ELLIPTIC LAW FOR REAL RANDOM MATRICES

ALEXEY NAUMOV[1]

Abstract. We prove elliptic law for real random matrices under assumption of finite forth moment of the matrix entries.

Contents

1. Introduction 2
2. Proof of the main result 4
3. Least singular value 5
3.1. The small ball probability via central limit theorem 6
3.2. Decomposition of the sphere 7
3.3. Invertibility for the compressible vectors 7
3.4. Invertibility for the incompressible vectors 9
3.5. Distance via the small ball probability 10
4. Uniform integrability of logarithm 15
5. Convergence of singular values 18
6. Acknowledgment 24
Appendix A. Appendix 24
Appendix B. Figures 30
References 32

Date: June 25, 2012.
Key words and phrases. Random matrices, elliptic law, logarithmic potential, least singular value, small ball probability.
[1]This research was supported by DAAD.

1. Introduction

Let us consider real asymmetric random matrix $\mathbf{X}_n(\omega) = \{X_{ij}(\omega)\}_{i,j=1}^n$ and assume that the following conditions (**C0**) hold
a) Pairs $(X_{ij}, X_{ji}), i \neq j$ are i.i.d. random vectors;
b) $\mathbb{E}\, X_{12} = \mathbb{E}\, X_{21} = 0, \mathbb{E}\, X_{12}^2 = \mathbb{E}\, X_{21}^2 = 1$ and $\max(\mathbb{E}\,|X_{12}|^4, \mathbb{E}\,|X_{21}|^4) \leq M_4$;
c) $\mathbb{E}(X_{12}X_{21}) = \rho$, $|\rho| \leq 1$;
d) The diagonal entries X_{ii} are i.i.d. random variables, independent of off-diagonal entries, $\mathbb{E}\, X_{11} = 0$ and $\mathbb{E}\, X_{11}^2 < \infty$.

Denote by $\lambda_1, ..., \lambda_n$ the eigenvalues of the matrix $n^{-1/2}\mathbf{X}_n$ and define empirical spectral measure by

$$\mu_n(B) = \frac{1}{n}\#\{1 \leq i \leq n : \lambda_i \in B\}, \quad B \in \mathcal{B}(\mathbb{C}).$$

We say that the sequence of random probability measures $m_n(\cdot)$ converges weakly in probability to probability measure $m(\cdot)$ if for all continues and bounded functions $f : \mathbb{C} \to \mathbb{C}$ and all $\varepsilon > 0$

$$\lim_{n\to\infty} \mathbb{P}\left(\left|\int_{\mathbb{C}} f(x) m_n(dz) - \int_{\mathbb{C}} f(x) m(dz)\right| > \varepsilon\right) = 0.$$

We denote weak convergence by symbol $\xrightarrow{weak}$.

In this paper we consider the case $|\rho| < 1$. The main result of the paper is following

Theorem 1.1. (Elliptic Law) *Let $\mathbf{X}_n$ satisfies condition (**C0**) and $|\rho| < 1$. Then $\mu_n \xrightarrow{weak} \mu$ in probability, and μ has a density g:*

$$g(x, y) = \begin{cases} \frac{1}{\pi(1-\rho^2)}, & x, y \in \mathcal{E}, \\ 0, & otherwise, \end{cases}$$

where

$$\mathcal{E} := \left\{x, y \in \mathbb{R} : \frac{x^2}{(1+\rho)^2} + \frac{y^2}{(1-\rho)^2} \leq 1\right\}.$$

Figure 1 illustates Elliptic law.

In 1985 Girko proved elliptic law for rather general ensembles of random matrices under assumption that matrix elements have a density, see [7] and [8]. Girko used method of characteristic functions. Using V-transform he reduced problem to the problem for Hermitian matrices $(n^{-1/2}\mathbf{X}_n - z\mathbf{I})^*(n^{-1/2}\mathbf{X}_n - z\mathbf{I})$ and established convergence of empirical spectral distribution of singular values of $n^{-1/2}\mathbf{X}_n - z\mathbf{I}$ to the limit which determines the elliptic law.

Let elements of real asymmetric random matrix $\mathbf{X}$ have Gaussian distribution with zero mean and correlations

$$\mathbb{E}\, X_{ij}^2 = 1 \text{ and } \mathbb{E}\, X_{ij} X_{ij} = \rho, \quad i \neq j, \quad |\rho| < 1.$$

The ensemble of such matrices can be specified by the probability measure

$$\mathbb{P}(dX) \sim \exp\left[-\frac{n}{2(1-\rho^2)} \operatorname{Tr}(XX^T - \rho X^2)\right].$$

It was proved that $\mu_n \xrightarrow{weak} \mu$, where μ has a density from Theorem 1.1, see [14]. We will use this result to prove Theorem 1.1 in the general case.

Remark 1.2. *This result can be generalized to an ensemble of Gaussian complex asymmetric matrices. In this case, the invariant measure is*

$$\mathbb{P}(dX) \sim \exp\left[-\frac{n}{1-|\rho|^2} \operatorname{Tr}(XX^T - 2\operatorname{Re}\rho X^2)\right]$$

and $\mathbb{E}\,|X_{ij}|^2 = 1, \mathbb{E}\, X_{ij} X_{ji} = |\rho| e^{2i\theta}$ *for* $i \neq j$. *Then the limit measure has a uniform density inside an ellipse which is centered at zero and has semiaxes* $1 + |\rho|$ *in the direction* θ *and* $1 - |\rho|$ *in the direction* $\theta + \pi/2$.

For the discussion of elliptic law in Gaussian case see also [6], [1, Chapter 18] and [10].

We repeat physical motivation of models of random matrices which satisfy condition (**C0**) from [14]: *"The statistical properties of random asymmetric matrices may be important in the understanding of the behavior of certain dynamical systems far from equilibrium. One example is the dynamics of neural networks. A simple dynamic model of neural network consists of n continues "scalar" degrees of freedom("neurons") obeying coupled nonlinear differential equations ("circuit equations"). The coupling between the neurons is given by a synaptic matrix* $\mathbf{X}$ *which, in general, is asymmetric and has a substantial degree of disorder. In this case, the eigenstates of the synaptic matrix play an important role in the dynamics particulary when the neuron nonlinearity is not big".*

It will be interesting to prove Theorem 1.1 only under assumption of finite second moment and prove sparse analogs. It is the direction of our further research.

If $\rho = 0$ we assume that all entries of $\mathbf{X}_n$ are independent random variables and Circular law holds (see [2], [15], [9]):

Theorem 1.3. (Circular law) *Let* $\mathbf{X}_n$ *be a random matrix with independent identically distributed entries,* $\mathbb{E}\, X_{ij} = 0$ *and* $\mathbb{E}\, X_{ij}^2 = 1$. *Then* $\mu_n \xrightarrow{weak} \mu$ *in probability, and* μ *has uniform density on the unit circular.*

See Figure 1 d) for illustration of Circulaw law.

If $\rho = 1$ then matrix $\mathbf{X}_n$ is symmetric and its eigenvalues are real numbers. In this case the next theorem is known as a Wigner's semi-circular law (see [2]):

Theorem 1.4. (Semi-circular law) *Let $\mathbf{X}_n$ be a symmetric random matrix with independent identically distributed entries for $i \geq j$, $\mathbb{E}\, X_{ij} = 0$, $\mathbb{E}\, X_{ij}^2 = 1$. Then $\mu_n \xrightarrow{weak} \mu$ in probability, and μ has a density g:*

$$g(x) = \begin{cases} \frac{1}{2\pi}\sqrt{4 - x^2}, & -2 \leq x \leq 2, \\ 0, & \textit{otherwise.} \end{cases}$$

Throughout this paper we assume that all random variables are defined on common probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and we will write almost surely (a.s) instead of $\mathbb{P}$-almost surely. For vector $x = (x_1, ..., x_n)$ denote $||x||_2 := (\sum_{i=1}^n x_i^2)^{1/2}$ and $||x||_3 := (\sum_{i=1}^3 |x_i|^3)^{1/3}$. We denote unit sphere and unit ball by $S^{n-1} := \{x : ||x||_2 = 1\}$ and $B_1^n := \{x : ||x||_2 \leq 1\}$ respectively. For matrix $\mathbf{A}$ define spectral norm by $||A|| := \sup_{x:||x||_2=1} ||Ax||_2$ and Hilbert-Schmidt norm by $||A||_{HS} := (\mathrm{Tr}(A^*A))^{1/2}$. By $[n]$ we mean the set $\{1, ..., n\}$.

2. Proof of the main result

Further we will need the definition of logarithmic potential (see [12]) and uniform integrability of function with respect to the sequence of probability measures.

Definition 2.1. *The logarithmic potential U_m of measure $m(\cdot)$ is a function $U_m : \mathbb{C} \to (-\infty, +\infty]$ defined for all $z \in \mathbb{C}$ by*

$$U_m(z) = -\int_{\mathbb{C}} \log|z - w| m(dw).$$

Definition 2.2. *(uniform integrability) The function $f : \mathbb{R} \to \mathbb{R}$ is uniformly integrable in probability with respect to the sequence of random measures $\{m_n\}_{n\geq 1}$ on $\mathbb{R}$ if for all $\varepsilon > 0$:*

$$\lim_{t\to\infty} \overline{\lim_{n\to\infty}}\, \mathbb{P}\left(\int_{|f|>t} |f(x)| m_n(dx) > \varepsilon \right) = 0.$$

Let $\sigma_1(n^{-1/2}\mathbf{X} - z\mathbf{I}) \geq \sigma_2(n^{-1/2}\mathbf{X} - z\mathbf{I}) \geq ... \geq \sigma_n(n^{-1/2}\mathbf{X} - z\mathbf{I})$ be singular values of $n^{-1/2}\mathbf{X}_n - z\mathbf{I}$ and

$$\nu_n(z, B) = \frac{1}{n}\#\{i \geq 1 : \sigma_i \in B\}, \quad B \in \mathcal{B}(\mathbb{R}) -$$

empirical spectral measure of singular values. We will omit argument z in notation of measure $\nu_n(z, B)$ if it doesn't confuse.

The convergence in the Theorem 1.1 will be proved via convergence of logarithmic potential of μ_n to the logarithmic potential of μ. We can rewrite logarithmic potential of μ_n via the logarithmic potential of measure ν_n by

$$U_{\mu_n}(z) = -\int_{\mathbb{C}} \log|z-w|\mu_n(dw) = -\frac{1}{n}\log\left|\det\left(\frac{1}{\sqrt{n}}\mathbf{X}_n - zI\right)\right|$$
$$= -\frac{1}{2n}\log\det\left(\frac{1}{\sqrt{n}}\mathbf{X}_n - zI\right)^*\left(\frac{1}{\sqrt{n}}\mathbf{X}_n - zI\right) = -\int_0^\infty \log x\nu_n(dx).$$

This allows us to consider Hermitian matrix $(n^{-1/2}\mathbf{X}_n - zI)^*(n^{-1/2}\mathbf{X}_n - zI)$ instead of asymmetric $n^{-1/2}\mathbf{X}$. To prove Theorem 1.1 we need the following

Lemma 2.3. *Let $(\mathbf{X}_n)_{n\geq 1}$ be a sequence of $n \times n$ random matrices. Suppose that for a.a. $z \in \mathbb{C}$ there exists a probability measure ν_z on $[0,\infty)$ such that*
a) $\nu_n \xrightarrow{weak} \nu_z$ as $n \to \infty$ in probability
b) log *is uniformly integrable in probability with respect to $\{\nu_n\}_{n\geq 1}$.*

Then there exists a probability measure μ such that
a) $\mu_n \to \mu$ as $n \to \infty$ in probability
b) for a.a. $z \in \mathbb{C}$

$$U_\mu(z) = -\int_0^\infty \log x\nu_z(dx).$$

Proof. See [3, Lemma 4.3] for the proof. □

Proof. (Proof of Theorem 1.1) Our aim is to prove convergence of ν_n to ν_z, uniform integrability of $\log(\cdot)$ with respect to $\{\nu_n\}_{n\geq 1}$ and show that ν_z determines elliptic law.

To prove uniform integrability we need bounds on the least singular value of $n^{-1/2}\mathbf{X} - z\mathbf{I}$. These bounds will follow from Theorem 3.1. By Theorem 4.1 we can conclude uniform integrability of $\log(\cdot)$.

In Theorem 5.2 it is proved that Stieltjes transform of $\mathbb{E}\,\nu_n$ converges to Stieltjes transform of some probability measure ν_z. By Lemma A.13 and Chebyshev's inequality we can conclude that Stieltjes transform of ν_n converges to Stieltjes transform of ν_z in probability.

