Products of independent non-Hermitian random matrices

Sean O'Rourke [*]   Alexander Soshnikov [†]

Abstract

We consider the product of a finite number of non-Hermitian random matrices with i.i.d. centered entries of growing size. We assume that the entries have a finite moment of order bigger than two. We show that the empirical spectral distribution of the properly normalized product converges, almost surely, to a non-random, rotationally invariant distribution with compact support in the complex plane. The limiting distribution is a power of the circular law.

Key words: Random matrices, Circular law.

AMS 2010 Subject Classification: Primary 60B20.

---

1 Introduction and Formulation of Results

Many important results in random matrix theory pertain to Hermitian random matrices. Two powerful tools used in this area are the moment method and the Stieltjes transform. Unfortunately, these two techniques are not suitable for dealing with non-Hermitian random matrices, [6].

1.1 The Circular Law

One of the fundamental results in the study of non-Hermitian random matrices is the circular law. We begin by defining the empirical spectral distribution (ESD).

Definition 1. Let X be a matrix of order n and let $\lambda_1, \ldots, \lambda_n$ be the eigenvalues of X. Then the *empirical spectral distribution* (ESD) μ_X of X is defined as

$$\mu_X(z, \bar{z}) = \frac{1}{n} \# \left\{ k \leq n : \operatorname{Re}\left(\lambda_k\right) \leq \operatorname{Re}(z); \operatorname{Im}\left(\lambda_k\right) \leq \operatorname{Im}(z) \right\}.$$

Let ξ be a complex random variable with finite non-zero variance σ^2 and let N_n be a random matrix of order n with entires being i.i.d. copies of ξ. We say that the circular law holds for ξ if, with probability 1, the ESD $\mu_{\frac{1}{\sigma\sqrt{n}}N_n}$ of $\frac{1}{\sigma\sqrt{n}}N_n$ converges (uniformly) to the uniform distribution over the unit disk as n tends to infinity.

The circular law was conjectured in the 1950's as a non-Hermitian counterpart to Wigner's semi-circle law. The circular law was first shown by Mehta in 1967 [22] when ξ is complex Gaussian. Mehta relied upon the joint density of the eigenvalues which was discovered by Ginibre [10] two years earlier.

Building on the work of Girko [11], Bai proved the circular law under the conditions that ξ has finite sixth moment and that the joint distribution of the real and imaginary parts of ξ has bounded density, [3]. In [6], the sixth moment assumption was weakened to $\mathbb{E}|\xi|^{2+\eta}$ for any specified $\eta > 0$, but the bounded density assumption still remained. Götze and Tikhomirov ([15]) proved the circular law in the case of i.i.d. sub-Gaussian matrix entries. Pan and Zhou proved the circular law for any distribution ξ with finite fourth moment [25] by building on [15] and utilizing the work of Rudelson and Vershynin in [27]. In an important development, Götze and Tikhomirov showed in [14] that the expected spectral distribution $\mathbb{E}\mu_{N_n}$ converges to the uniform distribution over the unit disk as n tends to infinity assuming that $\sup_{jk} \mathbb{E}|(N_n)_{jk}|^2 \phi((N_n)_{jk}) < \infty$, where $\phi(x) = (\ln(1+|x|))^{19+\eta}$, $\eta > 0$. In [28], Tao and Vu proved the circular law assuming a bounded $(2+\eta)^{\text{th}}$ moment, for any fixed $\eta > 0$. Finally, Tao and Vu have been able to remove the extra η in the moment condition. Namely, they proved the circular law in [29] assuming only that the second moment is bounded.

1.2 Main Results

In this paper, we study the ESD of the product

$$X^{(n)} = X_1^{(n)} X_2^{(n)} \cdots X_m^{(n)}$$

of m independent $n \times n$ non-Hermitian random matrices as n tends to infinity. Burda, Janik, and Waclaw [8] studied the mathematical expectation of the limiting ESD, $\lim_{n \to \infty} \mathbb{E}\mu_X^{(n)}$, in the case

that the entries of the matrices are Gaussian. Here we extend their results by proving the almost sure convergence of the ESD, $\mu_X^{(n)}$, for a class of non-Gaussian random matrices. Namely, we require that the entries of $X_i^{(n)}$, $i = 1, \ldots, m$, are i.i.d. random variables with a finite moment of order $2 + \eta$, $\eta > 0$.

Theorem 2. *Fix $m > 1$ and let ξ be a complex random variable with variance 1 such that $\text{Re}(\xi)$ and $\text{Im}(\xi)$ are independent each with mean zero and $\mathbb{E}|\xi|^{2+\eta} < \infty$ for some $\eta > 0$. Let $X_1^{(n)}, \ldots, X_m^{(n)}$ be independent random matrices of order n where the entries of $X_j^{(n)}$ are i.i.d. copies of $\sigma_j \frac{\xi}{\sqrt{n}}$ for some collection of positive constants $\sigma_1, \ldots, \sigma_m$. Then the ESD $\mu_X^{(n)}$ of $X^{(n)} = X_1^{(n)} X_2^{(n)} \cdots X_m^{(n)}$ converges, with probability 1, as $n \to \infty$ to the distribution whose density is given by*

$$\rho(z, \bar{z}) = \begin{cases} \frac{1}{m\pi} \sigma^{-\frac{2}{m}} |z|^{\frac{2}{m}-2} & \text{for } |z| \leq \sigma, \\ 0 & \text{for } |z| > \sigma, \end{cases} \tag{1}$$

where $\sigma = \sigma_1 \cdots \sigma_m$.

Remark 3. The almost sure convergence of $\mu_X^{(n)}$ implies the convergence of $\mathbb{E}\mu_X^{(n)}$ as well.

Remark 4. We refer the reader to [4] for bounds on powers of a square random matrix with i.i.d. entries. See also [1], [2], [9], [5], [7], and [24] for some other results on the spectral properties of products of random matrices.

2 Notation and Setup

The proof of Theorem 2 is divided into two parts and presented in Sections 3 and 4.

We note that without loss of generality, we may assume $\sigma_1 = \sigma_2 = \cdots = \sigma_m = 1$. Indeed, the spectrum for arbitrary $\sigma_1, \ldots, \sigma_m$ can be obtained by a trivial rescaling. Following Burda, Janik, and Waclaw in [8], we let $Y^{(n)}$ be a $(mn) \times (mn)$ matrix defined as

$$Y^{(n)} = \begin{pmatrix} 0 & X_1^{(n)} & & & 0 \\ 0 & 0 & X_2^{(n)} & & 0 \\ & & \ddots & \ddots & \\ 0 & & & 0 & X_{m-1}^{(n)} \\ X_m^{(n)} & & & & 0 \end{pmatrix}. \tag{2}$$

Section 4 will be devoted to proving that the ESD of $Y^{(n)}$ obeys the circular law as n tends to infinity. This statement is presented in the following Lemma.

Lemma 5 ($Y^{(n)}$ obeys the circular law)**.** *The ESD $\mu_{Y^{(n)}}$ of $Y^{(n)}$ converges, with probability 1, to the uniform distribution over the unit disk as $n \to \infty$.*

3 Proof of Theorem 2

With Lemma 5 above, we are ready to prove Theorem 2.

2221

Proof of Theorems 2. Using the definition of $Y^{(n)}$ in (2), we can compute

$$\left(Y^{(n)}\right)^m = \begin{pmatrix} Y_1 & & & 0 \\ & Y_2 & & \\ & & \ddots & \\ 0 & & & Y_m \end{pmatrix},$$

where $Y_k = X_k^{(n)} X_{k+1}^{(n)} \cdots X_m^{(n)} X_1^{(n)} \cdots X_{k-1}^{(n)}$ for $1 \leq k \leq m$. Notice that each Y_k has the same eigenvalues as $X^{(n)}$. Let $\lambda_1, \ldots, \lambda_n$ denote the eigenvalues of $X^{(n)}$ and let $\eta_1, \ldots, \eta_{mn}$ denote the eigenvalues of $Y^{(n)}$. Then it follows that each λ_k is an eigenvalue of $\left(Y^{(n)}\right)^m$ with multiplicity m.

Let $f : \mathbb{C} \to \mathbb{C}$ be a continuous, bounded function. Then we have

$$\int_{\mathbb{C}} f(z) \mathrm{d}\mu_{X^{(n)}}(z, \bar{z}) = \frac{1}{n} \sum_{k=1}^{n} f(\lambda_k) = \frac{1}{mn} \sum_{k=1}^{mn} f(\eta_k^m) = \int_{\mathbb{C}} f(z^m) \mathrm{d}\mu_{Y^{(n)}}(z, \bar{z}).$$

By Lemma 5,

$$\int_{\mathbb{C}} f(z^m) \mathrm{d}\mu_{Y^{(n)}}(z, \bar{z}) \longrightarrow \frac{1}{\pi} \int_{\mathbb{D}} f(z^m) \mathrm{d}z \mathrm{d}\bar{z} \qquad \text{a.s.}$$

as $n \to \infty$ where $\mathbb{D}$ denotes the unit disk in the complex plane. Thus, by the change of variables $z \mapsto z^m$ and $\bar{z} \mapsto \bar{z}^m$ we can write

$$\frac{1}{\pi} \int_{\mathbb{D}} f(z^m) \mathrm{d}z \mathrm{d}\bar{z} = \frac{m}{\pi} \int_{\mathbb{D}} f(z) \frac{1}{m^2} |z|^{\frac{2}{m} - 2} \mathrm{d}z \mathrm{d}\bar{z} = \frac{1}{m\pi} \int_{\mathbb{D}} f(z) |z|^{\frac{2}{m} - 2} \mathrm{d}z \mathrm{d}\bar{z}.$$

where the factor of m out front of the integral corresponds to the fact that the transformation maps the complex plane m times onto itself.

Therefore, we have shown that for all continuous, bounded functions f,

$$\int_{\mathbb{C}} f(z) \mathrm{d}\mu_{X^{(n)}}(z, \bar{z}) \longrightarrow \frac{1}{m\pi} \int_{\mathbb{D}} f(z) |z|^{\frac{2}{m} - 2} \mathrm{d}z \mathrm{d}\bar{z} \qquad \text{a.s.}$$

as $n \to \infty$ and the proof is complete. $\qquad \square$

4 Proof of Lemma 5

In order to prove that the ESD of $Y^{(n)}$ obeys the circular law, we follow the work of Bai in [3], Bai and Silverstein in [6], and use the results developed by Tao and Vu in [28]. To do so, we introduce the following notation. Let μ_n denoted the ESD of $Y^{(n)}$. That is,

$$\mu_n(x, y) = \frac{1}{mn} \# \left\{ k \leq mn : \mathrm{Re}(\lambda_k) \leq x; \mathrm{Im}(\lambda_k) \leq y \right\}$$

where $\lambda_1, \ldots, \lambda_{mn}$ are the eigenvalues of $Y^{(n)}$.

