

A note on the Marchenko-Pastur law for a class of random matrices with dependent entries

Sean O'Rourke*

Abstract

We consider a class of real random matrices with dependent entries and show that the limiting empirical spectral distribution is given by the Marchenko-Pastur law. Additionally, we establish a rate of convergence of the expected empirical spectral distribution.

Keywords: Random Matrix Theory; Marchenko-Pastur law; Stieltjes transform.

AMS MSC 2010: 60B20; 47A10; 15A18.

Submitted to ECP on May 11, 2012, final version accepted on July 12, 2012.

Supersedes arXiv:1201.3554.

1 Introduction and Main Results

Suppose M_n is a $n \times n$ matrix with real eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_n$. Then the empirical spectral distribution (ESD) of the matrix M_n is defined by

$$F^{M_n}(x) := \frac{\#\{1 \leq i \leq n : \lambda_i \leq x\}}{n}.$$

We will be interested in the case when $M_n := \frac{1}{n} A_n A_n^T$ and A_n is an $n \times N$ real random matrix.

If the entries of A_n are i.i.d. random variables with zero mean and variance one, we call M_n a sample covariance matrix. There are many results concerning the limiting behavior of the spectral distribution of sample covariance matrices. For example, Marchenko and Pastur ([17]) and Wachter ([21]) prove that the ESD $F^{\frac{1}{n} A_n A_n^T}(x)$ converges to $F_c(x)$ provided that $N/n \rightarrow c \in (0, \infty)$, where F_c is the distribution function for the Marchenko-Pastur law with parameter $c > 0$. That is, F_c has density

$$p_c(x) = \begin{cases} \frac{\sqrt{(x-a)(b-x)}}{2\pi x} & : a \leq x \leq b, \\ 0 & : \text{otherwise,} \end{cases}$$

and a point mass $1 - c$ at the origin if $c < 1$, where $a = (1 - \sqrt{c})^2$ and $b = (1 + \sqrt{c})^2$. The above convergence holds with probability 1 (see for example [5] and [8, Chapter 3]).

There are a number of results in which the independence assumption (on the entries of A_n) is weakened. In the seminal paper by Marchenko and Pastur [17], one considers independent rows rather than independent entries. In [22], Yin and Krishnaiah consider the case where the independent rows have a spherically symmetric distribution.

*Rutgers University, USA. E-mail: sdo21@math.rutgers.edu

More recently in 2006, Aubrun obtained the Marchenko-Pastur law for matrices with independent rows distributed uniformly on the l_p^n balls, [4]. This was generalized by Pajor and Pastur in [20] to matrices with independent rows distributed according to an arbitrary isotropic log-concave measure.

In [13] and [15], Götze and Tikhomirov study two classes of random matrices which generalize Wigner random matrices and sample covariance random matrices. In particular, these matrices satisfy certain martingale-type conditions without any assumption on the independence of the entries. In a similar setting, Adamczak studied a class of random matrices with uncorrelated entries in which each normalized row and normalized column converges to one in probability, [1].

Other random matrix ensembles with dependent entries that have been studied include random Markov matrices with independent rows and doubly stochastic random matrices (see [10, 9, 11] and references contained therein).

In this note, we study a class of random matrices with dependent entries and show that the limiting empirical distribution of the eigenvalues is given by the Marchenko-Pastur law. In particular, we consider a sequence of $n \times N$ random matrices A_n with the following properties.

Definition 1.1 (Condition **C0**). *Let $\{A_n\}_{n \geq 1}$ be a sequence of $n \times N$ real random matrices where $N = N(n)$ and $c_n := N/n$. We let $r_1^{(n)}, \dots, r_n^{(n)}$ denote the rows of $A_n = (\zeta_{ij}^{(n)})_{1 \leq i \leq n, 1 \leq j \leq N}$ and define the σ -algebra associated to row k as*

$$\mathcal{F}_k^{(n)} := \sigma(r_1^{(n)}, \dots, r_{k-1}^{(n)}, r_{k+1}^{(n)}, \dots, r_n^{(n)})$$

for all $k = 1, \dots, n$. Let $\mathbb{E}_k[\cdot]$ denote the conditional expectation with respect to the σ -algebra associated to row k . We then say that the sequence $\{A_n\}_{n \geq 1}$ obeys condition **C0** if the following hold:

(i) $\mathbb{E}_k[\zeta_{ki}^{(n)}] = 0$ for all i, k, n

(ii) One has

$$q_n := \sup_k \frac{1}{n} \sum_{i=1}^N \mathbb{E} |\mathbb{E}_k[(\zeta_{ki}^{(n)})^2] - 1| = o(1)$$

(iii) One has

$$\sup_{k, i \neq j} |\mathbb{E}_k[\zeta_{ki}^{(n)} \zeta_{kj}^{(n)}]| + \sup_{k, i, j \neq l} |\mathbb{E}_k[(\zeta_{ki}^{(n)})^2 \zeta_{kj}^{(n)} \zeta_{kl}^{(n)}]| = O(n^{-1/2} \gamma_n)$$

a.s., where $\gamma_n \rightarrow 0$ as $n \rightarrow \infty$.

(iv) $\sup |\mathbb{E}_k[\zeta_{ki}^{(n)} \zeta_{kj}^{(n)} \zeta_{kl}^{(n)} \zeta_{km}^{(n)}]| = O(n^{-1} \gamma_n)$ a.s where the supremum is over all k and all i, j, l, m distinct.