By relation between measures and their Stieltjes transforms if follows that ν_n converges to ν_z in probability. The crucial fact is that ν_z is non-random and doesn't depend on the distribution of matrix entries. So we can apply result in the Gaussian case and it follows that ν_z determines the elliptic law. □

3. Least singular value

Let $s_k(\mathbf{A})$ be singular values of $\mathbf{A}$ arranged in the non-increasing order. From properties of the largest and the smallest singular values

$$s_1(\mathbf{A}) = ||A|| = \sup_{x:||x||_2=1} ||\mathbf{A}x||_2, \quad s_n(\mathbf{A}) = \inf_{x:||x||_2=1} ||\mathbf{A}x||_2.$$

To prove uniform integrability of $\log(\cdot)$ we need to estimate probability of the event $\{s_n(\mathbf{A}) \le \varepsilon n^{-1/2}, ||\mathbf{X}|| \le K\sqrt{n}\}$, where $\mathbf{A} = \mathbf{X} - z\mathbf{I}$. We can assume that $\varepsilon n^{-1/2} \le K n^{1/2}$. If $|z| \ge 2K\sqrt{n}$ then probability of the event is automatically zero. So we can consider the case when $|z| \le 2Kn^{1/2}$. We have $||\mathbf{A}|| \le ||\mathbf{X}|| + |z| \le 3Kn^{1/2}$. In this section we prove theorem

Theorem 3.1. *Let* $\mathbf{A} = \mathbf{X} - z\mathbf{I}$, *where* $\mathbf{X}$ *is* $n \times n$ *random matrix satisfying* (**C0**). *Let* $K > 1$. *Then for every* $\varepsilon > 0$ *one has*

$$\mathbb{P}(s_n(\mathbf{A}) \le \varepsilon n^{-1/2}, ||\mathbf{A}|| \le 3K\sqrt{n}) \le C(\rho)\varepsilon^{1/8} + C_1(\rho) n^{-1/8},$$

where $C(\rho), C_1(\rho)$ *are some constants which can depend only on* ρ, K *and* M_4.

Remark 3.2. *Mark Rudelson and Roman Vershynin in [11] and Roman Vershynin in [16] found bounds for the least singular value of matrices with independent entries and symmetric matrices respectively. In this section we will follow their ideas.*

3.1. **The small ball probability via central limit theorem.** We recall definition of Levy concentration function

Definition 3.3. *Levy concentration function of random variable* Z *with values from* $\mathbb{R}^d$ *is a function*

$$\mathcal{L}(Z, \varepsilon) = \sup_{v \in \mathbb{R}^d} \mathbb{P}(||Z - v||_2 < \varepsilon).$$

The next statement gives the bound for Levy concentration function of sum of independent random variables in $\mathbb{R}$.

Statement 3.4. *Let* $\{a_i\xi_i + b_i\eta_i\}_{i \ge 1}$ *be independent random variables,* $\mathbb{E}\,\xi_i = \mathbb{E}\,\eta_i = 0$, $\mathbb{E}\,\xi_i^2 \ge 1, \mathbb{E}\,\eta_i^2 \ge 1$, $\mathbb{E}\,\xi_i\eta_i = \rho$, $\max(\mathbb{E}\,\xi_i^4, \mathbb{E}\,\eta_i^4) \le M_4$, $a_i^{-1}b_i = O(1)$. *We assume that* $c_1(2n)^{-1/2} \le |a_i| \le (c_0 n)^{-1/2}$, *where* c_0, c_1 *are some constants. Then*

$$\mathcal{L}\left(\sum_{i=1}^n (a_i\xi_i + b_i\eta_i), \varepsilon\right) \le \frac{C\varepsilon}{(1-\rho^2)^{1/2}} + \frac{C_1}{(1-\rho^2)^{3/2} n^{1/2}}.$$

Proof. Set $\sigma_{i1}^2 = \mathbb{E}\,\xi_i^2$ and $\sigma_{i2}^2 = \mathbb{E}\,\eta_i^2$. It is easy to see that

$$\sigma^2 = \mathbb{E}(\sum_{i=1}^n Z_i)^2 = \sum_{i=1}^n |a_i|^2(\sigma_{i1}^2 + 2\rho\sigma_{i1}\sigma_{i2}a_i^{-1}b_i + \sigma_{i2}^2(a_i^{-1}b_i)^2) \ge (1-\rho^2)\sum_{i=1}^n \sigma_{i1}^2|a_i|^2$$

and

$$\sum_{i=1}^n \mathbb{E}\,|a_i\xi_i + b_i\eta_i|^3 \le \sum_{i=1}^n |a_i|^3\,\mathbb{E}\,|\xi_i + a_i^{-1}b_i\eta_i|^3 \le C' M_4 ||a||_3^3,$$

where we have used the fact $a_i^{-1}b_i = O(1)$. By Central Limit Theorem A.1 for arbitrary vector $v \in \mathbb{R}$

$$\mathbb{P}\left(\left|\sum_{i=1}^n (a_i\xi_i + b_i\eta_i) - v\right| \le \varepsilon\right) \le \mathbb{P}\left(|g' - v| \le \varepsilon\right) + C'' \frac{\sum_{i=1}^n \mathbb{E}\,|a_i\xi_i + b_i\eta_i|^3}{\sigma^3},$$

where g' has gaussian distribution with zero mean and variance σ^2. The density of g' is uniformly bounded by $1/\sqrt{2\pi\sigma^2}$. We have

$$\mathbb{P}\left(|\sum_{i=1}^{n}(a_i\xi_i + b_i\eta_i) - v| \le \varepsilon \right) \le \frac{C\varepsilon}{(1-\rho^2)^{1/2}} + \frac{C_1}{(1-\rho^2)^{3/2}n^{1/2}}.$$

We can take maximum and conclude the statement. □

Remark 3.5. *Let us consider the case $b_i = 0$ for all $i \ge 1$. It is easy to show that*

$$\mathcal{L}\left(\sum_{i=1}^{n} a_i\xi_i, \varepsilon \right) \le C(\varepsilon + n^{-1/2}).$$

3.2. **Decomposition of the sphere.** To prove Theorem 3.1, we shall partition the unit sphere S^{n-1} into the two sets of compressible and incompressible vectors, and show the invertibility of A on each set separately.

Definition 3.6. *(Compressible and incompressible vectors) Let $c_0, c_1 \in (0,1)$. A vector $x \in \mathbb{R}^n$ is called sparse if $|supp(x)| \le c_0 n$. A vector $x \in S^{n-1}$ is called compressible if x is within Euclidian distance c_1 from the set of all sparse vectors. A vector $x \in S^{n-1}$ is called incompressible if it is not compressible. The sets of sparse, compressible and incompressible vectors will be denoted by Sparse = Sparse (c_0), Comp = Comp (c_0, c_1) and Incomp = Incomp(c_0, c_1) respectively.*

3.3. **Invertibility for the compressible vectors.** We first estimate $||Ax||$ for a fixed vector $x \in S^{n-1}$. The next two statements can be found in [16], but for the readers convenience we prove them here

Statement 3.7. *Let A be a random matrix from the Theorem 3.1. Then for all $x \in S^{n-1}$*

$$\mathcal{L}(\mathbf{A}x, c_3\sqrt{n}) \le 2e^{-c_3 n}.$$

Proof. Let us decompose the set of indices $[n]$ into two sets of roughly equal sizes, $\{1, ..., n_0\}$ and $\{n_0+1, ..., n\}$ where $n_0 = [n/2]$.

We fix arbitrary vector $u \in \mathbb{R}$ and decompose matrix $\mathbf{A}$ and vectors x, u into blocks

$$\mathbf{A} = \begin{pmatrix} \mathbf{E} & \mathbf{F} \\ \mathbf{G} & \mathbf{H} \end{pmatrix} \quad x = \begin{pmatrix} y \\ z \end{pmatrix} \quad u = \begin{pmatrix} v \\ w \end{pmatrix}. \tag{3.1}$$

We have $||\mathbf{A}x - u||_2^2 = ||\mathbf{E}y + \mathbf{F}z - v||_2^2 + ||\mathbf{G}y + \mathbf{H}z - w||_2^2$. Let us consider the first term $||\mathbf{E}y + \mathbf{F}z - v||_2^2$. Conditioning on an arbitrary realization of $\mathbf{E}$ and $\mathbf{H}$ we express

$$||\mathbf{E}y + \mathbf{F}z - v||_2^2 = ||\mathbf{F}z - v||_2^2 = \sum_{i=1}^{n_0}((\mathbf{F}_i, z) - d_i)^2,$$

where $\mathbf{F}_i$ is a row of $\mathbf{F}$ and d_i is a coordinate of $\mathbf{E}y-v$. For each i we can observe that $(\mathbf{F}_i, z)$ is a sum of independent random variables and by Lemma A.9

$$\sup_{d_i \in \mathbb{R}} \mathbb{P}(|(\mathbf{F}_i, z/||z||_2) - d_i| \le \frac{1}{2}) \le p_0 \in (0,1).$$

Now we can use Lemma A.10

$$\sup_{a \in \mathbb{R}^{n_0}} \mathbb{P}(||\mathbf{F}z - a||_2 \le \tau_0 \sqrt{n_0} ||z||_2) \le p_1^{n_0}.$$

Since $\mathbf{E}y - v$ is a fixed vector, this implies

$$\mathbb{P}(||\mathbf{E}y + \mathbf{F}z - v||_2 \le \tau_0 ||z||_2 \sqrt{n_0}) \le p_1^{n_0}.$$

Since this holds conditionally on an arbitrary realization of $\mathbf{E}$ and $\mathbf{H}$, it at also holds unconditionally.

By a similar argument we obtain that

$$\mathbb{P}(||\mathbf{G}y + \mathbf{H}z - w||_2 \le \tau_0 ||y||_2 \sqrt{n - n_0}) \le p_1^{n-n_0}.$$

Since $n_0 \ge n/2$ and $n - n_0 \ge n/3$ and $||y||_2^2 + ||z||_2^2 = 1$ we have $\tau_0^2 ||z||_2^2 n_0 + \tau_0^2 ||y||_2^2 (n - n_0) > \frac{1}{3}\tau_0^2 n$. By union bound, we conclude that

$$\mathbb{P}(||\mathbf{A}x - u||_2 \le \frac{1}{3}\tau_0^2 n^{1/2}) \le 2p_1^{n/3}.$$

□

Lemma 3.8. *Let $\mathbf{A}$ be a matrix from Theorem 3.1 and let $K > 1$. There exist constants $c_0, c_1, c_4 \in (0,1)$ that depend only on K and M_4 and such that the following holds. For every $u \in \mathbb{R}^n$, one has*

$$\mathbb{P}(\inf_{\frac{x}{||x||_2} \in Comp(c_0,c_1)} ||\mathbf{A}x - u||_2/||x||_2 \le c_4\sqrt{n}, ||\mathbf{A}|| \le 3K\sqrt{n}) \le 2e^{-c_4 n}. \tag{3.2}$$

Proof. Let us fix some small values of c_0, c_1 and c_4. According to Lemma A.4, there exists a $(2c_1)$-net $\mathcal{N}$ of the set $\mathrm{Comp}(c_0, c_1)$ such that

$$|\mathcal{N}| \le (9/c_0 c_1)^{c_0 n}.$$

Let $\mathcal{E}$ denote the event in the left hand side of (3.2) whose probability we would like to bound. Assume that $\mathcal{E}$ holds. Then there exist vectors $x_0 := x/||x||_2 \in Comp(c_0, c_1)$ and $u_0 := u/||x||_2 \in \mathrm{span}(u)$ such that

$$||\mathbf{A}x_0 - u||_2 \le c_4\sqrt{n}. \tag{3.3}$$

By definition of $\mathcal{N}$, there exists $y_0 \in \mathcal{N}$ such that

$$||x_0 - y_0||_2 \le 2c_1. \tag{3.4}$$

By norm inequality

$$||\mathbf{A}y_0||_2 \le ||\mathbf{A}|| \le 3K\sqrt{n}. \tag{3.5}$$

From (3.3) and (3.4) it follows

$$||\mathbf{A}y_0 - u_0||_2 \le ||\mathbf{A}|| ||x_0 - y_0||_2 + ||\mathbf{A}x_0 - u_0||_2 \le 6c_1 K\sqrt{n} + c_4\sqrt{n}. \tag{3.6}$$

This and (3.5) yield

$$||u||_0 \le 3K\sqrt{n} + 6c_1K\sqrt{n} + c_4\sqrt{n} \le 10K\sqrt{n}.$$

We see that

$$u_0 \in \mathrm{span}(u) \cap 10K\sqrt{n}B_1^n =: E.$$

Let $\mathcal{M}$ be some fixed $(c_1K\sqrt{n})$-net of the interval E, such that

$$|\mathcal{M}| \le \frac{20K\sqrt{n}}{c_1K\sqrt{n}} = \frac{20}{c_1}.$$

Let us choose a vector $v_0 \in \mathcal{M}$ such that $||u_0 - v_0|| \le c_1K\sqrt{n}$. It follows from (3.6) that

$$||\mathbf{A}y_0 - v_0||_2 \le 6c_1K\sqrt{n} + c_4\sqrt{n} + c_1K\sqrt{n} = (7c_1K + c_4)\sqrt{n}.$$

Choose values of $c_1, c_4 \in (0,1)$ so that $7c_1K + c_4 \le c_3$, where c_3 is a constant from Statement 3.7. We have that the event $\mathcal{E}$ implies the existence of vectors $y_0 \in \mathcal{N}$ and $v_0 \in \mathcal{M}$ such that $||\mathbf{A}y_0 - v_0||_2 \le c_3\sqrt{n}$. Taking the union bound over $\mathcal{N}$ and $\mathcal{M}$, we conclude that

$$\mathbb{P}(\mathcal{E}) \le |\mathcal{N}||\mathcal{M}| \max_{y_0\in\mathcal{N}, v_0\in\mathcal{M}} \mathbb{P}(||Ay_0 - v_0||_2 \le c_3\sqrt{n}).$$

Applying Statement 3.7 and using on the cardinalities of the nets $\mathcal{N}, \mathcal{M}$ we obtain

$$\mathbb{P}(\mathcal{E}) \le \left(\frac{9}{c_0c_1}\right)^{c_0n} \frac{20}{c_1} \cdot 2e^{-c_3n}.$$

We can choose c_0 small enough to ensure that

$$\mathbb{P}(\mathcal{E}) \le 2e^{-\frac{c_3}{2}n}.$$

□

3.4. **Invertibility for the incompressible vectors.** For the incompressible vectors, we shall reduce the invertibility problem to a lower bound on the distance between a random vector and a random hyperplane. For this aim we recall Lemma 3.5 from [11]

Lemma 3.9. *Let* $\mathbf{A}$ *be a random matrix from theorem. Let* $A_1, ..., A_n$ *denote the column vectors of* A*, and let* H_k *denote the span of all columns except the* k*-th. Then for every* $c_0, c_1 \in (0,1)$ *and every* $\varepsilon > 0$*, one has*

$$\mathbb{P}\Big(\inf_{x\in Incomp(c_0,c_1)} ||\mathbf{A}x||_2 < \varepsilon n^{-1}\Big) \le \frac{1}{c_0n}\sum_{k=1}^{n} \mathbb{P}(\mathrm{dist}(A_k, H_k) < c_1^{-1}\varepsilon). \tag{3.7}$$

3.5. **Distance via the small ball probability.** Lemma 3.9 reduces the invertibility problem to a lower bound on the distance between a random vector and a random hyperplane.

