The Annals of Probability
2010, Vol. 38, No. 4, 1444–1491
DOI: 10.1214/09-AOP522

THE CIRCULAR LAW FOR RANDOM MATRICES

BY FRIEDRICH GÖTZE[1] AND ALEXANDER TIKHOMIROV[1,2]

University of Bielefeld and St. Petersburg State University

We consider the joint distribution of real and imaginary parts of eigenvalues of random matrices with independent 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 those results to a larger class of sparse matrices.

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

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

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 $\mathbf{E}G_n(x,y)$ to the distribution function $G(x,y)$ of the uniform distribution in the unit disc in $\mathbb{R}^2$.

The main result of our paper is the following:

THEOREM 1.1. *Let $\varphi(x)$ denote the function $(\ln(1+|x|))^{19+\eta}$, $\eta > 0$, arbitrary, small and fixed. Let X_{jk}, $j, k \in \mathbf{N}$, denote independent complex random variables with*

$$\mathbf{E}X_{jk} = 0, \qquad \mathbf{E}|X_{jk}|^2 = 1 \quad \textit{and} \quad \varkappa := \sup_{j,k \in \mathbf{N}} \mathbf{E}|X_{jk}|^2 \varphi(X_{jk}) < \infty.$$

Then $\mathbf{E}G_n(x,y)$ converges weakly to the distribution function $G(x,y)$ as $n \to \infty$.

Received November 2007; revised December 2008.

[1]Supported in part by DFG Project G0-420/5-1 and CRC 701.

[2]Supported in part by RFBF Grant N 09-01-12180-ofi_m, by RFBR–DFG Grant N 09-01-91331-NNIO_a.

AMS 2000 subject classifications. Primary 60K35, 60K35; secondary 60K35.

Key words and phrases. Circular law, random matrices.

We shall prove the same result for the following class of sparse matrices. Let ε_{jk}, $j, k = 1, \dots, n$, denote a triangular array of Bernoulli random variables (taking values 0, 1 only) which are independent in aggregate and independent of $(X_{jk})_{j,k=1}^n$ with common success probability $p_n := \Pr\{\varepsilon_{jk} = 1\}$ depending on n. Consider the sequence of matrices $\mathbf{X}^{(\varepsilon)} = \frac{1}{\sqrt{np_n}}(\varepsilon_{jk}X_{jk})_{j,k=1}^n$. 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 $\mathbf{X}^{(\varepsilon)}$, that is,

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

THEOREM 1.2. *For $\eta > 0$ define $\varphi(x) = (\ln(1 + |x|))^{19+\eta}$. Let X_{jk}, $j, k \in \mathbf{N}$, denote independent complex random variables with*

$$\mathbf{E}X_{jk} = 0, \qquad \mathbf{E}|X_{jk}|^2 = 1 \quad \text{and} \quad \varkappa := \sup_{j,k\in\mathbf{N}} \mathbf{E}|X_{jk}|^2\varphi(X_{jk}) < \infty.$$

Assume that there is a $\theta \in (0, 1]$ such that $p_n^{-1} = \mathcal{O}(n^{1-\theta})$ as $n \to \infty$. 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 $s_n(z)$, respectively, $s_n^{(\varepsilon)}(z)$ of the shifted matrices $\mathbf{X} - z\mathbf{I}$, respectively, $\mathbf{X}^{(\varepsilon)} - z\mathbf{I}$. (See also [5], page 1561.) These bounds are based on the results obtained by Rudelson and Vershynin in [18]. In a previous version of this paper [10] we have used the corresponding results of Rudelson [17] 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 $\sup_{j,k} \mathbf{E}|X_{jk}|^4 \le \varkappa_4 < \infty$ in view of the fact (explicitly stated by Rudelson in [17]) 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 $\Omega_n(z) = \{s_n(z) \le cn^{-3}; \|\mathbf{X}\| \le K\}$ for the investigation of the smallest singular values, the inequality $\Pr\{\Omega_n(z)^c\} \le cn^{-1/2}$ follows from the results of Rudelson [17] without the assumption of sub-Gaussian tails for the matrix $\mathbf{X}$. A similar result has been proved by Pan and Zhou in [13] based on results of Rudelson and Vershynin [18] and Bai and Silverstein [2].

The strong circular law assuming moment condition of order larger than 2 only and comparable sparsity assumptions was proved independently by Tao and Vu in [22] based on their results in [23] 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 [18], Section 3.2, page 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 (nonsymmetric and non-Hermitian) random matrices with independent entries has a long history. Ginibre's [7], in 1965, 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's formula, Mehta [15], in 1967, determined the density of the expected spectral distribution function of random matrices 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 [16], page 64). Using the Ginibre results, Edelman [4], in 1997, proved the circular law for the matrices with i.i.d. Gaussian real entries. Rider proved in [21] and [20] results about the spectral radius and about linear statistics of eigenvalues of non-Hermitian matrices with Gaussian entries.

Girko [6], in 1984, investigated the circular law for general matrices with independent entries assuming that the distribution of the entries has 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., [24]). 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

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

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 A.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*

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

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 the 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 A.5 in the Appendix), we conclude the result. $\square$

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 [14], page 73 (see also Section 3.8 in the Appendix), we obtain that under these assumptions

$$\liminf_{n\to\infty} U_{\mu_n}^{(r)}(z) = U^{(r)}(z),$$

quasi-everywhere in $\mathbb{C}$ (for the definition of "*quasi-everywhere*" see, e.g., [14], page 24). Here $U^{(r)}(z)$ denotes the logarithmic potential of the 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

$$L(\mu; z_0, v) = \frac{1}{2\pi}\int_{-\pi}^{\pi} U_\mu(z_0 + v\exp\{i\theta\})\,d\theta \tag{1.4}$$

and

$$U_\mu(z) = \int \ln|\zeta - z|\,d\mu(\zeta). \tag{1.5}$$

Applying Theorem 1.2 in [14], page 84, we get

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

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

Since $\mathbf{E}\frac{1}{n}\operatorname{Tr}\mathbf{X}\mathbf{X}^* = 1$ 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 A.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 [see Section 3, equality (3.15)], 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 $s_1^{(\varepsilon)}(z,r) \geq \cdots \geq s_n^{(\varepsilon)}(z,r)$ 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, e.g., [8], Lemma 3.3, page 35), we may represent the function $U_{\mu_n}^{(r)}(z)$ as follows:

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

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 the 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 the potential." We solve this problem using bounds for the minimal singular values of the matrices $\mathbf{X}^{(\varepsilon)}(z) := \mathbf{X}^{(\varepsilon)} - z\mathbf{I}$ based on techniques developed in Rudelson [17] and Rudelson and Vershynin [18]. The bounds of minimal singular values of matrices $\mathbf{X}^{(\varepsilon)}$ are given in Section 4 and in the Appendix, Theorem 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 indices) some general absolute constant which may be changed from line to line. To specify a constant we shall use subindices. 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.

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

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

where $s_1^{(\varepsilon)}(z) \geq \cdots \geq s_n^{(\varepsilon)}(z) \geq 0$ 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

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

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

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

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 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 nondecreasing we have*

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

Then

$$\sup_x \bigl|F_n^{(\varepsilon)}(x, z) - F(x, z)\bigr| \le C \varkappa \bigl(\varphi\bigl(\sqrt{n p_n}\bigr)\bigr)^{-1/6}. \tag{2.2}$$

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

$$\sup_x \bigl|F_n^{(\varepsilon)}(x, z) - F(x, z)\bigr| \le C (n p_n)^{-1/12}. \tag{2.4}$$

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. Using the inverse of the partial matrix (see, e.g., [11], Chapter 08, page 18) it follows that, for $\alpha = u + iv$, $v > 0$,

$$\begin{aligned} (\mathbf{W} - \alpha \mathbf{I}_{2n})^{-1} = \Biggl(& \begin{matrix} \alpha\bigl(\mathbf{X}^{(\varepsilon)}(z)\mathbf{X}^{(\varepsilon)}(z)^* - \alpha^2\mathbf{I}\bigr)^{-1} \\ \mathbf{X}^{(\varepsilon)}(z)^*\bigl(\mathbf{X}^{(\varepsilon)}(z)\mathbf{X}^{(\varepsilon)}(z)^* - \alpha^2\mathbf{I}\bigr)^{-1} \end{matrix} \\ & \qquad \begin{matrix} \mathbf{X}^{(\varepsilon)}(z)\bigl(\mathbf{X}^{(\varepsilon)}(z)^*\mathbf{X}^{(\varepsilon)}(z) - \alpha^2\mathbf{I}\bigr)^{-1} \\ \alpha\bigl(\mathbf{X}^{(\varepsilon)}(z)^*\mathbf{X}^{(\varepsilon)}(z) - \alpha^2\mathbf{I}\bigr)^{-1} \end{matrix} \Biggr), \end{aligned} \tag{2.5}$$

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

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

We introduce the notation

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

With this notation 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

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

In 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 obtained from $\mathbf{W}$ by replacing 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

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

Here, and in 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

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

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

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

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

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

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 A.3 in the Appendix. Using the last inequalities we obtain, that for $v>0$

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

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

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

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

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

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 $v>0$

$$\left|\frac{1}{n^2}\sum_{j=1}^n\sum_{k=1}^n\mathbf{E}R_{jj}^{(jk)}R_{k+n,k+n}^{(jk)}-\big(S_n^{(\varepsilon)}(\alpha,z)\big)^2\right|\le\frac{C}{nv^4}. \tag{2.21}$$

Denote by $r(\alpha,z)$ some generic function with $|r(\alpha,z)|\le 1$ which may vary from line to line. We may now rewrite equality (2.8) as follows:

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

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 $V(\alpha,z)=\frac{1}{n}\mathbf{E}\operatorname{Tr}\mathbf{D}$. 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

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

According to equality (2.7), we have

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

Using the resolvent equality (2.10) and Lemma A.3, we get, for $v>c\times\varphi(\sqrt{np_n})/n$

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

Similar to (2.21) we obtain

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

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

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

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

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

where

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

or equivalently

$$\begin{aligned} S_n^{(\varepsilon)}(\alpha,z)\big(\alpha+S_n^{(\varepsilon)}(\alpha,z)\big)^2 & \\ +\big(\alpha+S_n^{(\varepsilon)}(\alpha,z)\big)-|z|^2S_n^{(\varepsilon)}(\alpha,z)&=\widetilde{\delta}_n(\alpha,z),\end{aligned} \tag{2.28}$$

where $\widetilde{\delta}_n(\alpha,z)=\theta\frac{C\varkappa r(\alpha,z)}{\varphi(\sqrt{np_n})v^3}$.