An important idea in the proof is to analyze the Stieltjes transformation $s_n : \mathbb{C} \to \mathbb{C}$ of μ_n defined by

$$s_n(z) = \frac{1}{mn} \sum_{k=1}^{mn} \frac{1}{\lambda_k - z} = \int_{\mathbb{C}} \frac{1}{x + iy - z} \mathrm{d}\mu_n(x, y).$$

Since $s_n(z)$ is analytic everywhere except the poles, the real part determines the eigenvalues. Let $z = s + it$. Then we can write

$$\text{Re}(s_n(z)) = \frac{1}{mn}\sum_{k=1}^{mn} \frac{\text{Re}(\lambda_k) - s}{|\lambda_k - z|^2}$$

$$= -\frac{1}{2mn}\sum_{k=1}^{mn} \frac{\partial}{\partial s} \ln|\lambda_k - z|^2$$

$$= -\frac{1}{2}\frac{\partial}{\partial s}\int_0^\infty \ln x \nu_n(dx, z)$$

where $\nu_n(\cdot, z)$ is the ESD of the Hermitian matrix $H_n = (Y^{(n)} - zI)^*(Y^{(n)} - zI)$. This reduces the task to controlling the distributions ν_n.

The main difficulties arise from the two poles of the log function, at ∞ and 0. We will need to use the bounds developed in [3] and [28] to control the largest singular value and the least singular value of $Y^{(n)} - zI$.

A version of the following lemma was first presented by Girko, [11]. We present a slightly refined version by Bai and Silverstein, [6].

Lemma 6. *For any $uv \neq 0$, we have*

$$c_n(u, v) = \int\int e^{iux+ivy}\mu_n(dx, dy)$$

$$= \frac{u^2 + v^2}{4iu\pi}\int\int \frac{\partial}{\partial s}\left[\int_0^\infty \ln x \nu_n(dx, z)\right]e^{ius+ivt}dtds, \tag{3}$$

where $z = s + it$.

We note that the singular values of $Y^{(n)}$ are the union of the singular values of $X_k^{(n)}$ for $1 \leq k \leq n$. Thus, under the assumptions of Theorem 2, the ESD of $Y^{(n)*}Y^{(n)}$ converges to the Marchenko-Pastur Law (see [20] and [6, Theorem 3.7]). Thus by Lemma 8 it follows that, with probability 1, the family of distributions μ_n is tight. To prove the circular law we will show that the right-hand side of (3) converges to $c(u, v)$, its counterpart generated by the circular law, for all $uv \neq 0$. Several steps of the proof will follow closely the work of Bai in [3] and Bai and Silverstein in [6]. We present an outline of the proof as follows.

1. We reduce the range of integration to a finite rectangle in Section 4.2. We will show that the proof reduces to showing that, for every large $A > 0$ and small $\epsilon > 0$,

$$\int\int_T \left[\frac{\partial}{\partial s}\int_0^\infty \ln x \nu_n(dx, z)\right]e^{ius+ivt}dsdt$$

$$\rightarrow \int\int_T \left[\frac{\partial}{\partial s}\int_0^\infty \ln x \nu(dx, z)\right]e^{ius+ivt}dsdt$$

where $T = \{(s, t) : |s| \leq A, |t| \leq A^3, |\sqrt{s^2 + t^2} - 1| \geq \epsilon\}$ and $\nu(x, z)$ is the limiting spectral distribution of the sequence of matrices $H_n = (Y^{(n)} - zI)^*(Y^{(n)} - zI)$.

2223

2. We characterize the limiting spectrum $v(\cdot, z)$ of $v_n(\cdot, z)$.

3. We establish a convergence rate of $v_n(\cdot, z)$ to $v(\cdot, z)$ uniformly in every bounded region of z.

4. Finally, we show that for a suitably defined sequence ϵ_n, with probability 1,

$$\limsup_{n \to \infty} \int \int_T \left| \int_{\epsilon_n}^{\infty} \ln x (v_n(dx, z) - v(dx, z)) \right| = 0$$

and

$$\lim_{n \to \infty} \int_0^{\epsilon_n} \ln x v_n(dx, z) = 0.$$

4.1 Notation

In this section, we introduce some notation that we will use throughout the paper.

First, we will drop the superscript (n) from the matrices $Y^{(n)}$, $X^{(n)}$, $X_1^{(n)}, \ldots, X_m^{(n)}$ and simply write $Y, X, X_1, \ldots, X_m$.

We write $R = Y - zI$ where I is the identity matrix and $z = s + it \in \mathbb{C}$. We will continue to let $H_n = (Y - zI)^*(Y - zI) = R^*R$ and have $v_n(x, z)$ denote the empirical spectral distribution of H_n for each fixed z.

For a $(mn) \times (mn)$ matrix A, there are m^2 blocks each consisting of a $n \times n$ matrix. We let A_{ab} denote the $n \times n$ matrix in position a, b where $1 \le a, b \le m$. $A_{a,b;i,j}$ then refers to the element $(A_{ab})_{ij}$ where $1 \le i, j \le n$.

Finally, C will be used as some positive constant that may change from line to line.

4.2 Integral Range Reduction

To establish Lemma 5, we need to find the limiting counterpart to

$$g_n(s, t) = \frac{\partial}{\partial s} \int_0^{\infty} \ln x v_n(dx, z).$$

We begin by presenting the following lemmas.

Lemma 7 (Bai-Silverstein [6]). *For all $uv \ne 0$, we have*

$$c(u, v) = \frac{1}{\pi} \int \int_{x^2+y^2 \le 1} e^{iux+ivy} dx dy = \frac{u^2 + v^2}{4iu\pi} \int \left[\int g(s, t) e^{ius+ivt} dt \right] ds,$$

where

$$g(s, t) = \begin{cases} \frac{2s}{s^2+t^2}, & \text{if } s^2 + t^2 > 1 \\ 2s, & \text{otherwise} \end{cases}$$

2224

Lemma 8 (Horn-Johnson [17])**.** *Let λ_j and η_j denote the eigenvalues and singular values of an $n \times n$ matrix A, respectively. Then for any $k \leq n$,*

$$\sum_{j=1}^{k} |\lambda_j|^2 \leq \sum_{j=1}^{k} \eta^2$$

if η_j is arranged in descending order.

Lemma 9 (Bai-Silverstein [6])**.** *For any $uv \neq 0$ and $A > 2$, we have*

$$\left| \int_{|s| \geq A} \int_{-\infty}^{\infty} g_n(s,t) e^{ius+ivt} \mathrm{d}t \mathrm{d}s \right| \leq \frac{4\pi}{|v|} e^{-\frac{1}{2}|v|A} + \frac{2\pi}{n|v|} \sum_{k=1}^{mn} I\left(|\lambda_k| \geq \frac{A}{2} \right)$$

and

$$\left| \int_{|s| \leq A} \int_{t \geq A^3} g_n(s,t) e^{ius+ivt} \mathrm{d}t \mathrm{d}s \right| \leq \frac{8A}{A^2 - 1} + \frac{4\pi A}{n} \sum_{k=1}^{mn} I(|\lambda_k| > A)$$

where $\lambda_1, \ldots, \lambda_{mn}$ are the eigenvalues of Y. Furthermore, if the function $g_n(s,t)$ is replaced by $g(s,t)$, the two inequalities above hold without the second terms.

Now we note that under the assumptions of Theorem 2 and by Lemma 8 and the law of large numbers we have

$$\frac{1}{n} \sum_{k=1}^{mn} I(|\lambda_k| > A) \leq \frac{1}{nA^2} \mathrm{Tr}(Y^*Y) \longrightarrow \frac{m}{A^2} \qquad \text{a.s.}$$

Therefore, the right-hand sides of the inequalities in Lemma 9 can be made arbitrarily small by making A large enough. The same is true when $g_n(s,t)$ is replaced by $g(s,t)$. Our task is then reduced to showing

$$\int_{|s| \leq A} \int_{|t| \leq A^3} [g_n(s,t) - g(s,t)] e^{ius+ivt} \mathrm{d}s \mathrm{d}t \longrightarrow 0.$$

We define the sets

$$T = \left\{ (s,t) : |s| \leq A, |t| \leq A^3 \text{ and } ||z| - 1| \geq \epsilon \right\}$$

and

$$T_1 = \{ (s,t) : ||z| - 1| < \epsilon \},$$

where $z = s + it$.

Lemma 10 (Bai-Silverstein [6])**.** *For all fixed A and $0 < \epsilon < 1$,*

$$\int \int_{T_1} |g_n(s,t)| \mathrm{d}s \mathrm{d}t \leq 32\sqrt{\epsilon}. \tag{4}$$

Furthermore, if the function $g_n(s,t)$ is replaced by $g(s,t)$, the inequality above holds.

Since the right-hand side of (4) can be made arbitrarily small by choosing ϵ small, our task is reduced to showing

$$\int \int_{T} [g_n(s,t) - g(s,t)] e^{ius+ivt} \mathrm{d}s \mathrm{d}t \longrightarrow 0 \qquad \text{a.s.} \tag{5}$$

4.3 Characterization of the Circular Law

In this section, we study the convergence of the distributions $v_n(x, z)$ to a limiting distribution $v(x, z)$ as well as discuss properties of the limiting distribution $v(x, z)$. We begin with a standard truncation argument which can be found, for example, in [6].

4.3.1 Truncation

Let $\widehat{Y}$ and $\widetilde{Y}$ be the $(mn) \times (mn)$ matrices with entries

$$\widehat{Y}_{a,b;i,j} = Y_{a,b;i,j} I(\sqrt{n}|Y_{a,b;i,j}| \leq n^\delta) - \mathbb{E} Y_{a,b;i,j} I(\sqrt{n}|Y_{a,b;i,j}| \leq n^\delta)$$

and

$$\widetilde{Y}_{a,b;i,j} = \frac{\widehat{Y}_{a,b;i,j}}{\sqrt{n \mathbb{E} \left|\widehat{Y}_{a,b;i,j}\right|^2}}$$

where $\delta > 0$. We denote the ESD of $\widehat{H}_n = (\widehat{Y} - zI)^*(\widehat{Y} - zI)$ by $\widehat{v}_n(\cdot, z)$ and the ESD of $\widetilde{H}_n = (\widetilde{Y} - zI)^*(\widetilde{Y} - zI)$ by $\widetilde{v}_n(\cdot, z)$.

We will let $L(F_1, F_2)$ be the Levy distance between two distribution functions F_1 and F_2 defined by

$$L(F_1, F_2) = \inf\{\epsilon : F_1(x - \epsilon) - \epsilon \leq F_2(x) \leq F_1(x + \epsilon) + \epsilon \text{ for all } x \in \mathbb{R}\}.$$

We then have the following Lemma.

Lemma 11. *We have that*
$$L(v_n(\cdot, z), \widetilde{v}_n(\cdot, z)) = o(n^{-\eta \delta/4}) \text{ a.s.}$$

where the bound is uniform for $|z| \leq M$.