We decompose matrix $\mathbf{A} = \mathbf{X} - z\mathbf{I}$ into the blocks

$$\begin{pmatrix} a_{11} & V^T \\ U & \mathbf{B} \end{pmatrix} \tag{3.8}$$

where $\mathbf{B}$ is $(n-1)\times(n-1)$ matrix, $U, V \in \mathbb{R}^{n-1}$.

Let h be any unit vector orthogonal to $A_2, ..., A_n$. It follows that

$$0 = \begin{pmatrix} V^T \\ \mathbf{B} \end{pmatrix}^T h = h_1 V + \mathbf{B}^T g,$$

where $h = (h_1, g)$, and

$$g = -h_1 \mathbf{B}^{-T} V$$

From definition of h

$$1 = ||h||_2^2 = |h_1|^2 + ||g||_2^2 = |h_1|^2 + |h_1|^2 ||\mathbf{B}^{-T} V||_2^2$$

Using this equations we estimate distance

$$\operatorname{dist}(A_1, H) \geq |(A_1, h)| = \frac{|a_{11} - (\mathbf{B}^{-T} V, U)|}{\sqrt{1 + ||\mathbf{B}^{-T} V||_2^2}}$$

It is easy to show that $||\mathbf{B}|| \leq ||\mathbf{A}||$. Let vector $e_1 \in S^{n-2}$ be such that $||\mathbf{B}|| = ||\mathbf{B} e_1||_2$. Then we can take vector $e = (0, e_1)^T \in S^{n-1}$ and for this vector

$$||\mathbf{A}|| \geq ||\mathbf{A} e||_2 = ||(V^T e_1, \mathbf{B} e_1)^T||_2 \geq ||\mathbf{B} e_1||_2 = ||\mathbf{B}||.$$

The bound for right hand sand of (3.7) will follow from the

Lemma 3.10. *Let matrix* $\mathbf{A}$ *be from Theorem 3.1. Then for all* $\varepsilon > 0$

$$\sup_{v \in \mathbb{R}} \mathbb{P}\left(\frac{|(\mathbf{B}^{-T} V, U) - v|}{\sqrt{1 + ||\mathbf{B}^{-T} V||_2^2}} \leq \varepsilon, ||\mathbf{B}|| \leq 3K\sqrt{n} \right) \leq C(\rho)\varepsilon^{1/8} + C'(\rho) n^{-1/8}, \tag{3.9}$$

where $\mathbf{B}, U, V$ *are determined by* (3.8) *and* $C(\rho), C_1(\rho)$ *are some constants which can depend only on* ρ, K *and* M_4.

To get this bound we need several statements. We introduce matrix

$$\mathbf{Q} = \begin{pmatrix} \mathbf{O}_{n-1} & \mathbf{B}^{-T} \\ \mathbf{B}^{-1} & \mathbf{O}_{n-1} \end{pmatrix} \quad W = \begin{pmatrix} U \\ V \end{pmatrix}, \tag{3.10}$$

where $\mathbf{O}_{n-1}$ is $(n-1)\times(n-1)$ matrix with zero entries. Scalar product in (3.9) can be rewritten using definition of Q:

$$\sup_{v \in \mathbb{R}} \mathbb{P}\left(\frac{|(\mathbf{Q} W, W) - v|}{\sqrt{1 + ||\mathbf{B}^{-T} V||_2^2}} \leq 2\varepsilon \right). \tag{3.11}$$

Introduce vectors

$$W' = \begin{pmatrix} U' \\ V' \end{pmatrix} \quad Z = \begin{pmatrix} U \\ V' \end{pmatrix}, \tag{3.12}$$

where U', U' are independent copies of U, V respectively. We need the following

Statement 3.11.

$$\sup_{v \in \mathbb{R}} \mathbb{P}_W \left(|(\mathbf{Q}W, W) - v| \le 2\varepsilon \right) \le \mathbb{P}_{W,W'} \left(|(\mathbf{Q}\mathbf{P}_{J^c}(W - W'), \mathbf{P}_J W) - u| \le 2\varepsilon \right),$$

where u doesn't depend on $\mathbf{P}_J W = (\mathbf{P}_J U, \mathbf{P}_J V)^T$.

Proof. Let us fix v and denote

$$p := \mathbb{P}\left(|(\mathbf{Q}W, W) - v| \le 2\varepsilon \right).$$

We can decompose the set $[n]$ into union $[n] = J \cup J^c$. We can take $U_1 = \mathbf{P}_J U, U_2 = \mathbf{P}_{J^c}, V_1 = \mathbf{P}_J V$ and $V_2 = \mathbf{P}_{J^c} V$. By Lemma A.2

$$\begin{aligned} p^2 &\le \mathbb{P}\left(|(\mathbf{Q}W, W) - v| \le 2\varepsilon, |(\mathbf{Q}Z, Z) - v| \le 2\varepsilon \right) \\ &\le \mathbb{P}\left(|(\mathbf{Q}W, W) - (\mathbf{Q}Z, Z)| \le 4\varepsilon \right). \end{aligned} \tag{3.13}$$

Let us rewrite $\mathbf{B}^{-T}$ in the block form

$$\mathbf{B}^{-T} = \begin{pmatrix} \mathbf{E} & \mathbf{F} \\ \mathbf{G} & \mathbf{H} \end{pmatrix}.$$

We have

$$\begin{aligned} &(\mathbf{Q}W, W) = (\mathbf{E}V_1, U_1) + (\mathbf{F}V_2, U_1) + (\mathbf{G}V_1, U_2) + (\mathbf{H}V_2, U_2) \\ &+ (\mathbf{E}^T U_1, V_1) + (\mathbf{G}^T U_2, V_1) + (\mathbf{F}^T U_1, V_2) + (\mathbf{H}^T U_2, V_2) \\ &(\mathbf{Q}Z, Z) = (\mathbf{E}V_1, U_1) + (\mathbf{F}V_2', U_1) + (\mathbf{G}V_1, U_2') + (\mathbf{H}V_2', U_2') \\ &+ (\mathbf{E}^T U_1, V_1) + (\mathbf{G}^T U_2', V_1) + (\mathbf{F}^T U_1, V_2') + (\mathbf{H}^T U_2', V_2') \end{aligned}$$

and

$$\begin{aligned} (\mathbf{Q}W, W) - (\mathbf{Q}Z, Z) &= 2(\mathbf{F}(V_2 - V_2'), U_1) + 2(\mathbf{G}^T(U_2 - U_2'), V_1) \\ &+ 2(\mathbf{H}V_2, V_2) - 2(\mathbf{H}V_2', V_2'). \end{aligned} \tag{3.14}$$

The last two terms in (3.14) depend only on U_2, U_2', V_2, V_2' and we conclude that

$$p_1^2 \le \mathbb{P}\left(|(\mathbf{Q}P_{J^c}(W - W'), P_J W) - u| \le 2\varepsilon \right),$$

where $u = u(U_2, V_2, U_2', V_2', \mathbf{F}, \mathbf{G}, \mathbf{H})$. □

Statement 3.12. *For all $u \in \mathbb{R}^{n-1}$*

$$\mathbb{P}\left(\frac{\mathbf{B}^{-T}u}{||\mathbf{B}^{-T}u||_2} \in Comp(c_0, c_1) \text{ and } ||\mathbf{B}|| \le 3Kn^{1/2} \right) \le 2e^{-c_4' n}.$$

Proof. Let $x = \mathbf{B}^{-T}u$. It is easy to see that

$$\left\{ \frac{\mathbf{B}^{-T}u}{||\mathbf{B}^{-T}u||_2} \in Comp(c_0, c_1) \right\} \subseteq \left\{ \exists x : \frac{x}{||x||_2} \in Comp(c_0, c_1) \text{ and } \mathbf{B}^T x = u \right\}$$

Replacing matrix $\mathbf{A}$ with $\mathbf{B}^T$ one can easily check that the proof of Lemma 3.8 remains valid for $\mathbf{B}^T$ as well as for $\mathbf{A}$. □

Remark 3.13. *The Statement 3.12 holds true for* $\mathbf{B}^{-T}$ *replaced with* $\mathbf{B}^{-1}$.

Statement 3.14. *Let* $\mathbf{A}$ *satisfies condition* $(\mathbf{C0})$ *and* $\mathbf{B}$ *be a matrix from decomposition* (3.8). *Assume that* $||\mathbf{B}|| \le 3K\sqrt{n}$. *Then with probability at least* $1 - e^{-cn}$ *matrix* $\mathbf{B}$ *has the following properties:*

a) $||\mathbf{B}^{-T}V||_2 \ge C$ *with probability* $1 - e^{-c_5 n}$ *in* W,

b) $||\mathbf{B}^{-T}V||_2 \le \varepsilon^{-1/2}||\mathbf{B}^{-T}||_{HS}$ *with probability* $1 - \varepsilon$ *in* V,

c) $||\mathbf{Q}W||_2 \ge \varepsilon||\mathbf{B}^{-T}||_{HS}$ *with probability* $1 - C(\varepsilon + n^{-1/2})$ *in* W.

Proof. Let $\{e_k\}_{k=1}^n$ be a standard basis in $\mathbb{R}^n$. For all $1 \le k \le n$ define vectors by

$$x_k := \frac{\mathbf{B}^{-1}e_k}{||\mathbf{B}^{-1}e_k||}.$$

By Statement 3.12 vector x_k is incompressible with probability $1 - e^{-cn}$. We fix matrix $\mathbf{B}$ with such property.

a) By norm inequality $||U||_2 \le ||\mathbf{B}||_2||\mathbf{B}^{-T}U||_2$. We know that $||\mathbf{B}|| \le 3K\sqrt{n}$. By Lemma A.8 and Lemma A.10 $||U|| \ge \sqrt{n}$. So we have that $||\mathbf{B}^{-1}U|| \ge C$ with probability $1 - e^{-c_5 n}$.

b) By definition

$$||\mathbf{B}^{-T}V||_2^2 = \sum_{i=1}^n (\mathbf{B}^{-1}e_k, V)^2 = \sum_{i=1}^n ||\mathbf{B}^{-1}e_i||_2^2 (x_k, V)^2.$$

It is easy to see that $\mathbb{E}(V, x_k)^2 = 1$. So

$$\mathbb{E}\,||\mathbf{B}^{-T}V||_2^2 = \sum_{i=1}^n ||\mathbf{B}^{-1}e_i||_2^2 = ||\mathbf{B}^{-1}||_{HS}^2.$$

By Markov inequality

$$\mathbb{P}(||\mathbf{B}^{-T}V||_2 \ge \varepsilon^{-1/2}||\mathbf{B}^{-1}||_{HS}) \le \varepsilon.$$

c) By Lemma A.3, Lemma A.5, Lemma A.7 and Remark 3.5

$$\mathbb{P}(||\mathbf{Q}W||_2 \le \varepsilon||\mathbf{B}^{-1}||_{HS}) \le \mathbb{P}(||\mathbf{B}^{-T}V||_2 \le \varepsilon||B^{-1}||_{HS})$$
$$= \mathbb{P}(||\mathbf{B}^{-T}V||_2^2 \le \varepsilon||\mathbf{B}^{-1}||_{HS}^2) = \mathbb{P}(\sum_{i=1}^n ||\mathbf{B}^{-1}e_i||_2 (x_i, V)^2 \le \varepsilon^2||\mathbf{B}^{-1}||_{HS}^2)$$
$$= \mathbb{P}(\sum_{i=1}^n p_i (x_i, V)^2 \le \varepsilon^2) \le 2\sum_{i=1}^n p_i\,\mathbb{P}((x_i, V) \le \sqrt{2}\varepsilon) \le C(\varepsilon + n^{-1/2}).$$

□

Proof. (proof of Lemma 3.10) We are going to construct random set J. We can consider independent Bernoulli random variables $\xi_1, ..., \xi_n$ with $\mathbb{E}\,\xi_i = c_{00}/2$. We then define $J := \{i : \xi_i = 0\}$. From large deviation inequalities we may conclude that

$$|J^c| \le c_{00} n \tag{3.15}$$

holds with probability at least $1 - 2\exp(-c_{00}^2 n/2)$. Now we find the probability of event

$$\varepsilon_0^{1/2}\sqrt{1 + ||\mathbf{B}^{-T}V||_2^2} \le ||\mathbf{B}^{-1}||_{HS} \le \varepsilon_0^{-1}||\mathbf{Q}\mathbf{P}_{J^c}(W - W')||_2. \tag{3.16}$$