Furthermore, we introduce the notation

$$\begin{aligned} &Q_n^{(\varepsilon)}(\alpha,z):=\big(\alpha+S_n^{(\varepsilon)}(\alpha,z)\big)^2-|z|^2 \quad \text{and}\\ &Q(\alpha,z):=\big(\alpha+S(\alpha,z)\big)^2-|z|^2,\\ &P(\alpha,z):=\alpha+S(\alpha,z) \quad \text{and} \quad P^{(\varepsilon)}(\alpha,z):=\alpha+S_n^{(\varepsilon)}(\alpha,z).\end{aligned} \tag{2.29}$$

We may rewrite the last equation as

$$S_n^{(\varepsilon)}(\alpha,z)=-\frac{P_n^{(\varepsilon)}(\alpha,z)}{Q_n^{(\varepsilon)}(\alpha,z)}+\widehat{\delta}_n(\alpha,z), \tag{2.30}$$

where

$$\widehat{\delta}_n(\alpha,z)=\frac{\widetilde{\delta}_n(\alpha,z)}{Q_n^{(\varepsilon)}(\alpha,z)}. \tag{2.31}$$

Furthermore, we prove the following simple lemma.

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

$$S(\alpha,z)=-\frac{P(\alpha,z)}{Q(\alpha,z)}, \tag{2.32}$$

and $\mathrm{Im}\{S(\alpha,z)\}>0$. *Then the 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:

$$S(\alpha, z) = -\frac{P(\alpha, z)}{Q(\alpha, z)}. \tag{2.33}$$

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

$$\operatorname{Im}\{P(\alpha, z)\} = \operatorname{Im}\{P(\alpha, z)\}\frac{|P(\alpha, z)|^2 + |z|^2}{|Q(\alpha, z)|^2} + v. \tag{2.34}$$

Equations (2.32) and (2.34) together imply

$$\operatorname{Im}\{\alpha + S(\alpha, z)\}\left(1 - \frac{|P(\alpha, z)|^2 + |z|^2}{|Q(\alpha, z)|^2}\right) = v. \tag{2.35}$$

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

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

In particular we have

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

Equality (2.35) and the last remark together imply

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

The proof is complete. □

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

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

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 provides 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

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

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

$$S(\alpha, z) = -\frac{P(\alpha, z)}{Q(\alpha, z)} \tag{2.36}$$

and

$$S_n^{(\varepsilon)}(\alpha, z) = -\frac{P_n^{(\varepsilon)}(\alpha, z)}{Q_n^{(\varepsilon)}(\alpha, z)} + \widehat{\delta}_n(\alpha, z), \tag{2.37}$$

respectively. These equations together imply

$$\begin{aligned} &S(\alpha, z) - S_n^{(\varepsilon)}(\alpha, z) \\ &\quad = \frac{(S(\alpha, z) - S_n^{(\varepsilon)}(\alpha, z))(P_n^{(\varepsilon)}(\alpha, z)P(\alpha, z) + |z|^2)}{Q(\alpha, z)Q_n^{(\varepsilon)}(\alpha, z)} + \widehat{\delta}_n(\alpha, z). \end{aligned} \tag{2.38}$$

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

$$\begin{aligned} \left|1 - \frac{P_n^{(\varepsilon)}(\alpha, z)P(\alpha, z) + |z|^2}{Q(\alpha, z)Q_n^{(\varepsilon)}(\alpha, z)}\right| &\ge \frac{1}{2}\left(1 - \frac{|P_n^{(\varepsilon)}(\alpha, z)|^2 + |z|^2}{|Q_n^{(\varepsilon)}(\alpha, z)|^2}\right) \\ &\quad + \frac{1}{2}\left(1 - \frac{|P(\alpha, z)|^2 + |z|^2}{|Q(\alpha, z)|^2}\right). \end{aligned}$$

The last inequality and Lemmas 2.2 and 2.3 together imply

$$\left|1 - \frac{P_n^{(\varepsilon)}(\alpha, z)P(\alpha, z) + |z|^2}{Q(\alpha, z)Q_n^{\varepsilon)}(\alpha, z)}\right| \ge \frac{v}{4}.$$

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 $v \ge v_0(n) = c(\varphi(\sqrt{np_n}))^{-1/6}$ the inequality*

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

holds.

PROOF. Note that

$$\left|Q_n^{(\varepsilon)}\right| \ge \left|P_n^{(\varepsilon)}(\alpha,z) - |z|\right|\left|P_n^{(\varepsilon)}(\alpha,z) + |z|\right| \ge v^2. \tag{2.39}$$

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 the 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

$$\int_{-\infty}^{\infty} |\widehat{\delta}_n(\alpha,z)|\, du \le \frac{c\varkappa}{v^3\varphi(\sqrt{np_n})}\int_{-\infty}^{\infty}\frac{du}{|Q_n^{(\varepsilon)}(\alpha,z))|}. \tag{2.40}$$

Furthermore, representation (2.30) implies that

$$\frac{1}{|Q_n^{(\varepsilon)}(\alpha,z)|} \le \frac{|S_n^{(\varepsilon)}(\alpha,z)|}{|P_n^{(\varepsilon)}(\alpha,z)|} + \frac{|\widehat{\delta}_n(\alpha,z)|}{|P_n^{(\varepsilon)}(\alpha,z)|}. \tag{2.41}$$

Note that, according to relation (2.27),

$$\frac{1}{|P_n^{(\varepsilon)}(\alpha,z)|} \le \frac{|z|^2|S_n^{(\varepsilon)}(\alpha,z)|}{|P_n^{(\varepsilon)}(\alpha,z)|^2} + \left|S_n^{(\varepsilon)}(\alpha,z)\right| + \frac{|\delta_n(\alpha,z)|}{|P_n^{(\varepsilon)}(\alpha,z)|^2}. \tag{2.42}$$

This inequality implies

$$\begin{aligned}\int_{-\infty}^{\infty}\frac{|S_n^{(\varepsilon)}(\alpha,z)|}{|P_n^{(\varepsilon)}(\alpha,z)|}\,du &\le \frac{C(1+|z|^2)}{v^2}\int_{-\infty}^{\infty}\left|S_n^{(\varepsilon)}(\alpha,z)\right|^2 du \\ &\quad + \int_{-\infty}^{\infty}|\delta_n(\alpha,z)|\frac{|S_n^{(\varepsilon)}(\alpha,z)|}{|P_n^{(\varepsilon)}(\alpha,z)|}\,du.\end{aligned} \tag{2.43}$$

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

$$|\delta_n(\alpha,z)| \le \frac{C\varkappa}{(\varphi(\sqrt{np_n}))v^4} < 1/2. \tag{2.44}$$

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

$$\int_{-\infty}^{\infty} \frac{|S_n^{(\varepsilon)}(\alpha, z)|}{|P_n^{(\varepsilon)}(\alpha, z)|}\,du \le \frac{C(1+|z|^2)}{v^2}\int_{-\infty}^{\infty}\big|S_n^{(\varepsilon)}(\alpha, z)\big|^2\,du \le \frac{C(1+|z|^2)}{v^3}. \tag{2.45}$$

Inequalities (2.42), (2.40) and the definition of $\widehat{\delta}_n(\alpha, z)$ together imply

$$\int_{-\infty}^{\infty} |\widehat{\delta}_n(\alpha, z)|\,du \le \frac{C(1+|z|^2)}{v^6\varphi(\sqrt{np_n})} + \frac{C\varkappa}{v^4\varphi(\sqrt{np_n})}\int_{-\infty}^{\infty}|\widehat{\delta}_n(\alpha, z)|\,du. \tag{2.46}$$

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

$$\int_{-\infty}^{\infty} |\widehat{\delta}_n(\alpha, z)|\,du \le \frac{C(1+|z|^2)}{\varphi(\sqrt{np_n})v^6}. \tag{2.47}$$

$\square$

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

$$\sup_x \big|F_n^{(\varepsilon)}(x, z) - F(x, z)\big| \le C\big(\varphi\big(\sqrt{np_n}\big)\big)^{-1/6}. \tag{2.48}$$

$\square$

3. Properties of the measure $\widetilde{\boldsymbol{\nu}}(\cdot, z)$. In this section we investigate the properties of the measure $\widetilde{\nu}(\cdot, z)$. 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 $v > 0$,

$$\operatorname{Im}\{S(\alpha, z)\} \ge 0$$

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 (2.36) it follows immediately that

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

Set $y = S(x, z) + x$ and consider equation (2.36) 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

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

It is straightforward to check that $\sqrt{3(1-|z|^2)} \le |x_1|$ 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 $|z| \le 1$ equation (3.4) has one real root for $|x| \le |x_1|$ and three real roots for $|x| > |x_1|$. In the case $|z| > 1$ equation (3.4) has one real root for $|x_2| \le x \le |x_1|$ and three real roots for $|x| \le |x_2|$ or for $|x| \ge |x_1|$.*

PROOF. Set

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

We consider the roots of the 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 $|z| \le 1$ and for

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

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

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

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

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

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

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

These relations imply that for $|z| \le 1$ the function $L(y)$ has three real roots for $|x| \ge |x_1|$ 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

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

for $|x| \le |x_2|$ and for $|x| \ge |x_1|$ 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 $|x| \le |x_2|$ or for $|x| \ge |x_1|$ the function $L(y)$ has three real roots. The lemma is proved. $\square$

REMARK 3.1. From Lemma 3.1 it follows that the measure $\widetilde{\nu}(x, z)$ has a density $p(x, z) = \lim_{v\to 0} \operatorname{Im} S(\alpha, z)$ and:

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

Introduce the function

$$g(s,t) := \begin{cases} \dfrac{2s}{s^2+t^2}, & \text{if } s^2+t^2 > 1, \\ 2s, & \text{otherwise}. \end{cases} \tag{3.7}$$

