The Circular Law for Random Matrices

F. Götze
Faculty of Mathematics
University of Bielefeld
Germany

A. Tikhomirov[1]
Faculty of Mathematics and Mechanics
Sankt-Peterburg State University
S.-Peterburg, Russia

November 1, 2007

Abstract

We consider the joint distribution of real and imaginary parts of eigenvalues of random matrices with independent real entries with mean zero and unit variance. We prove the convergence of this distribution to the uniform distribution on the unit disc without assumptions on the existence of a density for the distribution of entries. We assume that the entries have a finite moment of order larger than two and consider the case of sparse matrices. The results are based on previous work of Bai, Rudelson and the authors extending results to a larger class of sparse matrices.

1 Introduction

Let $X_{jk}, 1 \le j, k < \infty$, be complex random variables with $\mathbb{E}X_{jk} = 0$ and $\mathbb{E}|X_{jk}|^2 = 1$. For a fixed $n \ge 1$, denote by $\lambda_1, \dots, \lambda_n$ the eigenvalues of the $n \times n$ matrix

$$\mathbf{X} = (X_n(j,k))_{j,k=1}^n, \quad X_n(j,k) = \frac{1}{\sqrt{n}} X_{jk}, \text{ for } 1 \le j, k \le n, \tag{1.1}$$

and define its empirical spectral distribution function by

$$G_n(x,y) = \frac{1}{n} \sum_{j=1}^n I_{\{\mathrm{Re}\{\lambda_j\} \le x, \, \mathrm{Im}\{\lambda_j\} \le y\}}, \tag{1.2}$$

where $I_{\{B\}}$ denotes the indicator of an event B. We investigate the convergence of the expected spectral distribution function $\mathbb{E}G_n(x,y)$ to the distribution function $G(x,y)$ of the uniform distribution over the unit disc in $\mathbb{R}^2$.

The main result of our paper is the following

[1]Partially supported by RFBF grant N 07-01-00583-a, by RF grant of the leading scientific schools NSh-4222.2006.1. Partially supported by DFG Project G0-420/5-1

Theorem 1.1. *Let X_{jk} be independent random variables with*

$$\mathbf{E}\,X_{jk} = 0, \qquad \mathbf{E}\,|X_{jk}|^2 = 1, \qquad \text{and} \quad \mathbf{E}\,|X_{jk}|^2\varphi(X_{jk}) \le \varkappa,$$

where $\varphi(x) = (\ln(1+|x|))^{19+\eta}$, for some $\eta > 0$. Then $\mathbf{E}\,G_n(x,y)$ converges weakly to the distribution function $G(x,y)$ as $n \to \infty$.

We shall prove the same result for the follows class of sparse matrices. Let ε_{jk}, $j,k = 1,\dots,n$ denote Bernoulli random variables which are independent in aggregate and independent of $(X_{jk})_{j,k=1}^n$ with $p_n := \Pr\{\varepsilon_{jk} = 1\}$. Consider the matrix $\mathbf{X}^{(\varepsilon)} = \frac{1}{\sqrt{np_n}}(\varepsilon_{jk}X_{jk})_{j,k=1}^n$. Let $\lambda_1^{(\varepsilon)},\dots,\lambda_n^{(\varepsilon)}$ denote the (complex) eigenvalues of the matrix $\mathbf{X}^{(\varepsilon)}$ and denote by $G_n^{(\varepsilon)}(x,y)$ the empirical spectral distribution function of the matrix $\mathbf{X}^{(\varepsilon)}$, i. e.

$$G_n^{(\varepsilon)}(x,y) := \frac{1}{n}\sum_{j=1}^n I_{\{\mathrm{Re}\,\{\lambda_j^{(\varepsilon)}\}\le x,\ \mathrm{Im}\,\{\lambda_j^{(\varepsilon)}\}\le y\}}. \tag{1.3}$$

Theorem 1.2. *Let X_{jk} be independent random variables with*

$$\mathbf{E}\,X_{jk} = 0, \qquad \mathbf{E}\,|X_{jk}|^2 = 1, \qquad \text{and} \quad \mathbf{E}\,|X_{jk}|^2\varphi(X_{jk}) \le \varkappa,$$

where $\varphi(x) = (\ln(1+|x|))^{19+\eta}$, for some $\eta > 0$. Assume that $p_n^{-1} = \mathcal{O}(n^{1-\theta})$ for some $1 \ge \theta > 0$. Then $\mathbf{E}\,G_n^{(\varepsilon)}(x,y)$ converges weakly to the distribution function $G(x,y)$ as $n \to \infty$.

Remark 1.3. The crucial problem of the proofs of Theorems 1.1 and 1.2 is to bound the smallest singular values $s_1(z)$ of the shifted matrices $\mathbf{X} - z\mathbf{I}$ and $\mathbf{X}^{(\varepsilon)} - z\mathbf{I}$. These bounds are based on the results obtained by Rudelson and Vershynin in [21]. In our preprint [10] we have used the corresponding results of Rudelson [20] proving the circular law in the case of i.i.d. sub-Gaussian random variables. In fact, the results in [10] actually imply the circular law for i.i.d. random variables with $\mathbf{E}\,|X_{jk}|^4 \le \varkappa_4 < \infty$ in view of the fact (explicitly stated by Rudelson in [20]) that in his results the sub-Gaussian condition is needed for the proof of $\Pr\{\|\mathbf{X}\| > K\} \le C\exp\{-cn\}$ only. Restricting oneself to the set $\Omega_n(z) = \{s_1(z) \le cn^{-3};\quad \|\mathbf{X}\| \le K\}$ for the investigation of the smallest singular values, the bound $\Pr\{\Omega_n^{(c)}\} \le c\,n^{-\frac{1}{2}}$ follows from the results of Rudelson [20] *without* the assumption of sub-Gaussian tails for the matrix $\mathbf{X}$. A similar result has been proved by Pan and Zhou in [15] based on results of Rudelson and Vershynin [21] and Bai and Silverstein [3].

The circular law assuming less restrictive moment condition of order larger than 2 only and comparable sparsity assumptions was proved independently by T. Tao and V. Vu in [25] based on the results of [26] in connection with the multivariate Littlewood Offord problem.

The approach in this paper though is based on the fruitful idea of Rudelson and Vershynin to characterize the vectors leading to small singular values of matrices with independent entries via 'compressible' and 'incompressible' vectors, see [21], Section 3.2,

p. 15. For the approximation of the distribution of singular values of $\mathbf{X} - z\mathbf{I}$ we use a scheme different from the approach used in Bai [1].

The investigation of the convergence the spectral distribution functions of real or complex (non-symmetric and non-Hermitian) random matrices with independent entries has a long history. Ginibre's in 1965, [7], studied the real, complex and quaternion matrices with i. i. d. Gaussian entries. He derived the joint density for the distribution of eigenvalues of matrix. Applying Ginibre formula Mehta in 1967, [17] determined the density of the expected spectral distribution function of random matrix with Gaussian entries with independent real and imaginary parts and deduced the circle law. Pastur suggested in 1973 the circular law for the general case (see [18], p. 64). Using the Ginibre results, Edelman in 1997, [5] proved the circular law for the matrices with i. i. d. Gaussian entries. Rider proved in [24] and [23] results about the spectral radius and about linear statistics of eigenvalues of non-Hermitian matrices with Gaussian entries.

Girko in 1984, [6], investigated the circular law for general matrices with independent entries assuming that the distribution of the entries have densities. As pointed out by Bai [1], Girko's proof had serious gaps. Bai in [1] gave a proof of the circular law for random matrices with independent entries assuming that the entries had bounded densities and finite sixth moments. His result does not cover the case of the Wigner ensemble and in particular ensembles of matrices with Rademacher entries. These ensembles are of some interest in various applications, see e.g. [27]. Girko's [6] approach using families of spectra of Hermitian matrices for a characterization of the circular-law based on the so-called *V-transform* was fruitful for all later work. See, for example, Girko's Lemma 1 in [1]. In fact, Girko [6] was the first who used the logarithmic potential to prove the circular law. We shall outline his approach using logarithmic potential theory. Let ξ denote a random variable uniformly distributed over the unit disc and independent of the matrix $\mathbf{X}$. For any $r > 0$, consider the matrix,

$$\mathbf{X}(r) = \mathbf{X} - r\xi\mathbf{I},$$

where $\mathbf{I}$ denotes the identity matrix of order n. Let $\mu_n^{(r)}$ (resp. μ_n) be empirical spectral measure of matrix $\mathbf{X}(r)$ (resp. $\mathbf{X}$) defined on the complex plane as empirical measure of the set of eigenvalues of matrix. We define a logarithmic potential of the expected spectral measure $\mathbf{E}\,\mu_n^{(r)}(ds, dt)$ as

$$U_n^{(r)}(z) = -\frac{1}{n}\mathbf{E}\,\log|\det(\mathbf{X}(r) - z\mathbf{I})| = -\frac{1}{n}\sum \mathbf{E}\,\log|\lambda_j - z - r\xi|,$$

where $\lambda_1, \dots, \lambda_n$ are the eigenvalues of the matrix $\mathbf{X}$. Note that the expected spectral measure $\mathbf{E}\,\mu_n^{(r)}$ is the convolution of the measure $\mathbf{E}\,\mu_n$ and the uniform distribution on the disc of radius r (see Lemma 6.4 in the Appendix for details).

Lemma 1.1. *Assume that the sequence $\mathbf{E}\,\mu_n^{(r)}$ converges weakly to a measure μ as $n \to \infty$ and $r \to 0$. Then*

$$\mu = \lim_{n\to\infty} \mathbf{E}\,\mu_n. \tag{1.4}$$

Proof. Let J be a random variable which is uniformly distributed on the set $\{1,\dots,n\}$ and independent of the matrix $\mathbf{X}$. We may represent the measure $\mathbf{E}\,\mu_n^{(r)}$ as distribution of a random variable $\lambda_J + r\xi$ where λ_J and ξ are independent. Computing the characteristic function of this measure and passing first to the limit with respect to $n \to \infty$ and then with respect to $r \to 0$ (see also Lemma 6.5 in the Appendix), we conclude the result. ☐

Now we may fix $r > 0$ and consider the measures $\mathbf{E}\,\mu_n^{(r)}$. They have bounded densities. Assume that the measures $\mathbf{E}\,\mu_n$ have supports in a fixed compact set and that $\mathbf{E}\,\mu_n$ converges weakly to a measure μ. Applying Theorem 6.9 (Lower Envelope Theorem) from [16], p. 73 (see also Subsection 6.1 in the Appendix), we obtain that under these assumptions

$$\liminf_{n\to\infty} U_n^{(r)}(z) = U^{(r)}(z), \tag{1.5}$$

for quasi-everywhere in $\mathbb{C}$ (for the definition of "*quasi-everywhere*" see for example [16], p 24 and Subsection 6.1 in the Appendix). Here $U^{(r)}(z)$ denotes the logarithmic potential of measure $\mu^{(r)}$ which is the convolution of a measure μ and of the uniform distribution on the disc of radius r. Furthermore, note that $U^{(r)}(z)$ may be represented as

$$U^{(r)}(z_0) = \frac{2}{r^2}\int_0^r vL(\mu; z_0, v)dv,$$

where

$$L(\mu; z_0, v) = \frac{1}{2\pi}\int_{-\pi}^{\pi} U^{(\mu)}(z_0 + v\exp\{i\theta\})d\theta. \tag{1.6}$$

Applying Theorem 1.2 in [16], p. 84, (Theorem 6.2 in Subsection 6.1 in the Appendix) we get

$$\lim_{r\to 0} U_\mu^{(r)}(z) = U_\mu(z).$$

Let $s_1(\mathbf{X}) \ge \ldots \ge s_n(\mathbf{X})$ denote the singular values of the matrix $\mathbf{X}$.

Since $\mathbf{E}\,\mathrm{Tr}\,(\mathbf{X}\mathbf{X}^*) = 1$ the sequence of measures $\mathbf{E}\,\mu_n$ is weakly relatively compact. These results imply that for any $\eta > 0$ we may restrict the measures $\mathbf{E}\,\mu_n$ to some compact set K_η such that $\sup_n \mathbf{E}\,\mu_n(K_\eta^{(c)}) < \eta$. Moreover, Lemma 6.2 implies the existence of a compact K such that $\lim_{n\to\infty}\sup_n \mathbf{E}\,\mu_n(K^{(c)})$=0. If we take some subsequence of the sequence of restricted measures $\mathbf{E}\,\mu_n$ which converges to some measure μ, then $\liminf_{n\to\infty} U_{\mu_n}^{(r)}(z) = U_\mu^{(r)}(z)$, $r > 0$ and $\lim_{r\to 0} U_\mu^{(r)}(z) = U_\mu(z)$. If we prove that $\liminf_{n\to\infty} U_{\mu_n}^{(r)}(z)$ exists and $U_\mu(z)$ is equal to the logarithmic potential corresponding the uniform distribution on the unit disc then the sequence of measures $\mathbf{E}\,\mu_n$ weakly converges to the uniform distribution on the unit disc. Moreover, it is enough to prove that for some sequence $r = r(n) \to 0$, $\lim_{n\to\infty} U_{\mu_n}^{(r)}(z) = U_\mu(z)$.

Furthermore, let $s_1^{(\varepsilon)}(z,r) \ge \ldots \ge s_n^{(\varepsilon)}(z,r)$ denote the singular values of matrix $\mathbf{X}^{(\varepsilon)}(z,r) = \mathbf{X}^{(\varepsilon)}(r) - z\mathbf{I}$. We shall investigate the logarithmic potential $U_{\mu_n}^{(r)}(z)$. Using elementary properties of singular values (see for instance Lemma 3.3 [8], p.35), we may represent the function $U^{(r)}_{\mu_n}(z)$ as follows

$$U^{(r)}_{\mu_n}(z) = -\frac{1}{n}\sum_{j=1}^n \mathbf{E}\,\log s_j^{(\varepsilon)}(z,r) = -\frac{1}{2}\int_0^\infty \log x\ \nu_n^{(\varepsilon)}(dx,z,r),$$

where $\nu_n^{(\varepsilon)}(\cdot,z,r)$ denotes the expected spectral measure of the matrix $\mathbf{H}_n^{(\varepsilon)}(z,r) = (\mathbf{X}^{(\varepsilon)}(r) - z\mathbf{I})(\mathbf{X}^{(\varepsilon)}(r) - z\mathbf{I})^*$, which is the expectation of the counting measure of the set of eigenvalues of the matrix $\mathbf{H}_n^{(\varepsilon)}(z,r)$).

In Section 2 we investigate convergence of measure $\nu_n^{(\varepsilon)}(\cdot,z) = \nu^{(\varepsilon)}(\cdot,z,0)$. In Section 3 we study the properties of the limit measures $\nu(\cdot,z)$. But the crucial problem for the proof of the circular law is the so called "regularization of potential" problem. We solve this problem using bounds for the minimal singular values of matrices $\mathbf{X}^{(\varepsilon)}(z) := \mathbf{X}^{(\varepsilon)} - z\mathbf{I}$ based on techniques developed in Rudelson [20] and Rudelson and Vershynin [21]. These bounds are given in Section 4 and in the Appendix, Subsection 1.2. In Section 5 we give the proof of the main Theorem. In the Appendix we combine precise statements of relevant results from potential theory and some auxiliary inequalities for the resolvent matrices.

In the what follows we shall denote by C and c or $\alpha,\beta,\delta,\rho,\eta$ (without indexes) some general absolute constant which may be change from line to line. To specify a constant we shall use subindexes. By I_A we shall denote the indicator of an event A. For any matrix $\mathbf{G}$ we denote the Frobenius norm by $\|\mathbf{G}|_2$ and we denote by $\|\mathbf{G}\|$ the operator norm.

Acknowledgment. The authors would like to thank Terence Tao for drawing our attention to a gap in a first version of the paper. The authors would like to thank Dmitry Timushev for careful reading of the manuscript.

2 Convergence of $\nu_n^{(\varepsilon)}(\cdot,z)$

Denote by $F_n^{(\varepsilon)}(x,z)$ the distribution function of the measure $\nu_n^{(\varepsilon)}(\cdot,z)$,

$$F_n^{(\varepsilon)}(x,z) = \frac{1}{n}\sum_{j=1}^n \mathbf{E}\,I_{\{(s_j^{(\varepsilon)}(z))^2<x\}},$$

where $s_1^{(\varepsilon)}(z) \ge \ldots \ge s_n^{(\varepsilon)}(z) \ge 0$ denote the singular values of the matrix $\mathbf{X}^{(\varepsilon)}(z) = \mathbf{X}^{(\varepsilon)} - z\mathbf{I}$. For a positive random variable ξ and a Rademacher random variable (r. v.) κ consider the transformed r. v. $\widetilde{\xi} = \kappa\sqrt{\xi}$. If ζ has distribution function $F_n^{(\varepsilon)}(x,z)$ the variable $\widetilde{\zeta}$ has distribution function $\widetilde{F}_n^{(\varepsilon)}(x,z)$, given by

$$\widetilde{F}_n^{(\varepsilon)}(x,z) = \frac{1}{2}(1 + \operatorname{sgn}\{x\}F_n^{(\varepsilon)}(x^2,z))$$

for all real x. Note that this induces a one-to-one corresponds between the respective measures $\nu_n^{(\varepsilon)}(\cdot,z)$ and $\widetilde{\nu}_n^{(\varepsilon)}(\cdot,z)$. The limit distribution function of $F_n^{(\varepsilon)}(x,z)$ as $n\to\infty$, is denoted by $F(\cdot,z)$. The corresponding symmetrization $\widetilde{F}(x,z)$ is the limit of $\widetilde{F}_n^{(\varepsilon)}(x,z)$ as $n\to\infty$. We have

$$\sup_x |F_n^{(\varepsilon)}(x,z)-F(x,z)| = 2\sup_x |\widetilde{F}_n^{(\varepsilon)}(x,z)-\widetilde{F}(x,z)|.$$

Denote by $s_n^{(\varepsilon)}(\alpha,z)$ (resp. $s(\alpha,z)$) and $S_n^{(\varepsilon)}(x,z)$ (resp. $S(x,z)$) the Stieltjes transforms of the measures $\nu_n^{(\varepsilon)}(\cdot,z)$ (resp. $\nu(\cdot,z)$) and $\widetilde{\nu}_n^{(\varepsilon)}(\cdot,z)$ (resp. $\widetilde{\nu}(\cdot,z)$) correspondingly. Then we have

$$S_n^{(\varepsilon)}(\alpha,z)=\alpha s_n^{(\varepsilon)}(\alpha^2,z), \qquad S(\alpha,z)=\alpha s(\alpha^2,z).$$

Remark 2.1. As is shown in Bai [1], the measure $\nu(\cdot,z)$ has a density $p(x,z)$ with bounded support. More precisely, $p(x,z)\le C\max\{1,\frac{1}{\sqrt{x}}\}$. Thus the measure $\widetilde{\nu}(\cdot,z)$ has bounded support and bounded density $\widetilde{p}(x,z)=|x|p(x^2,z)$.