Proof. By [6, Corollary A.42] we have that

$$L^4(v(\cdot, z), \widehat{v}_n(\cdot, z)) \leq \frac{2}{n^2} \text{Tr}(H_n - \widehat{H}_n) \text{Tr}[(Y - \widehat{Y})^*(Y - \widehat{Y})].$$

By the law of large numbers it follows that, with probability 1,

$$\frac{1}{n} \text{Tr} H_n = \frac{1}{n} \sum_{a=1}^{m} \sum_{1 \leq i,j \leq n} |Y_{a,a+1;i,j}|^2 + m|z|^2 \longrightarrow m(1 + |z|^2).$$

Similarly, $\frac{1}{n} \text{Tr}(\widehat{H}_n) \to m(1 + |z|^2)$ a.s.

2226

For any $L > 0$, we have

$$\frac{n^{\delta\eta}}{n}\mathrm{Tr}[(Y-\widehat{Y})^*(Y-\widehat{Y})] = \frac{n^{\delta\eta}}{n}\sum_{a=1}^{m}\sum_{1\leq i,j\leq n}|(Y-\widehat{Y})_{a,a+1;i,j}|^2$$

$$\leq \frac{n^{\delta\eta}}{n^2}\sum_{a=1}^{m}\sum_{1\leq i,j\leq n}\left|\sqrt{n}Y_{a,a+1;i,j}I(\sqrt{n}|Y_{a,a+1;i,j}|>n^\delta) - \mathbb{E}\sqrt{n}Y_{a,a+1;i,j}I(\sqrt{n}|Y_{a,a+1;i,j}|>n^\delta)\right|^2$$

$$\leq 2n^{\delta\eta}\left(\frac{1}{n^2}\sum_{a=1}^{m}\sum_{1\leq i,j\leq n}|\sqrt{n}Y_{a,a+1;i,j}|^2 I(\sqrt{n}|Y_{a,a+1;i,j}|>n^\delta) + \mathbb{E}|\xi|^2 I(|\xi|>n^\delta)\right)$$

$$\leq \frac{2}{n^2}\sum_{a=1}^{m}\sum_{1\leq i,j\leq n}|\sqrt{n}Y_{a,a+1;i,j}|^{2+\eta}I(\sqrt{n}|Y_{a,a+1;i,j}|>L) + \mathbb{E}|\xi|^{2+\eta}I(|\xi|>L)$$

and hence

$$\limsup_{n\to\infty}\frac{n^{\delta\eta}}{n}\mathrm{Tr}[(Y-\widehat{Y})^*(Y-\widehat{Y})] \leq 4m\mathbb{E}|\xi|^{2+\eta}I(|\xi|>L)\text{ a.s.}$$

which can be made arbitrarily small by making L large. Thus we have that

$$L(v(\cdot,z),\widehat{v}_n(\cdot,z)) = o(n^{-\eta\delta/4})\text{ a.s.}$$

where the bound is uniform for $|z| \leq M$.

By [6, Corollary A.42] we also have that

$$L^4(\widehat{v}(\cdot,z),\widetilde{v}_n(\cdot,z)) \leq \frac{2}{n^2}\mathrm{Tr}(\widehat{H}_n+\widetilde{H}_n)\mathrm{Tr}(\widehat{Y}^*\widehat{Y})\left(1-\frac{1}{\sqrt{\mathbb{E}|\sqrt{n}\widehat{Y}_{1,2;1,1}|^2}}\right).$$

A similar argument shows that $1-\sqrt{\mathbb{E}|\sqrt{n}\widehat{Y}_{1,2;1,1}|^2} = o(n^{-\eta\delta})$ and the proof is complete.

□

Remark 12. For the remainder of the subsection, we will assume the conditions of Theorem 2 hold. Also, by Lemma 11 we additionally assume that $|Y_{a,a+1;i,j}| \leq n^\delta$.

4.3.2 Useful tools and lemmas

We begin by denoting the Stieltjes transform of $v_n(\cdot,z)$ by

$$\Delta_n(\alpha,z) = \int \frac{v_n(\mathrm{d}x,z)}{x-\alpha},$$

where $\alpha = x + iy$ with $y > 0$. We also note that $\Delta_n(\alpha,z) = \frac{1}{mn}\mathrm{Tr}(G)$ where $G = (H_n - \alpha I)^{-1}$ is the resolvent matrix. For brevity, the variable z will be suppressed when there is no confusion and we will simply write $\Delta_n(\alpha)$.

We first present a number of lemmas that we will need to study $\Delta_n(\alpha)$. We remind the reader that $R = Y - zI$ and $\alpha = x + iy$.

Lemma 13. *If $y > 0$ and $x \in K$ for some compact set K, then we have the following bounds,*

$$\|Y\|^2 \leq \max_{1 \leq k \leq m} \|X_k\|^2 \leq \sum_{k=1}^{m} \|X_k\|^2, \tag{6}$$

$$\|G\| \leq \frac{1}{y}, \tag{7}$$

$$\|RG\| \leq C \sqrt{\frac{1}{y^2} + \frac{1}{y}}, \tag{8}$$

$$\|GR^*\| \leq C \sqrt{\frac{1}{y^2} + \frac{1}{y}}, \tag{9}$$

for some constant $C > 0$ which depends on K. Moreover, there exists a constant C which depends only on K such that

$$\sup \left\{ \|RG\| : x \in K, y \geq y_n, z \in \mathbb{C} \right\} \leq C \sqrt{\frac{1}{y_n^2} + \frac{1}{y_n}}, \tag{10}$$

$$\sup \left\{ \|GR^*\| : x \in K, y \geq y_n, z \in \mathbb{C} \right\} \leq C \sqrt{\frac{1}{y_n^2} + \frac{1}{y_n}}, \tag{11}$$

for any sequence $y_n > 0$.

Proof. The first inequality in (6) follows from the definition of the norm and the second inequality is trivial. The resolvent bound in (7) follows immediately because H_n is a Hermitian matrix.

To prove (8), we use polar decomposition to write $R = U|R|$ where U is a partial isometry and $|R| = \sqrt{R^*R}$. Then

$$\|RG\| = \|U|R|(R^*R - \alpha)^{-1}\| \leq \||R|(R^*R - \alpha)^{-1}\|$$

$$\leq \sup_{t \in \mathrm{Sp}(R^*R)} |\sqrt{t}(t - \alpha)^{-1}| \leq \sup_{t \geq 0} |\sqrt{t}(t - \alpha)^{-1}| \leq C \sqrt{\frac{1}{y^2} + \frac{1}{y}}.$$

A similar argument verifies (9). (10) and (11) follow from (8) and (9) by using that $y \geq y_n$. $\square$

Lemma 14. *We have that*

$$\mathbb{E} \left[\frac{1}{n} \mathrm{Tr} G_{a,a} \right] = \mathbb{E} \left[\frac{1}{mn} \mathrm{Tr} G \right]$$

for any $1 \leq a \leq m$.

Proof. Fix $1 \leq a \leq m$ and $1 \leq i \leq n$. We will show that

$$\mathbb{E} G_{a,a;i,i} = \mathbb{E} G_{a+1,a+1;i,i}.$$

Using the adjoint formula for the inverse of a matrix, we can write that for any $1 \leq b \leq m$

$$G_{b,b;i,i} = \frac{\det(R^*R - \alpha I)^{(b,i)}}{\det(R^*R - \alpha I)}$$

2228

where $(R^*R - \alpha I)^{(b,i)}$ is the matrix $R^*R - \alpha I$ with the entries in the row and column that contain the element $(R^*R - \alpha I)_{b,b;i,i}$ replaced by zeroes except for the diagonal element which is replaced by a 1.

We will write $Q_b = X_b^* X_b + |z|^2 I - \alpha I$ and then note that $R^*R - \alpha I$ has the form

$$
\begin{pmatrix}
Q_m & -\bar{z}X_1 & 0 & \cdots & 0 & -zX_m^* \\
-zX_1^* & Q_1 & -\bar{z}X_2 & 0 & \cdots & 0 \\
0 & -zX_2^* & Q_2 & \ddots & 0 & \vdots \\
\vdots & 0 & \ddots & \ddots & -\bar{z}X_{m-2} & 0 \\
0 & \cdots & 0 & -zX_{m-2}^* & Q_{m-2} & -\bar{z}X_{m-1} \\
-\bar{z}X_m & 0 & \cdots & 0 & -zX_{m-1}^* & Q_{m-1}
\end{pmatrix},
\tag{12}
$$

where $Q_m, Q_1, \ldots, Q_{m-1}$ appear along the diagonal.

Let $\sigma = (1\ 2\ 3 \ldots m) \in S_m$. We now construct two bijective maps. Let T_σ be the map that takes matrices of the form (12) into the matrix where each occurrence of X_b is replaced by $X_{\sigma(b)}$ and each occurrence of Q_b is replaced by $Q_{\sigma(b)}$. Also, let

$$
\Omega = \underbrace{\mathbb{C}^{n^2} \times \mathbb{C}^{n^2} \times \cdots \times \mathbb{C}^{n^2}}_{m \text{ times}}
$$

denote the probability space. Then we write $\omega \in \Omega$ as $\omega = (X_1, X_2, \ldots, X_m)$. We now define $T_\sigma' : \Omega \to \Omega$ by $T_\sigma'(X_1, \ldots, X_m) = (X_2, X_3, \ldots, X_m, X_1)$. Since each $X_1, \ldots, X_m$ is an independent and identically distributed random matrix, T_σ' is a measure preserving map.

We claim that $\det(R^*R - \alpha I) = \det(T_\sigma(R^*R - \alpha I))$. Indeed, if λ is an eigenvalue of $(R^*R - \alpha I)$ with eigenvector $v = (v_m, v_1, \ldots, v_{m-1})^{\mathrm{T}}$ where v_b is an n-vector, then a simple computation reveals that $w = (v_{\sigma(m)}, v_{\sigma(1)}, \ldots, v_{\sigma(m-1)})^{\mathrm{T}} = (v_1, \ldots, v_m)^{\mathrm{T}}$ is an eigenvector of $T_\sigma(R^*R - \alpha I)$ with eigenvalue λ.

Similarly, $\det(R^*R - \alpha I)^{(b,i)} = \det\left(T_\sigma\left((R^*R - \alpha I)^{(b,i)}\right)\right)$. Define $f_{a,i}(\omega)$ to be $\det(R^*R - \alpha I)^{(b,i)}(\omega)$ for each realization $\omega \in \Omega$. Then we have that

$$
\begin{aligned}
f_{a+1,i}(\omega) &= \det\left(R^*R - \alpha I\right)^{(a+1,i)}(\omega) \\
&= \det\left(T_\sigma\left((R^*R - \alpha I)^{(a+1,i)}\right)\right)(\omega) \\
&= \det\left(T_\sigma\left(R^*R - \alpha I\right)\right)^{(a,i)}(\omega) \\
&= \det\left(R^*R - \alpha I\right)^{(a,i)}(T_\sigma'(\omega)) = f_{a,i}(T_\sigma'(\omega))
\end{aligned}
$$

and

$$
\det(R^*R - \alpha I)(\omega) = \det(R^*R - \alpha I)(T_\sigma'(\omega)).
$$

Thus $G_{a,a;i,i}(T_\sigma'(\omega)) = G_{a+1,a+1;i,i}(\omega)$ for each $\omega \in \Omega$. Since T_σ' is measure preserving, the proof is complete. $\qquad\square$

Next, we present the decoupling formula, which can be found, for example, in [18]. If ξ is a real-valued random variable such that $\mathbb{E}|\xi|^{p+2} < \infty$ and if $f(t)$ is a complex-valued function of a real variable such that its first $p+1$ derivatives are continuous and bounded, then

$$
\mathbb{E}[\xi f(\xi)] = \sum_{a=0}^{p} \frac{\kappa_{a+1}}{a!} \mathbb{E}[f^{(a)}(\xi)] + \epsilon,
\tag{13}
$$

where κ_a are the cumulants of ξ and $|\epsilon| \leq C \sup_t |f^{(p+1)}(t)| \mathbb{E}|\xi|^{p+2}$ where C depends only on p.