From Statement 3.14 we can conclude that

$$\mathbb{P}_{\mathbf{B},W,W',J}(3.16 \text{ holds} \cup ||B|| \ge 3K\sqrt{n}) \ge 1 - C'(\varepsilon_0 + n^{-1/2}) - 2e^{-c'n}.$$

Consider the random vector

$$w_0 = \frac{1}{||\mathbf{Q}\mathbf{P}_{J^c}(W - W')||_2}\begin{pmatrix}\mathbf{B}^{-T}\mathbf{P}_{J^c}(V - V') \\ \mathbf{B}^{-1}\mathbf{P}_{J^c}(U - U')\end{pmatrix} = \begin{pmatrix} a \\ b \end{pmatrix}.$$

By Statement 3.12 we know that event

$$a \in incomp(c_0, c_1) \tag{3.17}$$

holds with probability

$$\mathbb{P}_{\mathbf{B}}(a \in incomp(c_0, c_1) \cup ||\mathbf{B}|| \ge 3K\sqrt{n}|W, W', J) \ge 1 - 2\exp(-c''n).$$

Combining these probabilities we have

$$\begin{aligned}&\mathbb{P}_{\mathbf{B},W,W',J}(\text{ (3.15), (3.16) and (3.17) hold} \cup ||\mathbf{B}|| \ge 3K\sqrt{n})\\ &\ge 1 - 2e^{-c_{00}^2 n/2} - C'(\varepsilon_0 + n^{-1/2}) - 2e^{-c'n} - 2e^{-c''n} := 1 - p_0.\end{aligned}$$

We may fix J that satisfies (3.15) and

$$\mathbb{P}_{\mathbf{B},W,W'}(\text{(3.16), (3.17) hold} \cup ||\mathbf{B}|| \ge 3K\sqrt{n}) \ge 1 - p_0.$$

By Fubini's theorem $\mathbf{B}$ has the following property with probability at least $1 - \sqrt{p_0}$

$$\mathbb{P}_{W,W'}(\text{(3.16), (3.17) hold} \cup ||\mathbf{B}|| \ge 3K\sqrt{n}|\mathbf{B}) \ge 1 - \sqrt{p_0}.$$

The event $\{||\mathbf{B}|| \ge 3K\sqrt{n}\}$ depends only on B. We may conclude that random matrix $\mathbf{B}$ has the following property with probability at least $1 - \sqrt{p_0}$: either $||\mathbf{B}|| \ge 3K\sqrt{n}$, or

$$||\mathbf{B}|| \le 3K\sqrt{n} \text{ and } \mathbb{P}_{W,W'}(\text{(3.16), (3.17) hold }|\mathbf{B}) \ge 1 - \sqrt{p}_0 \tag{3.18}$$

The event we are interested in is

$$\mathcal{E} := \left(\frac{|(\mathbf{Q}W, W) - u|}{\sqrt{1 + ||B^{-T}V||_2^2}} \le 2\varepsilon\right).$$

We need to estimate probability

$$\mathbb{P}_{\mathbf{B},W}(\mathcal{E} \cap ||\mathbf{B}|| \le 3K\sqrt{n}) \le \mathbb{P}_{\mathbf{B},W}(\mathcal{E} \cap \text{ (3.18) holds}) + \mathbb{P}_{\mathbf{B},W}(\mathcal{E} \cap \text{ (3.18) fails}).$$

The last term is bounded by $\sqrt{p_0}$.

$$\mathbb{P}_{\mathbf{B},W}(\mathcal{E} \cap ||\mathbf{B}|| \le 3K\sqrt{n}) \le \sup_{B \text{ satisfies (3.18)}} \mathbb{P}_W(\mathcal{E}|\mathbf{B}) + \sqrt{p_0}.$$

We can conclude that

$$\mathbb{P}_{\mathbf{B},W}(\mathcal{E} \cap ||\mathbf{B}|| \le 3K\sqrt{n}) \le \sup_{B \text{ satisfies (3.18)}} \mathbb{P}_{W,W'}(\mathcal{E}, (3.16) \text{ holds}|\mathbf{B}) + 2\sqrt{p_0}.$$

Let us fix $\mathbf{B}$ that satisfies (3.18) and denote $p_1 := \mathbb{P}_{W,W'}(\mathcal{E}, (3.16) \text{ holds}|\mathbf{B})$. By Statement 3.11 and the first property in (3.16) we have

$$p_1^2 \le \mathbb{P}_{W,W'}\left(\underbrace{|(\mathbf{QP}_{J^c}(W - W'), \mathbf{P}_J W) - v| \le \frac{\varepsilon}{\sqrt{\varepsilon_0}}||\mathbf{B}^{-1}||_{HS}}_{\mathcal{E}_1}\right)$$

and

$$\mathbb{P}_{W,W'}(\mathcal{E}_1) \le \mathbb{P}_{W,W'}(\mathcal{E}_1, (3.16), (3.17) \text{ hold }) + \sqrt{p_0}.$$

Further

$$p_1^2 \le \mathbb{P}_{W,W'}(|(w_0, \mathbf{P}_J W) - v| \le 2\varepsilon_0^{-3/2}\varepsilon, (3.17) \text{ holds}) + \sqrt{p_0}.$$

By definition random vector w_0 is determined by the random vector $\mathbf{P}_{J^c}(W - W')$, which is independent of the random vector $\mathbf{P}_J W$. We fix $\mathbf{P}_{J^c}(W - W')$ and have

$$p_1^2 \le \sup_{\substack{w_0 = (a,b)^T: \\ a \in Incomp(c_0, c_1) \\ w \in \mathbb{R}}} \mathbb{P}_{\mathbf{P}_J W}\left(|(w_0, \mathbf{P}_J W) - w| \le \varepsilon_0^{-3/2}\varepsilon\right) + \sqrt{p_0}.$$

Let us fix a vector w_0 and a number w. We can rewrite

$$(w_0, P_J W) = \sum_{i \in J}(a_i x_i + b_i y_i), \tag{3.19}$$

where $||a||_2^2 + ||b||_2^2 = 1$. From Lemma A.5 and Remark A.6 we know that at least $2c_{00}n$ coordinates of vector $a \in Incomp(c_0, c_1)$ satisfy

$$\frac{c_1}{\sqrt{2n}} \le |a_k| \le \frac{1}{\sqrt{c_0 n}}.$$

We denote the set of coordinates of a with this property by $\text{spread}(a)$. By construction of J we can conclude that $|\text{spread}(a)| = [c_{00}n]$. By Lemma A.7 we can reduce our sum (3.19) to the set $\text{spread}(a)$. Now we will find the properties of $|b_i|$. We can decompose the set $\text{spread}(a)$ into two sets $\text{spread}(a) = I_1 \cup I_2$:
a) $I_1 = \{i \in \text{spread}(a) : |b_i|\sqrt{n} \to \infty \text{ as } n \to \infty\}$;
c) $I_2 = \{i \in \text{spread}(a) : |b_i| = O(n^{-1/2})\}$;
From $||b||_2^2 < 1$ it follows that $|I_1| = o(n)$. For $i \in I_2$ we have $|a_i^{-1}b_i| = O(1)$. By Lemma A.7 we have

$$\mathbb{P}(|\sum_{i \in \text{spread}(a)}(a_i x_i + b_i y_i) - w| < 2\varepsilon_0^{-3/2}\varepsilon) \le \mathbb{P}(|\sum_{i \in I_2}(a_i x_i + b_i y_i) - w'| < 2\varepsilon_0^{-3/2}\varepsilon).$$

We can apply Statement 3.4

$$\mathbb{P}(|\sum_{i \in I_2}(a_i x_i + b_i y_i) - w'| < 2\varepsilon_0^{-3/2}\varepsilon) \le \frac{C_1\varepsilon_0^{-3/2}\varepsilon}{(1-\rho^2)^{1/2}} + C_2(1-\rho^2)^{-3/2}n^{-1/2}.$$

It follows that

$$\mathbb{P}_{\mathbf{B},W}(\mathcal{E} \cap ||B|| \le 3K\sqrt{n}) \le$$
$$\le \left(\frac{C_1 \varepsilon_0^{-3/2} \varepsilon}{(1-\rho^2)^{1/2}} + C_2 (1-\rho^2)^{-3/2} n^{-1/2} \right)^{1/2} + p_0^{1/4} + 2\sqrt{p_0}.$$

We take $\varepsilon_0 = \varepsilon^{1/2}$ and conclude that

$$\mathbb{P}_{\mathbf{B},W}(\mathcal{E} \cap ||B|| \le 3K\sqrt{n}) \le C(\rho)\varepsilon^{1/8} + C'(\rho) n^{-1/8},$$

where $C(\rho), C'(\rho)$ are some constants which depend on ρ, K and M_4. □

Proof. (proof of Theorem 3.1) The result of the theorem follows from Lemmas 3.8, 3.9 and 3.10. □

Remark 3.15. *It not very difficult to show that we can change matrix $z\mathbf{I}$ in Theorem 3.1 by arbitrary non-random matrix $\mathbf{M}$ with $||\mathbf{M}|| \le K\sqrt{n}$. We can also assume that $\mathbb{E} X_{ij}^2 \ge 1$. Results of section 3.3 are based on Lemmas A.9 and A.10 which doesn't depend on shifts. It is easy to see that Statement 3.14 still holds true if we assume that $\varepsilon < n^{-Q}$ for some $Q > 0$. Then we can reformulate Theorem 3.1 in the following way: there exist some constants $A, B > 0$ such that*

$$\mathbb{P}(s_n(\mathbf{X}+\mathbf{M}) \le n^{-A}, ||\mathbf{X}+\mathbf{M}|| \le K\sqrt{n}) \le C(\rho) n^{-B}.$$

4. Uniform integrability of logarithm

In this section we prove the next result

Theorem 4.1. *Under the condition* (**C0**) $\log(\cdot)$ *is uniformly integrable in probability with respect to* $\{\nu_n\}_{n\ge1}$.

Before we need several lemmas about the behavior of the singular values

Lemma 4.2. *Let elements $\mathbf{X}_n$ depend on n, but satisfy conditions* (**C0**) *and $|x_{ij}| \le \delta_n \sqrt{n}, \mathbb{E} x_{ij}^2 \le 1$ and $\mathbb{E}|x_{ij}|^l \le b(\delta_n\sqrt{n})^{l-1}$ for some $b > 0, l \ge 3$ and $\delta_n \to 0$ with the convergence rate slower that any preassigned one as $n \to \infty$. Then for some $K > 0$*

$$\mathbb{P}(\sigma_1(\mathbf{X}) \ge K\sqrt{n}) \le o(n^{-l}).$$

Proof. We can decompose matrix $\mathbf{X}$ into symmetric and skew-symmetric matrices:

$$\mathbf{X} = \frac{\mathbf{X}+\mathbf{X}^T}{2} + \frac{\mathbf{X}-\mathbf{X}^T}{2} = \mathbf{X}_1 + \mathbf{X}_2.$$

In [2, Theorem 5.1] it is proved that for some $K_1 > \sqrt{2(1+\rho)}$

$$\mathbb{P}(\sigma_1(\mathbf{X}_1) \ge K_1\sqrt{n}) = o(n^{-l}). \tag{4.1}$$

and for some $K_2 > \sqrt{2(1-\rho)}$

$$\mathbb{P}(\sigma_1(\mathbf{X}_2) \ge K_2\sqrt{n}) = o(n^{-l}) \tag{4.2}$$

Set $K = \max(K_1, K_2)$. From (4.1), (4.2) and inequality

$$s_1(\mathbf{X}) \le s_1(\mathbf{X}_1) + s_1(\mathbf{X}_2)$$

it follows that

$$\mathbb{P}(\sigma_1(\mathbf{X}) \ge K\sqrt{n}) \le \mathbb{P}\left(\left\{\sigma_1(\mathbf{X}_1) \ge \frac{K\sqrt{n}}{2}\right\} \cup \left\{\sigma_1(\mathbf{X}_2) \ge \frac{K\sqrt{n}}{2}\right\}\right)$$
$$\le \mathbb{P}\left(\sigma_1(\mathbf{X}_1) \ge \frac{K\sqrt{n}}{2}\right) + \mathbb{P}\left(\sigma_1(\mathbf{X}_2) \ge \frac{K\sqrt{n}}{2}\right) = o(n^{-l})$$

□

Lemma 4.3. *There exists $c > 0$ and $0 < \gamma < 1$ such that a.s. for $n \gg 1$ and $n^{1-\gamma} \le i \le n-1$*

$$s_{n-i} \ge c\frac{i}{n}.$$

Proof. We follow the original proof of Tao and Vu [15]. Up to increasing γ, it is sufficient to prove the statement for all $2(n-1)^{1-\gamma} + 2 \le i \le n-1$ for some $\gamma \in (0,1)$ to be chosen later. We denote by $s_1 \ge s_2 \ge ... \ge s_n$ the singular values of $\mathbf{A} := n^{-1/2}\mathbf{X} - z\mathbf{I}$. We fix some $2(n-1)^{1-\gamma} + 2 \le i \le n-1$ and consider the matrix $\mathbf{A}'$ formed by the first $m := n - i/2$ rows of $\sqrt{n}\mathbf{A}$. Let $s_1' \ge ... \ge s_m'$ be the singular values of $\mathbf{A}'$. We get

$$n^{-1/2}s'_{n-i} \le s_{n-i}.$$

By R_i we denote the row of A' and $H_i = \mathrm{span}(R_j, j = 1, ..., m, j \ne i)$. By Lemma A.11 we obtain

$$s_1'^{-2} + ... + s_{n-i/2}'^{-2} = \mathrm{dist}_1^{-2} + ... + \mathrm{dist}_{n-i/2}^{-2}.$$