Theorem 2.2. *Let* $\mathbf{E}\,X_{jk}=0$, $\mathbf{E}\,|X_{jk}|^2=1$. *Assume for some function* $\varphi(x)>0$ *such that* $\varphi(x)\to\infty$ *as* $x\to\infty$ *and such that the function* $x/\varphi(x)$ *is non-decreasing we have*

$$\varkappa := \max_{1\le j,k<\infty} \mathbf{E}\,|X_{jk}|^2\varphi(X_{jk})<\infty. \tag{2.1}$$

Then

$$\sup_x |F_n^{(\varepsilon)}(x,z)-F(x,z)|\le C\varkappa(\varphi(\sqrt{np_n}))^{-\frac{1}{6}}. \tag{2.2}$$

Corollary 2.1. *Let* $\mathbf{E}\,X_{jk}=0$, $\mathbf{E}\,|X_{jk}|^2=1$, *and*

$$\varkappa = \max_{1\le j,k<\infty} \mathbf{E}\,|X_{jk}|^3<\infty. \tag{2.3}$$

Then

$$\sup_x |F_n^{(\varepsilon)}(x,z)-F(x,z)|\le C(np_n)^{-\frac{1}{12}}. \tag{2.4}$$

Proof. To bound the distance between the distribution functions $\widetilde{F}_n^{(\varepsilon)}(x,z)$ and $\widetilde{F}(x,z)$ we investigate the distance between their the Stieltjes transforms. Introduce the Hermitian $2n\times 2n$ matrix

$$\mathbf{W}=\begin{pmatrix}\mathbf{O}_n & (\mathbf{X}^{(\varepsilon)}-z\mathbf{I})\\ (\mathbf{X}^{(\varepsilon)}-z\mathbf{I})^* & \mathbf{O}_n\end{pmatrix},$$

where $\mathbf{O}_n$ denotes $n\times n$ matrix with zero entries. From Šur's complement formula (see for example [12], Ch. 08, p. 21) it follows that, for $\alpha=u+iv$, $v>0$,

$$(\mathbf{W}-\alpha\mathbf{I}_{2n})^{-1}=\begin{pmatrix}\alpha\left(\mathbf{X}^{(\varepsilon)}(z)(\mathbf{X}^{(\varepsilon)}(z))^*-\alpha^2\mathbf{I}\right)^{-1} & \mathbf{X}^{(\varepsilon)}(z)\left(\mathbf{X}^{(\varepsilon)}(z)(\mathbf{X}^{(\varepsilon)}(z))^*-\alpha^2\mathbf{I}\right)^{-1}\\ \left((\mathbf{X}^{(\varepsilon)}(z))^*\mathbf{X}^{(\varepsilon)}(z)-\alpha^2\mathbf{I}\right)^{-1}(\mathbf{X}^{(\varepsilon)}(z))^* & \alpha\left((\mathbf{X}^{(\varepsilon)}(z))^*\mathbf{X}^{(\varepsilon)}(z)-\alpha^2\mathbf{I}\right)^{-1}\end{pmatrix} \tag{2.5}$$

where $\mathbf{X}^{(\varepsilon)}(z) = \mathbf{X}^{(\varepsilon)} - z\mathbf{I}$ and $\mathbf{I}_{2n}$ denotes the unit matrix of order $2n$. By definition of $S_n^{(\varepsilon)}(\alpha, z)$, we have

$$S_n^{(\varepsilon)}(\alpha, z) = \frac{1}{2n}\mathbf{E}\,\mathrm{Tr}\,(\mathbf{W} - \alpha\mathbf{I}_{2n})^{-1}.$$

Set $\mathbf{R}(\alpha, z) := (R_{j,k}(\alpha, z))_{j,k=1}^{2n} = (\mathbf{W} - \alpha\mathbf{I}_{2n})^{-1}$. It is easy to check that

$$1 + \alpha S_n^{(\varepsilon)}(\alpha, z) = \frac{1}{2n}\mathbf{E}\,\mathrm{Tr}\,\mathbf{W}\mathbf{R}(\alpha, z).$$

We may rewrite this equality as

$$\begin{aligned}1 + \alpha S_n^{(\varepsilon)}(\alpha, z) &= \frac{1}{2n\sqrt{np_n}}\sum_{j,k=1}^{n}\mathbf{E}\,(\varepsilon_{jk}X_{jk}R_{k+n,j}(\alpha, z) + \varepsilon_{jk}\overline{X}_{jk}R_{j+n,k}(\alpha, z)) \\ &\quad - \frac{\overline{z}}{2n}\sum_{j=1}^{n}\mathbf{E}\,R_{j,j+n}(\alpha, z) - \frac{z}{2n}\sum_{j=1}^{n}\mathbf{E}\,R_{j+n,j}(\alpha, z).\end{aligned} \tag{2.6}$$

We introduce the notations

$$\begin{aligned}\mathbf{A} &= (\mathbf{X}^{(\varepsilon)}(z)(\mathbf{X}^{(\varepsilon)}(z))^* - \alpha^2\mathbf{I})^{-1}, \quad \mathbf{B} = \mathbf{X}^{(\varepsilon)}(z)\mathbf{A}, \\ \mathbf{C} &= ((\mathbf{X}^{(\varepsilon)}(z))^*\mathbf{X}^{(\varepsilon)}(z) - \alpha^2\mathbf{I})^{-1}, \quad \mathbf{D} = \mathbf{C}(\mathbf{X}^{(\varepsilon)}(z))^*.\end{aligned}$$

With these notations we rewrite equality (2.5) as follows

$$\mathbf{R}(\alpha, z) = (\mathbf{W} - \alpha\mathbf{I}_{2n})^{-1} = \begin{pmatrix}\alpha\mathbf{A} & \mathbf{B} \\ \mathbf{D} & \alpha\mathbf{C}\end{pmatrix}. \tag{2.7}$$

Equalities (2.7) and (2.6) together imply

$$\begin{aligned}1 + \alpha S_n^{(\varepsilon)}(\alpha, z) &= \frac{1}{2n\sqrt{np_n}}\sum_{j,k=1}^{n}\mathbf{E}\,(\varepsilon_{jk}X_{jk}R_{k+n,j}(\alpha, z) + \varepsilon_{jk}\overline{X}_{jk}R_{j,k+n}(\alpha, z)) \\ &\quad - \frac{z}{2n}\mathbf{E}\,\mathrm{Tr}\,\mathbf{D} - \frac{\overline{z}}{2n}\mathbf{E}\,\mathrm{Tr}\,\mathbf{B}.\end{aligned} \tag{2.8}$$

In the what follows we shall use a simple resolvent equality. For two matrices $\mathbf{U}$ and $\mathbf{V}$ let $\mathbf{R}_U = (\mathbf{U} - \alpha\mathbf{I})^{-1}$, $\mathbf{R}_{U+V} = (\mathbf{U} + \mathbf{V} - \alpha\mathbf{I})^{-1}$, then

$$\mathbf{R}_{U+V} = \mathbf{R}_U - \mathbf{R}_U\mathbf{V}\mathbf{R}_{U+V}.$$

Let $\{\mathbf{e}_1, \dots \mathbf{e}_{2n}\}$ denote the canonical orthonormal basis in $\mathbb{R}^{2n}$. Let $\mathbf{W}^{(jk)}$ denote the matrix is obtained from $\mathbf{W}$ by replacing the both entries $X_{j,k}$ and $\overline{X}_{j,k}$ by 0. In our notation we may write

$$\mathbf{W} = \mathbf{W}^{(jk)} + \frac{1}{\sqrt{np_n}}\varepsilon_{jk}X_{jk}\mathbf{e}_j\mathbf{e}_{k+n}^T + \frac{1}{\sqrt{np_n}}\varepsilon_{jk}\overline{X}_{jk}\mathbf{e}_{k+n}\mathbf{e}_j^T. \tag{2.9}$$

Using this representation and the resolvent equality, we get

$$\mathbf{R} = \mathbf{R}^{(j,k)} - \frac{1}{\sqrt{np_n}}\varepsilon_{jk}X_{jk}\mathbf{R}^{(j,k)}\mathbf{e}_j\mathbf{e}_{k+n}^T\mathbf{R} - \frac{1}{\sqrt{np_n}}\varepsilon_{jk}\overline{X}_{jk}\mathbf{R}^{(j,k)}\mathbf{e}_{k+n}\mathbf{e}_j^T\mathbf{R}. \tag{2.10}$$

Here and in the what follows we omit the arguments α and z in the notation of resolvent matrices. For any vector $\mathbf{a}$, let $\mathbf{a}^T$ denote the transposed vector $\mathbf{a}$. Applying the resolvent equality again, we obtain

$$\mathbf{R} = \mathbf{R}^{(j,k)} - \frac{1}{\sqrt{np_n}}\varepsilon_{jk}X_{jk}\mathbf{R}^{(j,k)}\mathbf{e}_j\mathbf{e}_{k+n}^T\mathbf{R}^{(j,k)} - \frac{1}{\sqrt{np_n}}\varepsilon_{jk}\overline{X}_{jk}\mathbf{R}^{(j,k)}\mathbf{e}_{k+n}\mathbf{e}_j^T\mathbf{R}^{(j,k)} + \mathbf{T}^{(jk)}, \tag{2.11}$$

where

$$\begin{aligned}\mathbf{T}^{(jk)} = {} & \frac{1}{\sqrt{np_n}}\varepsilon_{jk}X_{jk}\mathbf{R}^{(j,k)}\mathbf{e}_j\mathbf{e}_{k+n}^T(\mathbf{R}^{(j,k)} - \mathbf{R}) \\ & + \frac{1}{\sqrt{np_n}}\varepsilon_{jk}X_{jk}\mathbf{R}^{(j,k)}\mathbf{e}_j\mathbf{e}_{k+n}^T(\mathbf{R}^{(j,k)} - \mathbf{R}) \\ & + \frac{1}{\sqrt{np_n}}\varepsilon_{jk}(\overline{X}_{jk})\mathbf{R}^{(j,k)}\mathbf{e}_{k+n}\mathbf{e}_j^T(\mathbf{R}^{(j,k)} - \mathbf{R}) \\ & + \frac{1}{\sqrt{np_n}}\varepsilon_{jk}X_{jk}\mathbf{R}^{(j,k)}\mathbf{e}_{k+n}\mathbf{e}_j^T(\mathbf{R}^{(j,k)} - \mathbf{R}).\end{aligned} \tag{2.12}$$

This implies

$$\begin{aligned}\mathbf{R}_{j,k+n} &= \mathbf{R}_{j,k+n}^{(j,k)} - \frac{1}{\sqrt{np_n}}\varepsilon_{jk}X_{jk}\mathbf{R}_{j,j}^{(j,k)}\mathbf{R}_{k+n,k+n}^{(j,k)} - \frac{1}{\sqrt{np_n}}\varepsilon_{jk}\overline{X}_{jk}(\mathbf{R}_{j,k+n}^{(j,k)})^2 + \mathbf{T}_{j,k+n}^{(j,k)}, \\ \mathbf{R}_{k+n,j} &= \mathbf{R}_{k+n,j}^{(j,k)} - \frac{1}{\sqrt{np_n}}\varepsilon_{jk}X_{jk}\mathbf{R}_{k+n,j}^{(j,k)}\mathbf{R}_{j,k+n}^{(j,k)} - \frac{1}{\sqrt{np_n}}\varepsilon_{jk}\overline{X}_{jk}\mathbf{R}_{k+n,k+n}^{(j,k)}\mathbf{R}_{j,j}^{(j,k)} + \mathbf{T}_{k+n,j}^{(j,k)}.\end{aligned} \tag{2.13}$$

Applying these notations to the equality (2.8) and taking into account that X_{jk} and $\mathbf{R}^{(jk)}$ are independent, we get

$$\begin{aligned}1 + \alpha S_n^{(\varepsilon)}(\alpha, z) + \frac{z}{2n}\operatorname{Tr}\mathbf{D} + \frac{\overline{z}}{2n}\operatorname{Tr}\mathbf{B} = {} & -\frac{1}{n^2p_n}\sum_{j,k=1}^{n}\mathbf{E}\,\varepsilon_{jk}R_{j,j}^{(j,k)}R_{k+n,k+n}^{(j,k)} \\ & - \frac{1}{2n^2p_n}\sum_{j,k=1}^{n}\mathbf{E}\,\varepsilon_{jk}|X_{jk}|^2\mathbf{E}\,(R_{j,k+n}^{(j,k)})^2 \\ & - \frac{1}{2n\sqrt{np_n}}\sum_{j,k=1}^{n}\mathbf{E}\,(\varepsilon_{jk}X_{jk}T_{k+n,j}^{(j,k)} + \varepsilon_{jk}\overline{X}_{jk}T_{j,k+n}^{(j,k)}).\end{aligned} \tag{2.14}$$

From (2.10) it follows immediately that for any $p, q = 1, \ldots, 2n,\ j, k = 1, \ldots n$,

$$|R_{p,p} - R^{(j,k)}_{p,p}| \le \frac{C\varepsilon_{jk}|X_{jk}|}{\sqrt{np_n}}(|R^{jk}_{pj}||R_{k+n,p}| + |R^{jk}_{p,k+n}||R_{jp}|). \tag{2.15}$$

Since $\sum_{m,l=1}^{n} |R_{m,l}|^2 \le n/v^2$ and $\sum_{m,l=1}^{n} |R^{jk}_{m,l}|^2 \le n/v^2$, equality (2.13) implies

$$\frac{1}{n^2}\sum_{j,k=1}^{n} \mathbf{E}\,|R^{(j,k)}_{j,k+n}|^2 \le \frac{C}{nv^4}. \tag{2.16}$$

By definition (2.12) of $\mathbf{T}^{(j,k)}$, applying standard resolvent properties, we obtain the following bounds, for any $z = u + iv,\ v > 0$,

$$\frac{1}{n\sqrt{np_n}}\sum_{j,k=1}^{n} \mathbf{E}\,\varepsilon_{jk}|X_{jk}||T^{(j,k)}_{j,k+n}| \le \frac{C\varkappa}{v^3\varphi(\sqrt{np_n})}. \tag{2.17}$$

For the proof of this inequality see Lemma 6.3 in the Appendix. Using the last inequalities we obtain, that for $v > 0$

$$\begin{aligned}\Bigg|\frac{1}{n}\sum_{j=1}^{n}\mathbf{E}\,R_{jj}\frac{1}{n}\sum_{k=1}^{n}R_{k+n,k+n} - \frac{1}{n^2}\sum_{j=1}^{n}\sum_{k=1}^{n}\mathbf{E}\,R^{(jk)}_{jj}R^{(jk)}_{k+n,k+n}\Bigg| &\\ \le \frac{C}{n^2\sqrt{np_n}v}\sum_{j=1}^{n}\sum_{k=1}^{n}\mathbf{E}\,\varepsilon_{jk}|X_{jk}|(|R^{(jk)}_{jj}||R_{k+n,j}| + |R^{(jk)}_{j,k+n}||R_{jj}|) &\\ \le \frac{C}{nv^4}. &\end{aligned} \tag{2.18}$$

Since $\frac{1}{n}\sum_{j=1}^{n} R_{jj} = \frac{1}{n}\sum_{k=1}^{n} R_{k+n,k+n} = \frac{1}{2n}\mathrm{Tr}\,\mathbf{R}(\alpha, z)$, we obtain

$$\Bigg|\frac{1}{n^2}\sum_{j=1}^{n}\sum_{k=1}^{n}\mathbf{E}\,R^{(jk)}_{jj}R^{(jk)}_{k+n,k+n} - \mathbf{E}\,(\frac{1}{2n}\mathrm{Tr}\,\mathbf{R}(\alpha, z))^2\Bigg| \le \frac{C}{nv^4}. \tag{2.19}$$

Note that for any Hermitian random matrix $\mathbf{W}$ with independent entries on and above the diagonal we have

$$\mathbf{E}\,\Bigg|\frac{1}{n}\mathrm{Tr}\,\mathbf{R}(\alpha, z) - \mathbf{E}\,\frac{1}{n}\mathrm{Tr}\,\mathbf{R}(\alpha, z)\Bigg|^2 \le \frac{C}{nv^2}. \tag{2.20}$$

The proof of this inequality is easy and due to a martingale type expansion already used by Girko. Inequalities (2.19) and (2.20) together imply that for $v > 0$

$$\Bigg|\frac{1}{n^2}\sum_{j=1}^{n}\sum_{k=1}^{n}\mathbf{E}\,R^{(jk)}_{jj}R^{(jk)}_{k+n,k+n} - (S^{(\varepsilon)}_n(\alpha, z))^2\Bigg| \le \frac{C}{nv^4}. \tag{2.21}$$

Denote by $r(\alpha,z)$ some generic function with $|r(\alpha,z)|\le 1$ not necessary the same from line to line. We may now rewrite equality (2.8) as follows

$$1+\alpha S_n^{(\varepsilon)}(\alpha,z)+(S_n^{(\varepsilon)}(\alpha,z))^2=-\frac{z}{2n}\mathbf{E}\operatorname{Tr}\mathbf{D}-\frac{\overline{z}}{2n}\mathbf{E}\operatorname{Tr}\mathbf{B}+\frac{r(\alpha,z)}{v^3\varphi(\sqrt{np_n})}. \tag{2.22}$$

where $v>c\varphi(\sqrt{np_n})/n$.

We now investigate the functions $T(\alpha,z)=\frac{1}{n}\mathbf{E}\operatorname{Tr}\mathbf{B}$ and $V(\alpha,z)=\frac{1}{n}\mathbf{E}\,\mathbf{D}$. Since the arguments for both functions are similar we provide it for the first one only. By definition of the matrix $\mathbf{B}$, we have

$$\operatorname{Tr}\mathbf{B}=\frac{1}{\sqrt{np_n}}\sum_{j,k=1}^{n}\varepsilon_{jk}X_{j,k}(\mathbf{X}^{(\varepsilon)}(z)(\mathbf{X}^{(\varepsilon)}(z))^*-\alpha^2)^{-1})_{kj}-z\operatorname{Tr}\mathbf{A}.$$

According to equality (2.7), we have

$$\operatorname{Tr}\mathbf{B}=\frac{1}{\alpha\sqrt{np_n}}\sum_{j,k=1}^{n}\varepsilon_{jk}X_{j,k}R_{kj}-z\operatorname{Tr}\mathbf{A}.$$

Using the resolvent equality (2.10) and Lemma 6.3, we get, for $v>c\varphi(\sqrt{np_n})/n$

$$T(\alpha,z)=-\frac{1}{\alpha n^2}\sum_{j,k=1}^{n}\mathbf{E}\,R_{k,k+n}^{(jk)}R_{jj}^{(jk)}-\frac{z}{\alpha}S_n^{(\varepsilon)}(\alpha,z)+\frac{C\varkappa r(\alpha,z)}{v^3\varphi(\sqrt{np_n})}. \tag{2.23}$$

Similar to (2.21) we obtain

$$\Big|\frac{1}{n^2}\sum_{j,k=1}^{n}\mathbf{E}\,R_{jj}^{(jk)}R_{k,k+n}^{(jk)}-V(\alpha,z)S_n^{(\varepsilon)}(\alpha,z)\Big|\le\frac{C}{nv^4}. \tag{2.24}$$

Inequalities (2.23) and (2.24) together imply, for $v>c\varphi(\sqrt{np_n})/n$,

$$T(\alpha,z)=-\frac{zS_n^{(\varepsilon)}(\alpha,z)}{\alpha+S_n^{(\varepsilon)}(\alpha,z)}+\frac{C\varkappa r(\alpha,z)}{\varphi(\sqrt{np_n})v^3|\alpha+S_n^{(\varepsilon)}(\alpha,z)|}. \tag{2.25}$$

Analogously we get

$$V(\alpha,z)=-\frac{\overline{z}S_n^{(\varepsilon)}(\alpha,z)}{\alpha+S_n^{(\varepsilon)}(\alpha,z)}+\theta\frac{C}{\varphi(\sqrt{np_n})v^3|\alpha+S_n^{(\varepsilon)}(\alpha,z)|}. \tag{2.26}$$

Inserting (2.25) and (2.26) in (2.14), we get

$$(S_n^{(\varepsilon)}(\alpha,z))^2+\alpha S_n^{(\varepsilon)}(\alpha,z)+1-\frac{|z|^2S_n^{(\varepsilon)}(\alpha,z)}{\alpha+S_n^{(\varepsilon)}(\alpha,z)}=\delta_n(z), \tag{2.27}$$

where

$$|\delta_n(\alpha,z)| \le \frac{C\varkappa}{\varphi(\sqrt{np_n})v^3|S_n^{(\varepsilon)}(\alpha,z)+\alpha|}.$$

or equivalently

$$S_n^{(\varepsilon)}(\alpha,z)\left(\alpha+S_n^{(\varepsilon)}(\alpha,z)\right)^2+\left(\alpha+S_n(^{(\varepsilon)}\alpha,z)\right)-|z|^2S_n^{(\varepsilon)}(\alpha,z)=\widetilde{\delta}_n(\alpha,z), \tag{2.28}$$

where $\widetilde{\delta}_n(\alpha,z)=\theta\frac{C\varkappa r(\alpha,z)}{\varphi(\sqrt{np_n})v^3}$. We may rewrite the last equation as

$$S_n^{(\varepsilon)}(\alpha,z)=-\frac{\alpha+S_n^{(\varepsilon)}(\alpha,z)}{(\alpha+S_n^{(\varepsilon)}(\alpha,z))^2-|z|^2}+\widehat{\delta}_n(\alpha,z), \tag{2.29}$$

where

$$\widehat{\delta}_n(\alpha,z)=\frac{\widetilde{\delta}_n(\alpha,z)}{(\alpha+S_n^{(\varepsilon)}(\alpha,z))^2-|z|^2}. \tag{2.30}$$

Furthermore, we prove the following simple Lemma.

Lemma 2.2. *Let $\alpha=u+iv$, $v>0$. Let $S(\alpha,z)$ satisfy the equation*

$$S(\alpha,z)=-\frac{\alpha+S(\alpha,z)}{(\alpha+S(\alpha,z))^2-|z|^2}. \tag{2.31}$$

and $\operatorname{Im}\{S(\alpha,z)\}>0$. *Then the following inequality*

$$1-|S(\alpha,z)|^2-\frac{|z|^2|S(\alpha,z)|^2}{|\alpha+S(\alpha,z)|^2}\ge\frac{v}{v+1}.$$

holds.