If ξ is a Gaussian random variable with mean zero, then all the cumulants vanish except for κ_2 and the decoupling formula reduces to the exact equation

$$\mathbb{E}[\xi f(\xi)] = \mathbb{E}[\xi^2]\mathbb{E}[f'(\xi)].$$

Finally, to use (13), we need to compute the derivatives of the resolvent matrix G with respect to the various entries of Y. This can be done by utilizing the resolvent identity and we find

$$\frac{\partial G_{a,b;k,l}}{\partial \operatorname{Re}(Y_{c,c+1;q,p})} = -(GR^*)_{a,c;k,q}G_{c+1,b;p,l} - G_{a,c+1;k,p}(RG)_{c,b;q,l},$$

$$\frac{\partial G_{a,b;k,l}}{\partial \operatorname{Im}(Y_{c,c+1;q,p})} = -i(GR^*)_{a,c;k,q}G_{c+1,b;p,l} + iG_{a,c+1;k,p}(RG)_{c,b;q,l}.$$

4.3.3 Main Theorem

For the results below, we will consider $\alpha = x + iy$ where $y \geq y_n$ with $y_n = n^{-\eta\delta}$. Our goal is to establish the following result.

Theorem 15. *Under the conditions of Theorem 2 and the additional assumption that $|Y_{a,a+1;i,j}| \leq n^\delta$, we have*

$$\Delta_n^3(\alpha,z) + 2\Delta_n^2(\alpha,z) + \frac{\alpha + 1 - |z|^2}{\alpha}\Delta_n(\alpha,z) + \frac{1}{\alpha} = r_n(\alpha,z),$$

where if δ is chosen such that $\delta\eta \leq 1/32$ and $\delta \leq 1/32$, then the remainder term r_n satisfies

$$\sup\left\{|r_n(\alpha,z)| : |z| \leq M, |x| \leq N, y \geq y_n\right\} = O\left(\delta_n\right) \qquad a.s.$$

with $\delta_n = n^{-1/4}y_n^{-5}n^\delta$.

Remark 16. We note that the bounds presented here and in the rest of this section are not optimal and can be improved. The bounds given, however, are sufficient for our purposes.

In order to prove Theorem 15, we will need the following lemmas. The first lemma is McDiarmid's Concentration Inequality [21].

Lemma 17 (McDiarmid's Concentration Inequality). *Let $X = (X_1, X_2, \ldots, X_n)$ be a family of independent random variables with X_k taking values in a set A_k for each k. Suppose that the real-valued f defined on $\prod A_k$ satisfies*

$$|f(x) - f(x')| \leq c_k$$

whenever the vectors x and x' differ only in the kth coordinate. Let μ be the expected value of the random variable $f(X)$. Then for any $t \geq 0$,

$$\mathbb{P}\left(|f(X) - \mu| \geq t\right) \leq 2e^{-2t^2/\sum c_k^2}.$$

Remark 18. McDiarmid's Concentration Inequality also applies to complex-valued functions by applying Lemma 17 to the real part and imaginary part separately.

Lemma 19. *For $y \geq y_n$ and $|x| \leq N$ (where $\alpha = x + iy$),*

$$\mathbb{P}\left(|\Delta_n(\alpha, z) - \mathbb{E}\Delta_n(\alpha, z)| > t\right) \leq 4e^{-ct^2 n y_n^4} \tag{14}$$

for some absolute constant $c > 0$. Moreover,

$$\sup\left\{|\Delta_n(\alpha, z) - \mathbb{E}\Delta_n(\alpha, z)| : |z| \leq M, |x| \leq N, y \geq y_n\right\} = O\left(n^{-1/4} y_n^{-2}\right) \qquad \text{a.s.}$$

Proof. Let R_k denote the matrix R with the k-th column replaced by zeroes. Then $R^* R$ and $R_k^* R_k$ differ by a matrix with rank at most two. So by the resolvent identity

$$
\begin{aligned}
\left| \frac{1}{mn} \text{Tr} \left(R^* R - \alpha\right)^{-1} - \frac{1}{mn} \text{Tr} \left(R_k^* R_k - \alpha\right)^{-1} \right| \\
\leq \frac{2}{mn} \left\| \left(R^* R - \alpha\right)^{-1} \left(R_k^* R_k - R^* R\right) \left(R_k^* R_k - \alpha\right)^{-1} \right\| \\
\leq \frac{C}{n y_n} \sup_{t \geq 0} \left|(t - \alpha)^{-1} t\right| = C' \frac{1}{n y_n^2}
\end{aligned}
\tag{15}
$$

where the constant C' depends only on N. The mn columns of Y form an independent family of random variables. We now apply Lemma 17 to the complex-valued function $\frac{1}{mn} \text{Tr}(R^* R - \alpha)^{-1}$ with the bound $c_k = O(n^{-1} y_n^{-2})$ obtained in (15). This proves the bound (14). Thus, for any fixed point (α, z) in the region

$$\{(\alpha = x + iy, z = s + it) : |x| \leq N, y \geq y_n, |z| \leq M\} \tag{16}$$

one has

$$\mathbb{P}\left(|\Delta_n(\alpha, z) - \mathbb{E}\Delta_n(\alpha, z)| > n^{-1/4} y_n^{-2}\right) \leq 4e^{-cn^{1/2}}, \tag{17}$$

where we recall that $y_n = n^{-\eta \delta}$ and $\delta > 0$ could be chosen to be arbitrary small. If $y = \text{Im}\,\alpha > n^{1/4} y_n^2$, then

$$|\Delta_n(\alpha, z)| \leq \frac{1}{\text{Im}\,\alpha} < n^{-1/4} y_n^{-2}, \quad |\mathbb{E}\Delta_n(\alpha, z)| < n^{-1/4} y_n^{-2}. \tag{18}$$

Therefore, it is enough to bound the supremum of $|\Delta_n(\alpha) - \mathbb{E}\Delta_n(\alpha)|$ over the region

$$\mathscr{D} = \{(\alpha = x + iy, z = s + it) : |x| \leq N, \ y_n \leq y \leq n^{1/4} y_n^2, \ |z| \leq M\}. \tag{19}$$

To this end, we consider a finite n^{-C}-net of $\mathscr{D}$ where C is some sufficiently large positive constant to be chosen later. Clearly, one can construct such a net that contains at most $[4Mn^{4C} n^{1/4} y_n^2]$ points if n is sufficiently large, where $[k]$ denotes the integer part of k. Let us denote these points by $(\alpha_i, z_i), 1 \leq i \leq [4Mn^{4C} n^{1/4} y_n^2]$. It follows from (17) that one has

$$\mathbb{P}\left(\sup\{i : |\Delta_n(\alpha_i, z_i) - \mathbb{E}\Delta_n(\alpha_i, z_i)| > n^{-1/4} y_n^{-2}\right) \leq 16 M y_n^2 n^{4C+1/4} e^{-cn^{1/2}}, \tag{20}$$

where the supremum is taken over the points of the net. Appying the Borel-Cantelli lemma, we obtain that

$$\sup\left\{i : |\Delta_n(\alpha_i, z_i) - \mathbb{E}\Delta_n(\alpha_i, z_i)|\right\} = O\left(n^{-1/4} y_n^{-2}\right) \qquad \text{a.s.} \tag{21}$$

where the supremum is again taken over the points of the n^{-C}-net of $\mathscr{D}$. To extend the estimate (21) to the supremum over the whole region $\mathscr{D}$, we note that for $(\alpha, z) \in \mathscr{D}$,

$$\left| \frac{\partial \Delta_n(\alpha, z)}{\partial \operatorname{Re}\alpha} \right| \leq \frac{1}{y_n^2}, \tag{22}$$

$$\left| \frac{\partial \Delta_n(\alpha, z)}{\partial \operatorname{Im}\alpha} \right| \leq \frac{1}{y_n^2}, \tag{23}$$

$$\left| \frac{\partial \Delta_n(\alpha, z)}{\partial \operatorname{Re}z} \right| \leq const_m \frac{2(n^{1+\delta} + M)}{y_n^2}, \tag{24}$$

$$\left| \frac{\partial \Delta_n(\alpha, z)}{\partial \operatorname{Im}z} \right| \leq const_m \frac{2(n^{1+\delta} + M)}{y_n^2}, \tag{25}$$

where $const_m$ is a constant that depends only on m.

The bounds (22-23) are simple properties of the Stieltjes transform. Indeed, the l.h.s. of (22) and (23) are bounded from above by $\frac{1}{|\operatorname{Im}\alpha|^2}$. The proof of (24-25) follows from the resolvent identitity

$$(H_n(z_2) - \alpha I)^{-1} - (H_n(z_1) - \alpha I)^{-1} = (H_n(z_1) - \alpha I)^{-1}(H_n(z_2) - H_n(z_1))(H_n(z_2) - \alpha I)^{-1},$$

the formula $H_n(z) = (Y^{(n)} - zI)^*(Y^{(n)} - zI)$, the bound $|z| \leq M$, and the bound

$$\|Y^{(n)}\| \leq n^{1+\delta}. \tag{26}$$

We note that (26) follows from the fact that the matrix entries of $Y^{(n)}$ are bounded by n^δ.

Now, choosing C in the construction of the net sufficiently large, one extends the bound (21) to the whole region $\mathscr{D}$ by (22-25). This finishes the proof of the lemma. $\qquad\square$

Lemma 20. *For any $1 \leq a \leq m$,*

$$\sup\left\{ \operatorname{Var}\left(\frac{1}{n}\operatorname{Tr}G_{aa} \right) : |x| \leq N, y \geq y_n, z \in \mathbb{C} \right\} = O(n^{-1}y_n^{-2})$$

where $\alpha = x + iy$.

Proof. Let R_k denote the matrix R with the k-th column replaced by zeroes and let P_a be the orthogonal projector such that $\operatorname{Tr}G_{a,a} = \operatorname{Tr}(P_a G P_a)$. Following the same procedure as in the proof of Lemma 19, we have that

$$\left| \operatorname{Tr}(R^*R - \alpha)_{a,a}^{-1} - \operatorname{Tr}(R_k^*R_k - \alpha)_{a,a}^{-1} \right|$$
$$= \left| \operatorname{Tr}\left[P_a(R_k^*R_k - \alpha)^{-1}(R_k^*R_k - R^*R)(R^*R - \alpha)^{-1}P_a \right] \right| \tag{27}$$
$$\leq \frac{C}{y_n^2}.$$

where the constant C depends only on N.