We have

$$\frac{i}{2n}s_{n-i}^{-2} \le \frac{i}{2}s_{n-i}'^{-2} \le \sum_{j=n-i}^{n-i/2} s_j'^{-2} \le \sum_{j=1}^{n-i/2} \mathrm{dist}_j^{-2},$$

where $\mathrm{dist}_j := \mathrm{dist}(R_j, H_j)$. To estimate $\mathrm{dist}(R_j, H_j)$ we would like to apply Lemma A.12, but we can't do it directly, because R_j and H_j are not independent. Let's consider the case $j = 1$ only. To estimate distance dist_1 we decompose matrix A' into the blocks

$$A' = \begin{pmatrix} a_{1,1} & Y \\ X & B \end{pmatrix},$$

where $X \in \mathbb{R}^{m-1}, Y^T \in \mathbb{R}^{n-1}$ and B is an $m-1 \times n-1$ matrix formed by rows $B_1, ..., B_{m-1}$. We denote by $H_1' = \mathrm{span}(B_1, ..., B_{m-1})$. From definition of distance

$$\mathrm{dist}(R_1, H_1) = \inf_{v \in H_1} ||R_1 - v||_2 \ge \inf_{u \in H_1'} ||Y - u||_2 = \mathrm{dist}(Y, H_1')$$

and

$$\dim(H_1') \le \dim(H_1) \le n-1-i/2 \le n-1-(n-1)^{1-\gamma}.$$

Now vector Y and hyperplane H_1' are independent. Fixing realization of H_1', by Lemma A.12, with n, R, H replaced with $n-1, Y, H_1'$ respectively, we can obtain that

$$\mathbb{P}(\operatorname{dist}(Y, H_1') \le \frac{1}{2}\sqrt{n-1-\dim(H_1')}) \le \exp(-(n-1)^\delta).$$

Using this inequality it is easy to show that

$$\mathbb{P}\left(\bigcup_{n\gg 1}\bigcup_{i=2(n-1)^{1-\gamma}+2}^{n-1}\bigcup_{j=1}^{n-i/2}\left\{\operatorname{dist}(R_j, H_j) \le \frac{1}{2}\sqrt{\frac{i}{2}}\right\}\right) < \infty.$$

Now by Borel-Cantelli lemma we can conclude the statement of the lemma. □

Remark 4.4. *Lemma 4.3 holds true if we assume that* $\mathbb{E}\, X_{ij} \neq 0$ *and* $\mathbb{E}\, X_{ij}^2 = 1 + o(1)$.

Proof. (Proof of Theorem 4.1) To prove Theorem 4.1 we need to show that there exists $p > 0$ such that

$$\lim_{t\to\infty}\overline{\lim_{n\to\infty}}\, \mathbb{P}\left(\int_0^\infty x^p \nu_n(dx) > t\right) = 0 \tag{4.3}$$

and

$$\lim_{t\to\infty}\overline{\lim_{n\to\infty}}\, \mathbb{P}\left(\int_0^\infty x^{-p} \nu_n(dx) > t\right) = 0. \tag{4.4}$$

From strong law of large numbers it follows that

$$\int_0^\infty x^2 \nu_n(0, dx) \le \frac{1}{n^2}\sum_{i,j=1}^n X_{ij}^2 \to 1 \text{ as } n \to \infty.$$

Using this and the fact that $s_i(n^{-1/2}\mathbf{X} - z\mathbf{I}) \le s_i(n^{-1/2}\mathbf{X}) + |z|$ we conclude (4.3).

Now we prove (4.4). First we should transcate elements of matrix $\mathbf{X}$. Set $\hat{X}_{ij} = X_{ij}\,\mathbb{1}(|X_{ij}| \le \delta_n\sqrt{n})$, where $\delta_n \to 0$ with the convergence rate slower that any preassigned one as $n \to \infty$. Set $\Omega_0 = \{\omega \in \Omega : \hat{\mathbf{X}} \neq \mathbf{X} \text{ i.o.}\}$. It is easy to show that $\mathbb{P}(\Omega_0) = 0$. From condition b) of (**C0**) it also follows that $\max_{i,j} |\mathbb{E}\, \hat{X}_{ij}^2 - 1| = o(1)$. Without loss of generality, saving notation $\mathbf{X}$, we change $\mathbf{X}$ by $\hat{\mathbf{X}}$.

We denote $\Omega_1 = \{\omega \in \Omega : s_{n-i} > c\frac{i}{n}, n^{1-\gamma} \le i \le n-1\}$. From Lemma 4.3 and Remark 4.4 we know that $\mathbb{P}(\Omega_1) = 1$. Let us consider the set $\Omega_2 = \Omega_1 \cap \{\omega : s_n(\mathbf{X} - \sqrt{n}z\mathbf{I}) \ge n^{-A}\}$, with some $A > 0$.

$$\mathbb{P}\left(\int x^{-p}\nu_n(dx) > t\right) = \mathbb{P}\left(\int x^{-p}\nu_n(dx) > t, \Omega_2\right) + \mathbb{P}\left(\int x^{-p}\nu_n(dx) > t, \Omega_2^c\right).$$

We can estimate the second term by

$$\mathbb{P}\left(\int x^{-p}\nu_n(dx) > t, \Omega_2^c\right) \le \mathbb{P}(s_n(\mathbf{X} - \sqrt{n}z\mathbf{I}) \le n^{-A})$$

From $\mathbf{X} - z\mathbf{I} = \mathbf{X} - \mathbb{E}\,\mathbf{X} + \mathbb{E}\,\mathbf{X} - z\mathbf{I}$ it follows that we can centralize elemets of matrix $\mathbf{X}$. It is easy to show that $||\,\mathbb{E}\,\mathbf{X}|| \le C\sqrt{n}$. From Theorem 3.1, Remark 3.15 and Lemma 4.2 we conclude

$$\mathbb{P}(s_n(\mathbf{X} - \sqrt{n}z\mathbf{I}) \le n^{-A}) \le n^{-B}$$

for some $B > 0$.

To finish the proof of (4.4) it remains to bound the second term. From Markov inequality

$$\mathbb{P}\left(\int x^{-p}\nu_n(dx) > t, \Omega_2\right) \le \frac{1}{t}\,\mathbb{E}\left[\int x^{-p}\nu_n(dx)\,\mathbb{1}(\Omega_2)\right].$$

By definition of Ω_2

$$\mathbb{E}\left[\int x^{-p}\nu_n(dx)\,\mathbb{1}(\Omega_2)\right] \le \frac{1}{n}\sum_{i=n^{1-\gamma}}^{n-1} s_{n-i}^{-p} + \frac{1}{n}\sum_{i=n-n^{1-\gamma}}^{n} s_i^{-p}$$

$$\le 2n^{Ap-\gamma} + c^{-p}\frac{1}{n}\sum_{i=1}^{n}\left(\frac{n}{i}\right)^p \le 2n^{Ap-\gamma} + c_0^{-p}\int_0^1 s^{-p}ds.$$

If $0 \le p \le \min(1, \gamma/A)$ then the last integral is finite. □

5. Convergence of singular values

Let us recall definition of Stieltjes transform

Definition 5.1. *The Stieltjes transform of measure $m(\cdot)$ on $\mathbb{R}$ is*

$$S(\alpha) = \int_{\mathbb{R}} \frac{m(dx)}{x - \alpha}, \quad \alpha \in \mathbb{C}^+.$$

We denote Stieltjes transforms of $\mathbb{E}\,\nu_n$ and ν_z by $S_n(\alpha, z)$ and $S(\alpha, z)$ respectively. In this section we prove

Theorem 5.2. *Under condition* (**C0**) *$S_n(\alpha, z) \to S(\alpha, z)$ as $n \to \infty$, where $S(\alpha, z)$ is a Stieltjes transforms of some probability measure ν_z.*

Proof. Introduce the following $2n \times 2n$ matrices

$$\mathbf{V} = \begin{pmatrix} n^{-1/2}\mathbf{X} & \mathbf{O}_n \\ \mathbf{O}_n & n^{-1/2}\mathbf{X}^T \end{pmatrix}, \quad \mathbf{J}(z) = \begin{pmatrix} \mathbf{O}_n & z\mathbf{I} \\ \overline{z}\mathbf{I} & \mathbf{O}_n \end{pmatrix}, \quad \mathbf{J} := \mathbf{J}(1), \tag{5.1}$$

where $\mathbf{O}_n$ denotes $n \times n$ matrix with zero entries. Consider matrix

$$\mathbf{V}(z) := \mathbf{V}\mathbf{J} - \mathbf{J}(z).$$

It is known that eigenvalues of $\mathbf{V}(z)$ are singular values of $n^{-1/2}\mathbf{X} - z\mathbf{I}$ with signs $\pm$.

Set for $\alpha = u + iv, v > 0$

$$\mathbf{R}(\alpha, z) := (\mathbf{V}(z) - \alpha \mathbf{I}_{2n})^{-1}. \tag{5.2}$$

We introduce the following functions

$$s_n(\alpha, z) = \frac{1}{n}\sum_{i=1}^{n} \mathbb{E}[\mathbf{R}(\alpha, z)]_{ii} = \frac{1}{n}\sum_{i=1}^{n} \mathbb{E}[\mathbf{R}(\alpha, z)]_{i+n,i+n} = \frac{1}{2n}\sum_{i=1}^{2n} \mathbb{E}[\mathbf{R}(\alpha, z)]_{ii},$$

$$t_n(\alpha, z) = \frac{1}{n}\sum_{i=1}^{n} \mathbb{E}[\mathbf{R}(\alpha, z)]_{i+n,i}, \quad u_n(\alpha, z) = \frac{1}{n}\sum_{i=1}^{n} \mathbb{E}[\mathbf{R}(\alpha, z)]_{i,i+n}.$$

It is easy to see that s_n is a Stieltjes transform of

$$\mathbb{E}\,\hat{\nu}_{n^{-1/2}\mathbf{X}-z\mathbf{I}}(\cdot) = \frac{\mathbb{E}\,\nu_n(\cdot) + \mathbb{E}\,\nu_n(-\cdot)}{2},$$

which is the symmetrized version of a measure $\mathbb{E}\,\nu_n$. There is a relation between Stieltjes transform $S_n(\alpha, z)$ and $s_n(\alpha, z)$:

$$s_n(\alpha, z) = \alpha S_n(\alpha^2, z).$$

To prove Theorem 5.2 we need to show convergence of $s_n(\alpha, z)$ to Stieltjes transform of some symmetric measure $\nu_0(z, \cdot)$.

By resolvent equality we may write

$$1 + \alpha s_n(\alpha, z) = \frac{1}{2n}\,\mathbb{E}\,\mathrm{Tr}(\mathbf{VJR}(\alpha, z)) - z t_n(\alpha, z) - \bar{z} u_n(\alpha, z).$$

Introduce the notation

$$\mathbb{A} := \frac{1}{2n}\,\mathbb{E}\,\mathrm{Tr}(\mathbf{VJR})$$

and represent $\mathbb{A}$ as follows

$$\mathbb{A} = \frac{1}{2}\mathbb{A}_1 + \frac{1}{2}\mathbb{A}_2,$$

where

$$\mathbb{A}_1 = \frac{1}{n}\sum_{i=1}^{n} \mathbb{E}[\mathbf{VJR}]_{ii}, \quad \mathbb{A}_2 = \frac{1}{n}\sum_{i=1}^{n} \mathbb{E}[\mathbf{VJR}]_{i+n,i+n}.$$

First we consider $\mathbb{A}_1$. By definition of the matrix $\mathbf{V}$, we have

$$\mathbb{A}_1 = \frac{1}{n^{3/2}}\sum_{j,k=1}^{n} \mathbb{E}\,X_{jk} R_{k+n,j}.$$

Note that

$$\frac{\partial \mathbf{R}}{\partial X_{jk}} = -\frac{1}{\sqrt{n}}\mathbf{R}[e_j e_{k+n}^T]\mathbf{R}.$$

Applying Lemma A.15 we obtain

$$\mathbb{A}_1 = \mathbb{B}_1 + \mathbb{B}_2 + \mathbb{B}_3 + \mathbb{B}_4 + r_n(\alpha, z).$$

where

$$\mathbb{B}_1 = -\frac{1}{n^2}\sum_{j,k=1}^{n} \mathbb{E}[\mathbf{R}[e_j e_{k+n}^T]\mathbf{R}]_{k+n,j} = -\frac{1}{n^2}\sum_{j,k=1}^{n} \mathbb{E}(R_{k+n,j})^2$$
$$\mathbb{B}_2 = -\frac{1}{n^2}\sum_{j,k=1}^{n} \mathbb{E}[\mathbf{R}[e_{k+n} e_j^T]\mathbf{R}]_{k+n,j} = -\frac{1}{n^2}\sum_{j,k=1}^{n} \mathbb{E}\, R_{jj} R_{k+n,k+n}$$
$$\mathbb{B}_3 = -\frac{\rho}{n^2}\sum_{j,k=1}^{n} \mathbb{E}[\mathbf{R}[e_k e_{j+n}^T]\mathbf{R}]_{k+n,j} = -\frac{\rho}{n^2}\sum_{j,k=1}^{n} \mathbb{E}\, R_{k+n,k} R_{j+n,j}$$
$$\mathbb{B}_4 = -\frac{\rho}{n^2}\sum_{j,k=1}^{n} \mathbb{E}[\mathbf{R}[e_{j+n} e_k^T]\mathbf{R}]_{k+n,j} = -\frac{\rho}{n^2}\sum_{j,k=1}^{n} \mathbb{E}\, R_{kj} R_{k+n,j+n}.$$

Without loss of generality we can assume further that $\mathbb{E}\, X_{11}^2 = 1$ because the impact of diagonal is of order $O(n^{-1})$.