It is well known that for $z = s + it$ the logarithmic potential of uniform distribution on the unit disc is

$$U_0(z) := \iint \ln \frac{1}{|z - x + iy|}\, dG(x, y) = \begin{cases} \dfrac{1}{2}(1 - |z|^2), & \text{if } |z| \le 1, \\ -\ln |z|, & \text{if } |z| > 1, \end{cases} \tag{3.8}$$

and

$$\frac{\partial}{\partial s} \iint \ln \frac{1}{|z - x + iy|}\, dG(x, y) = -\frac{1}{2} g(s, t). \tag{3.9}$$

According to Lemma 4.4 in Bai [1], we have, for $z = s + it$,

$$\frac{\partial}{\partial s}\left(\int_0^\infty \log x\, \nu(dx, z)\right) = \frac{1}{2} g(s, t). \tag{3.10}$$

According to Remark 3.1, we have, for $|z| \ge 1$,

$$\ln(|x_2|/|z|) \le U_{\widetilde{\nu}}(z) + \ln |z| \le \ln(|x_1|/|z|). \tag{3.11}$$

This implies that

$$\lim_{|z|\to\infty} |U_{\widetilde{\nu}}(z) - U_0(z)| = 0. \tag{3.12}$$

Since

$$\int_{-\infty}^{\infty} \log|x| \widetilde{\nu}(dx, z) = \int_0^{\infty} \log x \, \nu(dx, z) \tag{3.13}$$

we get

$$\frac{\partial}{\partial s}\left(\int_{-\infty}^{\infty} \log|x| \widetilde{\nu}(dx, z)\right) = \frac{1}{2} g(s, t). \tag{3.14}$$

Comparing equalities (3.10) and (3.8) and using relation (3.12), we obtain

$$U_0(z) = -\int_0^{\infty} \ln x \, \nu(dx, z) = -\int_{-\infty}^{\infty} \ln|x| \widetilde{\nu}(dx, z) = U_{\mu}(z). \tag{3.15}$$

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 $\varepsilon_{jk} X_{jk}$, $j, k = 1, \dots, n$. Assume that $\mathbf{E} X_{jk} = 0$ and $\mathbf{E} X_{jk}^2 = 1$ and let ε_{jk} denote Bernoulli random variables with $p_n = \Pr\{\varepsilon_{jk} = 1\}$, $j, k = 1, \dots, n$. Denote by $s_1^{(\varepsilon)}(z) \geq \cdots \geq s_n^{(\varepsilon)}(z)$ 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}, $j, k \in \mathbf{N}$, be independent random complex variables with $\mathbf{E} X_{jk} = 0$ and $\mathbf{E}|X_{jk}|^2 = 1$, which are uniformly integrable, that is,*

$$\sup_{j,k} \mathbf{E}|X_{jk}|^2 I_{\{|X_{jk}| > M\}} \to 0 \qquad \text{as } M \to \infty. \tag{4.1}$$

Let ε_{jk}, $j, k = 1, \dots, n$, be independent Bernoulli random variables with $p_n := \Pr\{\varepsilon_{jk} = 1\}$. Assume that ε_{jk} are independent from X_{jk}, $j, k \in \mathbf{N}$, in aggregate. Let $p_n^{-1} = \mathcal{O}(n^{1-\theta})$ for some $0 < \theta \leq 1$. Let $K \geq 1$. 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

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

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 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

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

Simple calculations show that if $np_n \leq \ln n$ for all $n \geq 1$, then

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

Hence in the case $np_n \leq \ln n$ 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 [18], 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 the proof of the Lemma 4.1 below.

REMARK 4.5. We may 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 4.9 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 $0 < \tau \le \gamma_0$*

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

PROOF. Recall that $\mathbf{E}X_{ij} = 0$ and $\mathbf{E}|X_{ij}|^2 = 1$. 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

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

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

$$\begin{aligned}&\Pr\left\{\sum_{j=1}^n \zeta_j^2 < \tau^2 np_n\right\}\\ &\quad\le \exp\{np_n\tau^2 t^2/2\}\prod_{j=1}^n \mathbf{E}_{\xi_j}\prod_{k=1}^n \mathbf{E}_{\varepsilon_{jk}X_{jk}}\exp\{it\xi_j x_k\varepsilon_{jk}X_{jk}\},\end{aligned} \tag{4.8}$$

where ξ_j, $j = 1, \ldots, n$, 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:

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

(see [3], 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

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

Furthermore, we note that

$$\begin{aligned}
&|\mathbf{E}_{\varepsilon_{jk} X_{jk}} \exp\{it\xi_j x_k \varepsilon_{jk} X_{jk}\}| \\
&\quad \le \exp\left\{\frac{1}{2}(|\mathbf{E}_{\varepsilon_{jk} X_{jk}} \exp\{it\xi_j x_k \varepsilon_{jk} X_{jk}\}|^2 - 1)\right\} \\
&\quad \le \exp\left\{-p_n\left((1 - p_n)\big(1 - \operatorname{Re} f_{jk}(t x_k \xi_j)\big)\right.\right. \\
&\qquad\qquad \left.\left. + \frac{p_n}{2}\big(1 - |f_{jk}(t x_k \xi_j)|^2\big)\right)\right\},
\end{aligned} \tag{4.11}$$

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

$$\begin{aligned}
1 - \operatorname{Re} f_{jk}(t x_k \xi_j) &= \mathbf{E}_{X_{jk}}\big(1 - \cos(t x_k X_{jk} \xi_j)\big) \\
&\ge \tfrac{11}{24} t^2 x_k^2 \xi_j^2 \mathbf{E}|X_{jk}|^2 I_{\{|X_{jk}| \le M\}},
\end{aligned} \tag{4.13}$$

and similarly

$$\begin{aligned}
1 - |f_{jk}(t x_k \xi_j)|^2 &= \mathbf{E}_{X_{jk}}\big(1 - \cos(t x_k \widetilde{X}_{jk} \xi_j)\big) \\
&\ge \tfrac{11}{24} t^2 x_k^2 \xi_j^2 \mathbf{E}|\widetilde{X}_{jk}|^2 I_{\{|X_{jk}| \le M\}}.
\end{aligned} \tag{4.14}$$

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

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

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

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

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

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

We may choose C_1 large enough such that following inequalities hold:

$$\mathbf{E}_{\xi_j}(\exp\{-cp_n t^2 \xi_j^2\} | |\xi_j| \le C_1) \le \exp\{-ct^2 p_n/24\} \tag{4.18}$$

for all $|t| \le 1/(MC_1)$. Inequalities (4.8), (4.9), (4.11), (4.18) together imply that for any $\beta \in (0,1)$

$$\begin{aligned} &\Pr\left\{\sum_{j=1}^{n} \zeta_j^2 < \tau^2 np_n\right\} \\ &\qquad \le \exp\{np_n \tau^2 t^2/2\}\left(\exp\{-c\beta nt^2 p_n/24\} + \left(\frac{\beta}{\alpha}\right)^{n\beta/(1-\beta)}\right). \end{aligned} \tag{4.19}$$

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

$$\begin{aligned} &\Pr\left\{\sum_{j=1}^{n} \zeta_j^2 < \tau^2 np_n\right\} \\ &\qquad \le \exp\{np_n \tau^2 t^2/2\}\left(\exp\{-ct^2 np_n/60\} + \left(\frac{1}{2}\right)^{2n/3}\right). \end{aligned} \tag{4.20}$$

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

$$\Pr\left\{\sum_{j=1}^{n} \zeta_j^2 < \tau^2 np_n\right\} \le \exp\{-ct^2 np_n/120\}. \tag{4.21}$$

Inequality (4.21) implies that inequality (4.5) holds with some positive constant $c_0 > 0$. This completes 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

$$\begin{aligned}
&\Pr\{\|(\mathbf{X}^{(\varepsilon)} - z\mathbf{I})\mathbf{x}\|_2 \le \tau\} \\
&\quad \le \exp\{\tau^2 np_n t^2/2\} \\
&\qquad \times \min\Biggl\{\mathbf{E}\exp\Biggl\{-t^2\sum_{j=1}^n\Biggl|\sum_{k=1}^n(\xi_{jk}u_k - \eta_{jk}v_k)\varepsilon_{jk} \\
&\qquad\qquad - \sqrt{np_n}(uu_j - vv_j)\Biggr|^2\Bigg/2\Biggr\}, \\
&\qquad\qquad \mathbf{E}\exp\Biggl\{-t^2\sum_{j=1}^n\Biggl|\sum_{k=1}^n(\xi_{jk}v_k + \eta_{jk}u_k)\varepsilon_{jk} \\
&\qquad\qquad - \sqrt{np_n}(vu_j + uv_j)\Biggr|^2\Bigg/2\Biggr\}\Biggr\}.
\end{aligned} \tag{4.22}$$

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

$$\max\Biggl\{\sum_{k\in A}|x_k|^2, \sum_{k\in A^c}|x_k|^2\Biggr\} \ge 1/2. \tag{4.23}$$

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

$$A_j := \bigl\{k \in \{1, \dots, n\} : \mathbf{E}|\xi_{jk}u_k - \eta_{jk}v_k|^2 \ge |x_k|^2/2\bigr\}. \tag{4.24}$$

