ALEA, Lat. Am. J. Probab. Math. Stat. 18, 1047–1059 (2021)
DOI: 10.30757/ALEA.v18-39

A note on the universality of ESDs of inhomogeneous random matrices

Vishesh Jain and Sandeep Silwal

Stanford University
Department of Statistics,
Stanford CA 94305, USA.
E-mail address: visheshj@stanford.edu

Massachusetts Institute of Technology
Department of Computer Science,
Cambridge MA 02139, USA.
E-mail address: silwal@mit.edu

Abstract. In this short note, we extend the celebrated results of Tao and Vu, and Krishnapur on the universality of empirical spectral distributions to a wide class of *inhomogeneous* complex random matrices, by showing that a technical and hard-to-verify Fourier domination assumption may be replaced simply by a natural uniform anti-concentration assumption.

Along the way, we show that inhomogeneous complex random matrices, whose expected squared Hilbert-Schmidt norm is quadratic in the dimension, and whose entries (after symmetrization) are uniformly anti-concentrated at 0 and infinity, typically have smallest singular value $\Omega(n^{-1/2})$. The rate $n^{-1/2}$ is sharp, and closes a gap in the literature.

Our proofs closely follow recent works of Livshyts, and Livshyts, Tikhomirov, and Vershynin on inhomogeneous *real* random matrices. The new ingredient is an anti-concentration inequality for sums of independent, but not necessarily identically distributed, complex random variables, which may also be useful in other contexts.

1. Introduction

1.1. *The least singular value of inhomogeneous complex random matrices.* The (ordered) *singular values* of an $n \times n$ complex matrix A_n, denoted by $s_k(A_n)$ for $k \in [n]$, are defined to be the eigenvalues of $\sqrt{A_n^\dagger A_n}$ arranged in non-decreasing order. Recall that the extreme singular values $s_1(A_n)$ and $s_n(A_n)$ admit the following variational characterization:

$$s_1(A_n) := \sup_{x \in \mathbb{S}_{\mathbb{C}}^{n-1}} \|A_n x\|_2, \quad s_n(A_n) := \inf_{x \in \mathbb{S}_{\mathbb{C}}^{n-1}} \|A_n x\|_2, \tag{1.1}$$

Received by the editors June 11th, 2020-06-11; accepted February 26th, 2021.
2010 *Mathematics Subject Classification.* 60B20.
Key words and phrases. random matrices, universality, circular law.

where $\|\cdot\|_2$ denotes the standard Euclidean norm in $\mathbb{C}^n$ and $\mathbb{S}^{n-1}_{\mathbb{C}}$ denotes standard unit sphere in $\mathbb{C}^n$. In this short note, we will primarily be concerned with the non-asymptotic study of the smallest singular value $s_n(A_n)$ (for quite general random matrices A_n) – a subject which has its origins in numerical linear algebra, and which has attracted much attention in recent years (see, for instance, the references in Livshyts et al., 2021).

When the entries of A_n are i.i.d. complex Gaussians, Edelman (1988) showed that for any $\epsilon > 0$,

$$\Pr\left(s_n(A_n) \le \epsilon n^{-1/2}\right) \le \epsilon^2;$$

in particular, this shows that for any $\delta > 0$, with probability at least $1-\delta$, $s_n(A_n) = \Omega_\delta(n^{-1/2})$. In other words, the smallest singular value of a 'typical realization' of an i.i.d. complex Gaussian matrix is at least order $n^{-1/2}$ (which is known to be optimal).

As our first main result, we establish the optimal order of $s_n(A_n)$ for a typical realization of A_n for very general ensembles of random matrices – this is a complex analogue of a recent theorem of Livshyts (2021+) (see the discussion below).