(v) $\sup_{n, i, j} \mathbb{E} |\zeta_{ij}^{(n)}|^4 \leq M < \infty$

(vi) One has

$$\rho_n := \sup_k \frac{1}{n^2} \sum_{1 \leq i, j \leq N} \mathbb{E} |\mathbb{E}_k[(\zeta_{ki}^{(n)})^2 (\zeta_{kj}^{(n)})^2] - 1| = o(1).$$

(vii) There exists a non-negative integer sequence $\beta_n = o(\sqrt{n})$ such that $\sigma(r_{i_1}^{(n)}, \dots, r_{i_k}^{(n)})$ and $\sigma(r_{j_1}^{(n)}, \dots, r_{j_m}^{(n)})$ are independent σ -algebras whenever

$$\min_{1 \leq l \leq k, 1 \leq p \leq m} |i_l - j_p| > \beta_n.$$

Remark 1.2. Condition (i) implies that entries from different rows are uncorrelated while (iii) and (iv) allow for a weak correlation amongst entries in the same row. Condition (ii) is a requirement on the variance of the entries and (v) is a moment assumption on the entries. Condition (vi) is of a technical nature. In particular, (vi) (along with (ii)) allows one to control terms of the form

$$\sup_k \text{Var} \left(\frac{1}{n} |r_k^{(n)}|^2 \right)$$

where $|r_k^{(n)}|$ is the Euclidian norm of the vector $r_k^{(n)}$. In words, condition (vii) implies that rows, which are “far enough apart,” are independent.

Example 1.3. Let ξ be a real random variable with mean zero, variance one, and $\mathbb{E}|\xi^4| < \infty$. Let A_n be an $n \times N$ matrix where each entry is an i.i.d. copy of ξ . If $N/n \rightarrow c \in (0, \infty)$, then A_n satisfies Definition 1.1. All the results in this paper are already known for such matrices with i.i.d. entries. See for example [17], [8, Chapter 3], [5], and references contained therein.

Example 1.4. Let A_n be a $n \times (2n)$ matrix where the rows are i.i.d. random vectors such that the entries of $r_k^{(n)}$ are ± 1 symmetric Bernoulli random variables chosen uniformly such that the sum of the entries of each row is zero. Then the sequence $\{A_n\}_{n \geq 1}$ obeys condition **C0**. Indeed, one can compute

$$\begin{aligned} \mathbb{E}[\zeta_{ij}^{(n)}] &= 0, \\ \text{Var}[\zeta_{ki}^{(n)}] &= 1, \\ \mathbb{E}[\zeta_{ki}^{(n)} \zeta_{kj}^{(n)}] &= -\frac{1}{2N-1} \text{ for } i \neq j, \end{aligned}$$

and

$$\mathbb{E}[\zeta_{ki}^{(n)} \zeta_{kj}^{(n)} \zeta_{kl}^{(n)} \zeta_{km}^{(n)}] = \frac{12N^2 - 12N}{2N(2N-1)(2N-2)(2N-3)} = O\left(\frac{1}{N^2}\right)$$

for i, j, l, m distinct, where $N = 2n$. In particular, one finds that $\gamma_n = n^{-1/2}$ and $q_n, \rho_n, \beta_n = 0$.

Let us mention that the conditions in Definition 1.1 are similar to the assumptions of Theorem 1 in [17]. However, in [17], the authors require the rows of A_n to be independent.

Also, the sequence of random matrices defined in Example 1.4 satisfies condition **C0**, but does not satisfy the assumptions of the theorems provided in [4], [20], [13], or [1].

Let $\|M\|$ denote the spectral norm of the matrix M . In this paper, we shall prove the following theorems.

Theorem 1.5. Let $\{A_n\}_{n \geq 1}$ be a sequence of real random matrices that obey condition **C0** and assume $c_n := N/n \rightarrow c \in (0, \infty)$. Then

$$\|\mathbb{E}F_n^{\frac{1}{n}} A_n A_n^T - F_c\| := \sup_x |\mathbb{E}F_n^{\frac{1}{n}} A_n A_n^T - F_c| \rightarrow 0$$

as $n \rightarrow \infty$. Moreover, if there exists $p > 1$ such that

$$\sum_{n=1}^{\infty} \frac{(\beta_n + 1)^p}{n^{p/2}} < \infty \tag{1.1}$$

then

$$\|F_n^{\frac{1}{n}A_nA_n^T} - F_c\| \longrightarrow 0$$

almost surely as $n \rightarrow \infty$.

Theorem 1.6. Let $\{A_n\}_{n \geq 1}$ be a sequence of real random matrices that obey condition **C0** and assume $c_n := N/n \geq 1$ such that $c_n \rightarrow c \in [1, \infty)$. Additionally assume that

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \mathbb{E} \|A_n A_n^T\| < \infty. \tag{1.2}$$

Then we obtain that

$$\|\mathbb{E} F_n^{\frac{1}{n}A_nA_n^T} - F_{c_n}\| = O \left(\max \left(q_n^{1/22}, \gamma_n^{1/22}, \rho_n^{1/22}, \left(\frac{(\beta_n + 1)^2}{n} \right)^{1/22} \right) \right).$$

Remark 1.7. We stated Theorem 1.5 for a sequence of random matrices that obey condition **C0**. However, it is actually possible to prove the convergence of the expected ESD without condition (vii) from Definition 1.1. That is, if the sequence $\{A_n\}_{n \geq 1}$ satisfies conditions (i) - (vi) from Definition 1.1 with $c_n \rightarrow c \in (0, \infty)$, then

$$\|\mathbb{E} F_n^{\frac{1}{n}A_nA_n^T} - F_c\| \longrightarrow 0$$

as $n \rightarrow \infty$. The proof of this statement repeats the proof of Theorem 1.5 almost exactly. We detail the necessary changes in Remark 4.1. It should be noted that the almost sure convergence portion of Theorem 1.5 still requires condition (vii) from Definition 1.1 and (1.1).

Remark 1.8. Without any additional information on the convergence rate of c_n to c , we cannot obtain a rate of convergence of $\|\mathbb{E} F_n^{\frac{1}{n}A_nA_n^T} - F_c\|$. This is why F_{c_n} appears in Theorem 1.6.