From $||\mathbf{R}||_{HS} \le \sqrt{n}||\mathbf{R}|| \le \sqrt{n}v^{-1}$ it follows

$$|\mathbb{B}_1| \le \frac{1}{n^2}\sum_{j,k=1}^{n} \mathbb{E}\, X_{jk}^2\, \mathbb{E}(R_{k+n,j})^2 \le \frac{1}{nv^2}.$$

Similarly

$$|\mathbb{B}_4| \le \frac{1}{v^2 n}.$$

By Lemma A.13 $\mathbb{B}_2 = -s_n^2(\alpha, z) + \varepsilon(\alpha, z)$. By Lemma A.14 $\mathbb{B}_3 = -\rho t_n^2(\alpha, z) + \varepsilon(\alpha, z)$. We obtain that

$$\mathbb{A}_1 = -s_n^2(\alpha, z) - \rho t_n^2(\alpha, z) + \delta_n(\alpha, z).$$

Now we consider the term $\mathbb{A}_2$. By definition of the matrix $\mathbf{V}$, we have

$$\mathbb{A}_2 = \frac{1}{n^{3/2}}\sum_{j,k=1}^{n} \mathbb{E}\, X_{jk} R_{j,k+n}.$$

By Lemma A.15 we obtain that

$$\mathbb{A}_2 = \mathbb{C}_1 + \mathbb{C}_2 + \mathbb{C}_3 + \mathbb{C}_4 + r_n(\alpha, z). \tag{5.3}$$

where

$$\mathbb{C}_1 = -\frac{1}{n^2}\sum_{j,k=1}^{n} \mathbb{E}[\mathbf{R}[e_j e_{k+n}^T]\mathbf{R}]_{j,k+n} = -\frac{1}{n^2}\sum_{j,k=1}^{n} \mathbb{E}\, R_{jj}R_{k+n,k+n}$$

$$\mathbb{C}_2 = -\frac{1}{n^2}\sum_{j,k=1}^{n} \mathbb{E}[\mathbf{R}[e_{k+n} e_j^T]\mathbf{R}]_{j,k+n} = -\frac{1}{n^2}\sum_{j,k=1}^{n} \mathbb{E}(R_{j,k+n})^2$$

$$\mathbb{C}_3 = -\frac{\rho}{n^2}\sum_{j,k=1}^{n} \mathbb{E}[\mathbf{R}[e_k e_{j+n}^T]\mathbf{R}]_{j,k+n} = -\frac{\rho}{n^2}\sum_{j,k=1}^{n} \mathbb{E}\, R_{jk}R_{j+n,k+n}$$

$$\mathbb{C}_4 = -\frac{\rho}{n^2}\sum_{j,k=1}^{n} \mathbb{E}[\mathbf{R}[e_{j+n} e_k^T]\mathbf{R}]_{j,k+n} = -\frac{\rho}{n^2}\sum_{j,k=1}^{n} \mathbb{E}\, R_{j,j+n}R_{k,k+n}.$$

It is easy to show that

$$|\mathbb{C}_2| \le \frac{1}{v^2 n}, \quad |\mathbb{C}_3| \le \frac{1}{v^2 n}.$$

By Lemma A.13 $\mathbb{C}_1 = -s_n^2(\alpha, z) + \varepsilon_n(\alpha, z)$. By Lemma A.14 $\mathbb{C}_4 = -\rho u_n^2(\alpha, z) + \varepsilon_n(\alpha, z)$. We obtain that

$$\mathbb{A}_2 = -s_n^2(\alpha, z) - \rho u_n^2(\alpha, z) + \delta_n(\alpha, z).$$

So we have that

$$\mathbb{A} = -s_n^2(\alpha, z) - \frac{\rho}{2} t_n^2(\alpha, z) - \frac{\rho}{2} u_n^2(\alpha, z) + \varepsilon_n(\alpha, z).$$

No we will investigate the term $zt_n(\alpha, z)$ which we may represent as follows

$$\alpha t_n(\alpha, z) = \frac{1}{n}\sum_{j=1}^{n} \mathbb{E}[\mathbf{V}(z)\mathbf{R}]_{j+n,j} = \frac{1}{n}\sum_{j=1}^{n} \mathbb{E}[\mathbf{V}\mathbf{R}]_{j+n,j} - \overline{z} s_n(\alpha, z).$$

By definition of the matrix $\mathbf{V}$, we have

$$\alpha t_n(\alpha, z) = \frac{1}{n^{3/2}}\sum_{j,k=1}^{n} \mathbb{E}\, X_{jk}R_{j,k} - \overline{z} s_n(\alpha, z) =$$

$$\mathbb{D}_1 + \mathbb{D}_2 + \mathbb{D}_3 + \mathbb{D}_4 - \overline{z} s_n(\alpha, z) + r_n(\alpha, z),$$

where

$$\mathbb{D}_1 = -\frac{1}{n^2}\sum_{j,k=1}^{n} \mathbb{E}[\mathbf{R}[e_j e_{k+n}^T]\mathbf{R}]_{j,k} = -\frac{1}{n^2}\sum_{j,k=1}^{n} \mathbb{E}\, R_{j,j}R_{k+n,k}$$

$$\mathbb{D}_2 = -\frac{1}{n^2}\sum_{j,k=1}^{n} \mathbb{E}[\mathbf{R}[e_{k+n} e_j^T]\mathbf{R}]_{j,k} = -\frac{1}{n^2}\sum_{j,k=1}^{n} \mathbb{E}\, R_{j,k+n}R_{j,k}$$

$$\mathbb{D}_3 = -\frac{\rho}{n^2}\sum_{j,k=1}^{n} \mathbb{E}[\mathbf{R}[e_k e_{j+n}^T]\mathbf{R}]_{j,k} = -\frac{\rho}{n^2}\sum_{j,k=1}^{n} \mathbb{E}\, R_{j,k}R_{j+n,k}$$

$$\mathbb{D}_4 = -\frac{\rho}{n^2}\sum_{j,k=1}^{n} \mathbb{E}[\mathbf{R}[e_{j+n} e_k^T]\mathbf{R}]_{j,k} = -\frac{\rho}{n^2}\sum_{j,k=1}^{n} \mathbb{E}\, R_{j,j+n}R_{k,k}.$$

By similar arguments as before we can prove that

$$|\mathbb{D}_2| \le \frac{1}{v^2 n}, \quad |\mathbb{D}_3| \le \frac{1}{v^2 n}$$

and $\mathbb{D}_1 = -s_n(\alpha, z)t_n(\alpha, z) + \varepsilon_n(\alpha, z)$, $\mathbb{D}_4 = -\rho s_n(\alpha, z)u_n(\alpha, z) + \varepsilon_n(\alpha, z)$. We obtain that

$$\alpha t_n(\alpha, z) = -s_n(\alpha, z)t_n(\alpha, z) - \rho s_n(\alpha, z)u_n(\alpha, z) - \overline{z}s_n(\alpha, z) + \delta_n(\alpha, z).$$

Similar we can prove that

$$\alpha u_n(\alpha, z) = -s_n(\alpha, z)u_n(\alpha, z) - \rho s_n(\alpha, z)t_n(\alpha, z) - zs_n(\alpha, z) + \delta_n(\alpha, z).$$

So we have the system of equations

$$\begin{aligned}
&1 + \alpha s_n(\alpha, z) + s_n^2(\alpha, z) = \\
&= -\frac{\rho}{2}t_n^2(\alpha, z) - \frac{z}{2}t_n(\alpha, z) - \frac{\rho}{2}u_n^2(\alpha, z) - \frac{\overline{z}}{2}u_n(\alpha, z) + \delta_n(\alpha, z) \quad (5.4)\\
&\alpha t_n(\alpha, z) = \\
&= -s_n(\alpha, z)t_n(\alpha, z) - \rho s_n(\alpha, z)u_n(\alpha, z) - \overline{z}s_n(\alpha, z) + \delta_n(\alpha, z) \quad (5.5)\\
&\alpha u_n(\alpha, z) = \\
&= -s_n(\alpha, z)u_n(\alpha, z) - \rho s_n(\alpha, z)t_n(\alpha, z) - zs_n(\alpha, z) + \delta_n(\alpha, z). \quad (5.6)
\end{aligned}$$

It follows from (5.5) and (5.6) that

$$\begin{aligned}
&(\alpha + s_n)(zt_n + \rho t_n^2) = -s_n(z\rho u_n + \overline{z}\rho t) - \rho^2 s_n t_n u_n - |z|^2 s_n + \delta_n(\alpha, z)\\
&(\alpha + s_n)(\overline{z}u_n + \rho u_n^2) = -s_n(z\rho u_n + \overline{z}\rho t) - \rho^2 s_n t_n u_n - |z|^2 s_n + \delta_n(\alpha, z).
\end{aligned}$$

So, we can rewrite (5.4)

$$1 + \alpha s_n(\alpha, z) + s_n^2(\alpha, z) + \rho^2 t_n^2(\alpha, z) + zt_n(\alpha, z) = \delta_n(\alpha, z). \quad (5.7)$$

From equations (5.5) and (5.6) we can write equation for t_n

$$\left(\alpha + s_n - \frac{|\rho|^2 s_n^2}{\alpha + s_n}\right) t_n = \frac{\rho z s_n^2}{\alpha + s_n} - \overline{z}s_n + \delta_n(\alpha, z). \quad (5.8)$$

We denote

$$\Delta = \left(\alpha + s_n - \frac{|\rho|^2 s_n^2}{\alpha + s_n}\right).$$

After simple calculations we will have

$$\begin{aligned}
&(\alpha + s_n)(zt_n + \rho t_n^2) = \\
&- s_n\left(\frac{2\rho^2|z|^2 s_n^2}{(\alpha + s_n)\Delta} - \frac{\overline{z}^2\rho s_n}{\Delta} - \frac{z^2\overline{\rho}s_n}{\Delta}\right)\\
&- |\rho|^2 s_n\left(\frac{\rho z s_n^2}{(\alpha + s)\Delta} - \frac{\overline{z}s_n}{\Delta}\right)\left(\frac{\rho\overline{z}s_n^2}{(\alpha + s)\Delta} - \frac{zs_n}{\Delta}\right) - |z|^2 s_n + \delta_n(\alpha, z).
\end{aligned}$$

We denote $y_n := s_n$ and $w_n := \alpha + (\rho t_n^2 + z t_n)/y_n$. We can rewrite equations (5.4), (5.5) and (5.6)

$$1 + w_n y_n + y_n^2 = \delta_n(\alpha, z) \tag{5.9}$$

$$w_n = \alpha + \frac{\rho t_n^2 + z t_n}{y_n} \tag{5.10}$$

$$\begin{aligned}(\alpha + s_n)(z t_n + \rho t_n^2) = \\ &- s_n \left(\frac{2\rho^2 |z|^2 y_n^2}{(\alpha + y_n)\Delta} - \frac{\overline{z}^2 \rho y_n}{\Delta} - \frac{z^2 \rho y_n}{\Delta} \right) - |z|^2 y_n \\ &- |\rho|^2 y_n \left(\frac{\rho z y_n^2}{(\alpha + y_n)\Delta} - \frac{\overline{z} y_n}{\Delta} \right) \left(\frac{\rho \overline{z} y_n^2}{(\alpha + y_n)\Delta} - \frac{z y_n}{\Delta} \right) + \delta_n(\alpha, z).\end{aligned} \tag{5.11}$$

Remark 5.3. *If $\rho = 0$ then we can rewrite* (5.9), (5.10), *and* (5.11)

$$\begin{aligned} &1 + w_n y_n + y_n^2 = \delta_n(\alpha, z) \\ &w_n = \alpha + \frac{z t_n}{y_n} \\ &(w_n - \alpha) + (w_n - \alpha)^2 y_n - |z|^2 s_n = \delta_n(\alpha, z). \end{aligned}$$

This equations determine the Circular law, see [9].

We can see that the first equation (5.9) doesn't depend on ρ. So the first equation will be the same for all models of random matrices described in the introduction. On the Figure 2 we draw the distribution of eigenvalues of matrix $\mathbf{VJ}$ for $\rho = 0$ (Circular law case), $\rho = 0.5$ (Elliptic law case) and $\rho = 1$ (Semi-circular law). We mention here that eigenvalues of $\mathbf{VJ}$ are singular values of $n^{-1/2}\mathbf{X}$ with signs $\pm$.