Theorem 1.1. *Let A_n be an $n \times n$ complex random matrix whose entries $A_{i,j}$ are independent and satisfy the following two conditions:*

- $\sum_{i,j} \mathbb{E}|A_{i,j}|^2 \le Kn^2$ *for some $K > 0$, and*
- $\Pr\left(b^{-1} \ge |\widetilde{A_{i,j}}| \ge b\right) \ge b$ *for some $b \in (0,1)$ (here, $\widetilde{A_{i,j}}$ denotes the difference of two independent copies of $A_{i,j}$).*

Then, for all $\epsilon \in [0,1)$,

$$\Pr\left(s_n(A_n) \le \frac{\epsilon}{\sqrt{n}}\right) \le C\left(\epsilon + \exp(-c\epsilon^2 n)\right),$$

where C, c depend only on K and b.

Remark 1.2. Due to the presence of the term $\exp(-c\epsilon^2 n)$, we require that $\epsilon = \Omega(n^{-1/2})$ in order to obtain a non-trivial upper bound on the probability. The presence of this restriction is similar to the restriction in Rudelson (2008) and stems from the lack of consideration of the arithmetic structure of random hyperplanes. While it is believed that this term should be replaceable by $\exp(-cn)$ (see Equation (1.4) below), such a result is not even known in the case of i.i.d. complex random variables with finite non-zero variance (see Jain, 2020, 2019 for the best-known results in this direction).

In particular, Theorem 1.1 implies that for any fixed $\delta > 0$, with probability at least $1-\delta$, $s_n(A_n) \ge \Omega_\delta(n^{-1/2})$. The rate $n^{-1/2}$ is optimal, and to the best of our knowledge, all previous works considering general heavy-tailed complex random matrices miss this sharp rate. For instance, it was shown by Tao and Vu (2010b) that if the entries of $A_{i,j}$ are dominated (in a technical Fourier sense) by a complex random variable with κ-controlled second moment (see Tao and Vu, 2010b, Definition 5.1 for the definition), then for any $C, \alpha > 0$,

$$\Pr(s_n(A_n) \le n^{-C} \cdot n^{-1/2}) \lesssim_{C,\alpha,\kappa} n^{-C+\alpha+o_n(1)} + \Pr(\|A_n\| \ge n^{1/2}). \tag{1.2}$$

Here, $\lesssim_{C,\alpha,\kappa}$ hides a constant depending on C, α, κ. We also note that for the case when the $A_{i,j}$ are i.i.d., $\Pr\left(|\widetilde{A_{i,j}}| \ge b\right) \ge b$ for some $b \in (0,1)$, and $\mathbb{E}|A_{i,j}|^2 \le K$ for some $K > 0$, the first author showed Jain (2020) that for any $\epsilon, \alpha > 0$,

$$\Pr\left(s_n(A_n) \le \epsilon n^{-1/2-\alpha}\right) \lesssim_{K,b,\alpha} \epsilon + \exp(-cn^{1/50}), \tag{1.3}$$

where the constant c depends on K and b. Once again, this misses the correct rate.

The technical Fourier-domination condition needed for Equation (1.2) already implies that $\Pr(b^{-1} \ge |\widetilde{A_{i,j}}| \ge b) \ge b$ for some $b \in (0,1)$ (see Tao and Vu, 2010b, Corollary 6.3). On the other hand, there are natural examples of families of random variables which cannot be dominated by a random variable with κ-controlled second moment, but which nevertheless satisfy the uniform anti-concentration assumption of Theorem 1.1; one such example is provided by the family $\{e^{i\theta}\cdot\xi\}_{\theta\in[0,2\pi)}$, where ξ is a Rademacher random variable i.e. $\xi=\pm 1$ with probability $1/2$ each. Moreover, in order for the term $\Pr\left(\|A_n\|\geq n^{1/2}\right)$ to be bounded away from 1, one needs to further assume that $\sum_{i,j=1}^{n}\mathbb{E}|A_{i,j}|^4\leq Kn^2$ for some $K>0$, which is more restrictive than the assumption in Theorem 1.1.