Proof. For $\alpha=u+iv$ with $v>0$, the Stieltjes transform $S(\alpha,z)$ satisfies the following equation

$$S(\alpha,z)=-\frac{\alpha+S(\alpha,z)}{(\alpha+S(\alpha,z))^2-|z|^2}. \tag{2.32}$$

Comparing the imaginary parts of both sides of this equation, we get

$$\operatorname{Im}\{\alpha+S(\alpha,z)\}=\operatorname{Im}\{\alpha+S(\alpha,z)\}\frac{|\alpha+S(\alpha,z)|^2+|z|^2}{|(\alpha+S(\alpha,z))^2-|z|^2|^2}+v. \tag{2.33}$$

Equations (2.31) and (2.33) together imply

$$\operatorname{Im}\{\alpha+S(\alpha,z)\}\left(1-\frac{|\alpha+S(\alpha,z)|^2+|z|^2}{|(\alpha+S(\alpha,z))^2-|z|^2|^2}\right)=v. \tag{2.34}$$

Since $v>0$ and $\operatorname{Im}\{\alpha+S(\alpha,z)\}>0$, it follows that

$$1-\frac{|\alpha+S(\alpha,z)|^2+|z|^2}{|(\alpha+S(\alpha,z))^2-|z|^2|^2}>0.$$

In particular we have

$$|S(\alpha,z)| \le 1.$$

Inequality (2.34) and the last remark together imply

$$1 - \frac{|\alpha + S(\alpha,z)|^2 + |z|^2}{|(\alpha+S(\alpha,z))^2 - |z|^2|^2} = \frac{v}{\operatorname{Im}\{\alpha + S(\alpha,z)\}} \ge \frac{v}{v+1}.$$

The proof is completed. □

To compare the functions $S(\alpha,z)$ and $S_n(\alpha,z)$ we prove

Lemma 2.3. *Let*

$$|\widehat{\delta}_n(\alpha,z)| \le \frac{v}{2}.$$

Then the following inequality holds

$$1 - \frac{|\alpha + S_n^{(\varepsilon)}(\alpha,z)|^2 + |z|^2}{|(\alpha+S_n^{(\varepsilon)}(\alpha,z))^2 - |z|^2|^2} \ge \frac{v}{4}.$$

Proof. By the assumption, we have

$$\operatorname{Im}\{\widehat{\delta}_n(\alpha,z) + \alpha\} > \frac{v}{2}.$$

Repeating the arguments of Lemma 2.2 completes the proof. □

The next Lemma give us a bound for the distance between the Stieltjes transforms $S(\alpha,z)$ and $S_n^{(\varepsilon)}(\alpha,z)$.

Lemma 2.4. *Let*

$$|\widehat{\delta}_n(\alpha,z)| \le \frac{v}{8}.$$

Then

$$|S_n^{(\varepsilon)}(\alpha,z) - S(\alpha,z)| \le \frac{4|\widehat{\delta}_n(\alpha,z)|}{v}.$$

Proof. Note that $S(\alpha,z)$ and $S_n^{(\varepsilon)}(\alpha,z)$ satisfy the equations

$$S(\alpha,z) = -\frac{\alpha + S(\alpha,z)}{(\alpha+S(\alpha,z)^2 - |z|^2} \tag{2.35}$$

and

$$S_n^{(\varepsilon)}(\alpha,z) = -\frac{\alpha + S_n^{(\varepsilon)}(\alpha,z)}{(\alpha+S_n^{(\varepsilon)}(\alpha,z)^2 - |z|^2} + \widehat{\delta}_n(\alpha,z) \tag{2.36}$$

respectively. These equations together imply

$$S(\alpha,z) - S_n^{(\varepsilon)}(\alpha,z) = \frac{(\alpha + S_n^{(\varepsilon)}(\alpha,z))(\alpha+S(\alpha,z)) + |z|^2}{((\alpha+S(\alpha,z)^2 - |z|^2)((\alpha+S_n^{(\varepsilon)}(\alpha,z)^2 - |z|^2)} \times (S(\alpha,z) - S_n^{(\varepsilon)}(\alpha,z)) + \widehat{\delta}_n(\alpha,z). \tag{2.37}$$

Applying inequality $|ab| \le \frac{1}{2}(a^2+b^2)$, we get

$$\begin{aligned}\left|1-\frac{(\alpha+S_n^{(\varepsilon)}(\alpha,z))(\alpha+S(\alpha,z))+|z|^2}{((\alpha+S(\alpha,z)^2-|z|^2)((\alpha+S_n^{(\varepsilon)}(\alpha,z)^2-|z|^2)}\right| \\ \ge \frac{1}{2}\left(1-\frac{|\alpha+S_n(\alpha,z)|^2+|z|^2}{|(\alpha+S_n^{(\varepsilon)}(\alpha,z))^2-|z|^2|^2}\right) \\ +\frac{1}{2}\left(1-\frac{|\alpha+S(\alpha,z)|^2+|z|^2}{|(\alpha+S(\alpha,z))^2-|z|^2|^2}\right).\end{aligned}$$

The last inequality and Lemmas 2.2 and 2.3 together imply

$$\left|1-\frac{(\alpha+S_n^{(\varepsilon)}(\alpha,z))(\alpha+S(\alpha,z))+|z|^2}{((\alpha+S(\alpha,z)^2-|z|^2)((\alpha+S_n^{(\varepsilon)}(\alpha,z)^2-|z|^2)}\right| \ge \frac{v}{4}.$$

This completes the proof of the Lemma. □

To bound the distance between the distribution function $F_n(x,z)$ and the distribution function $F(x,z)$ corresponding the Stieltjes transforms $S_n(\alpha,z)$ and $S(\alpha,z)$ we use Corollary 2.3 from [9]. In the next lemma we give an integral bound for the distance between the Stieltjes transforms $S(\alpha,z)$ and $S_n^{(\varepsilon)}(\alpha,z)$.

Lemma 2.5. *For $v \ge v_0(n) = c(\varphi(\sqrt{np_n}))^{-\frac{1}{6}}$ the inequality*

$$\int_{-\infty}^{\infty} |S(\alpha,z)-S_n^{(\varepsilon)}(\alpha,z)|du \le \frac{C(1+|z|^2)\varkappa}{\varphi(\sqrt{np_n})v^7}.$$

holds.

Proof. Note that

$$|(\alpha+s_n^{(\varepsilon}(\alpha,z))^2-|z|^2| \ge |\alpha+s_n^{(\varepsilon}(\alpha,z)-|z|)||\alpha+s_n^{(\varepsilon}(\alpha,z)+|z|)|| \ge v^2. \tag{2.38}$$

It follows from here that $|\widehat{\delta}_n(\alpha,z)| \le \frac{C}{v^5\varphi(\sqrt{np_n})}$ and

$$|\widehat{\delta}_n(\alpha,z)| \le v/8$$

for $v \ge c(\varphi(\sqrt{np_n}))^{-1/6}$. Lemma 2.4 implies that it is enough to prove inequality

$$\int_{-\infty}^{\infty} |\widehat{\delta}_n(\alpha,z)|du \le C\gamma_n,$$

where $\gamma_n = \frac{C}{v^6\varphi(\sqrt{np_n})}$. By definition of $\widehat{\delta}(\alpha,z)$, we have

$$\int_{-\infty}^{\infty} |\widehat{\delta}_n(\alpha,z)|du \le \frac{c\varkappa}{v^3\varphi(\sqrt{np_n})}\int_{-\infty}^{\infty}\frac{du}{|(\alpha+S_n^{(\varepsilon)}(\alpha,z))^2-|z|^2|}. \tag{2.39}$$

Furthermore, the representation (2.29) implies that

$$\frac{1}{|(\alpha+S_n^{(\varepsilon)}(\alpha,z))^2-|z|^2|}\le\frac{|S_n^{(\varepsilon)}(\alpha,z)|}{|\alpha+S_n^{(\varepsilon)}(\alpha,z)|}+\frac{|\widehat{\delta}_n(\alpha,z)|}{|\alpha+S_n^{(\varepsilon)}(\alpha,z)|}. \tag{2.40}$$

Note that, according to the relation (2.27),

$$\frac{1}{|\alpha+S_n^{(\varepsilon)}(\alpha,z)|}\le\frac{|z|^2|S_n^{(\varepsilon)}(\alpha,z)|}{|\alpha+S_n^{(\varepsilon)}(\alpha,z)|^2}+|S_n^{(\varepsilon)}(\alpha,z)|+\frac{|\delta_n(\alpha,z)|}{|\alpha+S_n^{(\varepsilon)}(\alpha,z)|^2}. \tag{2.41}$$

This inequality implies

$$\int_{-\infty}^{\infty}\frac{|S_n^{(\varepsilon)}(\alpha,z)|}{|\alpha+S_n^{(\varepsilon)}(\alpha,z)|}du\le\frac{C(1+|z|^2)}{v^2}\int_{-\infty}^{\infty}|S_n^{(\varepsilon)}(\alpha,z)|^2du+\int_{-\infty}^{\infty}|\delta_n(\alpha,z)|\frac{|S_n^{(\varepsilon)}(\alpha,z)|}{|\alpha+S_n^{(\varepsilon)}(\alpha,z)|}du. \tag{2.42}$$

It follows from the relation (2.27), for $v>c(\varphi(\sqrt{np_n}))^{-\frac{1}{6}}$, that

$$|\delta_n(\alpha,z)|\le\frac{C\varkappa}{(\varphi(\sqrt{np_n}))v^4}<1/2. \tag{2.43}$$

The last two inequalities together imply that for sufficiently large n and $v>c(\varphi(\sqrt{np_n}))^{-\frac{1}{6}}$,

$$\int_{-\infty}^{\infty}\frac{|S_n^{(\varepsilon)}(\alpha,z)|}{|\alpha+S_n^{(\varepsilon)}(\alpha,z)|}du\le\frac{C(1+|z|^2)}{v^2}\int_{-\infty}^{\infty}|S_n^{(\varepsilon)}(\alpha,z)|^2du\le\frac{C(1+|z|^2)}{v^3}. \tag{2.44}$$

The inequalities (2.41), (2.39), and the definition of $\widehat{\delta}_n(\alpha,z)$ together imply

$$\int_{-\infty}^{\infty}|\widehat{\delta}_n(\alpha,z)|du\le\frac{C(1+|z|^2)}{v^6\varphi(\sqrt{np_n})}+\frac{C\varkappa}{v^4\varphi(\sqrt{np_n})}\int_{-\infty}^{\infty}|\widehat{\delta}_n(\alpha,z)|du. \tag{2.45}$$

If we choose v such that $\frac{C\varkappa}{v^4\varphi(\sqrt{np_n})}<\frac{1}{2}$ we obtain

$$\int_{-\infty}^{\infty}|\widehat{\delta}_n(\alpha,z)|du\le\frac{C(1+|z|^2)}{\varphi(\sqrt{np_n})v^6}. \tag{2.46}$$

□

In Section 3 we show that the measure $\nu(\cdot,z)$ has bounded support and bounded density for any z. To bound the distance between the distribution functions $F_n^{(\varepsilon)}(x,z)$ and $F(x,z)$ we may apply Corollary 3.2 from [9] (see also Lemma 6.6 in the Appendix). We take $V=1$ and $v_0=C(\varphi(\sqrt{np_n}))^{-\frac{1}{6}}$. Then Lemmas 2.2 and 2.3 together imply

$$\sup_x|F_n^{(\varepsilon)}(x,z)-F(x,z)|\le C(\varphi(\sqrt{np_n}))^{-\frac{1}{6}}. \tag{2.47}$$

□

3 Properties of the measure $\nu(\cdot, z)$

In this Section we investigate the properties of the measure $\nu(\cdot, z)$. At first note that there exists a solution $S(\alpha, z)$ of the equation

$$S(\alpha, z) = -\frac{S(\alpha, z) + \alpha}{(S(\alpha, z) + \alpha)^2 - |z|^2} \tag{3.1}$$

such that, for $v > 0$,

$$\mathrm{Im}\{S(\alpha, z)\} \geq 0$$

and $S(\alpha, z)$ is an analytic function in the upper half-plane $\alpha = u + iv$, $v > 0$. This follows from the relative compactness of the sequence of analytic functions $S_n(\alpha, z)$, $n \in \mathbb{N}$. From (2.35) it follows immediately that

$$|S(\alpha, z)| \leq 1. \tag{3.2}$$

Set $y = S(x, z) + x$ and consider the equation (2.35) on the real line

$$y = -\frac{y}{y^2 - |z|^2} + x, \tag{3.3}$$

or

$$y^3 - xy^2 + (1 - |z|^2)y + x|z|^2 = 0. \tag{3.4}$$

Set

$$x_1^2 = \frac{5 + 2|z|^2}{2} + \frac{(1 + 8|z|^2)^{\frac{3}{2}} - 1}{8|z|^2}, \quad x_2^2 = \frac{5 + 2|z|^2}{2} - \frac{(1 + 8|z|^2)^{\frac{3}{2}} + 1}{8|z|^2}. \tag{3.5}$$

It is straightforward to check that for $|z| \leq 1$ $\sqrt{3(1 - |z|^2)} \leq |x_1|$ and $x_2^2 < 0$ for $|z| < 1$ and $x_2^2 = 0$ for $|z| = 1$, and $x_2^2 > 0$ for $|z| > 1$.

Lemma 3.1. *In the case $|z| \leq 1$ equation (3.4) has one real root for $|x| \leq |x_1|$ and three real roots for $|x| > |x_1|$. In the case $|z| > 1$ equation (3.4) has one real root for $|x_2| \leq x \leq |x_1|$ and has tree real roots for $|x| \leq |x_2|$ or for $|x| \geq |x_1|$.*

Proof. Set

$$L(y) := y^3 - xy^2 + (1 - |z|^2)y + x|z|^2.$$

We consider the roots equation

$$L'(y) = 3y^2 - 2xy + (1 - |z|^2) = 0. \tag{3.6}$$

The roots of this equation are

$$y_{1,2} = \frac{x \pm \sqrt{x^2 - 3(1 - |z|^2)}}{3}.$$

This implies that, for $|z| \leq 1$ and for

$$|x| \leq \sqrt{3(1-|z|^2)},$$

the equation (3.4) has one real root. Furthermore, direct calculations show that

$$L(y_1)L(y_2) = \frac{1}{27}\left(-4|z|^2x^4 + (8|z|^4 + 20|z|^2 - 1)x^2 + 4(1-|z|^2)^3\right).$$

Solving the equation $L(y_1)L(y_2) = 0$ with respect to x, we get for $|z| \leq 1$ and $\sqrt{3(1-|z|^2)} \leq |x| \leq |x_1|$

$$L(y_1)L(y_2) \geq 0,$$

and for $|z| \leq 1$ and $|x| > \sqrt{\frac{20+8|z|^2}{8} + \frac{(1+8|z|^2)^{\frac{3}{2}}-1}{8|z|^2}}$

$$L(y_1)L(y_2) < 0,$$

These relations imply that for $|z| \leq 1$ the function $L(y)$ has three real roots for $|x| \geq |x_1|$ and one real root for $|x| < |x_1|$.

Consider the case $|z| > 1$ now. In this case $y_{1,2}$ are real for all x and $x_2^2 > 0$. Note that

$$L(y_1)L(y_2) \leq 0$$

for $|x| \leq |x_2|$ and for $|x| \geq |x_1|$ and

$$L(y_1)L(y_2) > 0$$

for $|x_2| < x < |x_1|$. These implies that for $|z| > 1$ and for $|x_2| < x < |x_1|$ the function $L(y)$ has one real root and for $|x| \leq |x_2|$ or for $|x| \geq |x_1|$ the function $L(y)$ has three real roots. The Lemma is proved. □

Remark 3.1. From Lemma 3.1 it follows that the measure $\nu(x,z)$ has a density $p(x,z)$ and

- $p(x,z) \leq 1$, for all x and z;
- for $|z| \leq 1$, if $|x| \geq x_1$ then $p(x,z) = 0$;
- for $|z| \geq 1$, if $|x| \geq x_1$ or $|x| \leq x_2$ then $p(x,z) = 0$;
- $p(x,z) > 0$ otherwise.

The next lemma is an analogue of Lemma 4.4 in Bai [1].

Lemma 3.2. *The following equality*

$$\frac{\partial}{\partial s}\left(\int_0^\infty \log x \nu(dx,z)\right) = \frac{1}{2}\Re\{g(x,z)\} \tag{3.7}$$

holds.

Proof. Following Bai [1] Lemma 4.4, we consider

$$I(C) := \int_0^C \frac{\partial y(x)}{\partial s} dx. \tag{3.8}$$

We have

$$y^3 + 2xy^2 + x^2 y - |z|^2 y + y + x = 0. \tag{3.9}$$

Taking the derivatives with respect to x and s correspondingly, we get

$$\frac{\partial y}{\partial x}\left(3y^2 + 4xy + (1 - |z|^2 + x^2)\right) = -1 - 2y(x+y) \tag{3.10}$$

and

$$\frac{\partial y}{\partial s}\left(3y^2 + 4xy + (1 - |z|^2 + x^2)\right) = 2sy. \tag{3.11}$$

These equalities together imply

$$\frac{\partial y}{\partial s} = -\frac{2sy}{1 + 2y(x+y)} \frac{\partial y}{\partial x}. \tag{3.12}$$

From equation (3.9) it follows that

$$1 + 2y(y+x) = \pm\sqrt{1 + 4|z|^2 y^2}. \tag{3.13}$$

Using the results of Remark 3.1, it is straightforward to check that for $|z| \le 1$

$$1 + 2y(y+x) = \sqrt{1 + 4|z|^2 y^2} \tag{3.14}$$

and for $|z| > 1$ there exists a number x_0 such that $\sqrt{1 + 4|z|^2 y^2} = 0$. Furthermore, we have for $-x_0 \le x \le 0$

$$1 + 2y(y+x) = \sqrt{1 + 4|z|^2 y^2} \tag{3.15}$$

and for $x < -x_0$ we obtain

$$1 + 2y(y+x) = -\sqrt{1 + 4|z|^2 y^2}. \tag{3.16}$$

Using these equalities, we get

$$\int_{-C}^0 \frac{\partial y}{\partial s} dx = -\int_{-C}^0 \frac{2sy}{1 + 2y(x+y)} \frac{\partial y}{\partial x} dx. \tag{3.17}$$

For $|z| \le 1$, we have

$$\int_{-C}^0 \frac{\partial y}{\partial s} dx = -\int_{-C}^0 \frac{2sy}{\sqrt{1 + 4|z|^2 y^2}} \frac{\partial y}{\partial x} dx = \frac{s}{4|z|^2}\left(\sqrt{1 + 4|z|^2 y^2(-C)} + \sqrt{1 + 4|z|^2(|z|^2 - 1)}\right). \tag{3.18}$$

In the limit $C \to \infty$, we get, for $|z| \le 1$,

$$\int_{-\infty}^{0} \frac{\partial y}{\partial s} dx = \frac{s}{2}. \tag{3.19}$$

For $|z| > 1$, we have

$$\int_{-\infty}^{0} \frac{\partial y}{\partial s} dx = \int_{-x_0}^{0} \frac{2sy}{\sqrt{1+4|z|^2y^2}} \frac{\partial y}{\partial x} dx - \int_{-\infty}^{-x_0} \frac{2sy}{\sqrt{1+4|z|^2y^2}} \frac{\partial y}{\partial x} dx = \frac{s}{2|z|^2}. \tag{3.20}$$

Similar to Bai [1] (equality (4.39)) we have

$$\begin{aligned}\int_{-C}^{0} y(x)dx = \int_{-C}^{0} y(x)dx &= \int_{0}^{C}\int_{0}^{\infty} \frac{1}{u+x}\nu(du,z)dx \\ &= \ln C + \int_{0}^{\infty} \left[\ln(u+C) - \ln u\right]\nu(du,z) \\ &= \ln C + \int_{0}^{\infty} \ln(1+\frac{u}{C})\nu(du,z) - \int_{0}^{\infty} \ln u\nu(du,z). \end{aligned} \tag{3.21}$$

After differentiation we get

$$\frac{\partial}{\partial s}\int_{0}^{\infty} \ln u\nu(du,z) = \frac{\partial}{\partial s}\int_{0}^{\infty} \ln(1+\frac{u}{C})\nu(du,z) - \int_{-C}^{0} \frac{\partial}{\partial s} y(x)dx. \tag{3.22}$$

Relations (3.19)–(3.22) together imply the result. □

4 The smallest singular value

Let $\mathbf{X}^{(\varepsilon)} = \frac{1}{\sqrt{np_n}}(\varepsilon_{jk}X_{jk})_{j,k=1}^{n}$ be an $n \times n$ matrix with independent entries $\varepsilon_{jk}X_{jk}$, $j,k = 1,\ldots,n$. Assume that $\mathbf{E}\,X_{jk} = 0$ and $\mathbf{E}\,X_{jk}^2 = 1$ and ε_{jk} denote Bernoulli random variables with $p_n = \Pr\{\varepsilon_{jk} = 1\}$, $j,k = 1,\ldots,n$. Denote by $s_1^{(\varepsilon)}(z) \ge \ldots \ge s_n^{(\varepsilon)}(z)$ the singular values of the matrix $\mathbf{X}^{(\varepsilon)}(z) := \mathbf{X}^{(\varepsilon)} - z\mathbf{I}$. In this Section we prove a bound for the minimal singular value of the matrices $\mathbf{X}^{(\varepsilon)}(z)$. We prove the following result.