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

$$\mathbf{E}|\eta_{jk}u_k + \xi_{jk}v_k|^2 \ge |x_k|^2/2. \tag{4.25}$$

According to inequality (4.23), for any $j = 1, \dots, n$, there exists a set B_j such that

$$\sum_{k\in B_j}|x_k|^2 \ge 1/2 \tag{4.26}$$

and for any $k \in B_j$

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

or

$$\mathbf{E}|\eta_{jk}u_k + \xi_{jk}v_k|^2 \ge |x_k|^2/2. \tag{4.28}$$

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

$$\widetilde{\zeta}_{jk} := \xi_{jk}u_k - \eta_{jk}v_k \tag{4.29}$$

and

$$\widehat{\zeta}_{jk} := \eta_{jk}u_k + \xi_{jk}v_k. \tag{4.30}$$

Inequalities (4.27) and (4.28) together imply that one of the following two inequalities

$$\operatorname{card}\{j : \text{for any } k \in B_j\ \mathbf{E}|\widehat{\zeta}_{jk}|^2 \geq |x_k|^2/2\} \geq n/2 \tag{4.31}$$

or

$$\operatorname{card}\{j : \text{for any } k \in B_j\ \mathbf{E}|\widetilde{\zeta}_{jk}|^2 \geq |x_k|^2/2\} \geq n/2 \tag{4.32}$$

holds. If (4.31) holds we shall bound the first term on the right-hand side of (4.22). 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 $K_n := Kn\sqrt{p_n}$, $\widehat{q}_n := q_n/(\ln(2/p_n)\ln K_n)$ and $\widehat{p}_n := p_n/(\ln(2/p_n)\ln K_n)$. Without loss of generality we shall assume that

$$\ln K_n/|\ln \gamma_0| \geq 1 \quad \text{and} \quad \ln K_n > 1. \tag{4.33}$$

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

$$\Pr\{\|\mathbf{X}^{(\varepsilon)}(z)\mathbf{x}\|_2 \leq \gamma_n \textit{ and } \|\mathbf{X}^{(\varepsilon)}(z)\| \leq K_n\} \leq \exp\{-cnq_n\} \tag{4.34}$$

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

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

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 [17]). By condition (4.34), we have for $\tau \leq \gamma_n$

$$\begin{aligned}\Pr\{&\text{there exists } \mathbf{x} \in \mathcal{N} : \|\mathbf{X}^{(\varepsilon)}(z)\mathbf{x}\|_2 < \tau \text{ and } \|\mathbf{X}^{(\varepsilon)}(z)\| \leq K_n\}\\ &\leq \left(\frac{3}{\eta}\right)^{2k} \exp\{-cnq_n\}.\end{aligned} \tag{4.35}$$

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

$$\|\mathbf{X}^{(\varepsilon)}(z)\mathbf{x}\|_2 \leq \|\mathbf{X}^{(\varepsilon)}(z)\mathbf{y}\|_2 + \|\mathbf{X}^{(\varepsilon)}(z)\|\,\|\mathbf{x} - \mathbf{y}\|_2 \leq \tfrac{1}{2}\tau + K_n\eta = \tau, \tag{4.36}$$

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

$$\Pr(V) \le \left(\left(\frac{3}{\eta}\right)^{2\delta_0/(\ln K_n \ln(2/p_n))} \exp\left\{-\frac{c_0}{4}\right\}\right)^{nq_n}. \tag{4.37}$$

Note that under assumption (4.33) we have

$$\frac{2\ln(3/\eta)}{\ln 2 \ln K_n} \le 10. \tag{4.38}$$

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

Following Rudelson and Vershynin [18], we shall partition the unit sphere $\mathcal{S}^{(n-1)}$ into the two sets of so-called 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

$$\mathrm{Sparse}(\delta), \qquad \mathrm{Comp}(\delta,\rho), \qquad \mathrm{Incomp}(\delta,\rho), \tag{4.39}$$

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 $\mathbf{x} \in \mathcal{C} \subset \mathcal{S}^{(n-1)}$*

$$\Pr\{\|\mathbf{X}^{(\varepsilon)}(z)\mathbf{x}\|_2 \le \gamma_n \text{ and } \|\mathbf{X}^{(\varepsilon)}(z)\| \le K_n\} \le \exp\{-cnq_n\} \tag{4.40}$$

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

$$\begin{aligned} &\Pr\Big\{\inf_{\mathbf{x}\in\mathrm{Comp}(\delta_1\widehat{q}_n,\rho_n)\cap\mathcal{C}} \|\mathbf{X}^{(\varepsilon)}(z)\mathbf{x}\|_2 \le \gamma_n \text{ and } \|\mathbf{X}^{(\varepsilon)}(z)\| \le K_n\Big\} \\ &\quad \le \exp\{-c_1 nq_n\}, \end{aligned} \tag{4.41}$$

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:

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

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

$$\frac{1}{n}\ln\binom{n}{k} \le -C\delta_1 \widehat{q}_n \ln(\delta\widehat{q}_n). \tag{4.42}$$

We may choose δ_1 small enough that

$$\frac{1}{n}\ln\binom{n}{k} \le c_0 q_n/16. \tag{4.43}$$

Thus we get

$$\Pr\Big\{\inf_{\mathbf{x}\in \mathrm{Sparse}(\delta_1 \widehat{p}_n)\cap \mathcal{C}} \|\mathbf{X}^{(\varepsilon)}(z)\mathbf{x}\|_2 \le \tau \text{ and } \|\mathbf{X}^{(\varepsilon)}(z)\| \le K_n\Big\} \le \exp\{-c_1 n q_n\}. \tag{4.44}$$

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 $\mathbf{y} \in \mathrm{Comp}(\delta_1 \widehat{p}_n, \rho K_n^{-1})$. Assume that V occurs and choose a point $\mathbf{x} \in \mathrm{Sparse}(\delta_1 \widehat{p}_n)$ such that $\|\mathbf{y} - \mathbf{x}\|_2 \le \rho K_n^{-1}$. Then

$$\|\mathbf{X}^{(\varepsilon)}(z)\mathbf{x}\|_2 \le \|\mathbf{X}^{(\varepsilon)}(z)\mathbf{y}\|_2 + \|\mathbf{X}^{(\varepsilon)}(z)\|\|\mathbf{x}-\mathbf{y}\|_2 \le \gamma_1 + \rho = \gamma_n/2. \tag{4.45}$$

Hence,

$$\Pr(V) \le \exp\Big\{-\frac{c_0}{8} n q_n\Big\}. \tag{4.46}$$

Thus the lemma is proved. $\square$

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

$$\sum_{k\in\sigma(\mathbf{x})} |x_k|^2 \ge \frac{1}{2}\rho^2 \tag{4.47}$$

and

$$\frac{\rho}{\sqrt{2n}} \le |x_k| \le \frac{1}{\sqrt{n\delta/2}} \qquad \text{for any } k \in \sigma(\mathbf{x}), \tag{4.48}$$

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

PROOF. See proof of Lemma 3.4 [18], page 16. For the reader's convenience we repeat this proof here. Consider the subsets of $\{1,\dots,n\}$ defined by

$$\sigma_1(\mathbf{x}) := \left\{k : |x_k| \le \frac{1}{\sqrt{\delta n/2}}\right\}, \qquad \sigma_2(\mathbf{x}) = \left\{k : |x_k| \ge \frac{\rho}{\sqrt{2n}}\right\} \tag{4.49}$$

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 $\mathbf{y} := P_{\sigma_1(\mathbf{x})^c}\mathbf{x} \in \mathrm{Sparse}(\delta)$, 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

$$\|P_{\sigma(\mathbf{x})}\mathbf{x}\|_2^2 \ge \|P_{\sigma_1(\mathbf{x})}\mathbf{x}\|_2^2 - \|P_{\sigma_2(\mathbf{x})}\mathbf{x}\|_2^2 \ge \rho^2/2. \tag{4.50}$$

Thus the lemma is proved. $\square$

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

$$\frac{\rho}{\sqrt{2n}} \le |x_k| \le \frac{1}{\sqrt{n\delta\widehat{p}_n/2}} \tag{4.51}$$

and

$$\|P_{\sigma(\mathbf{x})}\mathbf{x}\|_2^2 \ge \tfrac{1}{2}\rho^2. \tag{4.52}$$

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 $(X_{jk})_{j,k=1}^n$,

$$p_{\mathbf{x}}(\eta) = \max_{j\in\{1,\dots,n\}}\sup_{u\in\mathbb{C}}\Pr\left\{\left|\sum_{k=1}^n X_{jk}\varepsilon_{jk}x_k - u\right| \le \eta\right\}. \tag{4.53}$$

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 A.7. Let $\mathbf{x} \in \mathrm{Incomp}(\delta_n, \rho_n)$. 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*

$$p_{\mathbf{x}}(\eta\rho_n/\sqrt{2n}) \le 1 - r_2\delta_n n p_n \tag{4.54}$$

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

$$p_{\mathbf{x}}(\eta\rho_n/\sqrt{2n}) \le 1 - r_1 < 1 \tag{4.55}$$

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

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

$$
\begin{aligned}
&\sup_u \Pr\left\{\left|\sum_{k=1}^m X_{jk}\varepsilon_{jk}x_k - u\right| \le \eta\rho_n/\sqrt{2n}\right\}\\
&\qquad \le \Pr\left\{\sum_{k=1}^m \varepsilon_{jk} = 0\right\} \\
&\qquad\quad + \Pr\left\{\left|\sum_{k=1}^m X_{jk}\varepsilon_{jk}x_k - u\right| \le \eta\rho_n/\sqrt{2n}; \sum_{k=1}^m \varepsilon_{jk} \ge 1\right\}.
\end{aligned}
\tag{4.56}
$$

Introduce $\sigma(\mathbf{x}) := \{k \in \{1, \dots, n\} : \rho_n/\sqrt{2n} \le |x_k| \le 1/\sqrt{m/2}\}$. Since $\mathbf{x} \in \mathrm{Incomp}(\delta_n, \rho_n)$ 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

$$
\begin{aligned}
&\sup_u \Pr\left\{\left|\sum_{k=1}^m X_{jk}\varepsilon_{jk}x_k - u\right| \le \eta\rho_n/\sqrt{2n}\right\}\\
&\qquad \le (1 - p_n)^m + Q(\eta)\bigl(1 - (1 - p_n)^m\bigr).
\end{aligned}
\tag{4.57}
$$

According to Lemma A.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

$$
\begin{aligned}
\sup_u \Pr\left\{\left|\sum_{k=1}^m X_{jk}\varepsilon_{jk}x_k - u\right| \le \eta\rho_n/\sqrt{2n}\right\} &\le r_0 + (1 - r_0)e^{-mp_n}\\
&\le 1 - (1 - e^{-1/3})(1 - r_0)\\
&=: 1 - r_1 < 1.
\end{aligned}
\tag{4.58}
$$

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

$$
\begin{aligned}
\sup_u \Pr\left\{\left|\sum_{k=1}^m X_{jk}\varepsilon_{jk}x_k - u\right| \le \eta\rho_n/\sqrt{2n}\right\} &\le 1 - (1 - r_0)mp_n/3\\
&=: 1 - r_2 mp_n.
\end{aligned}
\tag{4.59}
$$

The lemma is proved. □

Now we state a tensorization lemma.