This somewhat dire situation in the complex case should be contrasted with the real case, where much more is known. The early breakthrough of Rudelson (2008) established that for an $n\times n$ matrix A_n whose entries are i.i.d. copies of a real centered sub-Gaussian random variable, and for any $\delta>0$, we have with probability at least $1-\delta$ that $s_n(A_n)=\Omega_\delta(n^{-1/2})$. A subsequent breakthrough of Rudelson and Vershynin (2008) refined this to the near-optimal tail bound

$$\Pr(s_n(A_n)\leq\epsilon\cdot n^{-1/2})\lesssim\epsilon+\exp(-cn),\tag{1.4}$$

where the implicit constant in $\lesssim$ and the constant c depend on the random variable. Extensions of the above tail bound to heavy-tailed and inhomogeneous random matrices has attracted much attention in recent years. Rebrova and Tikhomirov (2018) extended Rudelson and Vershynin's result to the case when the sub-Gaussian assumption is replaced by the finiteness of the second moment (the entries are still assumed to be identically distributed and centered). Livshyts (2021+) showed that if the entries $A_{i,j}$ are independent *real* random variables, $\Pr\left(|\widetilde{A_{i,j}}|\geq b\right)\geq b$ for some $b\in(0,1)$, and $\sum_{i,j=1}^{n}\mathbb{E}|A_{i,j}|^2\leq Kn^2$ for some $K>0$, then

$$\Pr\left(s_n(A_n)\leq\epsilon\cdot n^{-1/2}\right)\lesssim_{K,b}\epsilon+n^{-1/2}.\tag{1.5}$$

Finally, Livshyts, Rudelson, and Tikhomirov (Livshyts et al., 2021) obtained the near-optimal tail estimate Equation (1.4) under these assumptions (here, the implicit constant in $\lesssim$ and the constant c depend only on K and b).

Perhaps unsurprisingly, our proof makes use of tools introduced in Livshyts (2021+); Livshyts et al. (2021). The key new ingredient is an anti-concentration inequality for sums of independent complex random variables, which we will discuss in Section 2.

1.2. *Universality of ESDs of dense, inhomogeneous random matrices.* The *empirical spectral distribution* (ESD) μ_n of an $n\times n$ complex matrix A_n is defined on $\mathbb{R}^2$ by the expression

$$\mu_{A_n}(s,t):=\frac{1}{n}\cdot\left|\{k\in[n]\mid\Re(\lambda_k)\leq s;\Im(\lambda_k)\leq t\}\right|,$$

where $\lambda_1,\dots,\lambda_n$ denote the eigenvalues of A_n. A major highlight of random matrix theory is the celebrated *circular law* of Tao and Vu (2010a), which asserts that for *any* fixed complex random variable x of mean 0 and variance 1, the ESD of $A_n/\sqrt{n}$ – where A_n is an $n\times n$ random matrix each of whose entries is an independent copy of x – converges (as $n\to\infty$) uniformly to the distribution of the uniform measure on the unit disc in the complex plane,

$$\mu_\infty(s,t):=\frac{1}{\pi}\text{area}\{x\in\mathbb{C}\mid|x|\leq 1,\Re(x)\leq s,\Im(x)\leq t\}.$$

More generally, Tao and Vu showed that for any fixed complex random variables x and y of mean 0 and variance 1, and for any sequence of deterministic matrices M_n satisfying $\|M_n\|_{\text{HS}}^2=O(n^2)$ (here, $\|\cdot\|_{\text{HS}}$ denotes the Hilbert-Schmidt norm of a matrix), the difference of the ESDs of $(M_n+X_n)/\sqrt{n}$ and $(M_n+Y_n)/\sqrt{n}$ converges in probability to 0 as $n\to\infty$, where X_n is an $n\times n$ random matrix whose entries are i.i.d. copies of x, and Y_n is an $n\times n$ random matrix whose entries are i.i.d. copies of y. These results were extended by Krishnapur (Tao and Vu, 2010a) to independent, but not necessarily identically distributed random matrices, satisfying certain restrictions on the distributions of the entries.

Here, by using the arguments of Tao, Vu, and Krishnapur in conjunction with Theorem 1.1, we show the following.