Remark 1.9. The rates obtained in Theorem 1.6 are not optimal and are obtained as a simple corollary to Lemma 3.1 below.

Example 1.10. Let $\{A_n\}_{n \geq 1}$ be the sequence of random matrices defined in Example 1.4. Theorem 1.5 implies that

$$\|F_n^{\frac{1}{n}A_nA_n^T} - F_2\| \longrightarrow 0$$

almost surely as $n \rightarrow \infty$. We will now use Theorem 1.6 to obtain a rate of convergence for $\mathbb{E} F_n^{\frac{1}{n}A_nA_n^T}$. We must verify that (1.2) holds. By [2, Theorem 3.13]¹, there exists $C, C' > 0$ such that for any $0 < \epsilon < 1/3$,

$$\mathbb{P} (\|A_n A_n^T\| \geq Cn) \leq C' \frac{\log n}{n^{1+\epsilon}}.$$

Since we always have the bound

$$\|A_n A_n^T\| \leq \text{Tr}(A_n A_n^T) = n^2,$$

¹One technical assumption required by Theorem 3.13 is control over the ψ_1 -norm ($\|\cdot\|_{\psi_1}$) of the term $|\langle \xi, y \rangle|$ where ξ is a row of the matrix A_n and y is an arbitrary unit vector. In particular, one can show that

$$\|\langle \xi, y \rangle\|_{\psi_1} \leq \frac{\sqrt{n}}{\log n^{1+\epsilon}}.$$

The bound follows by applying Markov's inequality, which yields

$$\mathbb{P}(|\langle \xi, y \rangle| > t) = O(t^{-4}),$$

and then taking $t = n^{1/3}$.

it follows that

$$\mathbb{E}\|A_n A_n^T\| = O(n).$$

Therefore, Theorem 1.6 gives the rate of convergence

$$\|\mathbb{E}F_n^{\frac{1}{n}} A_n A_n^T - F_2\| = O(n^{-1/44}).$$

2 Stieltjes Transform

If $G(x)$ is a function of bounded variation on the real line, then its Stieltjes transform is defined by

$$S_G(z) = \int \frac{1}{x-z} dG(x)$$

for $z \in D := \{z \in \mathbb{C} : \text{Im}(z) > 0\}$.

Let $m_c(z)$ be the Stieltjes transform of F_c , the distribution function of the Marchenko-Pastur law with parameter c . One can then check (see for example [8]), that

$$m_c(z) = \frac{-(z+1-c) + \sqrt{(z+1-c)^2 - 4z}}{2z}.$$

Furthermore, $m_c(z)$ can be characterized uniquely as the solution to

$$m_c(z) = \frac{1}{c-1-z-zm_c(z)} \tag{2.1}$$

that satisfies $\text{Im}(zm_c(z)) \geq 0$ for all z with $\text{Im} z > 0$.

We will study the Stieltjes transform of the ESD of the random matrix $\frac{1}{n} A_n A_n^T$ in order to prove Theorems 1.5 and 1.6. In particular, the following lemma states that it suffices to show the convergence of the Stieltjes transform of the ESD to the Stieltjes transform of F_c .

Lemma 2.1 ([8, Theorem B.9]). *Assume that $\{G_n\}$ is a sequence of functions of bounded variation and $G_n(-\infty) = 0$ for all n . Then*

$$\lim_{n \rightarrow \infty} s_{G_n}(z) = s(z) \quad \forall z \in D$$

if and only if there is a function of bounded variation G with $G(-\infty) = 0$ and Stieltjes transform $s(z)$ and such that $G_n \rightarrow G$ vaguely.

We will also use the following lemma in order to establish the rate of convergence in Theorem 1.6.

Lemma 2.2 ([8, Theorem B.14]). *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 $s_F(z)$ and $s_G(z)$ respectively, where $z = u + iv \in D$. Then*

$$\begin{aligned} \|F - G\| \leq & \frac{1}{\pi(1-\xi)(2\rho-1)} \left(\int_{-A}^A |s_F(z) - s_G(z)| du \right. \\ & + 2\pi v^{-1} \int_{|x|>B} |F(x) - G(x)| dx \\ & \left. + v^{-1} \sup_x \int_{|u| \leq 2va} |G(x+u) - G(x)| du \right), \end{aligned}$$

where the constants $A > B > 0$, ξ , and a are restricted by $\rho = \frac{1}{\pi} \int_{|u| \leq a} \frac{1}{u^2+1} du > \frac{1}{2}$, and $\xi = \frac{4B}{\pi(A-B)(2\rho-1)} \in (0, 1)$.

3 Proof of Theorems 1.5 and 1.6

Let $\{A_n\}_{n \geq 1}$ be a sequence of real random matrices that obeys condition **C0** and assume $c_n := N/n \rightarrow c \in (0, \infty)$. We begin by introducing some notation. Let

$$s_n(z) := \frac{1}{n} \operatorname{Tr} \left(\frac{1}{n} A_n A_n^T - z I_n \right)^{-1} = \frac{1}{n} \operatorname{Tr} \left(\frac{1}{n} A_n A_n^T - z \right)^{-1}$$

where I_n is the identity matrix of order n and $z = u + iv$. Fix $\alpha > 0$ and let

$$D_{\alpha, n} := \{z = u + iv \in \mathbb{C} : |u| \leq \alpha, v_n \leq v \leq 1\}$$

where v_n is a sequence we will choose later such that $0 < v_n < 1$ for all n . We will eventually allow the sequence v_n to approach zero as n tends to infinity.

We will use the following lemma to prove Theorems 1.5 and 1.6.