We can write

$$\frac{1}{n}\operatorname{Tr}G_{a,a} - \mathbb{E}\left[\frac{1}{n}\operatorname{Tr}G_{a,a} \right] = \frac{1}{n}\sum_{k=1}^{mn}\gamma_k,$$

where γ_k is the martingale difference sequence

$$\gamma_k = \mathbb{E}_k\left[\frac{1}{n}\mathrm{Tr}G_{a,a}\right] - \mathbb{E}_{k-1}\left[\frac{1}{n}\mathrm{Tr}G_{a,a}\right]$$

and $\mathbb{E}_k$ denotes the conditional expectation with respect to the elements in the first k columns of Y. Then by the bound in (27) and [6, Lemma 2.12], we have

$$\mathbb{E}\left|\frac{1}{n}\mathrm{Tr}G_{a,a} - \mathbb{E}\frac{1}{n}\mathrm{Tr}G_{a,a}\right|^2 \leq \frac{C}{n^2}\sum_{k=1}^{mn}|\gamma_k|^2$$

$$\leq \frac{C}{ny_n^2}$$

where the constant C depends only on N. Since the bound holds for any $|x| \leq N, y \geq y_n$, and $z \in \mathbb{C}$ the proof is complete. $\qquad\square$

Remark 21. By Lemmas 14, 19, and 20, for every $1 \leq a, b, c \leq m$

$$\mathbb{E}\left[\frac{1}{n}\mathrm{Tr}G_{a,a}\right] = \mathbb{E}\left[\frac{1}{mn}\mathrm{Tr}G\right] = \frac{1}{mn}\mathrm{Tr}G + O(n^{-1/4}y_n^{-5}) \qquad \text{a.s.,}$$

$$\mathbb{E}\left[\frac{1}{n}\mathrm{Tr}G_{a,a}\frac{1}{n}\mathrm{Tr}G_{b,b}\right] = \mathbb{E}\left[\left(\frac{1}{mn}\mathrm{Tr}G\right)^2\right] + O(n^{-1/4}y_n^{-5})$$

$$= \left(\frac{1}{mn}\mathrm{Tr}G\right)^2 + O(n^{-1/4}y_n^{-5}) \qquad \text{a.s.,}$$

and

$$\mathbb{E}\left[\frac{1}{n}\mathrm{Tr}G_{a,a}\frac{1}{n}\mathrm{Tr}G_{b,b}\frac{1}{n}\mathrm{Tr}G_{c,c}\right] = \mathbb{E}\left[\left(\frac{1}{mn}\mathrm{Tr}G\right)^3\right] + O(n^{-1/4}y_n^{-5})$$

$$= \left(\frac{1}{mn}\mathrm{Tr}G\right)^3 + O(n^{-1/4}y_n^{-5}) \qquad \text{a.s.,}$$

where the bounds hold uniformly in the region $|x| \leq N, y \geq y_n$, and $|z| \leq M$.

We are now ready to prove Theorem 15.

Proof of Theorem 15. Fix $\alpha = x + iy$ with $|x| \leq N, y \geq y_n$ and $z \in \mathbb{C}$ with $|z| \leq M$. We will show that the remainder term $r_n(\alpha, z) = O(\delta_n)$ a.s. where the constants in the term $O(\delta_n)$ depend only on N and M. In particular, the remainder term will be estimated using Lemmas 13 and 19 and Remark 21 where the bounds all hold uniformly in the region. In the proof presented below, will use the notation $O_{N,M}(\cdot)$ to represent a term which is bounded uniformly in the region $|x| \leq N, y \geq y_n$, and $|z| \leq M$.

By applying the resolvent identity to G and replacing R and R^* with $Y - zI$ and $Y^* - \bar{z}I$, respectively, we obtain

$$\frac{1}{n}\mathrm{Tr}G_{a,a} = -\frac{1}{\alpha} + \frac{1}{\alpha n}\mathrm{Tr}[GY^*Y]_{a,a} - \frac{z}{\alpha n}\mathrm{Tr}[GY^*]_{a,a} - \frac{\bar{z}}{\alpha n}\mathrm{Tr}[GY]_{a,a} + \frac{|z|^2}{\alpha n}\mathrm{Tr}G_{a,a}.$$

We will let $Y^{(r)}$ be the $(mn) \times (mn)$ matrix containing the real entries of Y and $Y^{(i)}$ be the $(mn) \times (mn)$ matrix containing the imaginary entries of Y such that $Y = Y^{(r)} + iY^{(i)}$. By assumption, $Y^{(r)}$ and $Y^{(i)}$ are independent random matrices. Thus,

$$
\left(1 - \frac{|z|^2}{\alpha}\right) \mathbb{E}\frac{1}{n}\mathrm{Tr}G_{a,a} + \frac{1}{\alpha} = \frac{1}{\alpha n}\mathbb{E}\mathrm{Tr}(GY^*Y^{(r)})_{a,a} + \frac{i}{\alpha n}\mathbb{E}\mathrm{Tr}(GY^*Y^{(i)})_{a,a}
$$
$$
- \frac{z}{\alpha n}\mathbb{E}\mathrm{Tr}(GY^{(r)*})_{a,a} + \frac{iz}{\alpha n}\mathbb{E}\mathrm{Tr}(GY^{(i)*})_{a,a} \qquad (28)
$$
$$
- \frac{\bar{z}}{\alpha n}\mathbb{E}\mathrm{Tr}(GY^{(r)})_{a,a} - \frac{\bar{z}i}{\alpha n}\mathbb{E}\mathrm{Tr}(GY^{(i)})_{a,a}
$$

Let $\delta = \mathrm{Var}(\mathrm{Re}(\xi))$. Then $\mathrm{Var}(\mathrm{Im}(\xi)) = 1 - \delta$. To compute the expectation, we fix all matrix entries except one and integrate with respect to that entry. Thus, by applying the decoupling formula (13) with $p = 1$ and using the fact that $Y_{a,b;i,j} = 0$ whenever $b \neq a+1$, we obtain the following expansions for the terms on the right-hand side of (28),

$$
\frac{1}{\alpha n}\mathbb{E}\mathrm{Tr}(GY^*Y^{(r)})_{a,a} = \frac{1}{\alpha n}\mathbb{E}\sum_{1 \leq j,k,l \leq n} G_{a,a;j,k}\bar{Y}_{a-1,a;l,k}\mathrm{Re}\left(Y_{a-1,a;l,j}\right)
$$
$$
= \frac{\delta}{\alpha n}\mathbb{E}\mathrm{Tr}G_{a,a} - \frac{\delta}{\alpha n^2}\mathbb{E}\sum_{1 \leq j,k,l \leq n} \bar{Y}_{a-1,a;l,k}\left((GR^*)_{a,a-1;j,l}G_{a,a;j,k}\right)
$$
$$
- \frac{\delta}{\alpha n^2}\mathbb{E}\sum_{1 \leq j,k,l \leq n} \bar{Y}_{a-1,a;l,k}\left(G_{a,a;j,j}(RG)_{a-1,a;l,k}\right) + O_{N,M}\left(\frac{n^\delta}{n^{1/2}y_n^4}\right)
$$
$$
= \frac{\delta}{\alpha n}\mathbb{E}\mathrm{Tr}G_{a,a} - \frac{\delta}{\alpha n^2}\mathbb{E}\left[\mathrm{Tr}G_{a,a}\mathrm{Tr}(RGY^*)_{a-1,a-1}\right] + O_{N,M}\left(\frac{n^\delta}{n^{1/2}y_n^4}\right).
$$

Here we use that the ϵ error term in (13) contains the second derivative

$$
\frac{\partial^2(GY^*)_{a,a;j,l}}{\partial\,\mathrm{Re}\left(Y_{a-1,a;l,j}\right)^2} = O_{N,M}\left(\frac{1}{y_n^3}\right)
$$

which consists of several terms each bounded by Lemma 13. After summing over $1 \leq j, l \leq n$ and utilizing the fact that the third moment of $\mathrm{Re}\left(Y_{a-1,a;l,j}\right)$ is of order $n^{\delta-3/2}$, we obtain an error

bound of $O_{N,M}(n^{\delta-1/2}y_n^{-4})$. By following the same procedure for the other terms, we obtain

$$\frac{i}{\alpha n}\mathbb{E}\mathrm{Tr}(GY^*Y^{(i)})_{a,a} = \frac{i}{\alpha n}\mathbb{E}\sum_{1\leq j,kl\leq n}G_{a,a;j,k}\bar{Y}_{a-1,a;l,k}\,\mathrm{Im}\left(Y_{a-1,a;l,j}\right)$$

$$= \frac{1-\delta}{\alpha n}\mathbb{E}\mathrm{Tr}G_{a,a} - \frac{1-\delta}{\alpha n^2}\mathbb{E}\left[\mathrm{Tr}G_{a,a}\mathrm{Tr}(RGY^*)_{a-1,a-1}\right] + O_{N,M}\left(\frac{n^\delta}{n^{1/2}y_n^4}\right),$$

$$\frac{z}{\alpha n}\mathbb{E}\mathrm{Tr}(GY^{(r)*})_{a,a} = \frac{z}{\alpha n}\mathbb{E}\sum_{1\leq j,k\leq n}G_{a,a;j,k}\,\mathrm{Re}\left(Y_{a,a+1;j,k}\right)$$

$$= \frac{z\delta}{\alpha n^2}\mathbb{E}\left[\mathrm{Tr}G_{a+1}a + 1\mathrm{Tr}(GR^*)_{a,a}\right] + O_{N,M}\left(\frac{n^\delta}{n^{1/2}y_n^4}\right),$$

$$\frac{iz}{\alpha n}\mathbb{E}\mathrm{Tr}(GY^{(i)*})_{a,a} = \frac{iz}{\alpha n}\mathbb{E}\sum_{1\leq j,k\leq n}G_{a,a+1;j,k}\,\mathrm{Im}\left(Y_{a,a+1;j,k}\right)$$

$$= \frac{z(1-\delta)}{\alpha n^2}\mathbb{E}\left[\mathrm{Tr}G_{a+1,a+1}\mathrm{Tr}(GR^*)_{a,a}\right] + O_{N,M}\left(\frac{n^\delta}{n^{1/2}y_n^4}\right),$$

$$\frac{\bar{z}}{\alpha n}\mathbb{E}\mathrm{Tr}(GY^{(r)})_{a,a} = \frac{\bar{z}}{\alpha n}\mathbb{E}\sum_{1\leq j,k\leq n}G_{a,a-1;j,k}\,\mathrm{Re}\left(Y_{a-1,a;k,j}\right)$$

$$= \frac{\bar{z}\delta}{\alpha n^2}\mathbb{E}\left[\mathrm{Tr}G_{a,a}\mathrm{Tr}(RG)_{a-1,a-1}\right] + O_{N,M}\left(\frac{n^\delta}{n^{1/2}y_n^4}\right),$$

and

$$\frac{\bar{z}i}{\alpha n}\mathbb{E}\mathrm{Tr}(GY^{(i)})_{a,a} = \frac{i\bar{z}}{\alpha n}\mathbb{E}\sum_{1\leq j,k\leq n}G_{a,a-1;j,k}\,\mathrm{Im}\left(Y_{a-1,a;k,j}\right)$$

$$= \frac{\bar{z}(1-\delta)}{\alpha n^2}\mathbb{E}\left[\mathrm{Tr}G_{a,a}\mathrm{Tr}(RG)_{a-1,a-1}\right] + O_{N,M}\left(\frac{n^\delta}{n^{1/2}y_n^4}\right).$$