LEMMA 4.5. *Let $\zeta_1, \dots, \zeta_n$ be independent nonnegative random variables. Assume that*

$$
\Pr\{\zeta_j \le \lambda_n\} \le 1 - q_n \tag{4.60}
$$

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*

$$\Pr\left\{\sum_{j=1}^n \zeta_j^2 \le K_1^2 n q_n \lambda_n^2\right\} \le \exp\{-K_2 n q_n\}. \tag{4.61}$$

PROOF. We repeat the proof of Lemma 4.4 in [12]. Let $t = K_1\sqrt{q_n}\lambda_n$. For any $\tau > 0$ we have

$$\Pr\left\{\sum_{j=1}^n \zeta_j^2 \le nt^2\right\} \le e^{n\tau} \prod_{j=1}^n \mathbf{E}\exp\{-\tau\zeta_j^2/t^2\}. \tag{4.62}$$

Furthermore,

$$\begin{aligned}
\mathbf{E}\exp\{-\tau\zeta_j^2/t^2\} &= \int_0^\infty \Pr\{\exp\{-\tau\zeta_j^2/t^2\} > s\}\,ds \\
&= \int_0^1 \Pr\{1/s > \exp\{\tau\zeta_j^2/t^2\}\}\,ds \\
&\le \int_0^{\exp\{-\tau\lambda_n^2/t^2\}} ds + \int_{\exp\{-\tau\lambda_n^2/t^2\}}^1 (1-q_n)\,ds \qquad (4.63)\\
&\le 1 - q_n(1-\exp\{-\tau\lambda_n^2/t^2\}) \\
&= 1 - q_n\bigl(1-\exp\{-\tau/(K_1^2 q_n)\}\bigr).
\end{aligned}$$

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

$$\Pr\left\{\sum_{j=1}^n \zeta_j^2 \le nt^2\right\} \le \exp\{-nq_n/2\}. \tag{4.64}$$

Thus the lemma is proved. $\square$

Recall that we assume $p_n^{-1} = O(n^{1-\theta})$, $1 \ge \theta > 0$. 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 $\widehat{\mathcal{C}}_1 := \mathcal{C}_0 \cap \mathrm{Comp}(\delta_1 p_{n,1}, \rho_{n,1})$, where $p_{n,1} = \widehat{p}_n$, $\rho_{n,1} = \gamma_0/(4K_n)$, and the constants γ_0 and δ_1 are chosen as in Lemmas 4.1 and 4.2, respectively. Then Lemma 4.1 implies inequality (4.40) 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 $\mathbf{x}$ in $\mathcal{C}_0$ lie in $\mathcal{C}_1 := \mathrm{Incomp}(\delta_1 p_{n,1}, \rho_{n,1})$. According to Lemmas 4.4, 4.5 inequality (4.40) holds again 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 sparse ones, that is, $\widehat{\mathcal{C}}_2 := \mathcal{C}_1 \cap \mathrm{Comp}(\delta_2 p_{n,2}, \rho_{n,2})$ and the remaining ones, that is, $\mathcal{C}_2 := \mathcal{C}_1 \cap \mathrm{Incomp}(\delta_2 p_{n,2}, \rho_{n,2})$. Iterating this procedure L times we arrive at the incompressible set $\mathcal{C}_L$ of vectors $\mathbf{x}$ where Lemmas 4.4, 4.5 and Proposition 4.6 yield the required bound of order $\exp\{-\delta n\}$, for a sufficiently small absolute constant $\delta > 0$.

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

$$\mathcal{C}_l := \bigcap_{i=1}^{l} \mathrm{Incomp}(\delta_i p_{n,i}, \rho_{n,i}) \tag{4.65}$$

and

$$\widehat{\mathcal{C}}_l := \mathcal{C}_{l-1} \cap \mathrm{Comp}(\delta_l p_{n,l}, \rho_{n,l}) \qquad \text{with } \mathcal{C}_0 = \mathcal{S}^{(n-1)}. \tag{4.66}$$

Note that

$$\mathcal{S}^{(n-1)} = \bigcup_{l=1}^{L-1} \widehat{\mathcal{C}}_l \cup \mathcal{C}_L. \tag{4.67}$$

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

$$\Pr\{\|\mathbf{X}^{(\varepsilon)}(z)\mathbf{x}\|_2 \le \tau\} \le \exp\{-c_1 n((p_n n \delta_n) \wedge 1)\} \tag{4.68}$$

with

$$\gamma_n := c_2 \rho_n \sqrt{\delta_n}, \tag{4.69}$$

where $a \wedge b$ denotes the minimum of 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 $j = 1, \dots, n$,

$$\sup_{u \in \mathbb{C}} \Pr\left\{\left|\sum_{k=1}^{n} X_{jk}\varepsilon_{jk}x_k - u\right| \le \eta_0 \rho_n/\sqrt{2n}\right\} \le 1 - r_1 \delta_n n p_n. \tag{4.70}$$

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

$$\Pr\{\|\mathbf{X}^{(\varepsilon)}(z)\mathbf{x}\|_2 \le \gamma_n/2 \text{ and } \|\mathbf{X}^{(\varepsilon)}(z)\| \le K_n\} \le \exp\{-cn\delta_n n p_n\}. \tag{4.71}$$

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

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

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

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

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*

$$\begin{aligned}\Pr\Big\{\inf_{\mathbf{x}\in\widehat{\mathcal{C}}_l}\|\mathbf{X}^{(\varepsilon)}(z)\mathbf{x}\|_2 \le \gamma_{n,l} \text{ and } \|\mathbf{X}^{(\varepsilon)}(z)\| \le K_n\Big\}\\ \le \exp\{-\overline{c}_l n\big(((n\widehat{p}_n)^{l-1}p_n)\wedge 1\big)\}\end{aligned} \tag{4.74}$$

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

$$\gamma_{n,l} = \widehat{c}_l\rho_{n,l-1}\sqrt{\delta_{l-1}p_{n,l-1}} \tag{4.75}$$

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

$$\rho_{n,l} := \gamma_{n,l}/(4K_n), \tag{4.76}$$

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

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

$$\gamma_{n,L} \ge cn^{-L/2} \quad \text{and} \quad \rho_{n,L} \ge cn^{-(L+3)/2}. \tag{4.77}$$

PROOF. Note that $p_{n,l}^{-1} = \mathcal{O}(n^{1-l\theta})$. This implies that

$$\gamma_{n,L}^{-1} = \rho_{n,1}^{-1}\mathcal{O}(n^{L-L^2\theta/2}). \tag{4.78}$$

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

$$\gamma_{n,L}^{-1} = \mathcal{O}(n^{L/2}). \tag{4.79}$$ □

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 [18]. From $\mathbf{x} \in \mathcal{C}_l$ it follows that $\mathbf{x} \in \mathrm{Incomp}(\delta_{l-1}p_{n,l-1}, \rho_{n,l-1})$. Applying Lemma 4.6 with $\delta_n = p_{n,l-1}$ and $\rho_n = \rho_{n,l-1}$, we get

$$\begin{aligned}\Pr\{\|\mathbf{X}^{(\varepsilon)}(z)\mathbf{x}\|_2 \le \gamma_{n,l} \text{ and } \|\mathbf{X}^{(\varepsilon)}(z)\| \le K_n\}\\ \le \exp\{-c_1 n\big((np_n\widehat{p}_{n,l-1})\wedge 1\big)\}\end{aligned} \tag{4.80}$$

with

$$\gamma_{n,l} = c_2\rho_{n,l-1}\sqrt{\delta_{l-1}p_{n,l-1}}. \tag{4.81}$$

Inequality (4.80) and Lemma 4.2 together imply

$$\Pr\Big\{\inf_{\mathbf{x}\in\mathcal{C}_l}\|\mathbf{X}^{(\varepsilon)}(z)\mathbf{x}\|_2 \le \gamma_{n,l} \text{ and } \|\mathbf{X}^{(\varepsilon)}(z)\| \le K_n\Big\} \le \exp\{-c_1 n\widehat{p}_{n,l}\} \tag{4.82}$$

with δ_l defined in Lemma 4.2 and

$$\rho_{n,l} := \gamma_{n,l}/(4K_n). \tag{4.83}$$

Thus the lemma is proved. □

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

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

$$L(M) := \max_{n\ge 1}\max_{1\le k\le n} \mathbf{E}|X_k|^2 I_{\{|X_k|>M\}} \to 0 \qquad \text{as } M\to\infty. \tag{4.84}$$

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

$$p_{\mathbf{x}}(\varepsilon\rho_{n,L}/\sqrt{2n}) := \sup_v \Pr\Bigg\{\Bigg|\sum_{k=1}^n x_k\varepsilon_k X_k - v\Bigg| \le \varepsilon\rho_{n,L}/\sqrt{2n}\Bigg\} \le \frac{C\sqrt{\ln n}}{\sqrt{np_n}}. \tag{4.85}$$

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

$$\frac{\rho_{n,L}}{\sqrt{2n}} \le \frac{1}{2^{L_1+1/2}\sqrt{n\delta}} \le \frac{2\rho_{n,L}}{\sqrt{2n}}. \tag{4.86}$$

According to Remark 4.9, we have $\rho_{n,L} \ge cn^{-L/2}$. This implies $L_1 \le C\ln n$. Let $\sigma(\mathbf{x})$ denote the spread set of the vector $\mathbf{x}$, that is,

$$\sigma(\mathbf{x}) := \left\{k : \rho_{n,L}/\sqrt{2n} \le |x_k| \le \sqrt{\frac{2}{n\delta}}\right\}. \tag{4.87}$$

By Lemma 4.3, we have

$$|\sigma(\mathbf{x})| \ge n\delta/2. \tag{4.88}$$

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

$$\Delta_0 := \left\{k : \frac{\rho_{n,L}}{\sqrt{2n}} \le |x_k| \le \frac{1}{2^{L_1+1/2}\sqrt{n\delta}}\right\}, \tag{4.89}$$

$$\Delta_l := \left\{k : \frac{\sqrt{2}}{2^l\sqrt{n\delta}} \le |x_k| \le \frac{\sqrt{2}}{2^{l-1}\sqrt{n\delta}}\right\}, \qquad l = 1, \dots, L_1+1. \tag{4.90}$$