Now we prove convergence of s_n to some limit s_0. Let $\alpha = u + iv, v > 0$. Using (5.7) we write

$$\alpha(s_n - s_m) = -(s_n - s_m)(s_n + s_m) - \rho^2 (t_n - t_m)(t_n + t_m) - z(t_m - t_m) + \varepsilon_{n,m}.$$

By triangle inequality and the fact that $|s_n| \le v^{-1}$

$$|s_n - s_m| \le \frac{2|s_n - s_m|}{v^2} + \frac{\rho^2 |t_n - t_m||t_n + t_m|}{v} + \frac{|z||t_n - t_m|}{v} + \frac{|\varepsilon_{n,m}|}{v}. \tag{5.12}$$

From (5.8) it follows that

$$((\alpha + s_n)^2 - \rho^2 s_n^2) t_n = \rho z s_n^2 - \overline{z} \alpha s_n - \overline{z} s_n^2 + \varepsilon_n.$$

We denote $\Delta_n := ((\alpha + s_n)^2 - \rho^2 s_n^2)$. By triangle inequality

$$\begin{aligned} |\Delta_m||t_n - t_m| &\le |t_m||\Delta_n - \Delta_m| \\ &+ \frac{2|\rho||s_n - s_m| + 2|z||s_n - s_m|}{v} + |z||\alpha||s_n - s_m| + |\varepsilon_{n,m}|. \end{aligned} \tag{5.13}$$

We can find lower bound for $|\Delta_m|$:

$$\begin{aligned} |\Delta_m| &= |\alpha + (1 - \rho) s_m||\alpha + (1 + \rho) s_m| \\ &\ge \operatorname{Im}(\alpha + (1 - \rho) s_m) \operatorname{Im}(\alpha + (1 + \rho) s_m) \ge v^2, \end{aligned} \tag{5.14}$$

where we have used the fact that $\operatorname{Im} s_m \geq 0$. From definition of Δ_n it is easy to see that

$$|\Delta_n - \Delta_m| \leq 2|\alpha||s_n - s_m| + \frac{2(1+\rho^2)|s_n - s_m|}{v}. \tag{5.15}$$

We can take $|u| \leq C$, then $|\alpha| \leq v + C$. From (5.12),(5.13),(5.14) and (5.15) it follows that there exists constant C', which depends on ρ, C, z, such that

$$|s_n - s_m| \leq \frac{C'}{v}|s_n - s_m| + |\varepsilon'_{n,m}(\alpha, z)|.$$

We can find v_0 such that

$$\frac{C'}{v} < 1 \quad \text{for all } v \geq v_0.$$

Since $\varepsilon'_{n,m}(\alpha, z)$ converges to zero uniformly for all $v \geq v_0$, $|u| \leq C$ and s_n, s_m are locally bounded analytic functions in the upper half-plane we may conclude by Montel's Theorem (see [4, Theorem 2.9]) that there exists an analytic function s_0 in the upper half-plane such that $\lim s_n = s_0$. Since s_n are Nevanlinna functions, (that is analytic functions mapping the upper half-plane into itself) s_0 will be a Nevanlinna function too and there exists $\nu_0(z, \cdot)$ such that

$$s_0(\alpha) = \int \frac{\nu_0(z, dx)}{x - \alpha}.$$

The function s_0 satisfies the equations (5.9), (5.10), and (5.11). □

6. Acknowledgment

The author would like to thank Prof. Dr. Friedrich Götze and Prof. Dr. Alexander Tikhomirov for insightful comments and discussions. This research was conducted during my participation in SFB 701, Bielefeld University

Appendix A. Appendix

Theorem A.1. *(Central Limit Theorem) Let $Z_1, ..., Z_n$ be independent random variables with $\mathbb{E}\, Z_i = 0$ and finite third moment, and let $\sigma^2 = \sum_{i=1}^n \mathbb{E}\,|Z_i|^2$. Consider a standard normal variable g. The for every $t > 0$:*

$$\left| \mathbb{P}\left(\frac{1}{\sigma} \sum_{i=1}^n Z_i \leq t \right) - \mathbb{P}\left(g \leq t\right) \right| \leq C\sigma^{-3} \sum_{i=1}^n \mathbb{E}\,|Z_i|^3,$$

where C is an absolute constant.

Lemma A.2. *Let event $E(X, Y)$ depends on independent random vectors X and Y then*

$$\mathbb{P}(E(X, Y)) \leq (\mathbb{P}(E(X, Y), E(X, Y')))^{1/2},$$

where Y' is an independent copy of Y.

Proof. See in [5]. □

Lemma A.3. *Let $Z_1, ..., Z_n$ be a sequence of random variables and $p_1, ..., p_n$ be non-negative real numbers such that*

$$\sum_{i=1}^{n} p_i = 1,$$

then for every $\varepsilon > 0$

$$\mathbb{P}(\sum_{i=1}^{n} p_i Z_i \le \varepsilon) \le 2 \sum_{i=1}^{n} p_i \, \mathbb{P}(Z_i \le 2\varepsilon).$$

Proof. See in [16]. □

Lemma A.4. *Let $\mathcal{N}(T, \varepsilon)$ be an ε-net of set$T$. One has*

$$\mathcal{N}(Comp(c_0, c_1), 2c_1) \le (9/c_0 c_1)^{c_0 n}.$$

Proof. See in [16]. □

Lemma A.5. *If $x \in Incomp(c_0, c_1)$ then at least $\frac{1}{2} c_0 c_1^2 n$ coordinates x_k of x satisfy*

$$\frac{c_1}{\sqrt{2n}} \le |x_k| \le \frac{1}{\sqrt{c_0 n}}.$$

Remark A.6. *We can fix some constant c_{00} such that*

$$\frac{1}{4} c_0 c_1^2 \le c_{00} \le \frac{1}{4}.$$

Then for every vector $x \in Incomp(c_0, c_1)$ $|\operatorname{spread}(x)| = [c_{00} n]$.

Proof. See in [11]. □

Lemma A.7. *Let $S_J = \sum_{i \in J} \xi_i$, where $J \subset [n]$, and $I \subset J$ then*

$$\sup_{v \in \mathbb{R}} \mathbb{P}(|S_J - v| \le \varepsilon) \le \sup_{v \in \mathbb{R}} \mathbb{P}(|S_I - v| \le \varepsilon).$$

Proof. Let us fix arbitrary v. From independence of ξ_i we conclude

$$\mathbb{P}(|S_J - v| \le \varepsilon) \le \mathbb{E}\,\mathbb{P}(|S_I + S_{J/I} - v| \le \varepsilon | \{\xi_i\}_{i \in I}) \le \sup_{u \in \mathbb{R}} \mathbb{P}(|S_I - u| \le \varepsilon).$$

□

Lemma A.8. *Let Z be a random variable with $\mathbb{E}\, Z^2 \ge 1$ and with finite fourth moment, and put $M_4^4 := \mathbb{E}(Z - \mathbb{E}\, Z)^4$. Then for every $\varepsilon \in (0, 1)$ there exists $p = p(M_4, \varepsilon)$ such that*

$$\sup_{v \in \mathbb{R}} \mathbb{P}(|Z - v| \le \varepsilon) \le p.$$

Proof. See in [11]. □

Lemma A.9. *Let $\xi_1, ..., \xi_n$ be independent random variables with $\mathbb{E}\,\xi_i^2 \ge 1$ and $\mathbb{E}(\xi_k - \mathbb{E}\,\xi)^4 \le M_4^4$, where M_4 is some finite number. Then for every $\varepsilon \in (0, 1)$ there exists $p = p(M_4, \varepsilon) \in (0, 1))$ such that the following holds: for every vector $x = (x_1, ..., x_n) \in S^{n-1}$, the sum $S = \sum_{i=1}^n x_k \xi_k$ satisfies*

$$\sup_{v \in \mathbb{R}} \mathbb{P}(|S - v| \le \varepsilon) \le p.$$

Proof. See in [11]. □

Lemma A.10. *Let $X = (X_1, ..., X_n)$ be a random vector in $\mathbb{R}^n$ with independent coordinates X_k.*
1. Suppose there exists numbers $\varepsilon_0 \ge 0$ and $L \ge 0$ such that

$$\sup_{v \in \mathbb{R}} \mathbb{P}(|X_k - v| \le \varepsilon) \le L\varepsilon \quad \text{for all } \varepsilon \ge \varepsilon_0 \text{ and all } k.$$

Then

$$\sup_{v \in \mathbb{R}^n} \mathbb{P}(||X - v||_2 \le \varepsilon\sqrt{n}) \le (CL\varepsilon)^n \quad \text{for all } \varepsilon \ge \varepsilon_0,$$

where C is an absolute constant.
2. Suppose there exists numbers $\varepsilon > 0$ and $p \in (0, 1)$ such that

$$\sup_{v \in \mathbb{R}} \mathbb{P}(|X_k - v| \le \varepsilon) \le L\varepsilon \quad \text{for all } k.$$

Then there exists numbers $\varepsilon_1 = \varepsilon_1(\varepsilon, p) > 0$ and $p_1 = p_1(\varepsilon, p) \in (0, 1)$ such that

$$\sup_{v \in \mathbb{R}^n} \mathbb{P}(||X - v||_2 \le \varepsilon_1\sqrt{n}) \le (p_1)^n.$$

Proof. See [16, Lemma 3.4]. □

Lemma A.11. *Let $1 \le m \le n$. If A has full rank, with rows $R_1, ..., R_m$ and $H = \operatorname{span}(R_j, j \ne i)$, then*

$$\sum_{i=1}^m s_i(A)^{-2} = \sum_{i=1}^m \operatorname{dist}(R_i, H_i)^{-2}.$$

Proof. See [15, Lemma A.4]. □

Lemma A.12. *There exist $\gamma > 0$ and $\delta > 0$ such that for all $n \gg 1$ and $1 \le i \le n$, any deterministic vector $v \in \mathbb{C}$ and any subspace H of $\mathbb{C}^n$ with $1 \le \dim(H) \le n - n^{1-\gamma}$, we have, denoting $R := (X_1, ..., X_n) + v$,*

$$\mathbb{P}(\operatorname{dist}(R, H) \le \frac{1}{2}\sqrt{n - \dim(H)}) \le \exp(-n^{\delta}).$$

Proof. See [15, Statement 5.1]. □

Lemma A.13. *Under the condition* (**C0**) *for $\alpha = u + iv, v > 0$*

$$\mathbb{E}\left|\frac{1}{n}\sum_{i=1}^n R_{ii}(\alpha, z) - \mathbb{E}\left(\frac{1}{n}\sum_{i=1}^n R_{ii}(\alpha, z)\right)\right|^2 \le \frac{C}{nv^2}.$$

Proof. To prove this lemma we will use Girko's method. Let $\mathbf{X}^{(j)}$ be a matrix $\mathbf{X}$ with j-th row and column removed. Define matrices $\mathbf{V}^{(j)}$ and $\mathbf{V}^{(j)}(z)$ as in (5.1) and $\mathbf{R}^{(j)}$ by (5.2). It is easy to see that

$$\operatorname{Rank}(\mathbf{V}(z) - \mathbf{V}^{(j)}(z)) = \operatorname{Rank}(\mathbf{VJ} - \mathbf{V}^{(j)}\mathbf{J}) \le 4.$$

Then

$$\frac{1}{n}|\operatorname{Tr}(\mathbf{V}(z) - \alpha\mathbf{I})^{-1} - \operatorname{Tr}(\mathbf{V}^{(j)}(z) - \alpha\mathbf{I})^{-1}| \le \frac{\operatorname{Rank}(\mathbf{V}(z) - \mathbf{V}^{(j)}(z))}{nv} \le \frac{4}{nv}. \tag{A.1}$$

We introduce the family of σ-algebras $\mathcal{F}_i = \sigma\{X_{j,k}, j, k > i\}$ and conditional mathematical expectation $\mathbb{E}_i = \mathbb{E}(\cdot|\mathcal{F}_i)$ with respect to this σ-algebras. We can write

$$\frac{1}{n}\operatorname{Tr}\mathbf{R} - \frac{1}{n}\mathbb{E}\operatorname{Tr}\mathbf{R} = \frac{1}{n}\sum_{i=1}^{n}\mathbb{E}_i\operatorname{Tr}\mathbf{R} - \mathbb{E}_{i-1}\operatorname{Tr}\mathbf{R} = \sum_{i=1}^{n}\gamma_i.$$

The sequence $(\gamma_i, \mathcal{F}_i)_{i\ge1}$ is a martingale difference. By (A.1)

$$|\gamma_i| = \frac{1}{n}|\mathbb{E}_i(\operatorname{Tr}\mathbf{R} - \operatorname{Tr}\mathbf{R}^{(i)}) - \mathbb{E}_{i-1}(\operatorname{Tr}\mathbf{R} - \operatorname{Tr}\mathbf{R}^{(i)})| \le \tag{A.2}$$

$$\le |\mathbb{E}_i(\operatorname{Tr}\mathbf{R} - \operatorname{Tr}\mathbf{R}^{(i)})| + |\mathbb{E}_{i-1}(\operatorname{Tr}\mathbf{R} - \operatorname{Tr}\mathbf{R}^{(i)})| \le \frac{C}{vn}. \tag{A.3}$$

From Burkholder inequality for martingale difference (see [13])

$$\mathbb{E}\left|\sum_{i=1}^{n}\gamma_i\right|^2 \le K_2\,\mathbb{E}\left(\sum_{i=1}^{n}|\gamma_i|^2\right)$$

and (A.2) it follows

$$\mathbb{E}\left|\frac{1}{n}\sum_{i=1}^{n}R_{ii}(\alpha,z) - \mathbb{E}\left(\frac{1}{n}\sum_{i=1}^{n}R_{ii}(\alpha,z)\right)\right|^2 \le K_2\,\mathbb{E}(\sum_{i=1}^{n}|\gamma_i|^2) \le K_2\frac{C}{nv^2}.$$

□

Lemma A.14. *Under the condition* (**C0**) *for* $\alpha = u + iv, v > 0$

$$\mathbb{E}\left|\frac{1}{n}\sum_{i=1}^{n}R_{i,i+n}(\alpha,z) - \mathbb{E}\left(\frac{1}{n}\sum_{i=1}^{n}R_{i,i+n}(\alpha,z)\right)\right|^2 \le \frac{C}{nv^4}.$$