Lemma 3.1. *Suppose $\{A_n\}_{n \geq 1}$ is a sequence of real random matrices that obey condition **C0** and assume $c_n := N/n \rightarrow c \in (0, \infty)$. Then for any $\alpha > 0$*

$$\sup_{z \in D_{\alpha, n}} \left| \mathbb{E} s_n(z) - \frac{1}{c_n - 1 - z - z \mathbb{E} s_n(z)} \right| = O \left(\frac{\sqrt{q_n} + \sqrt{\gamma_n} + \sqrt{\rho_n}}{v_n^3} + \frac{\beta_n + 1}{\sqrt{n} v_n^3} \right).$$

We prove Lemma 3.1 in Section 4. For the moment, assume this lemma. First, take $v_n > 0$ to be fixed. Then $D_{\alpha, n}$ does not change with n . Since $c_n \rightarrow c$, we obtain that

$$\mathbb{E} s_n(z) = \frac{1}{c - 1 - z - z \mathbb{E} s_n(z)} + o(1)$$

for all $z \in D_{\alpha, n}$. Fix $z_0 = u_0 + iv_0 \in D_{\alpha, n}$. Since $|\mathbb{E} s_n(z_0)| \leq \frac{1}{v_0}$, one can use a compactness argument to obtain a convergent subsequence $\mathbb{E} s_{n_k}(z_0) \rightarrow s(z_0)$. Then $s(z_0)$ must satisfy the equation

$$s(z_0) = \frac{1}{c - 1 - z_0 - z_0 s(z_0)}.$$

Also, since the eigenvalues of $A_n A_n^T$ are non-negative, $\operatorname{Im}(z \mathbb{E} s_n(z)) \geq 0$ for all $\operatorname{Im}(z) > 0$ and hence $\operatorname{Im}(z_0 s(z_0)) \geq 0$. Thus, by the characterization (2.1), it follows that $s(z_0) = m_c(z_0)$. Since every convergent subsequence of $\{\mathbb{E} s_n(z_0)\}$ must converge to the same limit, we obtain that

$$\lim_{n \rightarrow \infty} \mathbb{E} s_n(z_0) = m_c(z_0)$$

and since $z_0 \in D_{\alpha, n}$ was arbitrary, one obtains

$$\lim_{n \rightarrow \infty} \mathbb{E} s_n(z) = m_c(z) \tag{3.1}$$

for all $z \in D_{\alpha, n}$. Finally, since $|\mathbb{E} s_n(z)| \leq \frac{1}{v}$, Vitali's Convergence Theorem implies that (3.1) holds for all $z \in D$. Therefore,

$$\|\mathbb{E} F \frac{1}{n} A_n A_n^T - F_c\| := \sup_x |\mathbb{E} F \frac{1}{n} A_n A_n^T - F_c| \rightarrow 0$$

as $n \rightarrow \infty$.

To obtain the almost sure convergence in Theorem 1.5, one repeats the argument above and then applies the Borel-Cantelli lemma, since

$$\mathbb{P}(|s_n(z) - \mathbb{E} s_n(z)| \geq \epsilon) \leq \frac{C_p (\beta_n + 1)^p}{n^{p/2} v_n^p \epsilon^p}$$

by Lemma A.1 from Appendix A.

To prove Theorem 1.6, we will apply Lemma 2.2. Under assumption (1.2), there exists $B > 0$ such that

$$\mathbb{E}F_n^{\frac{1}{n}A_nA_n^T}(x) - F_{c_n}(x) = 0$$

for all $|x| > B$ and n sufficiently large.

By [8, Lemma 8.15], it follows that

$$\sup_x \int_{|u| < v_n} |F_{c_n}(x+u) - F_{c_n}(x)| = O\left(v_n^{3/2}\right)$$

for $c_n \geq 1$.

From Lemma 3.1, we have that

$$\mathbb{E}s_n(z) = \frac{1}{c_n - 1 - z - z\mathbb{E}s_n(z)} + \delta_n$$

for all $z \in D_{\alpha,n}$. Thus

$$z(\mathbb{E}s_n(z))^2 + \mathbb{E}s_n(z)(z + 1 - c_n) + 1 = \delta_n(1 - c_n - z - z\mathbb{E}s_n(z)).$$

By subtracting the quadratic equation for $m_{c_n}(z)$ obtained from (2.1), one finds that

$$|\mathbb{E}s_n(z) - m_{c_n}(z)| = \frac{|\delta_n||1 - c_n + z + z\mathbb{E}s_n(z)|}{|z\mathbb{E}s_n(z) + zm_{c_n}(z) + z + 1 - c_n|} = O\left(\frac{\delta_n}{v_n^2}\right).$$

Therefore, from Lemma 2.2, one obtains that

$$\|\mathbb{E}F_n^{\frac{1}{n}A_nA_n^T} - F_{c_n}\| = O\left(\frac{\sqrt{q_n} + \sqrt{\gamma_n} + \sqrt{\rho_n}}{v_n^5} + \frac{\beta_n + 1}{\sqrt{nv_n^5}} + \sqrt{v_n}\right)$$

and hence we can take

$$v_n = \max\left(q_n^{1/11}, \gamma_n^{1/11}, \rho_n^{1/11}, \left(\frac{(\beta_n + 1)^2}{n}\right)^{1/11}\right).$$

The proof of Theorem 1.6 is complete.

It only remains to prove Lemma 3.1.