Theorem 4.1. *Let X_{jk} be independent random complex variables with $\mathbf{E}\,X_{jk} = 0$ and $\mathbf{E}\,|X_{jk}|^2 = 1$, which are uniformly integrable , i.e.*

$$\max_{j,k} \mathbf{E}\,|X_{jk}|^2 I_{\{|X_{jk}|>M\}} \to 0 \quad as \quad M \to 0. \tag{4.1}$$

Let ε_{jk}, $j,k = 1,\ldots,n$ be independent Bernoulli random variables with $p_n := \Pr\{\varepsilon_{jk} = 1\}$. Assume that ε_{jk} are independent from X_{jk} in aggregate. Let $p_n^{-1} = \mathcal{O}(n^{1-\theta})$ for some $0 < \theta \le 1$. Let $K \ge 1$. Then there exist constants $c, C, B > 0$ depending on θ and K such that for any $z \in \mathbb{C}$ and positive ε we have

$$\Pr\{s_n^{(\varepsilon)}(z) \le \varepsilon/n^B;\; s_1^{(\varepsilon)}(z) \le Kn\sqrt{p_n}\} \le \exp\{-c\,p_n\,n\} + \frac{C\sqrt{\ln n}}{\sqrt{np_n}}. \tag{4.2}$$

Remark 4.2. Let X_{jk} be i.i.d. random variables with $\mathbf{E}\,X_{jk}=0$ and $\mathbf{E}\,|X_{jk}|^2=1$. Then the condition (4.1) holds.

Remark 4.3. Consider the event A that there exists at least one row with zero entries only. Its probability is given by

$$\Pr\{A\} \ge 1-(1-(1-p_n)^n)^n. \tag{4.3}$$

Simple calculations show that if $np_n \le \ln n$ for all $n \ge 1$, then

$$\Pr\{A\} \ge \delta > 0. \tag{4.4}$$

Hence in the case $np_n \le \ln n$ and $np_n \to \infty$ we have no invertibility with positive probability

Remark 4.4. The proof of Theorem 4.1 uses ideas of Rudelson and Vershynin [21], to classify with high probability vectors $\mathbf{x}$ in the $(n-1)$-dimensional unit sphere $\mathcal{S}^{n-1}$ such that $||\mathbf{X}^{(\varepsilon)}(z)\mathbf{x}||_2$ is extremely small into two classes called compressible and incompressible vectors.

We develop our approach for shifted sparse and normalized matrices $\mathbf{X}^{(\varepsilon)}(z)$. The generalization to the case of complex sparse and shifted matrices $\mathbf{X}^{(\varepsilon)}(z)$ is straightforward. For details see for example the paper of Götze and Tikhomirov [10] and proof of the Lemma 4.1 below.

Remark 4.5. We can relax the condition $p_n^{-1}=\mathcal{O}(n^{1-\theta})$ to $p_n^{-1}=o(n/\ln^2 n)$. The quantity B in Theorem 4.1 should be of order $\ln n$ in this case. See Remark 4.10 for details.

Lemma 4.1. *Let $\mathbf{x}=(x_1,\dots,x_n)\in\mathcal{S}^{n-1}$ be a fixed unit vector and $\mathbf{X}^{(\varepsilon)}(z)$ be a matrix as in Theorem 4.1. Then there exist some positive absolute constants γ_0 and c_0 such that for any $0<\tau\le\gamma_0$*

$$\Pr\{\|\mathbf{X}^{(\varepsilon)}(z)\mathbf{x}\|_2 \le \tau\} \le \exp\{-c_0 np_n\} \vee \exp\{-c_0 n\}, \tag{4.5}$$

where $x \vee y$ denotes the larger of x and y

Proof of Lemma 4.1. Recall that $\mathbf{E}\,X_{ij}=0$ and $\mathbf{E}\,|X_{ij}|^2=0$. Assume first that X_{ij} are real independent r.v. with mean zero, and variance at least 1. Let $X_{ij}^{(\varepsilon)}=X_{ij}\,\varepsilon_{ij}$ with independent Bernoulli variables which are independent of X_{ij} in aggregate and let $z=0$. Assume also that $\mathbf{x}$ is a real vector. Then

$$||\mathbf{X}^{(\varepsilon)}\mathbf{x}||_2^2 = \frac{1}{np_n}\sum_{j=1}^n \left|\sum_{k=1}^n x_k X_{jk}\varepsilon_{jk}\right|^2 =: \frac{1}{np_n}\sum_{k=1}^n \zeta_j^2. \tag{4.6}$$

By Chebyshev's inequality we have

$$\Pr\{\sum_{j=1}^n \zeta_j^2 < \tau^2 np_n\} = \Pr\{\frac{\tau^2 np_n}{2} - \frac{1}{2}\sum_{j=1}^n \zeta_j^2 > 0\} \le \exp\{np_n\tau^2t^2/2\}\prod_{j=1}^n \mathbf{E}\,\exp\{-t^2\zeta_j^2/2\}. \tag{4.7}$$

Using $e^{-t^2/2} = \mathbf{E}\exp\{it\xi\}$ where ξ is a standard Gaussian random variable, we obtain

$$\Pr\{\sum_{j=1}^n \zeta_j^2 < \tau^2 np_n\} \le \exp\{np_n\tau^2t^2/2\}\prod_{j=1}^n \mathbf{E}_{\xi_j}\prod_{k=1}^n \mathbf{E}_{\varepsilon_{jk}X_{jk}}\exp\{it\xi_jx_k\varepsilon_{jk}X_{jk}\}, \tag{4.8}$$

where ξ_j, $j = 1, \dots, n$ denote i.i.d. standard Gaussian r.v.'s and $\mathbf{E}_Z$ denotes expectation with respect to Z conditional on all other r. v.'s. For every $\alpha, x \in [0,1]$ and $\rho \in (0,1)$ the following inequality holds

$$\alpha x + 1 - \alpha \le x^\beta \vee \left(\frac{\rho}{\alpha}\right)^{\frac{\beta}{1-\beta}}, \tag{4.9}$$

(see [4], inequality (3.7)). Take $\alpha = \Pr\{|\xi_j| \le C_1\}$ for some absolute positive constant C_1 which will be chosen later. Then it follows from (4.8) that

$$\Pr\{\sum_{j=1}^n \zeta_j^2 < \tau^2 np_n\} \le \exp\{np_n\tau^2t^2/2\}$$
$$\times\prod_{j=1}^n\left(\alpha\left|\mathbf{E}_{\xi_j}\left(\prod_{k=1}^n \mathbf{E}_{\varepsilon_{jk}X_{jk}}\exp\{it\xi_jx_k\varepsilon_{jk}X_{jk}\}\Big||\xi_j| \le C_1\right)\right| + 1 - \alpha\right). \tag{4.10}$$

Furthermore, we note that

$$|\mathbf{E}_{\varepsilon_{jk}X_{jk}}\exp\{it\xi_jx_k\varepsilon_{jk}X_{jk}\}| \le \exp\{\frac{1}{2}(|\mathbf{E}_{\varepsilon_{jk}X_{jk}}\exp\{it\xi_jx_k\varepsilon_{jk}X_{jk}\}|^2 - 1)\}$$
$$\le \exp\{-p_n\left((1-p_n)(1-\operatorname{Re}f_{jk}(tx_k\xi_j)) + \frac{p_n}{2}(1-|f_{jk}(tx_k\xi_j)|^2)\right)\}, \tag{4.11}$$

where $f_{jk}(u) = \mathbf{E}\exp\{iuX_{jk}\}$. Assuming (4.1), choose a constant $M > 0$ such that

$$\sup_{jk}\mathbf{E}|X_{jk}|^2 I_{\{|X_{jk}|>M\}} \le 1/2. \tag{4.12}$$

Since $1 - \cos x \ge 11/24x^2$ for $|x| \le 1$, conditioning on the event $|\xi_j| \le C_1$, we get for $0 < t \le 1/(MC_1)$

$$1 - \operatorname{Re}f_{jk}(tx_k\xi_j) = \mathbf{E}_{X_{jk}}(1 - \cos(tx_kX_{jk}\xi_j)) \ge \frac{11}{24}t^2x_k^2\xi_j^2\mathbf{E}|X_{jk}|^2I_{\{|X_{jk}|\le M\}}, \tag{4.13}$$

and similarly

$$1 - |f_{jk}(tx_k\xi_j)|^2 = \mathbf{E}_{X_{jk}}(1 - \cos(tx_k\widetilde{X}_{jk}\xi_j)) \ge \frac{11}{24}t^2x_k^2\xi_j^2\mathbf{E}|\widetilde{X}_{jk}|^2I_{\{|X_{jk}|\le M\}} \tag{4.14}$$

It follows from (4.11) for $0 < t < 1/(MC_1)$ and for some constant $c > 0$

$$|\mathbf{E}_{\varepsilon_{jk}X_{jk}}\exp\{it\xi_jx_k\varepsilon_{jk}X_{jk}\}| \le \exp\{-cp_nt^2x_k^2\xi_j^2\}. \tag{4.15}$$

This implies that conditionally on $|\xi_j| \le C_1$ and for $0 < t \le 1/(MC_1)$

$$|\prod_{k=1}^{n} \mathbf{E}_{\varepsilon_{jk}X_{jk}} \exp\{it\xi_j x_k \varepsilon_{jk} X_{jk}\}| \le \exp\{-cp_n t^2 \xi_j^2\}. \tag{4.16}$$

Let $\Phi_0(x) := 2\Phi(x) - 1$, $x > 0$ where $\Phi(x)$ denotes the standard Gaussian distribution function. It is straightforward to show that

$$\begin{aligned} &\mathbf{E}_{\xi_j} \left(\exp\{-cp_n t^2 \xi_j^2\} \,\Big|\, |\xi_j| \le C_1\right) \\ &\qquad = \frac{1}{\sqrt{1+2ct^2 p_n}} \frac{\Phi_0\left(C_1\sqrt{1+2t^2 c p_n}\right)}{\Phi_0(C_1)}. \end{aligned} \tag{4.17}$$

Applying Taylor's formula, we obtain

$$\begin{aligned} \frac{\Phi_0\left(C_1\sqrt{1+2ct^2 p_n}\right)}{\Phi_0(C_1)} &= 1 + \left(\sqrt{1+2t^2 c p_n} - 1\right) \\ &\quad \times \frac{\Phi_0'\left(C_1(1+\sqrt{1+2ct^2 p_n}\right)}{\Phi_0(C_1)}. \end{aligned} \tag{4.18}$$

Using that for $0 < y < 8$ we have $y/4 \ge \sqrt{1+y} - 1 \le y/2$ and $\Phi_0'\left(C_1(1+\sqrt{1+2t^2 p_n c}\right) \le \Phi_0'(C_1)$, we get

$$\frac{\Phi_0\left(C_1\sqrt{1+2ct^2 p_n}\right)}{\Phi_0(C_1)} \le 1 + ct^2 p_n \frac{\Phi_0'(C_1)}{\Phi_0(C_1)}. \tag{4.19}$$

We may choose C_1 large enough such that following inequalities hold

$$\begin{aligned} \mathbf{E}_{\xi_j} \left(\exp\{-cp_n t^2 \xi_j^2\} \Big| |\xi_j| \le C_1\right) &\le \frac{1+ct^2 p_n/8}{1+ct^2 p_n/4} \\ &\le \exp\{-ct^2 p_n/24\} \end{aligned} \tag{4.20}$$

for all $|t| \le 1/(MC_1) < 8$. Inequalities (4.8), (4.9), (4.11), (4.20) together imply that for any $\beta \in (0,1)$

$$\Pr\Big\{\sum_{j=1}^{n} \zeta_j^2 < \tau^2 n p_n\Big\} \le \exp\{n p_n \tau^2 t^2/2\}\Big(\exp\{-c\beta n t^2 p_n/24\} + \Big(\frac{\beta}{\alpha}\Big)^{\frac{n\beta}{1-\beta}}\Big). \tag{4.21}$$

Without loss of generality we may take C_1 sufficiently large, such that $\alpha \ge 4/5$ and choose $\beta = 2/5$. Then we obtain

$$\Pr\{\sum_{j=1}^{n} \zeta_j^2 < \tau^2 n p_n\} \le \exp\{n p_n \tau^2 t^2/2\}\Big(\exp\{-ct^2 n p_n/60\} + \Big(\frac{1}{2}\Big)^{\frac{2n}{3}}\Big). \tag{4.22}$$

For $\tau < \frac{\sqrt{c}}{\sqrt{60}}$ we conclude from here that for $|t| \le 1/(MC_1)$

$$\Pr\{\sum_{j=1}^{n} \zeta_j^2 < \tau^2 n p_n\} \le \exp\{-ct^2 n p_n/120\}. \tag{4.23}$$

Inequality (4.23) implies that inequality (4.5) holds with some positive constant $c_0 > 0$. This concludes the proof in the real case.

Consider now the general case. Let $X_{jk} = \xi_{jk} + i\eta_{jk}$ with $i = \sqrt{-1}$ with $\mathbf{E}\,|X_{jk}|^2 = 1$ and $x_k = u_k + iv_k$ and $z = u + iv$. In this notation we have

$$\begin{aligned}
&\Pr\{\|(\mathbf{X}^{(\varepsilon)} - z\mathbf{I})\mathbf{x}\|_2 \le \tau\} \\
&\le \exp\{\tau^2 n p_n t^2/2\} \min\left\{\mathbf{E}\exp\left\{-t^2\sum_{j=1}^{n}\left|\sum_{k=1}^{n}(\xi_{jk}u_k - \eta_{jk}v_k)\varepsilon_{jk} - \sqrt{np_n}(uu_j - vv_j)\right|^2/2\right\},\right. \\
&\qquad \left.\mathbf{E}\exp\left\{-t^2\sum_{j=1}^{n}\left|\sum_{k=1}^{n}(\xi_{jk}v_k + \eta_{jk}u_k)\varepsilon_{jk} - \sqrt{np_n}(vu_j + uv_j)\right|^2/2\right\}\right\}.
\end{aligned} \tag{4.24}$$

Note that for $\mathbf{x} = (x_1, \ldots, x_n) \in S^{(n-1)}$ (the unit sphere in $\mathbb{C}^n$) and for any set $A \subset \{1, \ldots, n\}$

$$\max\{\sum_{k\in A}|x_k|^2, \sum_{k\in A^c}|x_k|^2\} \ge 1/2. \tag{4.25}$$

For any $j = 1, \ldots, n$ we introduce the set A_j as follows

$$A_j := \{k \in \{1, \ldots, n\} :\ \mathbf{E}\,|\xi_{jk}u_k - \eta_{jk}v_k|^2 \ge |x_k|^2/2\}. \tag{4.26}$$

It is straightforward to check that for any $k \notin A_j$

$$\mathbf{E}\,|\eta_{jk}u_k + \xi_{jk}v_k|^2 \ge |x_k|^2/2. \tag{4.27}$$

According to inequality (4.25), for any $j = 1, \ldots, n$, there exist a set B_j such that

$$\sum_{k\in B_j}|x_k|^2 \ge 1/2 \tag{4.28}$$

and for any $k \in B_j$

$$\mathbf{E}\,|\xi_{jk}u_k - \eta_{jk}v_k|^2 \ge |x_k|^2/2, \tag{4.29}$$

or

$$\mathbf{E}\,|\eta_{jk}u_k + \xi_{jk}v_k|^2 \ge |x_k|^2/2. \tag{4.30}$$

Introduce the following random variables for any $j, k = 1, \ldots, n$

$$\widetilde{\zeta}_{jk} := \xi_{jk}u_k - \eta_{jk}v_k, \tag{4.31}$$

and

$$\widehat{\zeta}_{jk} := \eta_{jk}u_k + \xi_{jk}v_k. \tag{4.32}$$

The inequalities (4.29) and (4.30) together imply that one of the following two inequalities

$$\operatorname{card}\Big\{j: \text{ for any } \quad k \in B_j \quad \mathbf{E}\,|\widehat{\zeta}_{jk}|^2 \geq |x_k|^2/2\Big\} \geq n/2 \tag{4.33}$$

or

$$\operatorname{card}\Big\{j: \text{ for any } \quad k \in B_j \quad \mathbf{E}\,|\widetilde{\zeta}_{jk}|^2 \geq |x_k|^2/2\Big\} \geq n/2 \tag{4.34}$$

holds. If (4.33) holds we shall bound the first term on the right hand side of (4.24). In the other case we shall bound the second term. In what follows we may repeat the arguments leading to inequalities (4.10)–(4.16). Thus the Lemma is proved. □

For any $q_n \in (0,1)$ and $K > 0$ to be chosen later we define $K_n := Kn\sqrt{p_n}$, $\widehat{q}_n := q_n/\big(\ln(2/p_n)\ln K_n\big)$ and $\widehat{p}_n := p_n/\big(\ln(2/p_n)\ln K_n\big)$. Without loss of generality we shall assume that

$$\ln K_n/|\ln\gamma_0| \geq 1 \quad \text{and} \quad \ln K_n > 1. \tag{4.35}$$

Proposition 4.6. *Assume there exist an absolute constant $c > 0$ and values $\gamma_n, q_n \in (0,1)$ such that for any $\mathbf{x} \in \mathcal{C} \subset \mathcal{S}^{(n-1)}$*

$$\Pr\{\|\mathbf{X}^{(\varepsilon)}(z)\mathbf{x}\|_2 \leq \gamma_n \quad \textit{and} \quad \|\mathbf{X}^{(\varepsilon)}(z)\| \leq K_n\} \leq \exp\{-c\,nq_n\} \tag{4.36}$$

holds. Then there exists a constant $\delta_0 > 0$ depending on K and c only such that, for $k < \delta_0\, n\,\widehat{q}_n$,

$$\Pr\{\inf_{\mathbf{x}\in\mathcal{S}^{k-1}\cap\mathcal{C}} \|\mathbf{X}^{(\varepsilon)}(z)\mathbf{x}\|_2 \leq \gamma_n/2 \textit{ and } \|\mathbf{X}^{(\varepsilon)}(z)\| \leq K_n\} \leq \exp\{-c\,nq_n/8\}.$$

Proof. Let $\eta > 0$ to be chosen later. There exists an η–net $\mathcal{N}$ in $\mathcal{S}^{k-1}\cap\mathcal{C}$ of cardinality $|\mathcal{N}| \leq (\frac{3}{\eta})^{2k}$ (see e.g. Lemma 3.4 in [20]). By condition (4.36), we have for $\tau \leq \gamma_n$

$$\Pr\{\text{ there exists } \mathbf{x}\in\mathcal{N}: \|\mathbf{X}^{(\varepsilon)}(z)\mathbf{x}\|_2 < \tau \text{ and } \|\mathbf{X}^{(\varepsilon)}(z)\| \leq K_n\} \leq (\frac{3}{\eta})^{2k}\exp\{-c\,n\,q_n\}. \tag{4.37}$$

Let V be the event that $\|\mathbf{X}^{(\varepsilon)}(z)\| \leq K_n$ and $\|\mathbf{X}^{(\varepsilon)}(z)\mathbf{y}\|_2 \leq \frac{1}{2}\tau$ for some point $\mathbf{y} \in \mathcal{S}^{(k-1)}\cap\mathcal{C}$. Assume that V occurs and choose a point $\mathbf{x}\in\mathcal{N}$ such that $\|\mathbf{y}-\mathbf{x}\|_2 \leq \eta$. Then

$$\|\mathbf{X}^{(\varepsilon)}(z)\mathbf{x}\|_2 \leq \|\mathbf{X}^{(\varepsilon)}(z)\mathbf{y}\|_2 + \|\mathbf{X}^{(\varepsilon)}(z)\|\|\mathbf{x}-\mathbf{y}\|_2 \leq \frac{1}{2}\tau + K_n\eta = \tau \tag{4.38}$$

if we set $\eta = \tau/(2K_n)$. Hence,

$$\Pr(V) \leq \Big(\Big(\frac{3}{\eta}\Big)^{2\delta_0/(\ln K_n \ln(2/p_n))}\exp\{-\frac{c_0}{4}\}\Big)^{nq_n}. \tag{4.39}$$

Note that under assumption (4.35) we have

$$\frac{2\ln(3/\eta)}{\ln 2\ln K_n} \leq 10. \tag{4.40}$$

Choosing $\delta_0 = \frac{c}{80}$ and $\tau = \gamma_n$, we conclude the proof. □

Following Rudelson and Vershynin [21], we shall partition the unit sphere $\mathcal{S}^{(n-1)}$ into the two sets of socalled compressible and incompressible vectors and we will show the invertibility of $\mathbf{X}$ on each set separately.

Definition 4.7. Let $\delta, \rho \in (0,1)$. A vector $\mathbf{x} \in \mathbb{R}^n$ is called *Sparse* if $|\mathrm{supp}(\mathbf{x})| \le \delta$n. A vector $\mathbf{x} \in \mathcal{S}^{(n-1)}$ is called *compressible* if $\mathbf{x}$ is within Euclidean distance ρ from the set of all sparse vectors. A vector $\mathbf{x} \in \mathcal{S}^{(n-1)}$ is called *incompressible* if it is not compressible.