Note that there exists an $l_0 \in \{0, \dots, L_1+1\}$ such that

$$|\Delta_{l_0}| \ge n\delta/\big(2(L_1+2)\big) \ge Cn/\ln n. \tag{4.91}$$

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

$$p_{\mathbf{x}}(\varepsilon\rho_{n,L}/\sqrt{2n}) \le p_{\mathbf{y}}(\varepsilon\rho_{n,L}/\sqrt{2n}) \le p_{\mathbf{y}}(Mb_{l_0}). \tag{4.92}$$

By definition of Δ_{l_0}, we have

$$\sum_{k\in\Delta_{l_0}} |x_k|^2 \ge a_{l_0}^2|\Delta_{l_0}| \ge \rho_{n,L}^2/(2n)|\Delta_{l_0}| \tag{4.93}$$

and

$$\frac{a_{l_0}}{b_{l_0}} \ge \frac{1}{2}. \tag{4.94}$$

Define

$$D(\xi,\lambda) = \lambda^{-2}\mathbf{E}|\xi|^2 I_{\{|\xi|<\lambda\}} \tag{4.95}$$

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:

$$p_{\mathbf{y}}(Mb_{l_0}) \le CMb_{l_0}\Big(\sum_{k\in\Delta_{l_0}} \lambda_k^2 D(\widetilde{\xi_k\varepsilon_k};\lambda_k)\Big)^{-1/2} \tag{4.96}$$

with $\lambda_k \le Mb_{l_0}$. See Petrov [19], page 43, Theorem 3. Put $\lambda_k = M|x_k|$. It is straightforward to check that

$$\sum_{k\in\Delta_{l_0}} \lambda_k^2 D(\widetilde{\xi_k\varepsilon_k};\lambda_k) \ge p_n\Big(\sum_{k\in\Delta_{l_0}} |x_k|^2\big(\mathbf{E}|X_k|^2 - L(M)\big)\Big). \tag{4.97}$$

This implies

$$\sum_{k\in\Delta_{l_0}} \lambda_k^2 D(\widetilde{\xi_k\varepsilon_k};\lambda_k) \ge \frac{p_n}{2}\sum_{k\in\Delta_{l_0}} |x_k|^2 \ge \frac{p_n}{2}|\Delta_{l_0}|a_{l_0}^2. \tag{4.98}$$

Combining this inequality with (4.96) and (4.92) we obtain

$$p_{\mathbf{x}}(\varepsilon\rho_{n,L}/\sqrt{2n}) \le \frac{CMb_{l_0}}{\sqrt{|\Delta_{l_0}|p_n}a_{l_0}} \le \frac{CM}{\sqrt{|\Delta_{l_0}|p_n}} \le \frac{C\sqrt{\ln n}}{\sqrt{np_n}}. \tag{4.99}$$

The last relation concludes the proof. $\square$

Invertibility for the incompressible vectors via distance.

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

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

PROOF. Note that

$$\Pr\Big\{\inf_{\mathbf{x}\in\widehat{\mathcal{C}}_L} \|\mathbf{X}^{(\varepsilon)}(z)\mathbf{x}\|_2 < \eta\big(\rho_{n,L}/\sqrt{n}\big)^2/\sqrt{np_n}\Big\}$$
$$\le \Pr\Big\{\inf_{\mathbf{x}\in\operatorname{Incomp}(\delta_L,\rho_{n,L})} \|\mathbf{X}^{(\varepsilon)}(z)\mathbf{x}\|_2 < \eta\big(\rho_{n,L}/\sqrt{n}\big)^2/\sqrt{np_n}\Big\}. \tag{4.100}$$

For the upper bound of the r.h.s. of (4.100) (see [18], proof of Lemma 3.5). For the reader's 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 kth. Writing $\mathbf{G}\mathbf{x} = \sum_{k=1}^n x_k\mathbf{X}_k$, we have

$$\|\mathbf{G}\mathbf{x}\| \ge \max_{k=1,\dots,n} \operatorname{dist}(x_k\mathbf{X}_k, \mathcal{H}_k) = \max_{k=1,\dots,n} |x_k| \operatorname{dist}(\mathbf{X}_k, \mathcal{H}_k). \tag{4.101}$$

Put

$$p_k := \Pr\big\{\operatorname{dist}(\mathbf{X}_k, \mathcal{H}_k) < \eta\rho_{n,L}/\sqrt{n}\big\}. \tag{4.102}$$

Then

$$\mathbf{E}\big|\{k : \operatorname{dist}(\mathbf{X}_k, \mathcal{H}_k) < \eta\rho_{n,L}/\sqrt{n}\}\big| = \sum_{k=1}^n p_k. \tag{4.103}$$

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

$$\Pr\{U^c\} \le \frac{1}{n\delta_L}\sum_{k=1}^n p_k. \tag{4.104}$$

On the other hand, for every incompressible vector $\mathbf{x}$, the set $\sigma_2(\mathbf{x}) := \{k : |x_k| \ge \rho_{n,L}/\sqrt{n}\}$ contains at least $n\delta_L$ elements. (Otherwise, since $\|P_{\sigma_2(\mathbf{x})^c}\mathbf{x}\|_2 \le \rho_{n,L}$, we have $\|\mathbf{x}-\mathbf{y}\|_2 \le \rho_{n,L}$ 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 (4.101) and by definitions of the sets σ_1 and $\sigma_2(\mathbf{x})$, we have

$$\|\mathbf{G}\mathbf{x}\|_2 \ge |x_k| \operatorname{dist}(\mathbf{X}_k, \mathcal{H}_k) \ge \eta(\rho_{n,L} n^{-1/2})^2. \tag{4.105}$$

Summarizing we have shown that

$$\Pr\Big\{\inf_{\mathbf{x}\in \operatorname{Incomp}(\delta_L,\rho_{n,L})} \|\mathbf{G}\mathbf{x}\|_2 \le \eta(\rho_{n,L} n^{-1/2})^2\Big\} \le \Pr\{U^c\} \le \frac{1}{n\delta_L}\sum_{k=1}^n p_k. \tag{4.106}$$

This completes the proof. $\square$

We now reformulate Lemma 3.6 from [18]. Let $\mathbf{X}_n^*$ 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 δ_L, $\rho_L, \overline{c}_L$ as in Lemma 4.7. Then there exists an absolute constant $\widehat{c}_L > 0$ such that*

$$\Pr\{\mathbf{X}^* \notin \mathcal{C}_L \text{ and } \|\mathbf{X}^{(\varepsilon)}(z)\| \le K_n\} \le \exp\{-\widehat{c}_L n p_n\}. \tag{4.107}$$

PROOF. Note that

$$\mathcal{S}^{(n-1)} = \bigcup_{l=1}^{L-1} \widehat{\mathcal{C}}_l \cup \mathcal{C}_L. \tag{4.108}$$

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

$$\mathcal{E} := \Big\{\inf_{\mathbf{x}\in\bigcup_{l=1}^{L-1}\widehat{\mathcal{C}}_l : \|\mathbf{x}\|_2=1} \|\mathbf{X}^{(\varepsilon)}(z)\mathbf{x}\|_2 \le c \text{ and } \|\mathbf{X}^{(\varepsilon)}(z)\| \le K_n\Big\} \tag{4.109}$$

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

$$\Pr\{\mathbf{X}^* \notin \mathcal{C}_L \text{ and } \|\mathbf{X}^{(\varepsilon)}(z)\| \le K_n\} \tag{4.110}$$

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

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

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

$$\Pr\{\operatorname{dist}(\mathbf{X}_n, \mathcal{H}_n) < \rho_{n,L}/\sqrt{n} \text{ and } \|\mathbf{X}^{(\varepsilon)}(z)\| \le K_n\} \le \frac{C\sqrt{\ln n}}{\sqrt{np_n}}. \tag{4.112}$$

PROOF. We repeat Rudelson and Vershynin's proof of Lemma 3.8 in [18]. Let $\mathbf{X}^*$ be any unit vector orthogonal to $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_{n-1}$. We can choose $\mathbf{X}^*$ so that it is a random vector that depends on $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_{n-1}$ only and is independent of $\mathbf{X}_n$. We have

$$\operatorname{dist}(\mathbf{X}_n, \mathcal{H}_n) \ge |\langle \mathbf{X}_n, \mathbf{X}^*\rangle|.$$

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

$$\begin{aligned}
&\Pr\{\operatorname{dist}(\mathbf{X}_n, \mathcal{H}_n) < \rho_{n,L}/\sqrt{n} \text{ and } \|\mathbf{X}^{(\varepsilon)}(z)\| \le K_n\} \\
&\quad \le \mathbf{E}_{1,\dots,n-1}\Pr_n\{|\langle \mathbf{X}^*, \mathbf{X}_n\rangle| \le \rho_{n,L}/\sqrt{n} \text{ and } \mathbf{X}^* \in \mathcal{C}_L\} \\
&\qquad + \Pr\{\mathbf{X}^* \notin \mathcal{C}_L \text{ and } \|\mathbf{X}^{(\varepsilon)}(z)\| \le K_n\}.
\end{aligned} \tag{4.113}$$

According to Lemma 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 $\mathbf{X}^* = (a_1, \dots, a_n) \in \mathcal{S}^{(n-1)}$ and $\mathbf{X}_n = (\varepsilon_1\xi_1, \dots, \varepsilon_n\xi_n)$ are independent, we may use small ball probability estimates. We have

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

Let σ denote the spread set 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 functions, we get

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

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

$$\Pr_n\{|\langle \mathbf{X}_n, \mathbf{X}^*\rangle| \le \rho_{n,L}/\sqrt{n}\} \le \frac{C\sqrt{\ln n}}{\sqrt{np_n}}. \tag{4.114}$$

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 $\mathbf{X}_1, \dots, \mathbf{X}_n$ 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*