4 Proof of Lemma 3.1

Let $\{A_n\}_{n \geq 1}$ be a sequence of real random matrices that obey condition **C0** and assume $c_n := N/n \rightarrow c \in (0, \infty)$. Fix $\alpha > 0$. In order to simplify notation, we drop the superscript (n) and write ζ_{ij} and r_k for the entries of A_n and the rows of A_n respectively. We define the resolvent

$$R_n(z) := \left(\frac{1}{n}A_nA_n^T - z\right)^{-1}.$$

Using the Schur complement, we obtain that

$$(R_n(z))_{kk} = \frac{1}{\frac{1}{n}|r_k|^2 - z - \frac{1}{n}r_kA_{n,k}^TR_{n,k}(z)A_{n,k}r_k^T} =: \frac{1}{a_k}$$

where $A_{n,k}$ is obtained from the matrix A_n by removing the k -th row and

$$R_{n,k} = \left(\frac{1}{n}A_{n,k}A_{n,k}^T - z\right)^{-1}.$$

Since $|(R_n(z))_{kk}| \leq \|R_n(z)\| \leq \frac{1}{v_n}$, we obtain that $|a_k| \geq v_n$. Thus,

$$\left| \mathbb{E}s_n(z) - \frac{1}{n} \sum_{k=1}^n \frac{1}{\mathbb{E}a_k} \right| \leq \frac{1}{nv_n^2} \sum_{k=1}^n \mathbb{E}|a_k - \mathbb{E}a_k|. \quad (4.1)$$

We now compute the expectation of a_k . By condition (ii) in Definition 1.1, we have that

$$\sup_k \left| \mathbb{E} \frac{1}{n} |r_k|^2 - c_n \right| \leq q_n.$$

For convenience, write

$$B_{n,k} := A_{n,k}^\top R_{n,k}(z) A_{n,k}.$$

We first note that $\sup_k \|B_{n,k}\| = O(v_n^{-1})$. Indeed, since $|z|^2 \leq \alpha^2 + 1 = O(1)$,

$$\|B_{n,k}\| = \|R_{n,k}(z) A_{n,k} A_{n,k}^\top\| = \|I_{n-1} + zR_{n,k}(z)\| \leq 1 + \frac{|z|}{v_n}$$

for all $k = 1, \dots, n$. Then we have that

$$\mathbb{E} \frac{1}{n} r_k B_{n,k} r_k^\top = \frac{1}{n} \sum_{i,j=1}^N \mathbb{E} [\mathbb{E}_k [\zeta_{ki} \zeta_{kj}] (B_{n,k})_{ij}] = \frac{1}{n} \mathbb{E} \text{Tr} B_{n,k} + \epsilon_{n,k} + O\left(\frac{q_n}{v_n}\right) \quad (4.2)$$

uniformly for all k (by condition (ii) in Definition 1.1) where

$$\epsilon_{n,k} := \frac{1}{n} \sum_{i \neq j} \mathbb{E} [\mathbb{E}_k [\zeta_{ki} \zeta_{kj}] (B_{n,k})_{ij}].$$

By condition (iii), we have that

$$\begin{aligned} |\epsilon_{n,k}| &\leq \left(\mathbb{E} \frac{\gamma_n^2}{n^3} \sum_{i,j,s,t=1}^N |(B_{n,k})_{ij}| |(B_{n,k})_{st}| \right)^{1/2} \\ &\leq \left(\mathbb{E} \frac{2\gamma_n^2}{n} \text{Tr}(B_{n,k} B_{n,k}^*) \right)^{1/2} = O(v_n^{-1} \gamma_n) \end{aligned}$$

uniformly in k .

Combining the above yields,

$$\sup_k \left| \mathbb{E}a_k - \left(c_n - z - \mathbb{E} \frac{1}{n} \text{Tr} B_{n,k} \right) \right| = O\left(\frac{q_n + \gamma_n}{v_n}\right) \quad (4.3)$$

We now note that $\frac{1}{n} \text{Tr} B_{n,k} = \frac{n-1}{n} + z \frac{1}{n} \text{Tr} R_{n,k}(z)$. By equation (3.11) in [7] (or alternatively, by Cauchy's Interlacing Theorem), one finds that

$$\left| \frac{1}{n} \text{Tr} R_{n,k}(z) - \frac{1}{n} \text{Tr} R_n(z) \right| = O\left(\frac{1}{nv_n}\right) \quad (4.4)$$

uniformly in k . Therefore, from (4.3) and the fact that $\frac{1}{n} \text{Tr} R_n(z) = s_n(z)$, we obtain that

$$\sup_k |\mathbb{E}a_k - (c_n - 1 - z - z \mathbb{E}s_n(z))| = O\left(\frac{q_n + \gamma_n}{v_n} + \frac{1}{nv_n}\right) \quad (4.5)$$

We now turn our attention to obtaining a bound for $\mathbb{E}|a_k - \mathbb{E}a_k|$. First we note that

$$\mathbb{E} \left| \frac{1}{n} |r_k|^2 - c_n \right| \leq q_n$$

by condition (ii) of Definition 1.1. Using (4.2) and the bounds obtained above for $\epsilon_{n,k}$, we have that

$$\mathbb{E} \left| \frac{1}{n} r_k B_{n,k} r_k^T - \mathbb{E} \frac{1}{n} \text{Tr} B_{n,k} \right|^2 = \frac{1}{n^2} \sum_{i,j,s,t=1}^N \mathbb{E} [\zeta_{ki} \zeta_{kj} \zeta_{ks} \zeta_{kt} (B_{n,k})_{ij} (B_{n,k})_{st}] - \left(\mathbb{E} \frac{1}{n} \text{Tr} B_{n,k} \right)^2 + O \left(\frac{q_n + \gamma_n}{v_n^2} \right).$$

For the sum

$$\frac{1}{n^2} \sum_{i,j,s,t=1}^N \mathbb{E} [\zeta_{ki} \zeta_{kj} \zeta_{ks} \zeta_{kt} (B_{n,k})_{ij} (B_{n,k})_{st}].$$

we consider several cases:

(a) When we sum over all i, j, s, t distinct, one finds

$$\frac{1}{n^2} \sum \mathbb{E} [\zeta_{ki} \zeta_{kj} \zeta_{ks} \zeta_{kt} (B_{n,k})_{ij} (B_{n,k})_{st}] = O \left(\frac{\gamma_n}{v_n^2} \right)$$

by condition (iv).