The sets of sparse, compressible and incompressible vectors depending on δ and ρ will be denoted by

$$\mathit{Sparse}(\delta), \qquad \mathit{Comp}(\delta, \rho), \qquad \mathit{Incomp}(\delta, \rho), \tag{4.41}$$

respectively.

Lemma 4.2. *Let $\mathbf{X}^{(\varepsilon)}(z)$ be a random matrix as in Theorem 1.2, and let $K_n = Kn\sqrt{p_n}$ with a constant $K \ge 1$. Assume there exist an absolute constant $c > 0$ and values $\gamma_n, q_n \in (0,1)$ such that for any $\mathbf{x} \in \mathcal{C} \subset \mathcal{S}^{(n-1)}$*

$$\Pr\{\|\mathbf{X}^{(\varepsilon)}(z)\mathbf{x}\|_2 \le \gamma_n \quad \textit{and} \quad \|\mathbf{X}^{(\varepsilon)}(z)\| \le K_n\} \le \exp\{-c\, nq_n\} \tag{4.42}$$

holds. Then there exist δ_1, c_1 that depend on K and c only , such that

$$\Pr\left\{\inf_{\mathbf{x} \in \mathit{Comp}(\delta_1 \widehat{q}_n,\, \rho_n) \cap \mathcal{C}} \|\mathbf{X}^{(\varepsilon)}(z)\mathbf{x}\|_2 \le \gamma_n \textit{ and } \|\mathbf{X}^{(\varepsilon)}(z)\| \le K_n\right\} \le \exp\{-c_1\, nq_n\}, \tag{4.43}$$

where $\rho_n := \gamma_n/(4K_n)$.

Proof. At first we estimate the invertibility for sparse vectors. Let $k = [\delta_1 n \widehat{q}_n]$ with some positive constant δ_1 which will be chosen later. According to Proposition 4.6 for any $\delta_1 \le \delta_0$ and for any $\tau \le \gamma_n/2$, we have the following inequality

$$\begin{aligned}
&\Pr\left\{\inf_{\mathbf{x} \in \mathit{Sparse}(\delta_1 \widehat{p}_n) \cap \mathcal{C}} \|\mathbf{X}^{(\varepsilon)}(z)\mathbf{x}\|_2 \le \tau \text{ and } \|\mathbf{X}^{(\varepsilon)}(z)\| \le K_n\right\} \\
&\quad = \Pr\left\{\text{ there exist } \sigma, |\sigma| = k : \inf_{\mathbf{x} \in \mathbb{R}^\sigma \cap \mathcal{C}, \|\mathbf{x}\|_2 = 1} \|\mathbf{X}^{(\varepsilon)}(z)\mathbf{x}\|_2 \le \tau \text{ and } \|\mathbf{X}^{(\varepsilon)}(z)\| \le K_n\right\} \\
&\qquad \le \binom{n}{k} \exp\{-c_0 n q_n/8\}.
\end{aligned}$$

Using Stirling's formula, we get for some absolute positive constant C

$$\frac{1}{n} \ln \binom{n}{k} \le -C\delta_1 \widehat{q}_n \ln(\delta \widehat{q}_n). \tag{4.44}$$

We may choose δ_1 small enough that

$$\frac{1}{n} \ln \binom{n}{k} \le c_0 q_n/16. \tag{4.45}$$

Thus we get

$$\Pr\left\{\inf_{\mathbf{x}\in Sparse(\delta_1\widehat{p}_n)\cap\mathcal{C}}\|\mathbf{X}^{(\varepsilon)}(z)\mathbf{x}\|_2\le\tau \text{ and } \|\mathbf{X}^{(\varepsilon)}(z)\|\le K_n\right\}\le\exp\{-c_1nq_n\}. \tag{4.46}$$

Choose $\rho:=\gamma:=\gamma_n/4$. Let V be the event that $\|\mathbf{X}^{(\varepsilon)}(z)\|\le K_n$ and $\|\mathbf{X}^{(\varepsilon)}(z)\mathbf{y}\|_2\le\gamma_1$ for some point $\mathbf{y}\in Comp(\delta_1\widehat{p}_n,\rho K_n^{-1})$. Assume that V occurs and choose a point $\mathbf{x}\in Sparse(\delta_1\widehat{p}_n)$ such that $\|\mathbf{y}-\mathbf{x}\|_2\le\rho K_n^{-1}$. Then

$$\|\mathbf{X}^{(\varepsilon)}(z)\mathbf{x}\|_2\le\|\mathbf{X}^{(\varepsilon)}(z)\mathbf{y}\|_2+\|\mathbf{X}^{(\varepsilon)}(z)\|\|\mathbf{x}-\mathbf{y}\|_2\le\gamma_1+\rho=\gamma_n/2. \tag{4.47}$$

Hence,

$$\Pr(V)\le\exp\{-\frac{c_0}{8}nq_n\}. \tag{4.48}$$

Thus the Lemma is proved. □

Lemma 4.3. *Let $\delta,\rho\in(0,1)$. Let $\mathbf{x}\in Incomp(\delta,\rho)$. Then there exists a set $\sigma(\mathbf{x})\subset\{1,\dots,n\}$ of cardinality $|\sigma(\mathbf{x})|\ge\frac{1}{2}n\delta$ such that*

$$\sum_{k\in\sigma(\mathbf{x})}|x_k|^2\ge\frac{1}{2}\rho^2 \tag{4.49}$$

and

$$\frac{\rho}{\sqrt{2n}}\le|x_k|\le\frac{1}{\sqrt{n\delta/2}},\quad \textit{for any}\quad k\in\sigma(\mathbf{x}) \tag{4.50}$$

which we shall call "spread set of x" henceforth.

Proof. See proof in [21], p. 16, proof of Lemma 3.4. For the readers convenience we repeat this proof here. Consider the subsets of $\{1,\dots,n\}$ defined by

$$\sigma_1(\mathbf{x}):=\{k:\ |x_k|\le\frac{1}{\sqrt{\delta n/2}}\},\quad \sigma_2(\mathbf{x})=\{k:\ |x_k|\ge\frac{\rho}{\sqrt{2n}}\}, \tag{4.51}$$

and put $\sigma(\mathbf{x})=\sigma_1(\mathbf{x})\cap\sigma_2(\mathbf{x})$. Denote by $P_{\sigma(\mathbf{x})}$ the orthogonal projection onto $\mathbb{R}^{\sigma(\mathbf{x})}$ in $\mathbb{R}^n$. By Chebyshev's inequality $|\sigma_1(\mathbf{x})^c|\le\delta n/2$. Then $\mathbf{y}:=P_{\sigma_1(\mathbf{x})^c}\mathbf{x}\in Sparse(\delta)$, so the incompressibility of $\mathbf{x}$ implies that $\|P_{\sigma_1(\mathbf{x})}\mathbf{x}\|_2=\|\mathbf{x}-\mathbf{y}\|_2>\rho$. By the definition of $\sigma_2(\mathbf{x})$, we have $\|P_{\sigma_2(\mathbf{x})^c}\mathbf{x}\|^2\le n\frac{\rho^2}{2n}=\rho^2/2$. Hence

$$\|P_{\sigma(\mathbf{x})}\mathbf{x}\|_2^2\ge\|P_{\sigma_1(\mathbf{x})}\mathbf{x}\|_2^2-\|P_{\sigma_2(\mathbf{x})}\mathbf{x}\|_2^2\ge\rho^2/2. \tag{4.52}$$

Thus the Lemma is proved. □

Remark 4.8. If $\mathbf{x}\in Incomp(\delta\widehat{p}_n,\rho)$ then there exists a set $\sigma(\mathbf{x})$ with cardinality $|\sigma(\mathbf{x})|\ge\frac{1}{2}n\delta\widehat{p}_n$ such that

$$\frac{\rho}{\sqrt{2n}}\le|x_k|\le\frac{1}{\sqrt{n\delta\widehat{p}_n/2}} \tag{4.53}$$

and

$$\|P_{\sigma(\mathbf{x})}\mathbf{x}\|_2^2\ge\frac{1}{2}\rho^2. \tag{4.54}$$

Let $Q(\eta) = \sup_{jk}\sup_{u\in\mathbb{C}} \Pr\{|X_{jk} - u| \le \eta\}$. Introduce the maximal concentration function of the weighed sums of the rows of the matrix $(X_{jk})_{j,k=1}^{n}$,

$$p_{\mathbf{x}}(\eta) = \max_{j\in\{1,\dots,n\}} \sup_{u\in\mathbb{C}} \Pr\{|\sum_{k=1}^{n} X_{jk}\varepsilon_{jk}x_k - u| \le \eta\}. \tag{4.55}$$

We shall now bound this concentration function and prove a tensorization lemma for incompressible vectors.

Lemma 4.4. *Let δ_n and ρ_n be some functions of n such that $\rho_n, \delta_n \in (0,1)$. Let η_0 and r_0 as in Lemma 6.7. Let $\mathbf{x} \in Incomp(\delta_n, \rho_n)$. Then there exists positive constants r_1 and r_2 depending on r_0 such that for any $0 < \eta \le \eta_0$ we have*

$$p_{\mathbf{x}}(\eta\rho_n/\sqrt{2n}) \le 1 - r_2\delta_n\, np_n \tag{4.56}$$

for $n\delta_n p_n \le 1/3$ and

$$p_{\mathbf{x}}(\eta\rho_n/\sqrt{2n}) \le 1 - r_1 < 1 \tag{4.57}$$

for $n\delta_n p_n > 1/3$.

Proof. Put $m = n\delta_n$. We have

$$\sup_u \Pr\{|\sum_{k=1}^{m} X_{jk}\varepsilon_{jk}x_k - u| \le \eta\rho_n/\sqrt{2n}\} \le \Pr\{\sum_{k=1}^{m}\varepsilon_{jk} = 0\} + \Pr\{|\sum_{k=1}^{m} X_{jk}\varepsilon_{jk}x_k - u| \le \eta\rho_n/\sqrt{2n};\ \sum_{k=1}^{m}\varepsilon_{jk} \ge 1\}. \tag{4.58}$$

Introduce $\sigma(\mathbf{x}) := \{k \in \{1,\dots,n\} :\ \rho_n/\sqrt{2n} \le |x_k| \le 1/\sqrt{m/2}\}$. Since $\mathbf{x} \in Incomp(\delta_n, \rho_n)$ the cardinality of $\sigma(\mathbf{x})$ is at least $m/2$. Using that the concentration function of sum of independent random variables is less then concentration function of its summands, we obtain

$$\sup_u \Pr\{|\sum_{k=1}^{m} X_{jk}\varepsilon_{jk}x_k - u| \le \eta\rho_n/\sqrt{2n}\} \le (1-p_n)^m + Q(\eta)(1-(1-p_n)^m). \tag{4.59}$$

According to Lemma 6.7 in the Appendix for any $\eta \le \eta_0$, we have $Q(\eta) \le r_0 < 1$. Assume that $mp_n \ge 1/3$. Then we have

$$\sup_u \Pr\{|\sum_{k=1}^{m} X_{jk}\varepsilon_{jk}x_k - u| \le \eta\rho_n/\sqrt{2n}\} \le r_0 + (1-r_0)e^{-mp_n} \le 1 - (1-e^{-1/3})(1-r_0) =: 1 - r_1 < 1. \tag{4.60}$$

If $mp_n \le 1/3$ then $(1-p_n)^m \le 1 - mp_n/3$ and

$$\sup_u \Pr\{|\sum_{k=1}^m X_{jk}\varepsilon_{jk}x_k - u| \le \eta\rho_n/\sqrt{2n}\} \le 1-(1-r_0)mp_n/3 =: 1-r_2mp_n. \quad (4.61)$$

The Lemma is proved. □

Now we state a tensorization lemma.

Lemma 4.5. *Let $\zeta_1,\dots,\zeta_n$ be independent non-negative random variables. Assume that*

$$\Pr\{\zeta_j \le \lambda_n\} \le 1-q_n \quad (4.62)$$

for some positive $q_n \in (0,1)$ and $\lambda_n > 0$. Then there exists positive absolute constants K_1 and K_2 such that

$$\Pr\{\sum_{j=1}^n \zeta_j^2 \le K_1^2 nq_n\lambda_n^2\} \le \exp\{-K_2nq_n\}. \quad (4.63)$$

Proof. We repeat the proof of Lemma 4.4 in [13]. Let $t = K_1\sqrt{q_n}\lambda_n$. For any $\tau > 0$ we have

$$\Pr\{\sum_{j=1}^n \zeta_j^2 \le nt^2\} \le e^{n\tau}\prod_{j=1}^n \mathbf{E}\exp\{-\tau\zeta_j^2/t^2\}. \quad (4.64)$$

Furthermore,

$$\begin{aligned}\mathbf{E}\exp\{-\tau\zeta_j^2/t^2\} &= \int_0^\infty \Pr\{\exp\{-\tau\zeta_j^2/t^2\} > s\}ds \\ &= \int_0^1 \Pr\left\{1/s > \exp\{\tau\zeta_j/t^2\}\right\}ds \\ &\le \int_0^{\exp\{-\tau\lambda_n^2/t^2\}} ds + \int_{\exp\{-\tau\lambda_n^2/t^2\}}^1 (1-q_n)ds \\ &\le 1-q_n(1-\exp\{-\tau\lambda_n^2/t^2\}) = 1-q_n(1-\exp\{-\tau/(K_1^2q_n)\}). \quad (4.65)\end{aligned}$$

Choosing $\tau := q_n/4$ and $K_1^2 := \frac{1}{4\ln 2}$, we get

$$\Pr\{\sum_{j=1}^n \zeta_j^2 \le nt^2\} \le \exp\{-nq_n/2\}. \quad (4.66)$$

Thus the Lemma is proved. □

Recall that we assumed $p_n^{-1} = O(n^{1-\theta}), 1 \ge \theta > 0$. For this fixed θ consider $L := [\frac{1}{\theta}]$. Hence by definition $p_{n,l} := (n\widehat{p}_n)^l p_n \to 0,\ n \to \infty$ for $l = 1,\dots,L-1$ and $\limsup_{n\to\infty}(np_n)^L p_n > 0$. We put $p_{n,L} := 1$.

We shall assume that n is large enough such that $(np_n)^L p_n \ge q_1 > 0$ for some constant $q_1 > 0$. Starting with a decomposition of $\mathcal{C}_0 := \mathcal{S}^{(n-1)}$ into compressible vectors $\mathbf{x}$ in

$\widehat{\mathcal{C}}_1 := \mathcal{C}_0 \cap Comp(\delta_1 p_{n,1}, \rho_{n,1})$, where $p_{n,1} = \widehat{p}_n$, $\rho_{n,1} = \gamma_0/(4K_n)$, and the constants γ_0 and δ_1 as in Lemma 4.1 and Lemma 4.2 respectively. Then Lemma 4.1 implies inequality (4.42) with q_n replaced by p_n and γ_n replaced by γ_0. Hence, using Lemma 4.2, one obtains the claim for the subset of vectors $\widehat{\mathcal{C}}_1$. The remaining vectors $\mathbf{x}x$ in $\mathcal{C}_0$ lie in $\mathcal{C}_1 := Incomp(\delta_1 p_{n,1}, \rho_{n,1})$. According to Lemmas 4.4, 4.5 we again have inequality (4.42) for these vectors but with new parameters $q_n = np_n\, \delta_1\, p_{n,1}$ and $\gamma_n = c\rho_{n,1}\sqrt{\delta_1\, p_{n,1}}$. Thus we may again subdivide the vectors in $\mathcal{C}_1$ into the vectors within distance $\rho_{n,2}$ from these sparses ones i.e. $\widehat{\mathcal{C}}_2 := \mathcal{C}_1 \cap\ Comp(\delta_2\, p_{n,2}, \rho_{n,2})$ and the remaining ones, i.e. $\mathcal{C}_2 := \mathcal{C}_1 \cap Incomp(\delta_2\, p_{n,2}, \rho_{n,2})$. Iterating this procedure L times we arrive at the incompressible set $\mathcal{C}_L$ of vectors $\mathbf{x}x$ where Lemmas 4.4, 4.5 and Proposition 4.6 yield the required bound of order $\exp\{-\delta n\}$, for sufficiently small absolute constant $\delta > 0$.

Summarizing, we will determine iteratively constants δ_l, $\rho_{n,l}$, for $l = 1, \dots, L$ and the following sets of vectors

$$\mathcal{C}_l := \cap_{i=1}^{l} Incomp(\delta_i\, p_{n,i}, \rho_{n,i}) \tag{4.67}$$

and

$$\widehat{\mathcal{C}}_l := \mathcal{C}_{l-1} \cap Comp(\delta_l\, p_{n,l}, \rho_{n,l}) \quad \text{with} \quad \mathcal{C}_0 = \mathcal{S}^{(n-1)}. \tag{4.68}$$

Note that

$$\mathcal{S}^{(n-1)} = \cup_{l=1}^{L-1} \widehat{\mathcal{C}}_l \cup \mathcal{C}_L. \tag{4.69}$$

The main bounds to carry out this procedure are given in the following Lemmas 4.6 and 4.7.

Lemma 4.6. *Let $\delta_n, \rho_n \in (0,1)$ and let $\mathbf{x} \in Incomp(\delta_n, \rho_n)$ and $\mathbf{X}^{(\varepsilon)}(z)$ be a matrix as in Theorem 4.1. Then there exists some positive constants c_1 and c_2 depending on K, r_0, η_0 such that for any $0 < \tau \le \gamma_n$*

$$\Pr\{\|\mathbf{X}^{(\varepsilon)}(z)\mathbf{x}\|_2 \le \tau\} \le \exp\{-c_1 n((p_n\, n\delta_n) \wedge 1)\} \tag{4.70}$$

with

$$\gamma_n := c_2 \rho_n \sqrt{\delta_n}, \tag{4.71}$$

where $a \wedge b$ denotes the minimum from a and b.

Proof. Assume at first that $n\delta_n p_n \le 1/3$. According to Lemma 4.4, we have, for any $j = 1\dots, n$,

$$\sup_{u\in\mathbb{C}} \Pr\{|\sum_{k=1}^{n} X_{jk}\varepsilon_{jk}x_k - u| \le \eta_0\rho_n/\sqrt{2n}\} \le 1 - r_1\delta_n\, np_n. \tag{4.72}$$

Applying Lemma 4.5 with $q_n = r_1\delta_n\, np_n$, we get

$$\Pr\{\|\mathbf{X}^{(\varepsilon)}(z)\mathbf{x}\|_2 \le \gamma_n/2 \quad \text{and} \quad \|\mathbf{X}^{(\varepsilon)}(z)\| \le K_n\} \le \exp\{-cn\delta_n np_n\}. \tag{4.73}$$

Consider now the case $n\delta_n\, p_n \ge 1/3$. According to Lemma 4.4, we have

$$\sup_{u\in\mathbb{C}} \Pr\{|\sum_{k=1}^{n} X_{jk}\varepsilon_{jk}x_k - u| \le \eta_0\rho_n/\sqrt{2n}\} \le 1 - r_1. \tag{4.74}$$

Applying Lemma 4.5 with $q_n = r_1\delta_n\, np_n$, we get

$$\Pr\{\|\mathbf{X}^{(\varepsilon)}(z)\mathbf{x}\|_2 \le \gamma_n/2 \quad \text{and} \quad \|\mathbf{X}^{(\varepsilon)}(z)\| \le K_n\} \le \exp\{-cn\}. \tag{4.75}$$

This completes the proof of the Lemma. ☐

Lemma 4.7. *For $l = 2,\dots,L$ assume that $\delta_i, \rho_{n,i}$ have been already determined for $i = 1,\dots,l-1$. Then there exist absolute constants $\widehat{c}_l > 0$ and $\overline{c}_l > 0$ and $\delta_l > 0$ such that*

$$\Pr\{\inf_{\mathbf{x}\in\widehat{\mathcal{C}}_l} \|\mathbf{X}^{(\varepsilon)}(z)\mathbf{x}\|_2 \le \gamma_{n,l} \quad \textit{and} \quad \|\mathbf{X}^{(\varepsilon)}(z)\| \le K_n\} \le \exp\{-\overline{c}_l\, n(((n\widehat{p}_n)^{l-1}p_n)\wedge 1)\}, \tag{4.76}$$

with $\gamma_{n,l}$ defined by

$$\gamma_{n,l} = \widehat{c}_l\,\rho_{n,l-1}\sqrt{\delta_{l-1}\,p_{n,l-1}}, \tag{4.77}$$

and $\rho_{n,l}$ defined by

$$\rho_{n,l} := \gamma_{n,l}/(4K_n), \tag{4.78}$$

where $\widehat{\mathcal{C}}_l := \mathcal{C}_{l-1}\cap Comp(\delta_l\, p_{n,l}, \rho_{n,l})$.