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

PROOF. Note that

$$
\begin{aligned}
&\Pr\Big\{\inf_{\mathbf{x}\in\mathcal{C}_L}\|\mathbf{X}^{(\varepsilon)}(z)\mathbf{x}\|_2<\rho_{n,L}^2/n\Big\}\\
&\qquad\le\Pr\Big\{\inf_{\mathbf{x}\in\mathcal{C}_L}\|\mathbf{X}^{(\varepsilon)}(z)\mathbf{x}\|_2<\rho_{n,L}^2/n \text{ and } \|\mathbf{X}^{(\varepsilon)}(z)\|\le K_n\Big\}\\
&\qquad\quad+\Pr\{\|\mathbf{X}^{(\varepsilon)}(z)\|>K_n\}.
\end{aligned}\tag{4.115}
$$

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

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

Applying Lemma 4.11, we obtain

$$
\Pr\Big\{\inf_{\mathbf{x}\in\mathcal{C}_L}\|\mathbf{X}^{(\varepsilon)}(z)\mathbf{x}\|_2<\rho_{n,L}^2/n\Big\}\le\frac{C\sqrt{\ln n}}{\sqrt{np_n}}.\tag{4.116}
$$

Thus the lemma is proved. $\square$

PROOF OF THEOREM 4.1. By definition of the minimal singular value, we have

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

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

$$
\begin{aligned}
&\Pr\{s_n^{(\varepsilon)}(z)\le\rho_{n,L}^2/n \text{ and } s_1^{(\varepsilon)}(z)\le K_n\}\\
&\qquad\le\sum_{l=1}^{L-1}\Pr\Big\{\inf_{\mathbf{x}\in\widehat{\mathcal{C}}_l}\|\mathbf{X}^{(\varepsilon)}(z)\mathbf{x}\|_2\le\rho_{n,L}^2/n \text{ and } s_1^{(\varepsilon)}(z)\le K_n\Big\}\\
&\qquad\quad+\Pr\Big\{\inf_{\mathbf{x}\in\mathcal{C}_L}\|\mathbf{X}^{(\varepsilon)}(z)\mathbf{x}\|_2\le\rho_{n,L}^2/n \text{ and } s_1^{(\varepsilon)}(z)\le K_n\Big\}.
\end{aligned}\tag{4.117}
$$

According to Lemma 4.7, we have

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

Lemmas 4.12 and 4.7 together imply that

$$
\begin{aligned}
&\Pr\Big\{\inf_{\mathbf{x}\in\mathcal{C}_L}\|\mathbf{X}^{(\varepsilon)}(z)\mathbf{x}\|_2\le\rho_{n,L}^2/n \text{ and } s_1^{(\varepsilon)}(z)\le K_n\Big\}\\
&\qquad\le\Pr\Big\{\inf_{\mathbf{x}\in\operatorname{Incomp}(\delta_L,\rho_{n,L})}\|\mathbf{X}^{(\varepsilon)}(z)\mathbf{x}\|_2\le\rho_{n,L}^2/n \text{ and } s_1^{(\varepsilon)}(z)\le K_n\Big\}\\
&\qquad\le\frac{C\sqrt{\ln n}}{\sqrt{np_n}}+\exp\{-\widehat{c}_L n\}.
\end{aligned}\tag{4.118}
$$

The last two inequalities together imply the result. $\square$

REMARK 4.9. 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 put $L = \ln n$. Then the value L_1 in Lemma 4.8 is at most $C(\ln n)^2$, and hence we get the bound $C \ln n/\sqrt{np_n}$ in (4.85). This yields the bound $C \ln n/\sqrt{np_n} + \exp\{-\widehat{c}_L n\}$ in (4.118). 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 be defined as in Lemma A.2 with $q = 18$. Using the notation 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 A.1

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

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

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

According to Lemma A.2 with $q = 18$, we have

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

These inequalities imply

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

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

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

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*

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

we have

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

PROOF. By definition, we have

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

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

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

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

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

where

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

Note that

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

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 $b(1+\tau) < 2$,

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

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 $|\log x|^{1+\tau} \le \varepsilon^2 \times |\log\varepsilon|^{1+\tau}x^2$ for $x \ge \frac{1}{\varepsilon}$ and sufficiently small ε. Using this inequality, we obtain

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

Inequalities (5.6)–(5.9) together imply that

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

Furthermore, inequalities (5.2), (5.4), (5.5) and (5.10) together imply

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

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

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

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

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

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

where $\overline{F}_n^{(\varepsilon)}(\cdot,z,r)$ is the distribution function corresponding to the restriction of the measure $\nu_n^{(\varepsilon)}(\cdot,z,r)$ to 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

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

This implies that

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

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

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

Hence, we get

$$\begin{aligned}&\sup_x\bigl|\mathbf{E}F_n^{(\varepsilon)}(x,z,r) - F(x,z)\bigr|\\ &\quad \le \sup_x\bigl|\mathbf{E}F_n^{(\varepsilon)}(x,z) - F(x,z)\bigr| + \sup_x|F(x+r,z) - F(x,z)|.\end{aligned}\tag{5.16}$$

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

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

Choose $r = \frac{1}{\sqrt{np_n}}$. Inequalities (5.17) and (2.48) together imply

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

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

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

Note that

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

Let $\mathcal{K} = \{z \in \mathbb{C} : |z| \le R\}$ and let $\mathcal{K}^c$ denote $\mathbb{C} \setminus \mathcal{K}$. According to Lemma A.2 with $q = 18$, we have, for k_1 and R from Lemma A.2,

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

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

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

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

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

Similar to the proof of Lemma 5.1 we show that

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

This implies that

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

for all $z \in \mathbb{C}$. According to equality (3.15), $U_\mu(z)$ is equal to the potential of uniform distribution on the unit disc. This implies that the measure μ coincides with the uniform distribution on the unit disc. Since the measures $\overline{\mu}_n^{(r)}$ are compactly supported, Theorem 6.9 from [14] and Corollary 2.2 from [14] together imply that

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

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.

APPENDIX

In this appendix we collect some technical results.

The largest singular value. Recall that $|\lambda_1^{(\varepsilon)}| \ge \cdots \ge |\lambda_n^{(\varepsilon)}|$ denote the eigenvalues of the matrix $\mathbf{X}^{(\varepsilon)}$ ordered via decreasing absolute values, and let $s_1^{(\varepsilon)} \ge \cdots \ge s_n^{(\varepsilon)}$ denote the singular values of the matrix $\mathbf{X}^{(\varepsilon)}$.

We show the following:

LEMMA A.1. *Under condition of Theorem 1.1 for sufficiently large $K \ge 1$ we have*

$$\Pr\{s_1^{(\varepsilon)} \ge n\sqrt{p_n}\} \le C/np_n \tag{A.1}$$

for some positive constant $C > 0$.

PROOF. Using Chebyshev's inequality, we get

$$\Pr\{s_1^{(\varepsilon)} \ge n\sqrt{p_n}\} \le \frac{1}{n^2 p_n} \mathbf{E}\operatorname{Tr}\bigl(\mathbf{X}^{(\varepsilon)}(\mathbf{X}^{(\varepsilon)})^*\bigr) \le 1/(np_n). \tag{A.2}$$

Thus the lemma is proved. $\square$

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

$$\Pr\{|\lambda_{k_1}^{(\varepsilon)}| > R\} \le (\varphi(np_n))^{-(q-6)/(12q)}, \tag{A.3}$$

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

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

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

On the other hand,

$$\begin{aligned} \Pr\{|\lambda_{k_1}^{(\varepsilon)}| > R\} &\le \Pr\left\{\prod_{\nu=1}^{k_1} |\lambda_\nu^{(\varepsilon)}| > R^{k_1}\right\} \\ &\le \Pr\left\{\prod_{\nu=1}^{k_1} s_\nu^{(\varepsilon)} > R^{k_1}\right\} \le \Pr\left\{\frac{1}{k_1}\sum_{\nu=1}^{k_1} \ln s_\nu^{(\varepsilon)} > \ln R\right\}. \end{aligned} \tag{A.4}$$

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

$$\begin{aligned}
\Pr\left\{\frac{1}{k_1}\sum_{\nu=1}^{k_1}\ln s_\nu^{(\varepsilon)} > \ln R_1\right\}
&\le \Pr\{s_{k_0}^{(\varepsilon)} > R\} + \Pr\left\{\frac{k_0}{k_1}\ln s_1^{(\varepsilon)} + \ln R > \ln R_1\right\}\\
&\le \Delta_n^{(q-6)/(2q)} + \Pr\left\{\ln s_1^{(\varepsilon)} > \frac{k_1}{k_0}\ln\frac{R_1}{R}\right\}.
\end{aligned}$$

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

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

Taking into account Lemma A.1 and inequality (2.48) we obtain

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

for some positive constant $C > 0$, thus proving the lemma. $\square$

LEMMA A.3. *Let $\varkappa = \max_{j,k}\mathbf{E}|X_{jk}|^2\varphi(X_{jk})$. The following inequality holds*:

$$\frac{1}{n\sqrt{np_n}}\sum_{j,k=1}^{n}\mathbf{E}\varepsilon_{jk}|X_{jk}|\big(|T_{k+n,j}^{(jk)}| + |T_{j,k+n}^{(jk)}|\big) \le \frac{C}{v^3\varphi(\sqrt{np_n})}. \tag{A.5}$$

PROOF. Introduce the notation

$$B := \frac{1}{n\sqrt{np_n}}\sum_{j,k=1}^{n}\mathbf{E}\varepsilon_jk|X_{jk}|\big(|T_{k+n,j}^{(jk)}| + |T_{j,k+n}^{(jk)}|\big) \tag{A.6}$$

and

$$\begin{aligned}
B_1 &:= \frac{2}{n^2p_n}\sum_{j,k=1}^{n}\mathbf{E}\varepsilon_{jk}|X_{jk}|^2|R_{k+n,j}^{(jk)}||R_{k+n,j}^{(jk)} - R_{k+n,j}|,\\
B_2 &:= \frac{2}{n^2p_n}\sum_{j,k=1}^{n}\mathbf{E}\varepsilon_{jk}|X_{jk}|^2|R_{k+n,k+n}^{(jk)}||R_{j,j}^{(jk)} - R_{j,j}|,\\
B_3 &:= \frac{2}{n^2p_n}\sum_{j,k=1}^{n}\mathbf{E}\varepsilon_{jk}|X_{jk}|^2|R_{j,j}^{(jk)}||R_{k+n,k+n}^{(jk)} - R_{k+n,k+n}|,\\
B_4 &:= \frac{2}{n^2p_n}\sum_{j,k=1}^{n}\mathbf{E}\varepsilon_{jk}|X_{jk}|^2|R_{j,k+n}^{(jk)}||R_{j,k+n}^{(jk)} - R_{j,k+n}|.
\end{aligned} \tag{A.7}$$