(b) When we sum of all $i = j, s = t$, we obtain

$$\frac{1}{n^2} \sum \mathbb{E} [\zeta_{ki} \zeta_{kj} \zeta_{ks} \zeta_{kt} (B_{n,k})_{ij} (B_{n,k})_{st}] = \mathbb{E} \left[\left(\frac{1}{n} \text{Tr} B_{n,k} \right)^2 \right] + O \left(\frac{\rho_n}{v_n^2} \right)$$

by condition (vi).

(c) When we sum over all $i = s, j = t$ (or $i = t, j = s$), we have

$$\begin{aligned} \frac{1}{n^2} \sum \mathbb{E} [\zeta_{ki} \zeta_{kj} \zeta_{ks} \zeta_{kt} (B_{n,k})_{ij} (B_{n,k})_{st}] &= \mathbb{E} \left[\frac{1}{n^2} \text{Tr} (B_{n,k} B_{n,k}^*) \right] + O \left(\frac{\rho_n}{v_n^2} \right) \\ &= O \left(\frac{\rho_n}{v_n^2} + \frac{1}{n v_n^2} \right) \end{aligned}$$

by condition (vi).

(d) When $i = s, j \neq t$ (or $i = t, j \neq s$), one finds that

$$\frac{1}{n^2} \sum \mathbb{E} [\zeta_{ki} \zeta_{kj} \zeta_{ks} \zeta_{kt} (B_{n,k})_{ij} (B_{n,k})_{st}] = O \left(\frac{\gamma_n}{\sqrt{n} v_n^2} \right)$$

by condition (iii).

(e) When $i = j, s \neq t$, we find

$$\begin{aligned} \frac{1}{n^2} \sum \mathbb{E} [\zeta_{ki} \zeta_{kj} \zeta_{ks} \zeta_{kt} (B_{n,k})_{ij} (B_{n,k})_{st}] &\leq \frac{\gamma_n}{n^{3/2} v_n} \mathbb{E} \left(\sum_{s,t} |(B_{n,k})_{st}| \right) \\ &\leq \frac{\gamma_n}{n^{3/2} v_n} \mathbb{E} \left(\sum_{s,t,l,m} |(B_{n,k})_{st}| |(B_{n,k})_{lm}| \right)^{1/2} \\ &\leq \frac{\gamma_n}{n^{3/2} v_n} n \mathbb{E} (2 \text{Tr} (B_{n,k} B_{n,k}^*))^{1/2} = O \left(\frac{\gamma_n}{v_n^2} \right) \end{aligned}$$

by Cauchy-Schwarz and condition (iii).

Therefore, we obtain that

$$\mathbb{E} \left| \frac{1}{n} r_k B_{n,k} r_k^T - \mathbb{E} \frac{1}{n} \text{Tr} B_{n,k} \right|^2 = \text{Var} \left(\frac{1}{n} \text{Tr} B_{n,k} \right) + O \left(\frac{q_n + \gamma_n + \rho_n}{v_n^2} + \frac{1}{n v_n^2} \right). \quad (4.6)$$

The bound in (4.6) holds uniformly in k since the bounds in conditions (iii), (iv), and (vi) of Definition 1.1 hold uniformly in k .

By Lemma A.1 in Appendix A, we have that

$$\sup_k \text{Var} \left(\frac{1}{n} \text{Tr} B_{n,k} \right) = O \left(\frac{(\beta_n + 1)^2}{n v_n^2} \right)$$

and hence

$$\mathbb{E} \left| \frac{1}{n} r_k B_{n,k} r_k^T - \mathbb{E} \frac{1}{n} \text{Tr} B_{n,k} \right|^2 = O \left(\frac{q_n + \gamma_n + \rho_n}{v_n^2} + \frac{(\beta_n + 1)^2}{n v_n^2} \right). \quad (4.7)$$

Therefore

$$\sup_k \mathbb{E} |a_k - \mathbb{E} a_k| = O \left(\frac{\sqrt{q_n} + \sqrt{\gamma_n} + \sqrt{\rho_n}}{v_n} + \frac{\beta_n + 1}{\sqrt{n} v_n} \right). \quad (4.8)$$

One can also observe that $\text{Im}(z s_n(z)) \geq 0$ for all z with $\text{Im} z > 0$, since the eigenvalues of $A_n A_n^T$ are non-negative. Combining this fact with equations (4.1) and (4.3) and the estimates above, we obtain that

$$\left| \mathbb{E} s_n(z) - \frac{1}{1 - c_n - z - z s_n(z)} \right| = O \left(\frac{\sqrt{q_n} + \sqrt{\gamma_n} + \sqrt{\rho_n}}{v_n^3} + \frac{\beta_n + 1}{\sqrt{n} v_n^3} \right)$$

where the bound holds uniformly for $z \in D_{\alpha,n}$. The proof of Lemma 3.1 is complete.

Remark 4.1. As noted in Remark 1.7, it is possible to show that if the sequence $\{A_n\}_{n \geq 1}$ satisfies conditions (i) - (vi) from Definition 1.1 with $c_n \rightarrow c \in (0, \infty)$, then

$$\| \mathbb{E} F_n^{\frac{1}{n}} A_n A_n^T - F_c \| \rightarrow 0 \quad (4.9)$$

as $n \rightarrow \infty$. The proof of the above statement repeats the proof of Theorem 1.5 almost exactly; we now detail the necessary changes.

Since the Stieltjes transform is an analytic and bounded function, it suffices to prove the convergence of $\mathbb{E} s_n(z)$ to $m_c(z)$ for all z in a compact set in the upper-half plane with $\text{Im}(z) \geq \kappa$ for a sufficiently large constant κ to be chosen later.