Remark 4.9. There exists some absolute constant $c > 0$ that

$$\gamma_{n,L} \ge cn^{-L/2} \quad \text{and} \quad \rho_{n,L} \ge cn^{-(L+3)/2}. \tag{4.79}$$

Proof of the Remark. Note that $p_{n,l}^{-1} = \mathcal{O}(n^{-1+l\theta})$ This implies that

$$\gamma_{n,L}^{-1} = \rho_{n,1}^{-1}\mathcal{O}(n^{\frac{L(L-1)\theta}{2}+L(1-\theta)}) \tag{4.80}$$

According to Lemmas 4.1 and 4.2, we have $\rho_{n1}^{-1} = \mathcal{O}(n^{\frac{3-\theta}{2}})$. After simple calculations we get

$$\gamma_{n,L}^{-1} = \mathcal{O}(n^{L/2}). \tag{4.81}$$

☐

Proof of Lemma 4.7. To prove of this Lemma we may use arguments similar to those in the proofs of Lemmas 2.6 and 3.3 in [21]. From $\mathbf{x}\in\mathcal{C}_l$ it follows that $\mathbf{x}\in Incomp(\delta_{l-1}p_{n,l-1}, \rho_{n,l-1})$. Applying Lemma 4.6 with $\delta_n = p_{n,l-1}$ and $\rho_n = \rho_{n,l-1}$, we get

$$\Pr\{\|\mathbf{X}^{(\varepsilon)}(z)\mathbf{x}\|_2 \le \gamma_{n,l} \quad \text{and} \quad \|\mathbf{X}^{(\varepsilon)}(z)\| \le K_n\} \le \exp\{-c_1 n((np_n\,\widehat{p}_{n,l-1})\wedge 1)\} \tag{4.82}$$

with

$$\gamma_{n,l} = c_2\,\rho_{n,l-1}\sqrt{\delta_{l-1}\,p_{n,l-1}}. \tag{4.83}$$

Inequality (4.82) and Lemma 4.2 together imply

$$\Pr\{\inf_{\mathbf{x}\in\mathcal{C}_l} \|\mathbf{X}^{(\varepsilon)}(z)\mathbf{x}\|_2 \le \gamma_{n,l} \quad \text{and} \quad \|\mathbf{X}^{(\varepsilon)}(z)\| \le K_n\} \le \exp\{-c_1 n\,\widehat{p}_{n,l}\} \tag{4.84}$$

with δ_l defined in Lemma 4.2 and

$$\rho_{n,l} := \gamma_{n,l}/(4K_n). \tag{4.85}$$

Thus the Lemma is proved. ☐

The next Lemma gives an estimate of small ball probabilities adapted to our case.

Lemma 4.8. *Let* $\mathbf{x} \in$ *Incomp*$(\delta, \rho_{n,L})$. *Let* $X_1, \ldots, X_n$ *be random variables with zero mean and variance at least 1. Assume that the following condition holds,*

$$L(M) := \max_{n\ge 1} \max_{1\le k\le n} \mathbf{E}\,|X_k|^2 I_{\{|X_k|>M\}} \to 0 \quad \text{as } M \to \infty. \tag{4.86}$$

Then there exists some constants $C > 0$ *depending on* δ *such that for every* $\varepsilon > 0$

$$p_{\mathbf{x}}(\varepsilon\rho_{n,L}/\sqrt{2n}) := \sup_v \Pr\{|\sum_{k=1}^n x_k\varepsilon_k X_k - v| \le \varepsilon\rho_{n,L}/\sqrt{2n}\} \le \frac{C\sqrt{\ln n}}{\sqrt{np_n}}. \tag{4.87}$$

Proof. Put $L_1 := [-\log_2(\rho_{n,L}\sqrt{2\delta})]$. Note that

$$\frac{\rho_{n,L}}{\sqrt{2n}} \le \frac{1}{2^{L_1+1/2}\sqrt{n\delta}} \le \frac{2\rho_{n,L}}{\sqrt{2n}}. \tag{4.88}$$

According to Remark 4.9, we have $\rho_{n,L} \ge cn^{-L}$. This implies $L_1 \le C\ln n$. Let $\sigma(\mathbf{x})$ denote the spread set of the vector $\mathbf{x}$, i.e.

$$\sigma(\mathbf{x}) := \left\{k: \ \rho_{n,L}/\sqrt{2n} \le |x_k| \le \sqrt{\frac{2}{n\delta}}\right\}. \tag{4.89}$$

By Lemma 4.3, we have

$$|\sigma(\mathbf{x})| \ge n\delta/2. \tag{4.90}$$

We divide the spread interval of the vector $\mathbf{x}$ into L_1+2 intervals Δ_l, $l = 0, \ldots, L_1+1$ by

$$\Delta_0 := \left\{k: \ \frac{\rho_{n,L}}{\sqrt{2n}} \le |x_k| \le \frac{1}{2^{L_1+1/2}\sqrt{n\delta}}\right\}, \tag{4.91}$$

$$\Delta_l := \left\{k: \ \frac{\sqrt{2}}{2^l\sqrt{n\delta}} \le |x_k| \le \frac{\sqrt{2}}{2^{l-1}\sqrt{n\delta}}\right\}, \quad l = 1, \ldots, L_1+1. \tag{4.92}$$

Note that there exists an $l_0 = 0, \ldots, L_1+1$ such that

$$|\Delta_{l_0}| \ge n\delta/(2(L_1+2)) \ge Cn/\ln n. \tag{4.93}$$

Let $\mathbf{y} = P_{\Delta_{l_0}}\mathbf{x}$. Put $a_l := \min_{k\in\Delta_l}|x_k|$ and $b_l := \max_{k\in\Delta_l}|x_k|$. Choose a constant M such that $L(M) \le 1/2$. By the properties of concentration functions, we have

$$p_{\mathbf{x}}(\varepsilon\rho_{n,L}/\sqrt{2n}) \le p_{\mathbf{y}}(\varepsilon\rho_{n,L}/\sqrt{2n}) \le p_{\mathbf{y}}(Mb_{l_0}). \tag{4.94}$$

By definition of Δ_{l_0}, we have

$$\sum_{k\in\Delta_{l_0}} |x_k|^2 \ge a_{l_0}^2|\Delta_{l_0}| \ge \rho_{n,L}^2/(2n)|\Delta_{l_0}|, \tag{4.95}$$

and

$$\frac{a_{l_0}}{b_{l_0}} \ge \frac{1}{2}. \tag{4.96}$$

Define

$$D(\xi,\lambda) = \lambda^{-2}\mathbf{E}\,|\xi|^2 I_{\{|\xi|<\lambda\}} \tag{4.97}$$

and introduce for a random variable ξ, $\widetilde{\xi} := \xi - \widehat{\xi}$ where $\widehat{\xi}$ denotes an independent copy of ξ. Put $\xi_k := x_k\varepsilon_k X_k$. We use the following inequality for a concentration function of a sum of independent random variables

$$p_{\mathbf{y}}(Mb_{l_0}) \le CMb_{l_0}\left(\sum_{k\in\Delta_{l_0}} \lambda_k^2 D(\widetilde{\xi_k\varepsilon_k};\lambda_k)\right)^{-\frac{1}{2}} \tag{4.98}$$

with $\lambda_k \le Mb_{l_0}$. See Petrov [22], p.43, Theorem 3. Put $\lambda_k = M|x_k|$. It is straightforward to check that,

$$\sum_{k\in\Delta_{l_0}} \lambda_k^2 D(\widetilde{\xi_k\varepsilon_k};\lambda_k) \ge p_n\left(\sum_{k\in\Delta_{l_0}} |x_k|^2(\mathbf{E}\,|X_k|^2 - L(M))\right). \tag{4.99}$$

This implies

$$\sum_{k\in\Delta_{l_0}} \lambda_k^2 D(\widetilde{\xi_k\varepsilon_k};\lambda_k) \ge \frac{p_n}{2}\sum_{k\in\Delta_{l_0}} |x_k|^2 \ge \frac{p_n}{2}|\Delta_{l_0}|a_{l_0}^2. \tag{4.100}$$

Combining this inequality with (4.98) and (4.94) we obtain

$$p_{\mathbf{x}}(\varepsilon\rho_{n,L}/\sqrt{2n}) \le \frac{CMb_{l_0}}{\sqrt{|\Delta_{l_0}|p_n}a_{l_0}} \le \frac{CM}{\sqrt{|\Delta_{l_0}|p_n}} \le \frac{C\sqrt{\ln n}}{\sqrt{np_n}}. \tag{4.101}$$

The last relation concludes the proof. □

Invertibility for the incompressible vectors via distance.

Lemma 4.9. *Let $\mathbf{X}_1,\mathbf{X}_2,\ldots,\mathbf{X}_n$ denote the columns of $\sqrt{np_n}\mathbf{X}^{(\varepsilon)}(z)$, and let $\mathcal{H}_k$ denotes the span of all column vectors except k-th. Then for every $\delta,\rho\in(0,1)$ and every $\eta>0$ one has*

$$\Pr\left\{\inf_{\mathbf{x}\in\mathcal{C}_L}\|\mathbf{X}^{(\varepsilon)}(z)\mathbf{x}\|_2 < \eta(\rho_{n,L}/\sqrt{n})^2/\sqrt{np_n}\right\} \le \frac{1}{n\delta_L}\sum_{k=1}^n \Pr\{\operatorname{dist}(\mathbf{X}_k,\mathcal{H}_k) < \eta\rho_{n,L}/\sqrt{n}\}.$$

Proof. Note that

$$\Pr\left\{\inf_{\mathbf{x}\in\widehat{\mathcal{C}}_L}\|\mathbf{X}^{(\varepsilon)}(z)\mathbf{x}\|_2 < \eta(\rho_{n,L}/\sqrt{n})^2/\sqrt{np_n}\right\}$$
$$\le \Pr\left\{\inf_{\mathbf{x}\in Incomp(\delta_L,\rho_{n,L})}\|\mathbf{X}^{(\varepsilon)}(z)\mathbf{x}\|_2 < \eta(\rho_{n,L}/\sqrt{n})^2/\sqrt{np_n}\right\}. \tag{4.102}$$

For the upper bound of the r.h.s. of (4.102) see [21], proof of Lemma 3.5. For reader convenience we repeat this proof. Introduce the matrix $\mathbf{G} := \sqrt{np_n}\mathbf{X}^{(\varepsilon)}(z)$. Recall that $\mathbf{X}_1, \dots, \mathbf{X}_n$ denote the column vector of the matrix $\mathbf{G}$ and $\mathcal{H}_k$ denotes the span of all column vectors except the $k-th$. Writing $\mathbf{G}\mathbf{x} = \sum_{k=1}^n x_k\mathbf{X}_k$, we have

$$\|\mathbf{G}\mathbf{x}\| \geq \max_{k=1,\dots,n} \operatorname{dist}(x_k\mathbf{X}_k, \mathcal{H}_k) = \max_{k=1,\dots,n} |x_k|\operatorname{dist}(\mathbf{X}_k, \mathcal{H}_k). \tag{4.103}$$

Put

$$p_k := \Pr\left\{\operatorname{dist}(\mathbf{X}_k, \mathcal{H}_k) < \eta\rho_{n,L}/\sqrt{n}\right\}. \tag{4.104}$$

Then

$$\mathbf{E}\Big|\{k: \ \operatorname{dist}(\mathbf{X}_k, \mathcal{H}_k) < \eta\rho_{n,L}/\sqrt{n}\}\Big| = \sum_{k=1}^n p_k. \tag{4.105}$$

Denote by U the event that the set $\sigma_1 := \{k : \operatorname{dist}(\mathbf{X}_k, H_k) \geq \eta\rho_{n,L}/\sqrt{n}\}$ contains more than $(1-\delta_L)n$ elements. Then by Chebyshev's inequality

$$\Pr\{U^c\} \leq \frac{1}{n\delta_L}\sum_{k=1}^n p_k. \tag{4.106}$$

On the other hand, for every incompressible vector $\mathbf{x}$, the set $\sigma_2(\mathbf{x}) := \{k : \ |x_k| \geq \rho_{n,L}/\sqrt{n}\}$ contains at least $n\delta_L$ elements. (Otherwise, since $\|P_{\sigma_2(\mathbf{x})^c}\mathbf{x}\|_2 \leq \rho_{n,L}$, we have $\|\mathbf{x}-\mathbf{y}\|_2 \leq \rho_{n,L}$ for the sparse vector $\mathbf{y} := P_{\sigma_2(\mathbf{x})}\mathbf{x}$, which would contradict the incompressibility of $\mathbf{x}$).

Assume that the event U occurs. Fix any incompressible vector $\mathbf{x}$. Then $|\sigma_1| + |\sigma_2(\mathbf{x})| > (1-\delta_L)n + n\delta_L > n$, so the sets σ_1 and $\sigma_2(\mathbf{x})$ have nonempty intersection. Let $k \in \sigma_1 \cap \sigma_2(\mathbf{x})$. Then by (4.103) and by definitions of the sets σ_1 and $\sigma_2(\mathbf{x})$, we have

$$\|\mathbf{G}\mathbf{x}\|_2 \geq |x_k|\operatorname{dist}(\mathbf{X}_k, \mathcal{H}_k) \geq \eta\rho_{n,L}n^{-1/2}. \tag{4.107}$$

Summarizing we have shown that

$$\Pr\{\inf_{\mathbf{x}\in Incomp(\delta_L,\rho_{n,L})} \|\mathbf{G}\mathbf{x}\|_2 \leq \eta(\rho_{n,L}n^{-1/2})^2\} \leq \Pr\{U^c\} \leq \frac{1}{n\delta_L}\sum_{k=1}^n p_k. \tag{4.108}$$

This completes the proof.

□

We now reformulate Lemma 3.6 from [21]. Let $\mathbf{X}_n^*$ to be any unit vector orthogonal to $\mathbf{X}_1, \dots, \mathbf{X}_{n-1}$. Consider the subspace $\mathcal{H}_n = \operatorname{span}(\mathbf{X}_1, \dots, \mathbf{X}_{n-1})$.

Lemma 4.10. *Let δ_l, ρ_l, c_l, $l = 1, \dots, L-1$ be as in Lemma 4.2 and $\delta_L, \rho_L, \overline{c}_L$ as in Lemma 4.7. Then there exists an absolute constant $\widehat{c}_L > 0$ such that*

$$\Pr\Big\{\mathbf{X}^* \notin \mathcal{C}_L \ \textit{and} \ \|\mathbf{X}^{(\varepsilon)}(z)\| \leq K_n\Big\} \leq \exp\{-\widehat{c}_L np_n\}. \tag{4.109}$$

Proof. Note that

$$\mathcal{S}^{(n-1)} = \cup_{l=1}^{L-1} \widehat{\mathcal{C}}_l \cup \mathcal{C}_L. \tag{4.110}$$

The event $\{\mathbf{X}^* \notin \mathcal{C}_L \quad \text{and} \quad \|\mathbf{X}^{(\varepsilon)}(z)\| \le K_n\}$ implies that the event

$$\mathcal{E} := \{ \inf_{\mathbf{x} \in \cup_{l=1}^{L-1} \widehat{\mathcal{C}}_l:\ \|\mathbf{x}\|_2=1} \|\mathbf{X}^{(\varepsilon)}(z)\mathbf{x}\|_2 \le c \quad \text{and} \quad \|\mathbf{X}^{(\varepsilon)}(z)\| \le K_n\} \tag{4.111}$$

occurs for any positive c. This implies, for $c > 0$,

$$\Pr\{\mathbf{X}^* \notin \mathcal{C}_L \quad \text{and} \quad \|\mathbf{X}^{(\varepsilon)}(z)\| \le K_n\} \tag{4.112}$$

$$\le \sum_{l=1}^{L-1} \Pr\{ \inf_{\mathbf{x} \in \widehat{\mathcal{C}}_l:\ \|\mathbf{x}\|_2=1} \|\mathbf{X}^{(\varepsilon)}(z)\mathbf{x}\| \le c \quad \text{and} \quad \|\mathbf{X}^{(\varepsilon)}(z)\| \le K_n\}. \tag{4.113}$$

Now choose $c := \min\{\gamma_{n,l},\ l = 1, \ldots, L-1\}$. Applying Lemma 4.7 proves the claim. □

Lemma 4.11. *Let $\mathbf{X}^{(\varepsilon)}(z)$ be a random matrix as in Theorem 1.2 . Let $\mathbf{X}_1, \ldots, \mathbf{X}_n$ denote column vectors of matrix $\sqrt{np_n}\mathbf{X}^{(\varepsilon)}(z)$, and consider the subspace $\mathcal{H}_n = \mathrm{span}(\mathbf{X}_1, \ldots, \mathbf{X}_{n-1})$. Let $K_n = Kn\sqrt{p_n}$. Then we have*

$$\Pr\{\mathrm{dist}(\mathbf{X}_n, \mathcal{H}_n) < \rho_{n,L}/\sqrt{n} \quad \textit{and} \quad \|\mathbf{X}^{(\varepsilon)}(z)\| \le K_n\} \le \frac{C\sqrt{\ln n}}{\sqrt{np_n}}. \tag{4.114}$$

Proof. We repeat Rudelson and Vershynin's proof of Lemma 3.8 in [21]. Let $\mathbf{X}^*$ be any unit vector orthogonal to $\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_{n-1}$. We can choose $\mathbf{X}^*$ so that it is a random vector that depends on $\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_{n-1}$ only and is independent of $\mathbf{X}_n$. We have

$$\mathrm{dist}(\mathbf{X}_n, \mathcal{H}_n) \ge |< \mathbf{X}_n, \mathbf{X}^* >|.$$

We denote the probability with respect to $\mathbf{X}_n$ by $\Pr_n$ and the expectation with respect to $\mathbf{X}_1, \ldots, \mathbf{X}_{n-1}$ by $\mathbf{E}_{1,\ldots,n-1}$. Then

$$\begin{aligned} \Pr\{\mathrm{dist}(\mathbf{X}_n, \mathcal{H}_n) < \rho_{n,L}/\sqrt{n} \quad \text{and} \quad \|\mathbf{X}^{(\varepsilon)}(z)\| \le K_n\} \\ \le \mathbf{E}_{1,\ldots,n-1}\Pr_n\{|< \mathbf{X}^*, \mathbf{X}_n >| \le \rho_{n,L}/\sqrt{n} \quad \text{and} \quad \mathbf{X}^* \in \mathcal{C}_L\} \\ + \Pr\{\mathbf{X}^* \notin \mathcal{C}_L \quad \text{and} \quad \|\mathbf{X}^{(\varepsilon)}(z)\| \le K_n\}. \end{aligned} \tag{4.115}$$

According to Lemmas 4.10, the second term in the right hand side of the last inequality is less then $\exp\{-\widehat{c}_L n\}$. Since the vectors $\mathbf{X}^* = (a_1, \ldots, a_n) \in \mathcal{S}^{(n-1)}$ and $\mathbf{X}_n = (\varepsilon_1\xi_1, \ldots, \varepsilon_n\xi_n)$ are independent, we may use small ball probability estimates. We have

$$S =< \mathbf{X}_n, \mathbf{X}^* >= \sum_{k=1}^{n} a_k \varepsilon_k \xi_k.$$

Let σ denote the set of spread of coefficients of $\mathbf{X}^*$ as in Lemma 4.3. Let P_σ denote the orthogonal projection onto $\mathbb{R}^\sigma$ in $\mathbb{R}^n$. Denote by $S_\sigma = \sum_{k\in\sigma} \varepsilon_k a_k \xi_k$. Using the properties of concentration function, we get

$$\begin{aligned}\Pr_n\{| < \mathbf{X}_n, \mathbf{X}^* > | \le \rho_{n,L}/\sqrt{n}\} &\le \sup_v \Pr_n\{|S - v| \le \rho_{n,L}/\sqrt{n}\} \\ &\le \sup_v \Pr_n\{|S_\sigma - v| \le \rho_{n,L}/\sqrt{n}\}.\end{aligned}$$

By Lemma 4.8, we have for some absolute constant $C > 0$

$$\Pr_n\{| < \mathbf{X}_n, \mathbf{X}^* > | \le \rho_{n,L}/\sqrt{n}\} \le \frac{C\sqrt{\ln n}}{\sqrt{np_n}}. \tag{4.116}$$

Thus the Lemma is proved. □

Lemma 4.12. *Let $\mathbf{X}^{(\varepsilon)}(z)$ be a random matrix as in Theorem 4.1. Let $\delta_L, \rho_{n,L} \in (0,1)$. Let $\mathbf{X}_1, \ldots, \mathbf{X}_n$ denote column vectors of matrix $\sqrt{np_n}\mathbf{X}^{(\varepsilon)}(z)$. Let $K_n = Kn\sqrt{p_n}$ with $K \ge 1$. Then we have*

$$\Pr\{\inf_{\mathbf{x}\in\mathcal{C}_L} \|\mathbf{X}^{(\varepsilon)}(z)\mathbf{x}\|_2 < \rho_{n,L}^2/n\} \le \Pr\{\|\mathbf{X}^{(\varepsilon)}(z)\| > K_n\} + \frac{C\sqrt{\ln n}}{\sqrt{np_n}}.$$