Theorem 1.3. *Let $M_n = (\mu_{i,j}^{(n)})_{i,j\le n}$ and $C_n = (\sigma_{i,j}^{(n)})_{i,j\le n}$ be constant (i.e. deterministic) matrices satisfying*

(i) $\sup_n n^{-2}\|M_n\|_{\mathrm{HS}}^2 < \infty$;

(ii) $\alpha \le \sigma_{i,j}^{(n)} \le \beta$ *for all n, i, j, for some $0 < \alpha < \beta < \infty$.*

Given a matrix $\boldsymbol{X} = (x_{i,j})_{i,j\le n}$, set

$$A_n(\boldsymbol{X}) = M_n + C_n \cdot \boldsymbol{X} = (\mu_{i,j}^{(n)} + \sigma_{i,j}^{(n)} x_{i,j})_{i,j\le n},$$

where "$\cdot$" denotes the Hadamard product.

1. *Suppose that $x_{i,j}^{(n)}$ are independent complex-valued random variables with $\mathbb{E}[x_{i,j}^{(n)}] = 0$ and $\mathbb{E}[|x_{i,j}^{(n)}|^2] = 1$, and that $y_{i,j}^{(n)}$ are independent complex-valued random variables, also having zero mean and unit variance.*
2. *Assume that there exists some $b \in (0,1)$ such that $\Pr(|\widetilde{x_{i,j}}| \ge b) \ge b$ and similarly for $y_{i,j}$.*
3. *Also, assume Pastur's condition,*

$$\frac{1}{n^2}\sum_{i,j=1}^{n} \mathbb{E}\left[|x_{i,j}^{(n)}|^2 \mathbb{1}\{|x_{i,j}^{(n)}| \ge \epsilon\sqrt{n}\}\right] \xrightarrow{n\to\infty} 0 \quad \text{for all } \epsilon > 0,$$

and the same for $\boldsymbol{Y}$ in place of $\boldsymbol{X}$.

Then,

$$\mu_{n^{-1/2}\cdot A_n(\boldsymbol{X})} - \mu_{n^{-1/2}\cdot A_n(\boldsymbol{Y})} \xrightarrow{n\to\infty} 0$$

in the sense of probability.

Remark 1.4. In Tao and Vu (2010a), Krishnapur proved a similar result, except that the natural and mild anti-concentration assumption 2. was replaced by the stronger, technical, and hard-to-verify condition that $x_{i,j}, y_{i,j}$ have κ-controlled second moment.

Notation: Throughout the paper, we will omit floors and ceilings when they make no essential difference. We will use $\mathbb{S}_{\mathbb{C}}^{n-1}$ to denote the set of unit vectors in $\mathbb{C}^n$, $B(x,r)$ to denote the Euclidean ball of radius r centered at x, and $\Re(\boldsymbol{v}), \Im(\boldsymbol{v})$ to denote the real and imaginary parts of a complex vector $\boldsymbol{v} \in \mathbb{C}^n$. As is standard, we will use $[n]$ to denote the discrete interval $\{1,\dots,n\}$. We will also use the asymptotic notation $\lesssim, \gtrsim, \ll, \gg$ to denote $O(\cdot), \Omega(\cdot), o(\cdot), \omega(\cdot)$ respectively. For a matrix M, we will use $\|M\|$ to denote its standard $\ell^2 \to \ell^2$ operator norm and $\|M\|_{\mathrm{HS}}$ to denote the Hilbert-Schmidt norm. $\star$ denotes the Schur (entry-wise) product and $\mathrm{dist}(\cdot,\cdot)$ always denotes the Euclidean distance. All logarithms are natural unless noted otherwise.

2. Anti-concentration for sums of non-identically distributed independent complex random variables

The goal of the theory of anti-concentration is to obtain upper bounds on the Lévy concentration function, which is defined as follows.