A careful reading of the proof of Lemma 3.1 reveals that condition (vii) from Definition 1.1 is only used to invoke Lemma A.1 and obtain the variance bound (A.2). Thus, in order to prove (4.9), it suffices to show that

$$\text{Var} \left(\frac{1}{n} \text{Tr} R_n(z) \right) = o(1) \quad (4.10)$$

for all z in a compact set in the upper-half plane with $\text{Im}(z) \geq \kappa$.

We decompose

$$\begin{aligned} (R_n(z))_{kk} &= \frac{1}{\frac{1}{n} |r_k|^2 - z - \frac{1}{n} r_k A_{n,k}^T R_{n,k}(z) A_{n,k} r_k^T} \\ &= \frac{1}{c_n - 1 - z - z \mathbb{E} s_n(z) - \epsilon_k} \end{aligned}$$

where

$$\epsilon_k = c_n - 1 - z \mathbb{E} s_n(z) - \frac{1}{n} |r_k|^2 + \frac{1}{n} r_k A_{n,k}^T R_{n,k}(z) A_{n,k} r_k^T.$$

Thus

$$\begin{aligned} (R_n(z))_{kk} &= \frac{1}{c_n - 1 - z - z\mathbb{E}s_n(z)} \\ &\quad + \frac{\epsilon_k}{(c_n - 1 - z - z\mathbb{E}s_n(z))((c_n - 1 - z - z\mathbb{E}s_n(z)) - \epsilon_k)} \\ &= \frac{1}{c_n - 1 - z - z\mathbb{E}s_n(z)} [1 + (R_n(z))_{kk}\epsilon_k]. \end{aligned}$$

Taking $\text{Im}(z) \geq \kappa$ we obtain

$$\text{Var} \left(\frac{1}{n} \text{Tr} R_n(z) \right) \leq \frac{C}{\kappa^4 n} \sum_{k=1}^n \mathbb{E} |\epsilon_k|^2$$

for some absolute constant $C > 0$. Using condition (vi) from Definition 1.1, (4.4), and (4.6), we bound $\mathbb{E} |\epsilon_k|^2$ and obtain

$$\text{Var} \left(\frac{1}{n} \text{Tr} R_n(z) \right) \leq \frac{C|z|^2}{\kappa^4} \text{Var} \left(\frac{1}{n} \text{Tr} R_n(z) \right) + O \left(q_n + \gamma_n + \rho_n + \frac{1}{\sqrt{n}} \right).$$

Taking z in a compact set for which $|z|^2/\kappa^4$ is sufficiently small verifies (4.10).

A Estimate of Variance of Stieltjes Transform

Lemma A.1. Let A_n be an $n \times N$ real random matrix with rows $r_1^{(n)}, \dots, r_n^{(n)}$ that satisfy condition (vii) from Definition 1.1. Then for every $p > 1$ there exists a constant $C_p > 0$ (depending only on p) such that

$$\mathbb{E} \left| \frac{1}{n} \text{Tr} \left(\frac{1}{n} A_n A_n^\top - z I_n \right)^{-1} - \frac{1}{n} \mathbb{E} \text{Tr} \left(\frac{1}{n} A_n A_n^\top - z I_n \right)^{-1} \right|^p \leq \frac{C_p (\beta_n + 1)^p}{n^{p/2} |\text{Im } z|^p} \quad (\text{A.1})$$

for any z with $\text{Im } z \neq 0$. In particular, there exists an absolute constant $C > 0$ such that

$$\text{Var} \left[\frac{1}{n} \text{Tr} \left(\frac{1}{n} A_n A_n^\top - z I_n \right)^{-1} \right] \leq \frac{C (\beta_n + 1)^2}{n |\text{Im } z|^2} \quad (\text{A.2})$$

for any z with $\text{Im } z \neq 0$.

Proof. Since (A.1) implies (A.2) when $p = 2$, it suffices to prove (A.1) for arbitrary $p > 1$. Let $\mathbb{E}_{\leq k}$ denote the conditional expectation with respect to the σ -algebra generated by $r_1^{(n)}, \dots, r_k^{(n)}$. Define

$$R_n(z) := \left(\frac{1}{n} A_n A_n^\top - z I_n \right)^{-1}$$

and

$$Y_k := \mathbb{E}_{\leq k} \frac{1}{n} \text{Tr} R_n(z)$$

for $k = 0, 1, \dots, n$. Then $\{Y_k\}_{k=0}^n$ is a martingale since $\mathbb{E}_{\leq k} Y_{k+1} = Y_k$. Define the martingale difference sequence

$$\alpha_k := Y_k - Y_{k-1}$$

for $k = 1, 2, \dots, n$. We then note that

$$\sum_{k=1}^n \alpha_k = \frac{1}{n} \text{Tr} R_n(z) - \mathbb{E} \frac{1}{n} \text{Tr} R_n(z). \quad (\text{A.3})$$

We will bound the p -th moment of the sum in (A.3), but first we obtain a bound on the individual summands α_k .