Proof. Note that

$$\begin{aligned}&\Pr\{\inf_{\mathbf{x}\in\mathcal{C}_L} \|\mathbf{X}^{(\varepsilon)}(z)\mathbf{x}\|_2 < \rho_{n,L}^2/n\} \\ &\le \Pr\{\inf_{\mathbf{x}\in\mathcal{C}_L} \|\mathbf{X}^{(\varepsilon)}(z)\mathbf{x}\|_2 < \rho_{n,L}^2/n \quad \text{and} \quad \|\mathbf{X}^{(\varepsilon)}(z)\| \le K_n\} + \Pr\{\|\mathbf{X}^{(\varepsilon)}(z)\| > K_n\}.\end{aligned} \tag{4.117}$$

Applying Lemma 4.9 with $\eta = \sqrt{p_n}$, we get

$$\Pr\{\inf_{\mathbf{x}\in\mathcal{C}_L} \|\mathbf{X}^{(\varepsilon)}(z)\mathbf{x}\|_2 < \rho_{n,L}^2/n\} \le \frac{1}{n\,\delta_L}\sum_{k=1}^n \Pr\{\text{dist}(\mathbf{X}_k, \mathcal{H}_k) < \rho_{n,L}\sqrt{p_n}/\sqrt{n}\}.$$

Applying Lemma 4.11, we obtain

$$\Pr\{\inf_{\mathbf{x}\in\mathcal{C}_L} \|\mathbf{X}^{(\varepsilon)}(z)\mathbf{x}\|_2 < \rho_{n,L}^2/n\} \le \frac{C\sqrt{\ln n}}{\sqrt{n\,p_n}}. \tag{4.118}$$

Lemma is proved. □

Proof of Theorem 4.1. By definition of the minimal singular value, we have

$$\begin{aligned}&\Pr\{s_n^{(\varepsilon)}(z) \le \rho_{n,L}^2/n \quad \text{and} \quad s_1^{(\varepsilon)}(z) \le K_n\} \\ &\le \Pr\{\text{there exist } \mathbf{x} \in \mathcal{S}^{(n-1)} : \ \|\mathbf{X}^{(\varepsilon)}(z)\mathbf{x}\|_2 \le \rho_{n,L}^2/n \quad \text{and} \quad s_1^{(\varepsilon)}(z) \le K_n\}.\end{aligned}$$

Furthermore, using the decomposition of the sphere $\mathcal{S}^{(n-1)} = \cup_{l=1}^{L-1}\widehat{\mathcal{C}}_l \cup \mathcal{C}_L$ into compressible and incompressible vectors, we get

$$\begin{aligned}\Pr\{s_n^{(\varepsilon)}(z) \le \rho_{n,L}^2/n \quad \text{and} \quad s_1^{(\varepsilon)}(z) \le K_n\}\\ \le \sum_{l=1}^{L-1} \Pr\{\inf_{\mathbf{x}\in\widehat{\mathcal{C}}_l} \|\mathbf{X}^{(\varepsilon)}(z)\mathbf{x}\|_2 \le \rho_{n,L}^2/n \quad \text{and} \quad s_1^{(\varepsilon)}(z) \le K_n\}\\ + \Pr\{\inf_{\mathbf{x}\in\mathcal{C}_L} \|\mathbf{X}^{(\varepsilon)}(z)\mathbf{x}\|_2 \le \rho_{n,L}^2/n \quad \text{and} \quad s_1^{(\varepsilon)}(z) \le K_n\}. \quad (4.119)\end{aligned}$$

According to Lemma 4.7, we have

$$\Pr\{\inf_{\mathbf{x}\in\widehat{\mathcal{C}}_l} \|\mathbf{X}^{(\varepsilon)}(z)\mathbf{x}\|_2 \le \rho_{n,L}^2/n \quad \text{and} \quad s_1^{(\varepsilon)}(z) \le K_n\} \le \exp\{-c_l np_n(n\widehat{p}_n)^{l-1}\}.$$

Lemmas 4.12 and 4.7 together imply that

$$\begin{aligned}\Pr\{\inf_{\mathbf{x}\in\mathcal{C}_L} \|\mathbf{X}^{(\varepsilon)}(z)\mathbf{x}\|_2 \le \rho_{n,L}^2/n \quad \text{and} \quad s_1^{(\varepsilon)}(z) \le K_n\}\\ \le \Pr\{\inf_{\mathbf{x}\in Incomp(\delta_L,\rho_{n,L})} \|\mathbf{X}^{(\varepsilon)}(z)\mathbf{x}\|_2 \le \rho_{n,L}^2/n \quad \text{and} \quad s_1^{(\varepsilon)}(z) \le K_n\}\\ \le \frac{C\sqrt{\ln n}}{\sqrt{np_n}} + \exp\{-\widehat{c}_L n\}. \quad (4.120)\end{aligned}$$

The last two inequalities together imply the result. □

Remark 4.10. To relax the condition $p_n^{-1} = \mathcal{O}(n^{1-\theta})$ of Theorem 4.1 to $p_n^{-1} = o(n/\ln^2 n)$ we should to put $L = \ln n$. Then the value L_1 of Lemma 4.8 is at most $C(\ln n)^2$ and hence we have the bound $C\ln n/\sqrt{np_n}$ in (4.87). The last yields the bound $C\ln n/\sqrt{np_n} + \exp\{-\widehat{c}_L n\}$ in (4.120). Thus Theorem 4.1 holds with B chosen to be of order $C\ln n$.

5 Proof of the main Theorem

In this Section we give the proof of Theorem 1.2. Theorem 1.1 follows from Theorem 1.2 with $p_n = 1$. Let $\gamma := \frac{1}{3}$ and let $R > 0$ and k_1 define in Lemma 6.2 with $q = 18$. Using the notations of Theorem 4.1 we introduce for any $z \in \mathbb{C}$ and absolute constant $c > 0$ the set $\Omega_n(z) = \{\omega \in \Omega: \ c/n^B \le s_n^{(\varepsilon)}(z), \ s_1(\varepsilon) \le n\sqrt{p_n}, \ |\lambda_{k_1}^{(\varepsilon)}| \le R\}$. According to Lemma 6.1

$$\Pr\{s_1^{(\varepsilon)}(\mathbf{X}) \ge n\sqrt{p_n}\} \le C(np_n)^{-1}.$$

According to Theorem 4.1, with $\varepsilon = c$,

$$\Pr\{c/n^B \ge s_n^{(\varepsilon)}(z)\} \le \frac{C\sqrt{\ln n}}{\sqrt{np_n}} + \Pr\{s_1^{(\varepsilon)} \ge n\sqrt{p_n}\}.$$

According to Lemma 6.2 with $q = 18$, we have

$$\Pr\{|\lambda_{k_1}^{(\varepsilon)}| \le R\} \le C\Delta_n^{\gamma} \le C[\varphi(\sqrt{np_n})]^{-\frac{1}{18}}. \quad (5.1)$$

These inequalities imply

$$\Pr\{\Omega_n(z)^c\} \le (\varphi(\sqrt{np_n}))^{-\frac{1}{18}}. \tag{5.2}$$

Let $r = r(n)$ be such that $r(n) \to 0$ as $n \to \infty$. A more specific choice will be made later. Consider the potential $U_{\mu_n}^{(r)}$. We have

$$\begin{aligned}
U_{\mu_n}^{(r)} &= -\frac{1}{n}\mathbf{E}\,\log|\det(\mathbf{X}^{(\varepsilon)} - z\mathbf{I} - r\xi\mathbf{I})| \\
&= -\frac{1}{n}\sum_{j=1}^{n}\mathbf{E}\,\log|\lambda_j^{(\varepsilon)} - r\xi - z|I_{\Omega_n(z)} - \frac{1}{n}\sum_{j=1}^{n}\mathbf{E}\,\log|\lambda_j^{(\varepsilon)} - r\xi - z|I_{\Omega_n^{(c)}(z)} \\
&= \overline{U}_{\mu_n}^{(r)} + \widehat{U}_{\mu_n}^{(r)},
\end{aligned}$$

where I_A denotes an indicator function of an event A and $\Omega_n(z)^c$ denotes the complement of $\Omega_n(z)$.

Lemma 5.1. *Assuming the conditions of Theorem 4.1, for r such that*

$$\ln(1/r)\,(\varphi(\sqrt{np_n}))^{-1/19} \to \infty \quad as \quad n \to \infty$$

we have

$$\widehat{U}_{\mu_n}^{(r)} \to 0, \;\; as\ n \to \infty. \tag{5.3}$$

Proof. By definition, we have

$$\widehat{U}_{\mu_n}^{(r)} = -\frac{1}{n}\sum_{j=1}^{n}\mathbf{E}\,\log|\lambda_j^{(\varepsilon)} - r\xi - z|I_{\Omega_n^{(c)}(z)}. \tag{5.4}$$

Applying Cauchy's inequality, we get, for any $\tau > 0$,

$$\begin{aligned}
|\widehat{U}_{\mu_n}^{(r)}| &\le \frac{1}{n}\sum_{j=1}^{n}\mathbf{E}^{\frac{1}{1+\tau}}|\log|\lambda_j^{(\varepsilon)} - r\xi - z||^{1+\tau}\left(\Pr\{\Omega_n^{(c)}\}\right)^{\frac{\tau}{1+\tau}} \\
&\le \left(\frac{1}{n}\sum_{j=1}^{n}\mathbf{E}\,|\log|\lambda_j^{(\varepsilon)} - r\xi - z||^{1+\tau}\right)^{\frac{1}{1+\tau}}\left(\Pr\{\Omega_n^{(c)}\}\right)^{\frac{\tau}{1+\tau}}.
\end{aligned} \tag{5.5}$$

Furthermore, since ξ is uniformly distributed in the unit disc and independent of λ_j, we may write

$$\mathbf{E}\,|\log|\lambda_j - r\xi - z||^{1+\tau} = \frac{1}{2\pi}\mathbf{E}\int_{|\zeta|\le 1}|\log|\lambda_j^{(\varepsilon)} - r\zeta - z||^{1+\tau}d\zeta = \mathbf{E}\,J_1^{(j)} + \mathbf{E}\,J_2^{(j)} + \mathbf{E}\,J_3^{(j)},$$

where

$$J_1^{(j)} = \frac{1}{2\pi}\int_{|\zeta|\le 1,\ |\lambda_j^{(\varepsilon)}-r\zeta-z|\le\varepsilon} |\log|\lambda_j^{(\varepsilon)} - r\zeta - z||^{1+\tau} d\zeta,$$
$$J_2^{(j)} = \frac{1}{2\pi}\int_{|\zeta|\le 1,\ \frac{1}{\varepsilon}>|\lambda_j^{(\varepsilon)}-r\zeta-z|>\varepsilon} |\log|\lambda_j^{(\varepsilon)} - r\zeta - z||^{1+\tau} d\zeta,$$
$$J_3^{(j)} = \frac{1}{2\pi}\int_{|\zeta|\le 1,\ |\lambda_j-r\zeta-z|>\frac{1}{\varepsilon}} |\log|\lambda_j^{(\varepsilon)} - r\zeta - z||^{1+\tau} d\zeta.$$

Note that

$$|J_2^{(j)}| \le \log\left(\frac{1}{\varepsilon}\right).$$

Since for any $b > 0$, the function $-u^b \log u$ is not decreasing on the interval $[0, \exp\{-\frac{1}{b}\}]$, we have for $0 < u \le \varepsilon < \exp\{-\frac{1}{b}\}$,

$$-\log u \le \varepsilon^b u^{-b} \log\left(\frac{1}{\varepsilon}\right).$$

Using this inequality, we obtain, for $b(1+\tau) < 2$,

$$|J_1^{(j)}| \le \frac{1}{2\pi}\varepsilon^{b(1+\tau)}\left(\log\left(\frac{1}{\varepsilon}\right)\right)^{1+\tau}\int_{|\zeta|\le 1,\ |\lambda_j^{(\varepsilon)}-r\zeta-z|\le\varepsilon} |\lambda_j^{(\varepsilon)} - r\zeta - z|^{-b(1+\tau)} d\zeta \tag{5.6}$$
$$\le \frac{1}{2\pi r^2}\varepsilon^b \log\left(\frac{1}{\varepsilon}\right)\int_{|\zeta|\le\varepsilon} |\zeta|^{-b(1+\tau)} d\zeta \le C(\tau, b)\varepsilon^2 r^{-2}\left(\log\left(\frac{1}{\varepsilon}\right)\right)^{1+\tau}. \tag{5.7}$$

If we choose $\varepsilon = r$, then we get

$$|J_1^{(j)}| \le C(\tau, b)\left(\log\left(\frac{1}{r}\right)\right)^{1+\tau}. \tag{5.8}$$

The following bound holds for $\frac{1}{n}\sum_{j=1}^n \mathbf{E}\, J_3^{(j)}$. Note that $|\log x|^{1+\tau} \le \varepsilon^2|\log\varepsilon|^{1+\tau}x^2$ for $x \ge \frac{1}{\varepsilon}$ and sufficiently small ε. Using this inequality, we obtain

$$\frac{1}{n}\sum_{j=1}^n \mathbf{E}\, J_3^{(j)} \le C(\tau)\varepsilon^2|\log\varepsilon|\frac{1}{n}\sum_{j=1}^n \mathbf{E}\,|\lambda_j^{(\varepsilon)} - r\zeta - z|^2 \le C(\tau)(1+|z|^2+r^2)\varepsilon^2|\log\varepsilon|$$
$$\le C(\tau)(2+|z|^2)r^2|\log r|. \tag{5.9}$$

The inequalities (5.6)–(5.9) together imply that

$$|\frac{1}{n}\sum_{j=1}^n \mathbf{E}\,|\log|\lambda_j^{(\varepsilon)} - r\xi - z||^{1+\tau}| \le C\left(\log\left(\frac{1}{r}\right)\right)^{1+\tau}. \tag{5.10}$$

Furthermore, the inequalities (5.4), (5.5), and (5.10) together imply

$$|\widehat{U}_{\mu_n}^{(r)}| \le C\left(\log\left(\frac{1}{r}\right)\right)\left(C(\varphi(\sqrt{np_n}))^{-\frac{1}{18}}\right)^{\frac{\tau}{1+\tau}}.$$

We choose $\tau = 18$ and rewrite the last inequality as follows

$$|\widehat{U}_{\mu_n}^{(r)}| \le C\left(\log\left(\frac{1}{r}\right)\right)(\varphi(\sqrt{np_n}))^{-\frac{1}{19}} \le C\left(\log\left(\frac{1}{r}\right)\right)(\varphi(\sqrt{np_n}))^{-\frac{1}{19}}$$

If we choose $r = \frac{1}{\sqrt{np_n}}$ we obtain $\log(1/r)((\varphi(\sqrt{np_n}))^{-\frac{1}{19}} \to 0$, then (5.3) holds and the Lemma is proved. □

We shall investigate $\overline{U}_{\mu_n}^{(r)}$ now. We may write

$$\begin{aligned}\overline{U}_{\mu_n}^{(r)} &= -\frac{1}{n}\sum_{j=1}^{n}\mathbf{E}\,\log|\lambda_j^{(\varepsilon)} - z - r\xi| I_{\Omega_n(z)} = -\frac{1}{n}\sum_{j=1}^{n}\mathbf{E}\,\log(s_j(\mathbf{X}^{(\varepsilon)}(z,r)) I_{\Omega_n(z)}\\ &= -\int_{n^{-B}}^{K_n+|z|}\log x d\mathbf{E}\,\overline{F}_n(x,z,r),\end{aligned} \tag{5.11}$$

where $\overline{F}_n(\cdot,z,r)$ is the distribution function corresponding to the restriction of the measure $\nu_n(\cdot,z,r)$ on the set $\Omega_n(z)$. Introduce the notation

$$\overline{U}_\mu = -\int_{n^{-B}}^{K_n+|z|}\log x dF(x,z). \tag{5.12}$$

Integrating by parts, we get

$$\begin{aligned}\overline{U}_{\mu_n}^{(r)} - \overline{U}_\mu = &-\int_{n^{-B}}^{K_n+|z|}\frac{\mathbf{E}\,F_n(x,z,r) - F(z,r)}{x}dx\\ &+ C\sup_x|\mathbf{E}\,F_n(x,z,r) - F(z,r)||\log(n^{B+1})|.\end{aligned} \tag{5.13}$$

This implies that

$$|\overline{U}_{\mu_n}^{(r)} - \overline{U}_\mu| \le C\ln n \sup_x|\mathbf{E}\,F_n(x,z,r) - F(x,z)|. \tag{5.14}$$

Note that, for any $r > 0$, $|s_j^{(\varepsilon)}(z) - s_j^{(\varepsilon)}(z,r)| \le r$. This implies that

$$\mathbf{E}\,F_n(x-r,z) \le \mathbf{E}\,F_n(x,z,r) \le \mathbf{E}\,F_n(x+r,z). \tag{5.15}$$

Hence, we get

$$\sup_x|\mathbf{E}\,F_n(x,z,r) - F(x,z)| \le \sup_x|\mathbf{E}\,F_n(x,z) - F(x,z)| + \sup_x|F(x+r,z) - F(x,z)|. \tag{5.16}$$

Since the distribution function $F(x,z)$ has a density $p(x,z)$ which is bounded (see Remark 3.1) we obtain

$$\sup_x |\mathbf{E}\, F_n(x,z,r) - F(x,z)| \le \sup_x |\mathbf{E}\, F_n(x,z) - F(x,z)| + Cr. \tag{5.17}$$

Choose $r = \frac{1}{\sqrt{np_n}}$. Inequalities (5.17) and (2.47) together imply

$$\sup_x |\mathbf{E}\, \overline{F}_n(x,z,r) - \overline{F}(x,z)| \le C(\varphi(\sqrt{np_n}))^{-\frac{1}{18}} + \frac{1}{\sqrt{np_n}}). \tag{5.18}$$

From inequalities (5.18) and (5.14) it follows that

$$|\overline{U}^{(r)}_{\mu_n} - \overline{U}_\mu| \le C(\varphi(\sqrt{np_n}))^{-\frac{1}{18}} + \frac{1}{\sqrt{np_n}}) \log(n^B).$$

Note that

$$|\overline{U}^{(r)}_{\mu_n} - U_\mu| \le |\int_0^{n^{-B}} \log x dF(x,z)| \le Cn^{-B}|\ln(n^{-B})|.$$

Let $\mathcal{K} = \{z \in \mathbb{C} : \ |z| \le R\}$ and let $\mathcal{K}^c$ denote $\mathbb{C} \setminus \mathcal{K}$. According to Lemma 6.2 with $q = 18$, we have, for k_1 and R from Lemma 6.2,

$$1 - q_n := \mathbf{E}\, \mu^{(r)}_n(\mathcal{K}^c) \le \frac{k_1}{n} + \Pr\{|\lambda_{k_1}| > R\} \le C(\varphi(np_n))^{-\frac{1}{18}}. \tag{5.19}$$

Furthermore, let $\overline{\mu}^{(r)}_n$ and $\widehat{\mu}^{(r)}_n$ be probability measures supported on the compact set K and $K^{(c)}$ respectively, such that

$$\mathbf{E}\, \mu^{(r)}_n = q_n \overline{\mu}^{(r)}_n + (1-q_n)\widehat{\mu}^{(r)}_n. \tag{5.20}$$

Introduce the logarithmic potential of the measure $\overline{\mu}^{(r)}_n$,

$$U_{\overline{\mu}^{(r)}_n} = -\int \log|z-\zeta| d\overline{\mu}^{(r)}_n(\zeta).$$

Similar to the proof of Lemma 5.1 we show that

$$\lim_{n\to\infty} |U^{(r)}_{\mu_n} - U_{\overline{\mu}^{(r)}_n}| \le C \ln n (\varphi(np_n))^{-\frac{1}{19}}.$$

This implies that

$$\lim_{n\to\infty} U_{\overline{\mu}^{(r)}_n}(z) = U_\mu(z)$$

for all $z \in \mathbb{C}$. Since the measures $\overline{\mu}^{(r)}_n$ are compactly supported, Theorem 6.9 from [16] and Corollary 2.2 from [16] (see also the Appendix, Theorem 6.1 and Corollary 6.8) together imply that

$$\lim_{n\to\infty} \overline{\mu}^{(r)}_n = \mu \tag{5.21}$$

in the weak topology. Inequality (5.19) and relations (5.20) and (5.20) together imply that

$$\lim_{n\to\infty} \mathbf{E}\,\mu_n^{(r)} = \mu$$

in the weak topology. Finally, by Lemma 1.1 we get

$$\lim_{n\to\infty} \mathbf{E}\,\mu_n = \mu \tag{5.22}$$

in the weak topology. Thus Theorem 1.2 is proved.

6 Appendix

In this Section we collect some technical results.