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

$$\big|R_{l,m}^{(jk)} - R_{l,m}\big| \le \frac{1}{v} I_{\{|X_{jk}|>\sqrt{np_n}\}} + \frac{1}{v^2\varphi(\sqrt{np_n})}\varphi(X_{jk}). \tag{A.8}$$

It is easy to check that

$$\max\{B_k, k=1,\dots,8\} \le \frac{C\varkappa}{v^3\varphi(\sqrt{np_n})}. \tag{A.9}$$

This implies that

$$B \le \frac{C\varkappa}{v^3\varphi(\sqrt{np_n})}. \tag{A.10}$$

□

LEMMA A.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, that is,*

$$\mathbf{E}\mu_n^{(r)} = (\mathbf{E}\mu_n) * (\nu_r). \tag{A.11}$$

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

$$\mu_n = \frac{1}{n}\sum_{j=1}^n \delta_{\lambda_j} \tag{A.12}$$

and

$$\mu_n^{(r)} = \frac{1}{n}\sum_{j=1}^n \delta_{\lambda_j + r\xi}. \tag{A.13}$$

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

$$\mathbf{E}\mu_n = \frac{1}{n}\sum_{j=1}^n \mu_{nj} \tag{A.14}$$

and

$$\mathbf{E}\mu_n^r = \frac{1}{n}\sum_{j=1}^n \mu_{nj} * \nu_r = \left(\frac{1}{n}\sum_{j=1}^n \mu_{nj}\right) * (\nu_r) = (\mathbf{E}\mu_n) * (\nu_r). \tag{A.15}$$

Thus the lemma is proved. □

Let

$$f_n^{(r)}(t,v) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \exp\{itx + ivy\}\, dG_n^{(r)}(x,y) \tag{A.16}$$

and

$$f_n(t,v) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \exp\{itx + ivy\}\, dG_n(x,y), \tag{A.17}$$

where

$$G_n^{(r)}(x,y) = \frac{1}{n}\sum_{j=1}^{n} \Pr\{\operatorname{Re}\lambda_j + r\xi \le x, \operatorname{Im}\lambda_j + r\xi \le y\} \tag{A.18}$$

and

$$G_n(x,y) = \frac{1}{n}\sum_{j=1}^{n} \Pr\{\operatorname{Re}\lambda_j \le x, \operatorname{Im}\lambda_j \le y\}. \tag{A.19}$$

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

$$h(t,v) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \exp\{iux + ivy\}\, dG(x,y). \tag{A.20}$$

LEMMA A.5. *The following relations hold*

$$f_n^{(r)}(t,v) = f_n(t,v)h(rt,rv). \tag{A.21}$$

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

$$\begin{aligned}\lim_{r\to 0}\lim_{n\to\infty} f_n^{(r)}(t,v) &= \lim_{n\to\infty}\lim_{r\to 0} f_n^{(r)}(t,v)\\ &= \lim_{n\to\infty} f_n(t,v).\end{aligned} \tag{A.22}$$

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 A.6 ([9], Lemma 2.1). *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

$$
\begin{aligned}
\sup_x |F(x) - G(x)| \le C_1 \sup_{x \in J} \int_{-\infty}^{x} |S_F(u + iV) - S_G(u + iV)|\, du \\
+ \sup_{u \in J} \int_{v_0}^{V} |S_F(u + iv) - S_G(u + iv)|\, dv + C_3 v_0.
\end{aligned}
\tag{A.23}
$$

LEMMA A.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*

$$\max_{j,k} \mathbf{E}|X_{jk}|^2 I_{\{|X_{jk}| > M\}} \to 0 \qquad \text{for } M \to +\infty.$$

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

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

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

$$
\begin{aligned}
\Pr\{|X_{jk} - u| \ge \eta_0\} \ge \Pr\{|X_{jk}| \le M\} \ge \frac{1}{M^2} \mathbf{E}\big(|X_{jk}|^2 I_{\{|X_{jk}| \le M\}}\big) \\
> \frac{7}{8M^2}.
\end{aligned}
\tag{A.24}
$$

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

$$
\begin{aligned}
\Pr\{|X_{jk} - u| \ge \eta_0\} &\ge \mathbf{E}\big(I_{\{2M+\eta_0 \ge |X_{jk} - u| \ge \eta_0\}}\big) \\
&\ge \frac{1}{4M^2} \mathbf{E}\big(|X_{jk} - u|^2 I_{\{2M+\eta_0 \ge |X_{jk} - u| \ge \eta_0\}}\big) \\
&\ge \frac{1}{4M^2} \big(1 - \mathbf{E}\big(|X_{jk} - u|^2 I_{\{|X_{jk} - u| < \eta_0\}}\big) \\
&\qquad - \mathbf{E}\big(|X_{jk} - u|^2 I_{\{|X_{jk} - u| > 2M + \eta_0\}}\big)\big) \\
&\ge \frac{1}{4M^2} \big(1 - \eta_0^2 - \mathbf{E}\big(|X_{jk} - u|^2 I_{\{|X_{jk}| > M\}}\big)\big) \\
&\ge \frac{1}{4M^2} \left(\frac{3}{4} - \eta_0^2 - \frac{|u|^2}{4M^2}\right) \\
&\ge \frac{1}{16M^2} \left(3 - 4\eta_0^2 - \left(1 + \frac{\eta_0^2}{M}\right)^2\right).
\end{aligned}
\tag{A.25}
$$

Combining inequalities (A.24) and (A.25) we obtain the claim. $\square$

Acknowledgments. The authors would like to thank Terence Tao for drawing their attention to a gap in a previous version of the paper and Dmitry Timushev for a careful reading of this manuscript.

REFERENCES

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

[2] BAI, Z. D. and SILVERSTEIN, J. (2006). *Spectral Analysis of Large Dimensional Random Matrices. Mathematics Monograph Series* **2**. Sciences Press, Beijing.

[3] BICKEL, P. J., GÖTZE, F. and VAN ZWET, W. R. (1986). The Edgeworth expansion for U-statistics of degree two. *Ann. Statist.* **14** 1463–1484. MR868312

[4] EDELMAN, A. (1997). The probability that a random real Gaussian matrix has k real eigenvalues, related distributions, and the circular law. *J. Multivariate Anal.* **60** 203–232. MR1437734

[5] FRIEDLAND, S., RIDER, B. and ZEITOUNI, O. (2004). Concentration of permanent estimators for certain large matrices. *Ann. Appl. Probab.* **14** 1559–1576. MR2071434

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

[7] GINIBRE, J. (1965). Statistical ensembles of complex, quaternion, and real matrices. *J. Math. Phys.* **6** 440–449. MR0173726

[8] GOHBERG, I. C. and KREIN, M. G. (1991). *Introduction to the Theory of Linear Operator*. Cambridge Univ. Press, New York.

[9] GÖTZE, F. and TIKHOMIROV, A. (2003). Rate of convergence to the semi-circular law. *Probab. Theory Related Fields* **127** 228–276. MR2013983

[10] GÖTZE, F. and TIKHOMIROV, A. On the circular law. Available at http://arxiv.org/abs/math/0702386.

[11] HORN, R. A. and JOHNSON, C. R. (1990). *Matrix Analysis*. Cambridge Univ. Press, Cambridge. MR1084815

[12] LITVAK, A. E., PAJOR, A., RUDELSON, M. and TOMCZAK-JAEGERMANN, N. (2005). Smallest singular value of random matrices and geometry of random polytopes. *Adv. Math.* **195** 491–523. MR2146352

[13] PAN, G. and ZHOU, W. (2010). Circular law, extreme singular values and potential theory. *J. Multivariate Anal.* **101** 645–656. MR2575411

[14] SAFF, E. B. and TOTIK, V. (1997). *Logarithmic Potentials with External Fields. Grundlehren der Mathematischen Wissenschaften* [*Fundamental Principles of Mathematical Sciences*] **316**. Springer, Berlin. MR1485778

[15] MEHTA, M. L. (1991). *Random Matrices*, 2nd ed. Academic Press, Boston, MA. MR1083764

[16] PASTUR, L. A. (1973). Spectra of random selfadjoint operators. *Uspehi Mat. Nauk* **28** 3–64. MR0406251

[17] RUDELSON, M. (2008). Invertibility of random matrices: Norm of the inverse. *Ann. of Math.* (2) **168** 575–600. MR2434885

[18] RUDELSON, M. and VERSHYNIN, R. (2008). The Littlewood–Offord problem and invertibility of random matrices. *Adv. Math.* **218** 600–633. MR2407948

[19] PETROV, V. V. (1975). *Sums of Independent Random Variables*. Springer, New York. MR0388499

[20] RIDER, B. and VIRÁG, B. (2007). The noise in the circular law and the Gaussian free field. *Int. Math. Res. Not. IMRN* **2** 33. MR2361453

[21] RIDER, B. (2003). A limit theorem at the edge of a non-Hermitian random matrix ensemble: Random matrix theory. *J. Phys. A* **36** 3401–3409. MR1986426

[22] TAO, T. and VU, V. (2008). Random matrices: The circular law. *Commun. Contemp. Math.* **10** 261–307. MR2409368

[23] TAO, T. and VU, V. H. (2009). Inverse Littlewood–Offord theorems and the condition number of random discrete matrices. *Ann. of Math.* (2) **169** 595–632. MR2480613

[24] TIMME, M., WOLF, F. and GEISEL, T. (2004). Topological speed limits to network synchronization. *Phys. Rev. Lett.* **92** 074101-1–4.

FACULTY OF MATHEMATICS
UNIVERSITY OF BIELEFELD
BIELEFELD
GERMANY
E-MAIL: goetze@math.uni-bielefeld.de

FACULTY OF MATHEMATICS AND MECHANICS
SANKT-PETERBURG STATE UNIVERSITY
ST. PETERSBURG
RUSSIA
E-MAIL: tichomir@math.uni-bielefeld.de