Combining these terms yields,

$$\left(1-\frac{|z|^2}{\alpha}\right)\mathbb{E}\frac{1}{n}\mathrm{Tr}G_{a,a} = -\frac{1}{\alpha}+\frac{1}{\alpha n}\mathbb{E}\mathrm{Tr}G_{a,a}-\frac{1}{\alpha n^2}\mathbb{E}\left[\mathrm{Tr}G_{a,a}\mathrm{Tr}(RGR^*)_{a-1,a-1}\right]$$

$$+\frac{z}{\alpha n^2}\mathbb{E}\left[\mathrm{Tr}G_{a+1,a+1}\mathrm{Tr}(GR^*)_{a,a}\right]+O_{N,M}\left(\frac{n^\delta}{n^{1/2}y_n^4}\right)$$

$$= -\frac{1}{\alpha}+\frac{1}{\alpha n}\mathbb{E}\mathrm{Tr}G_{a,a}-\frac{1}{\alpha n^2}\mathbb{E}\left[\mathrm{Tr}G_{a,a}\mathrm{Tr}(RGR^*)_{a-1,a-1}\right]$$

$$+\frac{z}{\alpha n^2}\mathbb{E}\left[\mathrm{Tr}G_{a+1,a+1}\mathrm{Tr}(GR^*R)_{a,a}\right]$$

$$+\frac{1}{\alpha n^2}\mathbb{E}\left[\mathrm{Tr}G_{a+1,a+1}\mathrm{Tr}(GR^*Y)_{a,a}\right]+O_{N,M}\left(\frac{n^\delta}{n^{1/2}y_n^4}\right).$$

We note that by Remark 21, we have that

$$\frac{-1}{\alpha n^2}\mathbb{E}\left[\text{Tr}G_{a+1,a+1}\text{Tr}(GR^*R)_{a,a}\right] = -\frac{1}{\alpha n}\mathbb{E}\text{Tr}G_{a+1,a+1}$$
$$-\frac{1}{mn^2}\mathbb{E}\text{Tr}G_{a,a}\mathbb{E}\text{Tr}G + O_{N,M}(\delta_n)$$

and

$$\frac{-1}{\alpha n^2}\mathbb{E}\left[\text{Tr}G_{a,a}\text{Tr}(RGR^*)_{a-1,a-1}\right] = -\frac{1}{\alpha n^2}\mathbb{E}\left[\text{Tr}G_{a,a}\text{Tr}(R^*RG)_{a,a}\right]$$
$$-\frac{|z|^2}{\alpha n^2}\mathbb{E}\left[\text{Tr}G_{a,a}(\text{Tr}G_{a,a} - \text{Tr}G_{a-1,a-1})\right]$$
$$= -\frac{1}{\alpha n}\mathbb{E}\text{Tr}G_{a,a} - \frac{1}{mn^2}\mathbb{E}\text{Tr}G_{a,a}\mathbb{E}\text{Tr}G + O_{N,M}(\delta_n).$$

Finally, we expand Y in terms of $Y^{(r)}$ and $Y^{(i)}$ and again apply the decoupling formula (13) to obtain

$$\frac{1}{\alpha n^2}\mathbb{E}\left[\text{Tr}G_{a+1,a+1}\text{Tr}(GR^*Y)_{a,a}\right] = -\frac{1}{n^3}\mathbb{E}\left[\text{Tr}G_{a+1,a+1}\text{Tr}G_{a,a}\text{Tr}G_{a,a}\right] + O_{N,M}\left(\frac{n^\delta}{n^{1/2}y_n^4}\right)$$
$$= -\frac{1}{n^3 m^2}\left(\mathbb{E}\text{Tr}G\right)^2\mathbb{E}\text{Tr}G_{a,a} + O_{N,M}(\delta_n),$$

where the last equality comes from Remark 21. Therefore, we have that

$$\left(1 - \frac{|z|^2}{\alpha}\right)\mathbb{E}\frac{1}{n}\text{Tr}G_{a,a} = -\frac{1}{\alpha} - \frac{1}{\alpha n}\mathbb{E}\text{Tr}G_{a+1,a+1} - \frac{2}{mn^2}\mathbb{E}\text{Tr}G_{a,a}\mathbb{E}\text{Tr}G$$
$$-\frac{1}{n^3 m^2}\left(\mathbb{E}\text{Tr}G\right)^2\mathbb{E}\text{Tr}G_{a,a} + O_{N,M}(\delta_n).$$

By summing over a and dividing by m, we obtain

$$\left(\mathbb{E}\Delta_n(\alpha)\right)^3 + 2\left(\mathbb{E}\Delta_n(\alpha)\right)^2 + \frac{1+\alpha-|z|^2}{\alpha}\mathbb{E}\Delta_n(\alpha) + \frac{1}{\alpha} = O_{N,M}(\delta_n).$$

Thus, the proof is complete by Lemma 19. $\qquad\square$

Consider the cubic equation

$$\Delta^3 + 2\Delta^2 + \frac{\alpha+1-|z|^2}{\alpha}\Delta + \frac{1}{\alpha} = 0 \tag{29}$$

where $\alpha = x + iy$. The solution of the equation has three analytic branches when $\alpha \neq 0$ and when there is no multiple root. Below we show that the Stieltjes transform of $\nu_n(\cdot, z)$ converges to a root of (29). Following the argument of Bai and Silverstein in [6], we have that there is only one of the three analytic branches, denoted by $\Delta(\alpha)$, to which the Stieltjes transforms are converging to. We let $m_2(\alpha)$ and $m_3(\alpha)$ denote the other two branches and note that Δ, m_2, and m_3 are also functions of $|z|$.

By [6, Theorem B.9], there exists a distribution function $\nu(\cdot, z)$ such that

$$\Delta(\alpha) = \int \frac{1}{u-\alpha}\nu(\mathrm{d}u, z).$$

Then we use the following Lemmas due to Bai and Silverstein, [6].

2236

Lemma 22. *The limiting distribution function $v(x,z)$ satisfies*

$$|v(w+u,z) - v(w,z)| \leq \frac{2}{\pi}\max\{2\sqrt{3|u|}, |u|\}$$

for all z. Also, the limiting distribution function $v(u,z)$ has support in the interval $[x_1, x_2]$ when $|z| > 1$ and $[0, x_2]$ when $|z| \leq 1$, where

$$x_1 = \frac{1}{8|z|^2}\left[-1 + 20|z|^2 + 8|z|^4 - (\sqrt{1+8|z|^2})^3\right],$$

$$x_2 = \frac{1}{8|z|^2}\left[(\sqrt{1+8|z|^2})^3 - 1 + 20|z|^2 + 8|z|^4\right].$$

Lemma 23. *For any given constants $N > 0$, $A > 0$, and $\epsilon \in (0,1)$ (recall that A and ϵ are used to define the region T), there exist positive constants ϵ_1 and ϵ_0 (ϵ_0 may depend on ϵ_1) such that for all large n,*

(i) *for $|\alpha| \leq N$, $y \geq 0$, and $z \in T$,*

$$\max_{j=2,3}|\Delta(\alpha) - m_j(\alpha)| \geq \epsilon_0,$$

(ii) *for $|\alpha| \leq N$, $y \geq 0$, $|\alpha - x_2| \geq \epsilon_1$ (and $|\alpha - x_1| \geq \epsilon_1$ if $|z| \geq 1 + \epsilon$), and $z \in T$,*

$$\min_{j=2,3}|\Delta(\alpha) - m_j(\alpha)| \geq \epsilon_0,$$

(iii) *for $z \in T$ and $|\alpha - x_2| < \epsilon_1$,*

$$\min_{j=2,3}|\Delta(\alpha) - m_j(\alpha)| \geq \epsilon_0\sqrt{|\alpha - x_2|},$$

(iv) *for $|z| > 1 + \epsilon$, $z \in T$, and $|\alpha - x_1| < \epsilon_1$,*

$$\min_{j=2,3}|\Delta(\alpha) - m_j(\alpha)| \geq \epsilon_0\sqrt{|\alpha - x_1|}.$$

Remark 24. Lemma 23 shows that away from the real line, Δ is distinct from the branches m_2 and m_3.

Lemma 25. *We have*

$$\frac{\partial}{\partial s}\int_0^\infty \ln x v(\mathrm{d}x, z) = g(s, t).$$

Remark 26. Lemma 25 shows that $v(\cdot, z)$ is the distribution which corresponds to the circular law.

4.4 Rate of Convergence of $v_n(x,z)$

For this subsection, we return to the original assumptions on the entries of Y. Before we prove Lemma 5, we need to establish a rate of convergence of $v_n(x,z)$ to $v(x,z)$. We remind the reader that $v_n(\cdot,z)$ is the ESD of $H_n = (Y - zI)^*(Y - zI)$ and $\widetilde{v}_n(\cdot,z)$ is the ESD of $\widetilde{H}_n = (\widetilde{Y} - zI)^*(\widetilde{Y} - zI)$.

Lemma 27. *For any $M_2 > M_1 \geq 0$,*

$$\sup_{M_1 \leq |z| \leq M_2} \|v_n(\cdot,z) - v(\cdot,z)\| = \sup_{x, M_1 \leq |z| \leq M_2} |v_n(x,z) - v(x,z)| = O(n^{-\delta\eta/8}).$$

Proof. We first note that it is enough to show

$$\sup_{M_1 \leq |z| \leq M_2} \|\widetilde{v}_n(\cdot,z) - v(\cdot,z)\| = O(\sqrt{y_n}). \tag{30}$$

Indeed, by Lemma 11, we have that

$$L(v_n(\cdot,z), v(\cdot,z)) \leq L(v_n(\cdot,z), \widetilde{v}_n(\cdot,z)) + \|\widetilde{v}_n(\cdot,z) - v(\cdot,z)\|$$
$$\leq \|\widetilde{v}_n(\cdot,z) - v(\cdot,z)\| + o(n^{-\eta\delta/4}).$$

and by Lemma 22,

$$\|v_n(\cdot,z) - v(\cdot,z)\| \leq C\sqrt{L(v_n(\cdot,z), v(\cdot,z))}$$

uniformly for $|z| \leq M$.