The largest singular value. We show the following

Lemma 6.1. *Under condition of Theorem 1.1 for sufficiently large $K \geq 1$ we have,*

$$\Pr\{s_1^{(\varepsilon)} \geq n\sqrt{p_n}\} \leq C/np_n \tag{6.1}$$

for some positive constant $C > 0$.

Proof. Using Chebyshev's inequality, we get

$$\Pr\{s_1^{(\varepsilon)}(\geq n\sqrt{p_n}\} \leq \frac{1}{n^2 p_n}\mathbf{E}\,\mathrm{Tr}\,(\mathbf{X}^{(\varepsilon)}(\mathbf{X}^{(\varepsilon)})^*) \leq 1/(np_n) \tag{6.2}$$

Thus the Lemma is proved. ☐

Recall that $|\lambda_1^{(\varepsilon)}| \geq \ldots \geq |\lambda_n^{(\varepsilon)}|$ denote the eigenvalues of the matrix $\mathbf{X}^{(\varepsilon)}$ ordered in decreasing of absolute values and $s_1^{(\varepsilon)} \geq \ldots \geq s_n^{(\varepsilon)}$ denote the singular values of the matrix $\mathbf{X}^{(\varepsilon)}$.

Lemma 6.2. *Assume that* $\max_{j,k} \mathbf{E}\,|X_{jk}|^2\varphi(X_{jk}) \leq C$ *with* $\varphi(x) := (\ln(1+|x|))^q$, $q \geq 7$, *and* $\Delta_n := \sup_x |F_n^{(\varepsilon)}(x,z) - F(x,z)|$. *Then there exists some absolute positive constant R such that*

$$\Pr\{|\lambda_{k_1}^{(\varepsilon)}| > R\} \leq (\varphi(np_n))^{-\frac{q-6}{12q}}, \tag{6.3}$$

where $k_1 := \left[\Delta_n^{(q+6)/(2q)} n \ln n\right]$.

Proof. Let us introduce $k_0 := \left[\Delta_n^{(q+6)/(2q)} n\right]$. Using Chebyshev's inequality we obtain, for sufficiently large $R > 0$,

$$\Pr\{s_{k_0}^{(\varepsilon)} > R\} \leq \frac{1 - \mathbf{E}\,F_n(R)}{k_0/n} \leq \Delta_n^{\frac{q-6}{2q}}.$$

On the other hand

$$\Pr\{|\lambda_{k_1}^{(\varepsilon)}| > R\} \le \Pr\{\prod_{\nu=1}^{k_1} |\lambda_\nu^{(\varepsilon)}| > R^{k_1}\} \le \Pr\{\prod_{\nu=1}^{k_1} s_\nu^{(\varepsilon)} > R^{k_1}\} \le \Pr\{\frac{1}{k_1}\sum_{\nu=1}^{k_1} \ln s_\nu^{(\varepsilon)} > \ln R\}.$$

Furthermore, for any value $R_1 \ge 1$, splitting into the events $s_{k_0}^{(\varepsilon)} > R$ and $s_{k_0}^{(\varepsilon)} \le R$, we get

$$\Pr\{\frac{1}{k_1}\prod_{\nu=1}^{k_1} \ln s_\nu^{(\varepsilon)} > \ln R_1\} \le \Pr\{s_{k_0}^{(\varepsilon)} > R\} + \Pr\{\frac{k_0}{k_1}\ln s_1^{(\varepsilon)} + \ln R > \ln R_1\}$$
$$\le \Delta_n^{\frac{q-6}{2q}} + \Pr\{\ln s_1^{(\varepsilon)} > \frac{k_1}{k_0}\ln\frac{R_1}{R}\}.$$

Now choose $R_1 := R^2$. Thus, since $k_1/k_0 \sim \ln n$,

$$\Pr\{|\lambda_{k_1}^{(\varepsilon)}| > R\} \le \Delta_n^{\frac{q-6}{2q}} + \Pr\{\ln s_1^{(\varepsilon)} > \ln R\, \ln n\}.$$

Taking into account Lemma 6.1 and inequality (2.47) we obtain

$$\Pr\{|\lambda_{k_1}^{(\varepsilon)}| > R\} \le \Delta_n^{\frac{q-6}{2q}} + \frac{C}{np_n} \le C(\varphi(np_n))^{-\frac{q-6}{12q}},$$

for some positive constant $C > 0$, thus proving the Lemma. □

Lemma 6.3. *Let* $\varkappa = \max_{j,k} \mathbf{E}\,|X_{jk}|^2\varphi(X_{jk})$. *The following inequality holds*

$$\frac{1}{n\sqrt{np_n}}\sum_{j,k=1}^{n} \mathbf{E}\,\varepsilon_{jk}|X_{jk}|(|T_{k+n,j}^{(jk)}| + |T_{j,k+n}^{(jk)}|) \le \frac{C}{v^3\varphi(\sqrt{np_n})}. \tag{6.4}$$

Proof. Introduce the notations

$$B := \frac{1}{n\sqrt{np_n}}\sum_{j,k=1}^{n} \mathbf{E}\,\varepsilon_j k|X_{jk}|(|T_{k+n,j}^{(jk)}| + |T_{j,k+n}^{(jk)}|) \tag{6.5}$$

and

$$B_1 := \frac{2}{n^2p_n}\sum_{j,k=1}^{n} \mathbf{E}\,\varepsilon_{jk}|X_{jk}|^2|R_{k+n,j}^{(jk)}||R_{k+n,j}^{(jk)} - R_{k+n,j}|,$$
$$B_2 := \frac{2}{n^2p_n}\sum_{j,k=1}^{n} \mathbf{E}\,\varepsilon_{jk}|X_{jk}|^2|R_{k+n,k+n}^{(jk)}||R_{j,j}^{(jk)} - R_{j,j}|,$$
$$B_3 := \frac{2}{n^2p_n}\sum_{j,k=1}^{n} \mathbf{E}\,\varepsilon_{jk}|X_{jk}|^2|R_{j,j}^{(jk)}||R_{k+n,k+n}^{(jk)} - R_{k+n,k+n}|,$$
$$B_4 := \frac{2}{n^2p_n}\sum_{j,k=1}^{n} \mathbf{E}\,\varepsilon_{jk}|X_{jk}|^2|R_{j,k+n}^{(jk)}||R_{j,k+n}^{(jk)} - R_{j,k+n}|. \tag{6.6}$$

Since the function $|x|/\varphi(x)$ not decreasing, it follows from inequality (2.10) that

$$|R^{(jk)}_{l,m} - R_{l,m}| \le \frac{1}{v} I_{\{|X_{jk}|>\sqrt{np_n}\}} + \frac{1}{v^2\varphi(\sqrt{np_n})}\varphi(X_{jk}). \tag{6.7}$$

It is easy to check that

$$\max\{B_k,\ k=1,\dots,8\} \le \frac{C\varkappa}{v^3\varphi(\sqrt{np_n})}. \tag{6.8}$$

This implies that

$$B \le \frac{C\varkappa}{v^3\varphi(\sqrt{np_n})}. \tag{6.9}$$

□

Lemma 6.4. *Let μ_n be the empirical spectral measure of the matrix $\mathbf{X}$ and ν_r be the uniform distribution on the disc of radius r. Let $\mu_n^{(r)}$ be the empirical spectral measure of the matrix $\mathbf{X}(r) = \mathbf{X} - r\xi\mathbf{I}$, where ξ is a random variable which is uniformly distributed on the unit disc. Then the measure $\mathbf{E}\,\mu_n^{(r)}$ is the convolution of the measures $\mathbf{E}\,\mu_n$ and ν_r, i. e.*

$$\mathbf{E}\,\mu_n^{(r)} = (\mathbf{E}\,\mu_n) * (\nu_r). \tag{6.10}$$

Proof. Let J be a random variable which is uniformly distributed on the set $\{1,\dots,n\}$. Let $\lambda_1,\dots,\lambda_n$ be the eigenvalues of the matrix $\mathbf{X}$. Then $\lambda_1+r\xi,\dots,\lambda_n+r\xi$ are eigenvalues of the matrix $\mathbf{X}(r)$. Let δ_x be denote the Dirac measure. Then

$$\mu_n = \frac{1}{n}\sum_{j=1}^n \delta_{\lambda_j} \tag{6.11}$$

and

$$\mu_n^{(r)} = \frac{1}{n}\sum_{j=1}^n \delta_{\lambda_j+r\xi}. \tag{6.12}$$

Denote by μ_{nj} the distribution of λ_j. Then

$$\mathbf{E}\,\mu_n = \frac{1}{n}\sum_{j=1}^n \mu_{nj} \tag{6.13}$$

and

$$\mathbf{E}\,\mu_n^r = \frac{1}{n}\sum_{j=1}^n \mu_{nj} * \nu_r = \left(\frac{1}{n}\sum_{j=1}^n \mu_{nj}\right) * (\nu_r) = (\mathbf{E}\,\mu_n) * (\nu_r). \tag{6.14}$$

Thus the Lemma is proved. □

Let

$$f_n^{(r)}(t,v) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \exp\{itx + ivy\} dG_n^{(r)}(x,y) \tag{6.15}$$

and

$$f_n(t,v) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \exp\{itx + ivy\} dG_n(x,y), \tag{6.16}$$

where

$$G_n^{(r)}(x,y) = \frac{1}{n}\sum_{j=1}^{n} \Pr\{\mathrm{Re}\,\lambda_j + r\xi \le x, \mathrm{Im}\,\lambda_j + r\xi \le y\}, \tag{6.17}$$

and

$$G_n(x,y) = \frac{1}{n}\sum_{j=1}^{n} \Pr\{\mathrm{Re}\,\lambda_j \le x, \mathrm{Im}\,\lambda_j \le y\}. \tag{6.18}$$

Denote by $h(t,v)$ the characteristic function of the joint distribution of the real and imaginary parts of ξ,

$$h(t,v) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \exp\{iux + ivy\} dG(x,y). \tag{6.19}$$

Lemma 6.5. *The following relations hold*

$$f_n^{(r)}(t,v) = f_n(t,v)h(rt,rv). \tag{6.20}$$

If for any t,v there exists $\lim_{n\to\infty} f_n(t,v)$, then

$$\lim_{r\to 0}\lim_{n\to\infty} f_n^{(r)}(t,v) = \lim_{n\to\infty}\lim_{r\to 0} f_n^{(r)}(t,v) = \lim_{n\to\infty} f_n(t,v). \tag{6.21}$$

Proof. The first equality follows immediately from the independence of the random variable ξ and the matrix $\mathbf{X}$. Since $\lim_{r\to 0} h(rt,rv) = h(0,0) = 1$ the first equality implies the second one. □

Lemma 6.6. *Let F and G be distribution functions with Stieltjes transforms $S_F(z)$ and $S_G(z)$ respectively. Assume that $\int_{-\infty}^{\infty} |F(x) - G(x)|dx < \infty$. Let $G(x)$ have a bounded support J and density bounded by some constant K. Let $V > v_0 > 0$ and a be positive numbers such that*

$$\gamma = \frac{1}{\pi}\int_{|u|\le a} \frac{1}{u^2+1}\, du > \frac{3}{4}.$$

Then there exist some constants C_1, C_2, C_3 depending on J and K only such that

$$\sup_x |F(x) - G(x)| \le C_1 \sup_{x\in J} \int_{-\infty}^{x} |S_F(u+iV) - S_G(u+iV)|\, du$$
$$+ \sup_{u\in J}\int_{v_0}^{V} |S_F(u+iv) - S_G(u+iv)|dv + C_3\ v_0. \tag{6.22}$$

Lemma 6.7. *Let X_{jk}, $1 \le j,k \le n$, be independent complex random variables with $\mathbf{E}\,X_{j,k} = 0$ and $\mathbf{E}\,|X_{j,k}|^2 = 1$. Assume furthermore that $\max_{j,k} \mathbf{E}\,|X_{jk}|^2 I_{\{|X_{jk}|>M\}} \to 0$ for $M \to +\infty$. Then we have, for some positive r_0 and η_0,*

$$\sup_{u\in\mathbb{C}} \max_{j,k} \Pr\{|X_{jk} - u| < \eta_0\} \le r_0 < 1.$$

Proof. First we note, that there exists a positive number M such that

$$\min_{j,k} \mathbf{E}\,(|X_{jk}|^2\, I_{\{|X_{jk}|\le M\}}) > \frac{7}{8}.$$

Let η_0 be a small positive number. For $|u| > M + \eta_0$ we have

$$\Pr\{|X_{jk} - u| \ge \eta_0\} \ge \Pr\{|X_{jk}| \le M\} \ge \frac{1}{M^2}\mathbf{E}\,(|X_{jk}|^2\, I_{\{|X_{jk}|\le M\}}) > \frac{7}{8M^2}. \tag{6.23}$$

Consider now $|u| \le M + \eta_0$. Then

$$\begin{aligned}
&\Pr\{|X_{jk} - u| \ge \eta_0\} \ge \mathbf{E}\,(I_{\{2M+\eta_0\ge|X_{jk}-u|\ge\eta_0\}}) \ge \frac{1}{4M^2}\mathbf{E}\,(|X_{jk} - u|^2 I_{\{2M+\eta_0\ge|X_{jk}-u|\ge\eta_0\}})\\
&\ge \frac{1}{4M^2}(1 - \mathbf{E}\,(|X_{jk} - u|^2 I_{\{|X_{jk}-u|<\eta_0\}}) - \mathbf{E}\,(|X_{jk} - u|^2 I_{\{|X_{jk}-u|>2M+\eta_0\}}))\\
&\ge \frac{1}{4M^2}(1 - \eta_0 - \mathbf{E}\,(|X_{jk} - u|^2 I_{\{|X_{jk}|>M\}}) \ge \frac{1}{4M^2}\Big(\frac{3}{4} - \eta_0 - \frac{|u|^2}{4M^2}\Big)\\
&\ge \frac{1}{16M^2}\Big(3 - 4\eta_0 - \Big(1 + \frac{\eta_0}{M}\Big)^2\Big).
\end{aligned} \tag{6.24}$$

Combining inequalities 6.23 and 6.24 we obtain the claim. □

6.1 Some facts from logarithmic potential theory

We cite here some definitions and Theorems about logarithmic potentials, see [16]. Let $\Sigma \subset \mathbb{C}$ be a compact set of the complex plane and $\mathcal{M}(\Sigma)$ the collection of all positive Borel probability measures with support in Σ. The *logarithmic energy* of $\mu \in \mathcal{M}(\Sigma)$ is defined as

$$I(\mu) := \int\int \log \frac{1}{|z-t|} d\mu(z)d\mu(t), \tag{6.25}$$

and the energy of Σ by

$$V := \inf\{I(\mu)|\mu \in \mathcal{M}(\Sigma)\}. \tag{6.26}$$

The quantity

$$cap(\Sigma) := e^{-V} \tag{6.27}$$

is called the *logarithmic capacity* of Σ.

The *capacity* of an arbitrary Borel set E is defined as

$$\cap(E) := sup\{cap(K)|K \subset E, K \text{ compact}\}. \tag{6.28}$$

Note that every Borel set of capacity zero has zero two-dimensional Lebesgue measure. A property is said to hold *quasi-everywhere* (q. e.) on a set E if the set of exceptional points is of capacity zero. The next Theorem is called *Lower Envelope Theorem*

Theorem 6.1. *Let μ_n, $n = 1, 2 \ldots$, be a sequence of positive Borel probability measures having support in a fixed compact set. If $\mu_n \to \mu$ weakly, then*

$$\liminf_{n\to\infty} U^{\mu_n}(z) = U^{\mu}(z) \tag{6.29}$$

for quasi-every $z \in \mathbb{C}$.

The following fact is Corollary 2.2 from the Unicity Theorem of logarithmic potential theory (see [16], p. 98).

Corollary 6.8. *If μ and ν are compactly supported measures and the potentials U^{μ} and U^{ν} coincides almost everywhere with respect to two-dimensional Lebesgue measure, then $\mu = \nu$.*

For reader convenience we give here the statement of Theorem 1.2 from [16].

Theorem 6.2. *Let μ be a finite positive measure of compact support on the plane. Then for any z_0 and $r > 0$ the mean value*

$$L(U^{\mu}; z_0, r) := \frac{1}{2\pi}\int_{-\pi}^{\pi} U^{\mu}(z_0 + r\exp\{i\theta\})d\theta \tag{6.30}$$

exists as a finite number, and $L(U^{\mu}; z_0, r)$ is a non-increasing function of r that is absolutely continuous on any closed subinterval of $(0, \infty)$. Furthermore,

$$\lim_{r\to 0} L(U^{\mu}; z_0, r) = U^{\mu}(z_0). \tag{6.31}$$

References

[1] Bai, Z. D. *Circular law.*
Annals of Probab., **25** (1997), 494– 529

[2] Yin, Y. Q. and Bai, Z. D. and Krishnaiah, P. R., *On the Limit of largest Eigenvalue of Large Dimensional Sample Covariance Matrix.*
Probab. and related fields. **78**, (1988) 509–521.

[3] Bai, Z. D. and Silverstein J. W. *No Eigenvalues Outside the support of the limit specvtral distribution of large dimensional matrix.*
Annals of Probab., **26** (1998), 316– 345

[4] Bickel P.J., Götze, F. and van Zwet W.R. *The Edgeworth expansion for U-statistics of degree two.*
The Annals of Statistics, **14** (1986), 1463–1484.

[5] Edelman, A. *The probability that a random real Gaussian matrix has k real eigenvalues, related distributions, and circular law.*
Journ. Mult. Analysis. **60** (1997), 203–232.

[6] Girko, V. L. *Circular law.*
Theory Probab. Appl., **29** (1989), 694–706

[7] Ginibre, J. *Statistical ensembles of complex, quaterninon, and real matrices.*
J. Math. Phys., **6** (1965), 440–449

[8] Gohberg, I. C. and Krein, M. G. *Introduction to the Theory of Linear Operator.*
Cambridge University Press, New York 1991

[9] Götze, F. and Tikhomirov, A. N. *Rate of convergence to the semi-circular law.*
Probab. Theory Relat. Fields **127** (2003), 228–276

[10] Götze, F. and Tikhomirov, A. N. *On the circular law.*
http://www.arXiv.org/abs/math.PR/0702386

[11] Gradstein, I. S. and Ryzhik, I. M. *Table of Integrals, Series, and Products.*
Academic Press, Inc. New York, 1994

[12] Horn, R., Johnson, Ch. *Matrix analysis.*
Cambridge University Press, 1991, pp. 561

[13] Litvak, A. E. and Pajor, A. and Rudelson, M., Tomczak-Jaegermann N., *Smallest singular value of random matrices and geometry of random polytopes.*
Adv. Math., **195**, 491–523, 2005 New York 1991

[14] Marchenko and V., Pastur, L. *The eigenvalue distribution in some ensembles of random matrices.*
Math.USSR Sbornik, **1** (1967), 457-483

[15] G. Pan and W. Zhou, *Circular law, Extreme singular values and potential theory.*
http://arXiv.org/abs/math.PR/0705.3773.

[16] Saff, E. B., Totik, V. *Logarithmic potentials with external fields.*
Springer, Berlin, 1997, pp. 505

[17] Mehta, M. L. *Random Matrices.*
2nd ed., Academic Press, San Diego 1991

[18] Pastur, L.A.*Spectra of random self adjoint operators.*
Russian mathematical Surveys **28**, 1, 1–67, 1973

[19] Petrov, V. V. *Sums of Independent Random variables.*
Springer Verlag, Berlin – Heidelberg – New York, 1975, pp. 346

[20] Rudelson, Mark, *Invertiility of random matrices: Norm of the inverse.*
www.math.missouri.edu/ rudelson/papers/square-matrix.pdf, 1–25, 2006

[21] Rudelson, M. and Vershynin R., *The Littlewood–Offord Problem and Invertibility of Random Matrices.*
http://arXiv.org/abs/math.PR/0703307.

[22] Petrov, V. V. *Sums of independent random variables.*
Springer Verlag, Berlin 1975, 345 pp.

[23] Brian Rider and Bálint Virág, *The noise in the circular law and Gaussian free field.*
http://arXiv:math/0606663v2[math.PR], 24 Aug 2006

[24] Brian Rider, *A limit theorems at the edge of non-Hermitian random matrix ensemble.*
Journal of Physics A: Mathematical and General, **36**, 3401–3409, 2003

[25] Tao, T. and Vu, V. *Random Matrices: The circular Law.*
http://arXiv.org/abs/math.PR/0708.2895.

[26] T. Tao and V. Vu, *Inverse Littlewood-Offord theorems and the condition number of random discrete matrices.*
Annals of Mathematics, to appear.

[27] Timme, M. and Wolf, F. and Geisel, T., *Topological Speed Limits to Network Synchronization.*
Phys. Rev. Letters, **92**, (2004) no. 7, 074101-1–4

[28] Wigner, E. *Characteristic vectors of bordered matrices with infinite dimensions.*
Ann. of Math.**62** (1955), 548–564

[29] Wigner, E. *On the distribution of the roots of certain symmetric matrices.*
Ann. of Math.**67** (1958), 325–327