Definition 2.1 (Lévy concentration function)**.** Let $X := (X_1, \dots, X_n) \in \mathbb{C}^n$ be a complex random vector, and let $v := (v_1, \dots, v_n) \in \mathbb{C}^n$. We define the *Lévy concentration function of v at radius r with respect to X* by

$$\rho_{r,X}(v) := \sup_{x\in\mathbb{C}} \Pr\left(v_1X_1 + \cdots + v_nX_n \in B(x, r)\right).$$

Rudelson and Vershynin (2008) introduced the notion of the essential least common denominator (LCD) to control the Lévy concentration function. This notion was generalized in Livshyts et al. (2021) to the randomized least common denominator (RLCD) and used to handle non-i.i.d. real random variables. We give a generalization of this to non-i.i.d. complex random variables which will be useful for us.

Definition 2.2 (CRLCD)**.** For a complex random vector $X := (X_1, \dots, X_n) \in \mathbb{C}^n$, a deterministic vector $v := (v_1, \dots, v_n) \in \mathbb{C}^n$, and parameters $L > 0, u \in (0, 1)$, define

$$\mathrm{CRLCD}^X_{L,u}(v) := \inf_{\theta\in\mathbb{C}} \left\{|\theta| > 0 : \mathbb{E}[\mathrm{dist}^2(\theta v \star \widetilde{X}, (\mathbb{Z} + i\mathbb{Z})^n)] < \min(u|\theta|^2\|v\|_2^2, L^2)\right\},$$

where $\widetilde{X}$ denotes the symmetrization of X (i.e. $\widetilde{X} \sim X' - X''$, where X' and X'' are independent copies of X).

Before proceeding to the results of this section, we need a couple of additional definitions.

Definition 2.3 (Tao and Vu, 2008)**.** For a complex random vector

$$X := (X_1, \dots, X_n) \in \mathbb{C}^n$$

and a deterministic vector $v := (v_1, \dots, v_n) \in \mathbb{C}^n$, let

$$P_X(v) := \mathbb{E}\left[-\pi|\langle\widehat{X}, v\rangle|^2\right].$$

Here, $\widehat{X} := \widetilde{X} \star (x_1, \dots, x_n)$, where $x_1, \dots, x_n$ are mutually independent Ber(1/2) random variables, which are also independent of $\widetilde{X}$.

Definition 2.4 (Tao and Vu, 2008)**.** For a complex random variable $z \in \mathbb{C}$ and a fixed complex number $a \in \mathbb{C}$, let

$$\|a\|_z := \left(\mathbb{E}\left[\|\Re(a \cdot \widetilde{z})\|^2_{\mathbb{R}/\mathbb{Z}}\right]\right)^{1/2},$$

where $\widetilde{z}$ denotes the symmetrization of z and $\|\cdot\|_{\mathbb{R}/\mathbb{Z}}$ denotes the distance to the nearest integer.

Lemma 2.5 (Tao and Vu, 2008)**.** *For a complex random vector*

$$X := (X_1, \dots, X_n) \in \mathbb{C}^n$$

with independent coordinates, and deterministic vectors $v := (v_1, \dots, v_n), w := (w_1, \dots, w_n) \in \mathbb{C}^n$:

(1) $\rho_{r,X}(v) \le \exp(\pi r^2) \cdot P_X(v)$.

(2) $P_X(v)P_X(w) \le 2P_{XX}(vw)$. *Here, $vw \in \mathbb{C}^{2n}$ denotes the vector whose first n coordinates coincide with v and last n coordinates coincide with w, and $XX \in \mathbb{C}^{2n}$ denotes the complex random vector whose first n coordinates and last n coordinates are both independent copies of X.*

(3) $P_X(v) \le \int_{\mathbb{C}} \exp\left(-\sum_{i=1}^n \|\xi \cdot v_i\|^2_{X_i}/2\right) \exp(-\pi|\xi|^2)d\xi$.

Proof: 1. follows from Tao and Vu (2008, Lemma 4.3), 2. follows from Tao and Vu (2008, Lemma 4.5(iii)) and 3. follows from Tao and Vu (2008, Lemma 5.2). Actually, in Tao and Vu (2008), these results are stated only in the case when the coordinates of the random vector X are identically distributed, but exactly the same proof also works for our more general setting. ☐

Next, we need a small modification of a 'doubling trick' from the proof of Jain (2020, Theorem 2.11).