For each $1 \leq k \leq n$, define the set

$$J(k) := \{1 \leq j \leq n : k \leq j \leq k + \beta_n\}.$$

Now let $A_{n,J(k)}$ be obtained from the matrix A_n by removing row j if and only if $j \in J(k)$. Let

$$R_{n,J(k)}(z) = \left(\frac{1}{n} A_{n,J(k)} A_{n,J(k)}^T - z I_n \right)^{-1}.$$

A simple computation (see for instance [12, Example 5.1.5]) reveals that $\mathbb{E}_{\leq k} \text{Tr} R_{n,J(k)}(z) = \mathbb{E}_{\leq k-1} \text{Tr} R_{n,J(k)}(z)$ by condition (vii) of Definition 1.1. Thus

$$\alpha_k = \mathbb{E}_{\leq k} \frac{1}{n} (\text{Tr} R_n(z) - \text{Tr} R_{n,J(k)}(z)) - \mathbb{E}_{\leq k-1} \frac{1}{n} (\text{Tr} R_n(z) - \text{Tr} R_{n,J(k)}(z)).$$

Using the triangle inequality and equation (3.11) in [7], we have that

$$|\text{Tr} R_n(z) - \text{Tr} R_{n,J(k)}(z)| \leq \frac{\beta_n + 1}{|\text{Im } z|}$$

and hence

$$|\alpha_k| \leq \frac{2(\beta_n + 1)}{n |\text{Im } z|}.$$

We now apply the Burkholder inequality (see for example [8, Lemma 2.12] for a complex-valued version of the Burkholder inequality) and obtain that there exists a constant $C_p > 0$ such that

$$\begin{aligned} \mathbb{E} \left| \sum_{k=1}^n \alpha_k \right|^p &\leq C_p \mathbb{E} \left(\sum_{k=1}^n |\alpha_k|^2 \right)^{p/2} \\ &\leq C_p \left(\frac{4(\beta_n + 1)^2 n}{n^2 |\text{Im } z|^2} \right)^{p/2} \\ &\leq \frac{C_p 2^p (\beta_n + 1)^p}{n^{p/2} |\text{Im } z|^p}. \end{aligned}$$

□

References

- [1] Adamczak, R., *On the Marchenko-Pastur and circular laws for some classes of random matrices with dependent entries*, Electronic Journal of Probability, Vol. 16 (2011). MR-2820070
- [2] Adamczak, R.; Litvak, A.; Pajor, A.; Tomczak-Jaegermann, N., *Quantitative estimates of the convergence of the empirical covariance matrix in log-concave ensembles*, J. Amer. Math. Soc. 23 (2010), 535-561. MR-2601042
- [3] Anderson, G.W., Guionnet, A., and Zeitouni, O. *An Introduction to Random Matrices*, Cambridge Studies in Advanced Mathematics 118, Cambridge University Press, New York, 2010. MR-2760897
- [4] Aubrun, G., *Random points in the unit ball of l_p^n* , Positivity, 10(4):755-759, 2006. MR-2280648
- [5] Bai, Z. D. *Methodologies in spectral analysis of large-dimensional random matrices, a review*. Statist. Sinica 9, 611-677 (1999). MR-1711663
- [6] Bai, Z. D., Hu, J., Pan, G., Zhou, W., *A Note on Rate of Convergence in Probability to Semi-circular Law*, Electronic Journal of Probability, Vol. 16(2011). MR-2861680

- [7] Bai, Z.-D.; Convergence Rate of Expected Spectral Distributions of Large Random Matrices. Part I. Wigner Matrices, *Ann. Probab.* Volume 21, Number 2 (1993), 625–648. MR-1217559
- [8] Bai, Z.-D.; Silverstein, J.; Spectral analysis of large dimensional random matrices, *Mathematics Monograph Series 2*, Science Press, Beijing 2006.
- [9] Bordenave, B., Caputo, P., Chafaï, D., *Circular law theorem for random Markov matrices*, *Probability Theory and Related Fields* 152, 3-4 (2012) 751-779
- [10] Chafaï, D., *The Dirichlet Markov Ensemble*, *J. Multivariate Anal.* 101 (2010), no. 3, 555-567. MR-2575404
- [11] Chatterjee, S., Diaconis P., Sly A., *Properties of Uniform Doubly Stochastic Matrices*, arXiv:1010.6136v1 [math.PR]
- [12] Durrett R., *Probability. Theory and Examples*, 4th ed., Cambridge University Press, New York, 2010. MR-2722836
- [13] Götze, F.; Tikhomirov, A.; *Limit theorems for spectra of positive random matrices under dependence*, *Zap. Nauchn. Sem. S.-Petersburg. Otdel. Mat. Inst. Steklov. (POMI)*, Vol. 311 (2004), Veroyatn. i Stat.7, 92-123, 299. MR-2092202
- [14] Götze, F.; Tikhomirov, A.; *On the Rate of Convergence to the Marchenko–Pastur Distribution*, arXiv:1110.1284v1 [math.PR].
- [15] Götze, F.; Tikhomirov, A.; *Limit Theorems for spectra of random matrices with martingale structure*, *Stein’s Method and Applications*, Singapore Univ. Press (2005), pp. 181-195. MR-2205336
- [16] Horn, R.; Johnson, Ch.; *Topics in Matrix analysis*, Cambridge University Press, 1991. MR-1091716
- [17] V. Marchenko, L. Pastur, *Distribution of eigenvalues of some sets of random matrices*, *Math USSR-Sb.* **1**, (1967), 457–486.
- [18] M.L. Mehta, *Random Matrices*, 3rd ed., Elsevier/Academic Press, Amsterdam, 2004. MR-2129906
- [19] Mendelson, S.; Pajor, A. *On singular values of matrices with independent rows*, *Bernoulli* 12(5), 2006, 761-773. MR-2265341
- [20] Pajor, A.; Pastur, L., *On the Limiting Empirical Measure of the sum of rank one matrices with log-concave distribution*, *Studia Math.* 195 (2009) MR-2539559
- [21] Wachter, K. W., *The strong limits of random matrix spectra for sample matrices of independent elements*, *Ann. Probab.* 6, 1–18 (1978). MR-0467894
- [22] Y. Q. Yin, P. R. Krishnaiah, *Limit Theorem for the Eigenvalues of the Sample Covariance Matrix when the Underlying Distribution is Isotropic*, *Theory Probab. Appl.* 30, pp. 861-867. MR-0816299

Acknowledgments. The author is grateful to A. Litvak for pointing out Theorem 3.13 in [2]. The author would also like to thank A. Soshnikov and D. Renfrew for useful conversations and comments.