We now prove (30). Since $|\Delta_n(\alpha_0)| \leq (\operatorname{Im}\alpha_0)^{-1}$ for any fixed α_0 with $\operatorname{Im}\alpha_0 > 0$, there exists a convergent subsequence of $\{\Delta_n(\alpha_0)\}_{n=1}^{\infty}$. Since Δ is the only branch of (29) that defines a Stieltjes transform, the subsequence must converge to $\Delta(\alpha_0)$. Hence, $\Delta_n(\alpha_0) \to \Delta(\alpha_0)$ as $n \to \infty$ for any fixed α_0 with $\operatorname{Im}\alpha_0 > 0$. Let $m_1 = \Delta$ and m_2 and m_3 be the other two branches of the cubic equation (29).

We remind the reader that T is a bounded set and that the supports of $v(\cdot,z)$ are bounded for all $z \in T$. So by [6, Corollary B.15] there exists N and some absolute constant C such that

$$\|\widetilde{v}_n(\cdot,z) - v(\cdot,z)\|$$
$$\leq C\left(\int_{|x|\leq N} |\Delta_n(\alpha) - \Delta(\alpha)|dx + \frac{1}{y_n}\sup_x \int_{|y|\leq 2y_n} |v(x+y,z) - v(x,z)|dy\right)$$
$$\leq C\left(\int_{|x|\leq N} |\Delta_n(\alpha) - \Delta(\alpha)|dx + \sqrt{y_n}\right),$$

where $\alpha = x + iy_n$ and the last inequality follows from Lemma 22. So, to complete the proof we only need to estimate the integral in the last inequality above.

We first show that for $\alpha = x + iy$, $|x| \leq N$, $|x - x_2| \geq \epsilon_1$ ($|x - x_1| \geq \epsilon_1$ if $|z| < 1$), $y \geq y_n$, $M_1 \leq |z| \leq M_2$, and all large n,

$$|\Delta_n(\alpha) - \Delta(\alpha)| < \frac{C'\epsilon_0}{3}\delta_n \tag{31}$$

2238

where ϵ_0 and ϵ_1 come from Lemma 23 and C' is a positive constant. By Theorem 15, consider a realization where

$$|\Delta_n(\alpha) - \Delta(\alpha)||\Delta_n(\alpha) - m_2(\alpha)||\Delta_n(\alpha) - m_3(\alpha)| \leq C'\frac{4}{27}\epsilon_0^3\delta_n.$$

for some positive constant C'. Fix $\alpha_0 = x_0 + iy_0$ with $|x_0| \leq N, y_0 > 0$, and $\min_{k=1,2}|x_0 - x_k| \geq \epsilon_1$. Fix $z \in T$. Choose n large enough such that $|\Delta_n(\alpha_0) - \Delta(\alpha_0)| < \frac{\epsilon}{3}$. Then for $k \in \{1,2\}$,

$$\epsilon_0 \leq |\Delta(\alpha_0) - m_k(\alpha_0)| \leq |\Delta(\alpha_0) - \Delta_n(\alpha_0)| + |\Delta_n(\alpha_0) - m_k(\alpha_0)|$$

and hence

$$\min_{k=1,2}|\Delta_n(\alpha_0) - m_k(\alpha_0)| > \frac{2\epsilon_0}{3}.$$

Thus,

$$|\Delta_n(\alpha_0) - \Delta(\alpha_0)| \leq C'\epsilon_0\delta_n.$$

Next we show (31) is true for all $y \geq y_n$, $|x| \leq N$, and $\min_{k=1,2}|x - x_k| \geq \epsilon_1$. Suppose (31) is false. By continuity there exists a subsequence n_l, $z_l \in T$, and α_l with $|\text{Re}(\alpha_l)| \leq N$ and $\text{Im}(\alpha_l) \geq y_{n_l}$ such that

$$|\Delta_{n_l}(\alpha_l) - \Delta(\alpha_l)| = \frac{C'\epsilon_0}{3}\delta_{n_l}.$$

Then

$$|\Delta_{n_l}(\alpha_l) - \Delta(\alpha_l)| < \frac{\epsilon_0}{3}$$

for all l greater than some L. By a similar argument as above and Lemma 23, we have

$$\min_{k=1,2}|\Delta_{n_l}(\alpha_l) - m_k(\alpha_l)| > \frac{2\epsilon_0}{3}$$

for all $l > L$ and hence

$$|\Delta_{n_l}(\alpha_l) - \Delta(\alpha_l)| < \frac{C'\epsilon_0}{3}\delta_n,$$

a contradiction.

Finally, for the case where $|\alpha - x_k| \leq \epsilon_1$ for $k = 1$ or 2, we apply a similar argument and Lemma 23 to obtain

$$|\Delta_n(\alpha) - \Delta(\alpha)| = O\left(\frac{\delta_n}{\sqrt{|\alpha - x_k|}}\right) = O(\delta_n y_n^{-1/2}).$$

$\square$

4.5 Least Singular Value Bound

A key part of proving Lemma 5 is to control the least singular value of $Y - zI$. Equivalently, we wish to obtain control of the norm of the inverse $\|(Y - zI)^{-1}\|$.

We will obtain a bound using the results of Tao and Vu in [28]. We present Tao and Vu's bound on the least singular value below, which only requires a finite second moment assumption on the entries of the matrix.

Theorem 28 (Tao-Vu; Least singular value bound). *Let A, C_1 be positive constants, and let ξ be a complex-valued random variable with non-zero finite variance (in particular, the second moment is finite). Then there are positive constants B and C_2 such that the following holds: if N_n is the random matrix of order n whose entries are i.i.d. copies of ξ, and M is a deterministic matrix of order n with spectral norm at most n^{C_1}, then,*

$$\mathbb{P}\left(\|(M + N_n)^{-1}\| \geq n^B\right) \leq C_2 n^{-A}. \tag{32}$$

Remark 29. We note that the bound in (32) is independent of the matrix M. In particular, this bound holds for any deterministic matrix of order n with spectral norm at most n^{C_1}.

We will prove an analogous version of Theorem 28 for the matrix Y. We first need the following bounds for the norm of Y.

Lemma 30. *We have the following bounds for the norm of Y.*

$$\|Y\| = O(n) \text{ a.s.,}$$
$$\mathbb{E}\|Y\| = O(n).$$

We also have that for any $1 \leq a \leq m$,

$$\mathbb{E}\|X_a\| = O(n). \tag{33}$$

Proof. We note that

$$Y^*Y = \begin{pmatrix} X_m^* X_m & & & 0 \\ & X_1^* X_1 & & \\ & & \ddots & \\ 0 & & & X_{m-1}^* X_{m-1} \end{pmatrix},$$

and hence the singular values of Y are the the union of the singular values of X_k for $1 \leq k \leq m$. Let $s_1, \ldots s_{mn}$ denote the singular values of Y. Then

$$\frac{1}{mn}\|Y\| \leq \frac{1}{mn}\sum_{j=1}^{mn} s_j \leq \frac{1}{mn}\sum_{j=1}^{mn} s_j^2 + 1$$

$$\leq \frac{1}{mn}\mathrm{Tr}\,Y^*Y + 1 = \frac{1}{mn}\sum_{k=1}^{m}\sum_{1 \leq i,j \leq n}\left|(X_k)_{ij}\right|^2 + 1 \longrightarrow 2 \text{ a.s.}$$

as $n \to \infty$ by the law of large numbers. The same argument shows that

$$\frac{1}{mn}\mathbb{E}\|Y\| \leq \frac{1}{mn}\sum_{k=1}^{m}\sum_{1 \leq i,j \leq n}\mathbb{E}\left|(X_k)_{ij}\right|^2 + 1 = 2.$$

A similar argument verifies (33).

$\square$

Theorem 31 (Least singular value bound for Y). *Let Y be the $(mn) \times (mn)$ matrix defined in (2) and let A be a positive constant. Then, under the hypothesis of Theorem 2, there exists positive constants B and C (depending on both A and m) such that*

$$\mathbb{P}\left(\|(Y - zI)^{-1}\| \geq n^B\right) \leq Cn^{-A}$$

uniformly for $|z| \leq M_2$.

Proof. We remind the reader that $(Y - zI)^{-1}$ is an $(mn) \times (mn)$ matrix and again refer to the m^2 blocks $(Y - zI)^{-1}_{a,b}$ each of size $n \times n$. A simple computation reveals, that (when invertible) $(Y - zI)^{-1}_{a,b}$ has the form

$$z^\kappa X_{j_1} \cdots X_{j_l} \left(X_{i_1} \cdots X_{i_q} - z^r \right)^{-1}$$

where κ, l, q, r are nonnegative integers no bigger than m, the variables $\kappa, l, q, r, j_1, \ldots j_l, i_1, \ldots, i_q$ depend only on a and b, and the indices $i_1, \ldots, i_q$ are all distinct.

By the definition of the norm, we have that

$$\|(Y - zI)^{-1}\| \leq C_m \max_{1 \leq a, b \leq m} \|(Y - zI)^{-1}_{a,b}\| \leq C_m \sum_{1 \leq a, b \leq m} \|(Y - zI)^{-1}_{a,b}\|$$

where C_m is a constant that depends only on m. Thus, it is enough to show that given a positive constant A, there exists B and C such that

$$\mathbb{P} \left(\|z^\kappa X_{j_1} \cdots X_{j_l} \left(X_{i_1} \cdots X_{i_q} - z^r \right)^{-1}\| \geq n^B \right) \leq C n^{-A}$$

uniformly for $|z| \leq M_2$.