Lemma 2.6. *Let $X := (X_1, \dots, X_n) \in \mathbb{C}^n$ be a complex random vector with independent coordinates, and let $w := (w_1, \dots, w_n) \in \mathbb{C}^n$ be a deterministic vector. Then,*

$$\rho_{r,X}(w)^2 \le 2\exp(2\pi r^2) \cdot \int_{\mathbb{C}} \exp\left(-\frac{1}{2}\mathbb{E}\left[\operatorname{dist}^2\left(\xi w \star \widetilde{X}, (\mathbb{Z}+i\mathbb{Z})^n\right)\right]\right) \exp(-\pi|\xi|^2) d\xi.$$

Proof: Let $w_{\mathbb{C}} \in \mathbb{C}^{2n}$ denote the vector whose first n coordinates coincide with w and last n coordinates coincide with $i \cdot w$. Then, since $\rho_{r,X}(w) = \rho_{r,X}(i \cdot w)$, we have

$$\begin{aligned}
\rho_{r,X}(w)^2 &= \rho_{r,X}(w)\rho_{r,X}(i \cdot w) \\
&\le \exp(2\pi r^2) \cdot P_X(w) \cdot P_X(i \cdot w) \\
&\le 2\exp(2\pi r^2) \cdot P_{XX}(w_{\mathbb{C}}) \\
&\le 2\exp(2\pi r^2) \cdot \int_{\mathbb{C}} \exp\left(-\frac{1}{2}\sum_{i=1}^n \left(\|\xi \cdot w_i\|_{X_i}^2 + \|i\xi \cdot w_i\|_{X_i}^2\right)\right) e^{-\pi|\xi|^2} d\xi,
\end{aligned}$$

where the second, third and fourth inequalities follow from Lemma 2.5 parts 1, 2, and 3 respectively. Finally, note that

$$\begin{aligned}
\sum_{i=1}^n \left(\|\xi \cdot w_i\|_{X_i}^2 + \|i\xi \cdot w_i\|_{X_i}^2\right) &= \mathbb{E}\sum_{i=1}^n \left(\|\Re(\xi w_i \cdot \widetilde{X_i})\|_{\mathbb{R}/\mathbb{Z}}^2 + \|\Re(i\xi w_i \cdot \widetilde{X_i})\|_{\mathbb{R}/\mathbb{Z}}^2\right) \\
&= \mathbb{E}\sum_{i=1}^n \left(\|\Re(\xi w_i \cdot \widetilde{X_i})\|_{\mathbb{R}/\mathbb{Z}}^2 + \|\Im(\xi w_i \cdot \widetilde{X_i})\|_{\mathbb{R}/\mathbb{Z}}^2\right) \\
&= \mathbb{E}\left[\operatorname{dist}^2(\xi w \star \widetilde{X}, (\mathbb{Z}+i\mathbb{Z})^n)\right]. \qquad \square
\end{aligned}$$

The next proposition is the main result of this section.

Proposition 2.7. *Let $X := (X_1, \dots, X_n) \in \mathbb{C}^n$ be a complex random vector with independent coordinates and let $v := (v_1, \dots, v_n) \in \mathbb{C}^n$ be such that $\frac{1}{2} \le \|v\|_2 \le 2$. Then, for all parameters $L > 0, u \in (0,1)$, and for all $\epsilon > 0$,*

$$\rho_{\epsilon,X}(v) \le C_{2.7}\left(\epsilon u^{-1/2} + \exp\left(-\frac{1}{4}L^2\right) + \exp\left(-\frac{\pi}{4}\epsilon^2 \operatorname{CRLCD}_{L,u}^X(v)^2\right)\right),$$

where $C_{2.7}$ is an absolute constant.

