^{\bar{B}_N}}(z) \right) > 0 \quad a.s.$$

Write  $\bar{B}_N - zI = (x - z)I + \bar{X}T\bar{X}^t - xI$ , and then

$$(x - z)^{-1} - m_{\bar{B}_N}(z) = \frac{1}{N} \sum_{j=1}^n \frac{\tau_j}{1 + \tau_j m_{F^{\bar{B}_N}}(z)} d_j, \quad (27)$$

where

$$d_j = d_j^N = \frac{1 + \tau_j m_{F^{\bar{B}_N}}(z)}{1 + \tau_j \bar{q}_j^t (\bar{B}_{(j)} - zI)^{-1} \bar{q}_j} \bar{q}_j^t (\bar{B}_{(j)} - zI)^{-1} ((x - z)^{-1} I) \bar{q}_j \\ - \frac{1}{N} \text{tr} (\bar{B}_N - zI)^{-1} ((x - z)^{-1} I),$$

with  $\bar{q}_j$  denote the  $j$ th column of  $\bar{X}$ , and  $\bar{B}_{(j)}$ ,  $x$ ,  $x_{(j)}$  are defined by relations (8) and (9).

Lemma 3.1 of [8] and assumption (d) of the Theorem 1 permit us to obtain

$$\max_{j \leq n} \max [\beta_1, \beta_2, \beta_3] \rightarrow 0 \text{ a.s} \quad (28)$$

where

$$\beta_1 = \left| \|\bar{q}_j\|^2 - 1 \right|, \\ \beta_2 = \left| \bar{q}_j^t (\bar{B}_{(j)} - zI)^{-1} \bar{q}_j - \frac{1}{N} \text{tr} (\bar{B}_{(j)} - zI)^{-1} \right|, \\ \beta_3 = \left| \bar{q}_j^t (\bar{B}_{(j)} - zI)^{-1} ((x_{(j)} - z)I)^{-1} \bar{q}_j \right. \\ \left. - \frac{1}{N} \text{tr} (\bar{B}_{(j)} - zI)^{-1} ((x_{(j)} - z)I)^{-1} \right|.$$

Lemma 2.6 of [8] gives us,  $\max_{j \leq n} \max[|\gamma_1|, |\gamma_2|] \rightarrow 0$ , where  $\gamma_1 = m_{F^{\bar{B}_{(j)}}}(z) - m_{F^{\bar{B}_N}}(z)$ ,  $\gamma_2 = m_{F^{\bar{B}_N}}(z) - \bar{q}_j^t (\bar{B}_{(j)} - zI)^{-1} \bar{q}_j$ .

So that for  $N$  large enough, we have,  $\max_{j \leq n} \max[|\Im \gamma_1|, |\Im \gamma_2|] < \frac{\delta}{2}$ .

Then, for  $j, l \leq n$ ,

$$\left| \frac{1 + \tau_j m_{F^{\bar{B}_N}}(z)}{1 + \tau_j \bar{q}_j^t (\bar{B}_{(j)} - zI)^{-1} \bar{q}_j} - 1 \right| < \frac{2}{\delta} |\gamma_2|,$$

and

$$\left| \frac{\tau_l}{1 + \tau_l m_{F^{\bar{B}_N}}(z)} - \frac{\tau_l}{1 + \tau_l m_{F^{\bar{B}_{(j)}}}(z)} \right| \leq \frac{2}{\delta^2} |\gamma_1|.$$

Therefore,

$$\max_{j \leq n} \max \left[ \left| \frac{1 + \tau_j m_{F^{\bar{B}_N}}(z)}{1 + \tau_j \bar{q}_j^t (\bar{B}_{(j)} - zI)^{-1} \bar{q}_j} - 1 \right|, |x - x_{(j)}| \right] \rightarrow 0. \quad (29)$$

Using Lemmas 2.6, 2.7 of [8] and (28), (29), we may have

$$\max_{j \leq n} d_j \rightarrow 0.$$

Since

$$\left| \frac{\tau_j}{1 + \tau_j m_{F^{\bar{B}_N}}(z)} \right| \leq \frac{1}{\delta},$$

we may conclude from (27) that

$$(x - z)^{-1} - m_{\bar{B}_N}(z) \rightarrow 0.$$

Hence, the relation (11) is satisfied.

Finally, using (6), the proof of Theorem 1 is now complete.

**Acknowledgements** The authors would like to thank the Editor and anonymous referees for insightful comments improving the presentation of this paper.

## References

1. Bai, Z.D., Silverstein, J.W.: Spectral Analysis of Large Dimensional Random Matrices, 2nd edn. Springer (2010)
2. Bosq, D.: Linear processes in function spaces. Theory and Applications. Lecture Notes in Statistics, vol. 149. Springer (2000)
3. Boutet de Monvel, A., Khorunzhy, A.: Limit theorems for random matrices with correlated entries. Markov Process. Relat. Fields **4**(2), 175–197 (1998)
4. Grenander, U., Silverstein, J.W.: Spectral analysis of networks with random topologies. SIAM J. Appl. Math. **32**(2), 499–519 (1977)
5. Jonsson, D.: Some limit theorems for the eigenvalues of a sample covariance matrix. J. Multivar. Anal. **12**, 1–38 (1982)
6. Khettab, Z., Mourid, T.: Eigenvalues empirical distribution of covariance matrices of AR processes. A simulation study. Annales de l'ISUP, 60, fasc 1–2, 3–22 (2016)
7. Marcenko, V.A., Pastur, L.A.: Distribution of eigenvalues in certain sets of random matrices. Mat. SB (N. S.). **72**(114), 507–536 (1967)
8. Silverstein, J.W., Bai, Z.D.: On the empirical distribution of eigenvalues of a class of large dimensional random matrices. J. Multivar. Anal. **54**(2), 175–192 (1995)
9. Silverstein, J.W., Bai, Z.D., Couillet, R., Debbah, M.: Eigen-inference for energy estimation of multiple sources. IEEE Trans. Inf. Theor. **57**(4), 2420–2439 (2011)
10. Silverstein, J.W., Choi, S.I.: Analysis of the limiting spectral distribution of large dimensional random matrices. J. Multivar. Anal. **54**(2), 295–309 (1995)
11. Wachter, K.W.: The limiting empirical measure of multiple discriminant ratios. Ann. Stat. **8**, 937–957 (1980)
12. Yin, Y.Q.: Limiting spectral distribution for a class of random matrices. J. Multivar. Anal. **20**, 50–68 (1986)