So we have,

$$\mathbb{P} \left(\|z^\kappa X_{j_1} \cdots X_{j_l} \left(X_{i_1} \cdots X_{i_q} - z^r \right)^{-1}\| \geq n^B \right)$$

$$\leq m \mathbb{P} \left(\|X_1\| \geq n^{B/(m+2)} \right) + \mathbb{P} \left(\| \left(X_{i_1} \cdots X_{i_q} - z^r \right)^{-1}\| \geq n^{B/(m+2)} \right)$$

for $|z| \leq M_2$ and n large. The first term can be estimated by Markov's inequality

$$\mathbb{P} \left(\|X_1\| \geq n^{B/(m+2)} \right) \leq \frac{\mathbb{E}\|X_1\|}{n^{B/(m+2)}} = O(n^{-B/(m+2)+1})$$

since $\mathbb{E}\|X_1\| = O(n)$ by Lemma 30. Therefore, this term is order n^{-A} by taking $B > (m+2)(A+1)$. So, it is now enough to show that given a positive constant A, there exists B and C such that

$$\mathbb{P} \left(\| \left(X_{i_1} \cdots X_{i_q} - z^r \right)^{-1}\| \geq n^B \right) \leq C n^{-A}.$$

We note that,

$$\left(X_{i_1} \cdots X_{i_q} - z^r \right)^{-1} = X_{i_q}^{-1} \cdots X_{i_2}^{-1} \left(X_{i_1} - z^r X_{i_q}^{-1} \cdots X_{i_2}^{-1} \right)^{-1}.$$

By Theorem 28 there exists positive constants B and C such that

$$\mathbb{P} \left(\|X_{i_q}^{-1} \cdots X_{i_2}^{-1}\| \geq n^B \right) \leq m \mathbb{P} \left(\|X_1^{-1}\| \geq n^{B/m} \right) \leq C n^{-A}. \tag{34}$$

Thus, we only need to show that given A there exists B and C such that

$$\mathbb{P} \left(\left(X_{i_1} - z^r X_{i_q}^{-1} \cdots X_{i_2}^{-1} \right)^{-1} \geq n^B \right) \leq C n^{-A}.$$

2241

We then have that

$$\mathbb{P}\left(\left(X_{i_1} - z^r X_{i_q}^{-1} \cdots X_{i_2}^{-1}\right)^{-1} \geq n^B\right)$$

$$= \mathbb{P}\left(\left(X_{i_1} - z^r X_{i_q}^{-1} \cdots X_{i_2}^{-1}\right)^{-1} \geq n^B \ \middle| \ \|X_{i_q}^{-1} \cdots X_{i_2}^{-1}\| \leq n^{C_1}\right)$$

$$\times \mathbb{P}\left(\|X_{i_q}^{-1} \cdots X_{i_2}^{-1}\| \leq n^{C_1}\right)$$

$$+ \mathbb{P}\left(\left(X_{i_1} - z^r X_{i_q}^{-1} \cdots X_{i_2}^{-1}\right)^{-1} \geq n^B \ \middle| \ \|X_{i_q}^{-1} \cdots X_{i_2}^{-1}\| \geq n^{C_1}\right)$$

$$\times \mathbb{P}\left(\|X_{i_q}^{-1} \cdots X_{i_2}^{-1}\| \geq n^{C_1}\right)$$

$$\leq C n^{-A}$$

where the first term is controlled by Theorem 28 (in particular, see Remark 29) and the second term is estimated as in (34). This completes the proof of the Theorem. □

4.6 Proof of Lemma 5

Proof of Lemma 5. In order to finish the proof of Lemma 5 we need to show (5) holds. By integration by parts, we have

$$\left|\iint_{z \in T} (g_n(s,t) - g(s,t)) e^{isu+itv} \mathrm{d}t \mathrm{d}s\right|$$

$$= \left| - \int_{z \in T} iu\tau(s,t)\mathrm{d}t\mathrm{d}s + \int_{|t| \leq A^3} (\tau(A,t) - \tau(-A,t))\mathrm{d}t \right.$$

$$- \int_{|t| \leq 1+\epsilon} \left(\tau(\sqrt{(1+\epsilon)^2 - t^2}, t) - \tau(-\sqrt{(1+\epsilon)^2 - t^2}, t)\right) \mathrm{d}t$$

$$\left. + \int_{|t| \leq 1-\epsilon} \left(\tau(\sqrt{(1-\epsilon)^2 - t^2}, t) - \tau(-\sqrt{(1-\epsilon)^2 - t^2}, t)\right) \mathrm{d}t \right|,$$

where

$$\tau(s,t) = e^{ius+ivt} \int_0^\infty \ln x (\nu_n(\mathrm{d}x, z) - \nu(\mathrm{d}x, z)).$$

Let $\epsilon_n = e^{-n^{\eta\delta/16}}$. By Theorem 31 and the Borel-Cantelli lemma, with probability 1,

$$\lim_{n\to\infty} \int_{z \in T} \left| \int_0^{\epsilon_n} \ln x \nu_n(\mathrm{d}x, z) \right| \mathrm{d}t\mathrm{d}s = 0.$$

By Lemma 22,

$$\lim_{n\to\infty} \int_{z \in T} \left| \int_0^{\epsilon_n} \ln x \nu(\mathrm{d}x, z) \right| \mathrm{d}t\mathrm{d}s = 0.$$

2242

By Lemma 30, there exists $\kappa > 0$ such that the support of $v_n(\cdot, z)$ lies in $[0, \kappa n^2]$ for all $z \in T$. Thus, by Lemma 27

$$\int_{z \in T} \left| \int_{\epsilon_n}^{\infty} \ln x(v_n(\mathrm{d}x, z) - v(\mathrm{d}x, z)) \right| \mathrm{d}t \mathrm{d}s$$

$$= \int_{z \in T} \left| \int_{\epsilon_n}^{\kappa n^2} \ln x(v_n(\mathrm{d}x, z) - v(\mathrm{d}x, z)) \right| \mathrm{d}t \mathrm{d}s$$

$$\leq C \left[|\ln(\epsilon_n)| + \ln(\kappa n^2) \right] \max_{z \in T} \|v_n(\cdot, z) - v(\cdot, z)\| \longrightarrow 0 \qquad \text{a.s.}$$

Therefore, with probability 1,

$$\lim_{n \to \infty} iu \int_{z \in T} \tau(s, t) \mathrm{d}s \mathrm{d}t = 0.$$

In a similar fashion, we can show that the boundary terms satisfy the following

$$\lim_{n \to \infty} \int_{|t| \leq A^3} \tau(\pm A, t) \mathrm{d}t = 0 \qquad \text{a.s.},$$

$$\lim_{n \to \infty} \int_{|t| \leq 1+\epsilon} \tau(\pm \sqrt{(1+\epsilon)^2 - t^2}, t) \mathrm{d}t = 0 \qquad \text{a.s.},$$

and

$$\lim_{n \to \infty} \int_{|t| \leq 1-\epsilon} \tau(\pm \sqrt{(1-\epsilon)^2 - t^2}, t) \mathrm{d}t = 0 \qquad \text{a.s.}$$

The proof of Lemma 5 is complete. $\qquad \square$

Remark 32. After we finished our paper, we learned about a very recent preprint [16] where F. Götze and A. Tikhomirov proved the convergence of the expected spectral distribution $\mathbb{E}\mu_X$ to the limit defined by (1) under the assumption that the matrix entries are mutually independent centered complex random variables with variance one. Our approach is different from the one used in [16]. We are grateful to Z. Burda, T. Tao and A. Tikhomirov for useful comments regarding the results of the paper. In addition, we are grateful to unanimous referees for valuable and constructive criticism regarding the proofs of Theorem 15 and Lemma 19, and for bringing to our attention the reference [13] where a similar result was obtained for $m = 2$.

References

[1] N. Alexeev, F. Götze, A. Tikhomirov, On the singular spectrum of powers and products of random matrices, *Doklady Mathematics*, **82**, No. 1, 505-507, (2010).

[2] N. Alexeev, F. Götze, A. Tikhomirov, On the asymptotic distribution of singular values of power of random matrices, *Lithuanian Math. Journal*, **50**, No. 2, 121-132, (2010). MR2653641

[3] Z. D. Bai, Circular law, *Ann. Probab.* **25** (1997), 494–529. MR1428519

[4] Z. D. Bai, Y.Q. Yin, Limiting behavior of the norm of products of random matrices and two problems of Geman-Hwang, *Prob. Th. Rel. Fields*, **73**, 555-569, (1986). MR0863545

[5] Z. D. Bai, B. Miao, B. Jin, On the limit theorem for the eigenvalues of product of two matrices, *J. Multivariate Anal.*, **98**, No. 1, 76-101, (2007). MR2292918

[6] Z. D. Bai, J. Silverstein, Spectral analysis of large dimensional random matrices, Springer Series in Statistics, Springer, New York, 2010. MR2567175

[7] T. Banica, S. Belinschi, M. Capitaine, B. Collins, Free Bessel laws, available at arXiv:0710.5931. MR2779129

[8] Z. Burda, R. A. Janik,and B. Waclaw, Spectrum of the product of independent random Gaussian matrices, *Phys. Rev. E* **81**, 041132 (2010). MR2736204

[9] Z. Burda et.al., Eigenvalues and singular values of products of rectangular Gaussian random matrices, *Phys. Rev. E* **82**, 061114 (2010). MR2787469

[10] J. Ginibre, Statistical Ensembles of Complex, Quaternion, and Real Matrices, *Journal of Mathematical Physics* **6** (1965), 440–449. MR0173726

[11] V. L. Girko, Circular law, *Theory Probab. Appl.* (1984), 694–706.

[12] V. L. Girko, The strong circular law. Twenty years later. II. *Random Oper. Stochastic Equations* **12** (2004), no. 3, 255–312. MR2085255

[13] V. L. Girko, A.I. Vladimirova, Spectral analysis of stochastic recurrence systems of growing dimension under G-condition. Canonical equation K_{91}. *Random Oper. Stochastic Equations* **17** (2009), no. 3, 243–274. MR2574103

[14] F. Götze, A. Tikhomirov, The circular law for random matrices, *Annals of Probab.* **38**, No. 4, 1444-1491, (2010). MR2663633

[15] F. Götze, A. Tikhomirov, On the circular law, available at arXiv:0702386.

[16] F. Götze, A. Tikhomirov, On the asymptotic of spectrum of products of independent random matrices, available at arXiv:1012.2710v1 math.PR.

[17] R. Horn, Ch. Johnson, Topics in Matrix analysis, Cambridge University Press, 1991. MR1091716

[18] A.M.Khorunzhy, B.A. Khoruzhenko, and L.A. Pastur, Asymptotic properties of large random matrices with independent entries, *J. Math. Phys.* **37**, (1996), 5033–5059. MR1411619

[19] R. Latala, Some estimates of norms of random matrices, *Proc. Amer. Math. Soc.* 133 (2005), no. **5**, 1273?1282. MR2111932

[20] V. Marchenko, L. Pastur, *Distribution of eigenvalues of some sets of random matrices*, Math USSR-Sb. **1**, (1967), 457–486.

[21] C. McDiarmid, On the Method of Bounded Differences, *Surveys in Combinatorics* **141**: 148-188, 1989. MR1036755

[22] M.L. Mehta, Random Matrices and the Statistical Theory of Energy Levels, Academic Press, New York, NY, 1967. MR0220494

[23] M.L. Mehta, Random Matrices, 3rd ed., Elsevier/Academic Press, Amsterdam, 2004. MR2129906

[24] J.A. Mingo, R.Speicher, Sharp bounds for sums associated to graphs of matrices, available at arXiv:0909.4277.

[25] G. Pan, W. Zhou, Circular law, extreme singular values and potential theory, *J. Multivariate Anal.* (2010)*, **101**, 645-656. MR2575411

[26] M. Rudelson, Invertibility of random matrices: Norm of the inverse, *Ann. of Math.* **168**, No.2, 575-600 (2008) , 575?600. MR2434885

[27] M. Rudelson, R. Vershynin, The Littlewood-Offord problem and invertibility of random matrices, *Adv. Math.* (2008) **218**, 600-633. MR2407948

[28] T. Tao, V. Vu, Random Matrices: the Circular Law, *Commun. Contemp. Math.* **10**, No. 2, 261-307 (2008). MR2409368

[29] T. Tao, V. Vu, Random Matrices: Universality of ESDs and the Circular Law, *Annals of Probab.* **38**, No. 5, 2023-2065 (2010). MR2722794