Proof. As in Lemma A.13 we introduce matrices $\mathbf{V}^{(j)}$ and $\mathbf{R}^{(j)}$. We have

$$\mathbf{VJ} = \mathbf{V}^{(j)}\mathbf{J} + e_je_j^T\mathbf{VJ} + \mathbf{VJ}e_je_j^T + e_{j+n}e_{j+n}^T\mathbf{VJ} + \mathbf{VJ}e_{j+n}e_{j+n}^T$$

By resolvent equality $\mathbf{R}-\mathbf{R}^{(j)}=-\mathbf{R}^{(j)}(\mathbf{V}(z)-\mathbf{V}^{(j)}(z))\mathbf{R}$

$$\begin{aligned}&\frac{1}{n}\sum_{k=1}^{n}(\mathbf{R}_{k,k+n}-\mathbf{R}^{(j)}_{k,k+n})=\\&=\frac{1}{n}\sum_{k=1}^{n}[\mathbf{R}^{(j)}(e_je_j^T\mathbf{V}\mathbf{J}+e_{j+n}e_{j+n}^T\mathbf{V}\mathbf{J}+\mathbf{V}\mathbf{J}e_je_j^T+\mathbf{V}\mathbf{J}e_{j+n}e_{j+n}^T)\mathbf{R}]_{k,k+n}=\\&=\mathbb{T}_1+\mathbb{T}_2+\mathbb{T}_3+\mathbb{T}_4.\end{aligned}$$

Let us consider the first term. The arguments for other terms are similar.

$$\sum_{k=1}^{n}[\mathbf{R}^{(j)}e_je_j^T\mathbf{V}\mathbf{J}\mathbf{R}]_{k,k+n}=\operatorname{Tr}\mathbf{R}^{(j)}e_je_{j+n}^T\mathbf{V}\mathbf{J}\mathbf{R}=\sum_{i=1}^{2n}[\mathbf{R}^{(j)}\mathbf{R}]_{ij}[e_je_{j+n}^T\mathbf{V}\mathbf{J}]_{ji}.$$

From $\max(||\mathbf{R}^{(j)}||,||\mathbf{R}||)\le v^{-1}$ and Hölder inequality it follows that

$$\mathbb{E}\left|\sum_{k=1}^{n}[\mathbf{R}^{(j)}e_je_j^T\mathbf{V}\mathbf{J}\mathbf{R}]_{k,k+n}\right|^2\le\frac{C}{v^4}.$$

By similar arguments as in Lemma A.13 we can conclude the statement of the Lemma. □

Lemma A.15. *Under the condition* (**C0**) *for* $\alpha=u+iv, v>0$

$$\begin{aligned}&\frac{1}{n^{3/2}}\sum_{j,k=1}^{n}\mathbb{E}\,X_{jk}R_{k+n,j}=\\&=\frac{1}{n^2}\sum_{j,k=1}^{n}\mathbb{E}\left[\frac{\partial\mathbf{R}}{\partial X_{jk}}\right]_{k+n,j}+\frac{\rho}{n^2}\sum_{j,k=1}^{n}\mathbb{E}\left[\frac{\partial\mathbf{R}}{\partial X_{kj}}\right]_{k+n,j}+r_n(\alpha,z),\end{aligned}$$

where

$$|r_n(\alpha,z)|\le\frac{C}{\sqrt{n}v^3}$$

Proof. By Taylor's formula

$$\begin{aligned}\mathbb{E}\,Xf(X,Y)&=f(0,0)\,\mathbb{E}\,X+f'_x(0,0)\,\mathbb{E}\,X^2+f'_y(0,0)\,\mathbb{E}\,XY+\\&+\mathbb{E}(1-\theta)[X^3f''_{xx}(\theta X,\theta Y)+2X^2Yf''_{xy}(\theta X,\theta Y)+XY^2f''_{yy}(\theta X,\theta Y)]\end{aligned}\tag{A.4}$$

and

$$\begin{aligned}\mathbb{E}\,f'_x(X,Y)&=f'_x(0,0)+\mathbb{E}(1-\theta)[Xf''_{xx}(\theta X,\theta Y)+Yf''_{xy}(\theta X,\theta Y)]\\\mathbb{E}\,f'_y(X,Y)&=f'_y(0,0)+\mathbb{E}(1-\theta)[Xf''_{xy}(\theta X,\theta Y)+Yf''_{yy}(\theta X,\theta Y)],\end{aligned}\tag{A.5}$$

where θ has uniform distribution on $[0,1]$. From (A.4) and (A.5) for $j \neq k$

$$\left| \mathbb{E}\, X_{jk} R_{k+n,j} - \mathbb{E}\left[\frac{\partial \mathbf{R}}{\partial X_{jk}}\right]_{k+n,j} - \rho\, \mathbb{E}\left[\frac{\partial \mathbf{R}}{\partial X_{kj}}\right]_{k+n,j} \right| \leq$$

$$(|X_{jk}|^3 + |X_{jk}|) \left| \left[\frac{\partial^2 \mathbf{R}}{\partial X_{jk}^2}(\theta X_{jk}, \theta X_{kj})\right]_{k+n,j} \right| +$$

$$(|X_{kj}|^2 |X_{jk}| + |X_{kj}|) \left| \left[\frac{\partial^2 \mathbf{R}}{\partial X_{kj}^2}(\theta X_{jk}, \theta X_{kj})\right]_{k+n,j} \right| +$$

$$(2|X_{jk}|^2 |X_{kj}| + |X_{jk}| + |X_{kj}|) \left| \left[\frac{\partial^2 \mathbf{R}}{\partial X_{jk} \partial X_{kj}}(\theta X_{jk}, \theta X_{kj})\right]_{k+n,j} \right|.$$

Let us consider the first term in the sum. The bounds for the second and third terms can be obtained by similar arguments. We have

$$\frac{\partial^2 \mathbf{R}}{\partial X_{jk}^2} = \frac{1}{n} \mathbf{R}(e_j e_{n+k}^T + e_{n+k} e_j^T) \mathbf{R}(e_j e_{n+k}^T + e_{n+k} e_j^T) \mathbf{R} \quad = \mathbb{P}_1 + \mathbb{P}_2 + \mathbb{P}_3 + \mathbb{P}_4,$$

where

$$\mathbb{P}_1 = \frac{1}{n} \mathbf{R} e_j e_{n+k}^T \mathbf{R} e_j e_{n+k}^T \mathbf{R}$$
$$\mathbb{P}_2 = \frac{1}{n} \mathbf{R} e_j e_{n+k}^T \mathbf{R} e_{n+k} e_j^T \mathbf{R}$$
$$\mathbb{P}_3 = \frac{1}{n} \mathbf{R} e_{n+k} e_j^T \mathbf{R} e_j e_{n+k}^T \mathbf{R}$$
$$\mathbb{P}_4 = \frac{1}{n} \mathbf{R} e_{n+k} e_j^T \mathbf{R} e_{n+k} e_j^T \mathbf{R}.$$

From $|\mathbf{R}_{i,j}| \leq v^{-1}$ it follows that

$$\frac{1}{n^{5/2}} \sum_{j,k=1}^{n} \mathbb{E}\, |X_{jk}|^\alpha |[\mathbb{P}_i]_{n+k,j}| \leq \frac{C}{\sqrt{n} v^3}$$

for $\alpha = 1, 3$ and $i = 1, ..., 4$. For $j = k$

$$\frac{1}{n^2} \sum_{j=1}^{n} \mathbb{E}\left[\frac{\partial \mathbf{R}}{\partial X_{jj}}\right]_{j+n,j} = \frac{1}{n^2} \sum_{j=1}^{n} (\mathbb{E}\, R_{j+n,j}^2 + \mathbb{E}\, R_{j,j} R_{j+n,j+n}) \leq \frac{C}{n v^2}.$$

So we can add this term to the sum

$$\frac{\rho}{n^2} \sum_{\substack{j,k=1 \\ j \neq k}}^{n} \mathbb{E}\left[\frac{\partial \mathbf{R}}{\partial X_{kj}}\right]_{k+n,j}.$$

□

Appendix B. Figures

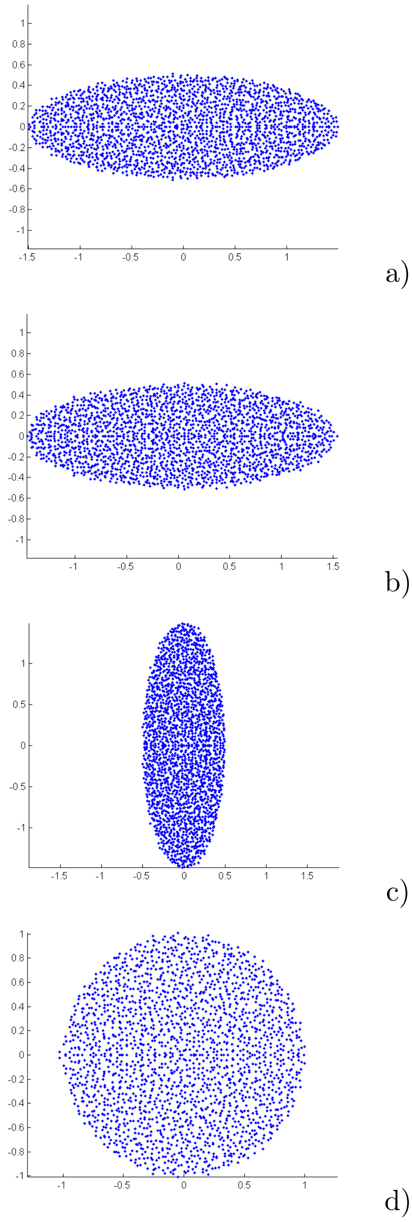
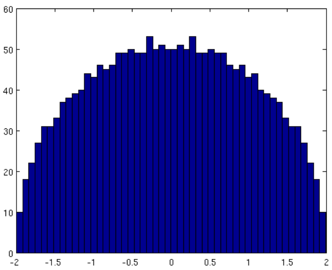
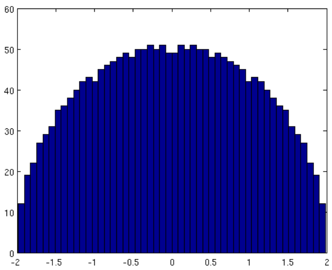


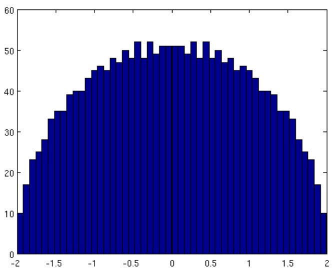
Figure 1. Spectrum of matrix $n^{-1/2}\mathbf{X}$ for $n = 2000$, a) entries are Gaussian random variables with $\rho = 0.5$ b) entries are Rademacher random variables with $\rho = 0.5$; c) entries are Gaussian random variables with $\rho = -0.5$ d) entries are Gaussian random variables with $\rho = 0$



a)



b)



c)

FIGURE 2. Histogram of eigenvalues of matrix $\mathbf{VJ}$ for $n = 1000$, entries are Gaussian random variables with a) $\rho = 0$ (Circular law case); b) $\rho = 0.5$ (Elliptic law case); c) $\rho = 1$ (Semi-circular law case)

References

[1] G. Akemann, J. Baik, and P. Di Francesco. *The Oxford Handbook of Random Matrix Theory*. Oxford Unversity Press, London, 2011.

[2] Z. Bai and J. W. Silverstein. *Spectral analysis of large dimensional random matrices*. Springer, New York, second edition, 2010.

[3] Charles Bordenave and Djalil Chafaï. Around the circular law. *arXiv:1109.3343*.

[4] John B. Conway. *Functions of one complex variable*, volume 11. Springer-Verlag, New York, second edition, 1978.

[5] Kevin Costello. Bilinear and quadratic variants on the Littlewood-Offord problem. *Submitted*.

[6] Yan V. Fyodorov, Boris A. Khoruzhenko, and Hans-Juergen. Sommers. Universality in the random matrix spectra in the regime of weak non-hermiticity. *Ann. Inst. Henri Poincare: Phys. Theor.*, 68(4):449–489, 1998.

[7] V. L. Girko. The elliptic law. *Teor. Veroyatnost. i Primenen.*, 30(4):640–651, 1985.

[8] V. L. Girko. The strong elliptic law. Twenty years later. *Random Oper. and Stoch. Equ.*, 14(1):59–102, 2006.

[9] Friedrich Götze and Alexander Tikhomirov. The circular law for random matrices. *Ann. Probab.*, 38(4):1444–1491, 2010.

[10] Michel Ledoux. Complex hermitian polynomials: from the semi-circular law to the circular law. *Commun. Stoch. Anal.*, 2(1):27–32, 2008.

[11] Mark Rudelson and Roman Vershynin. The Littlewood-Offord problem and invertibility of random matrices. *Adv. Math.*, 218(2):600–633, 2008.

[12] E. B. Saff and V. Totik. *Logarithmic potentials with external fields*, volume 316. Springer-Verlag, Berlin, 1997.

[13] Albert N. Shiryaev. *Probability*, volume 95. Springer-Verlag, New York, second edition, 1996.

[14] Hans-Juergen. Sommers, A. Crisanti, H. Sompolinsky, and Y. Stein. Spectrum of large random asymmetric matrices. *Phys. Rev. Lett.*, 60:1895–1898, May 1988.

[15] Terence Tao and Van Vu. Random matrices: universality of local eigenvalue statistics. *Acta Math.*, 206(1):127–204, 2011.

[16] Roman Vershynin. Invertibility of symmetric random matrices. *arXiv:1102.0300*.

A. Naumov, Faculty of Computational Mathematics and Cybernetics, Moscow State University, Moscow, Russia

E-mail address: naumovne@gmail.com, anaumov@math.uni-bielefeld.de