Proof: Let $w := v/\epsilon \in \mathbb{C}^n$. Then, $2^{-1}\epsilon^{-1} \le \|w\|_2 \le 2\epsilon^{-1}$ and $\rho_{\epsilon,X}(v) = \rho_{1,X}(w)$. Moreover,

$$\begin{aligned}
\rho_{1,X}(w)^2 &\le 2\exp(2\pi) \cdot \int_{\mathbb{C}} \exp\left(-\frac{1}{2}\mathbb{E}\left[\operatorname{dist}^2\left(\xi w \star \widetilde{X}, (\mathbb{Z}+i\mathbb{Z})^n\right)\right]\right) \exp(-\pi|\xi|^2) d\xi \\
&= 2\exp(2\pi)\epsilon^2 \cdot \int_{\mathbb{C}} \exp\left(-\frac{1}{2}\mathbb{E}\left[\operatorname{dist}^2\left(\eta v \star \widetilde{X}, (\mathbb{Z}+i\mathbb{Z})^n\right)\right]\right) e^{-\pi\epsilon^2|\eta|^2} d\eta
\end{aligned}$$

where the first line follows from Lemma 2.6 and the second line follows from the change of variables $\xi = \epsilon\eta$.

Let

$$F(\eta) = \exp\left(-\frac{1}{2}\mathbb{E}\left[\operatorname{dist}^2\left(\eta v \star \widetilde{X}, (\mathbb{Z}+i\mathbb{Z})^n\right)\right]\right) \exp(-\pi\epsilon^2|\eta|^2).$$

We break the above integral into two regions, $B(0, \operatorname{CRLCD}_{L,u}^X(v))$ and its complement.

For the first region, note that by the definition of CRLCD,

$$\begin{aligned}\int_{B(0,\mathrm{CRLCD}^X_{L,u}(v))} F(\eta)d\eta &\le \int_{B(0,\mathrm{CRLCD}^X_{L,u}(v))} \exp\left(-\frac{1}{2}\min\left(u|\eta|^2\|v\|_2^2, L^2\right) - \pi\epsilon^2|\eta|^2\right) d\eta \\ &\le \int_{\mathbb{C}} \exp\left(-\frac{1}{2}\min\left(u|\eta|^2\|v\|_2^2, L^2\right) - \pi\epsilon^2|\eta|^2\right) d\eta \\ &\le \int_{\mathbb{C}} \exp\left(-\frac{1}{2}u|\eta|^2\|v\|_2^2\right) d\eta + \int_{\mathbb{C}} \exp\left(-\frac{1}{2}L^2 - \pi\epsilon^2|\eta|^2\right) d\eta \\ &\le C\left(u^{-1} + \epsilon^{-2}\cdot\exp\left(-\frac{1}{2}L^2\right)\right),\end{aligned}$$

for some absolute constant $C > 0$. For the second region, note that

$$\begin{aligned}\int_{\mathbb{C}\setminus B(0,\mathrm{CRLCD}^X_{L,u}(v))} F(\eta)d\eta &\le \int_{\mathbb{C}\setminus B(0,\mathrm{CRLCD}^X_{L,u}(v))} \exp(-\pi\epsilon^2|\eta|^2)d\eta \\ &= \epsilon^{-2}\int_{\mathbb{C}\setminus B(0,\epsilon\,\mathrm{CRLCD}^X_{L,u}(v))} \exp(-\pi|\xi|^2)d\xi \\ &\le C\epsilon^{-2}\exp\left(-\frac{\pi}{2}\epsilon^2\,\mathrm{CRLCD}^X_{L,u}(v)^2\right),\end{aligned}$$

for some absolute constant $C > 0$. Putting everything together, we see that

$$\rho_{\epsilon,X}(v)^2 \le C\left(\epsilon^2u^{-1} + \exp\left(-\frac{1}{2}L^2\right) + \exp\left(-\frac{\pi}{2}\epsilon^2\,\mathrm{CRLCD}^X_{L,u}(v)^2\right)\right),$$

so that

$$\rho_{\epsilon,X}(v) \le C\left(\epsilon u^{-1/2} + \exp\left(-\frac{1}{4}L^2\right) + \exp\left(-\frac{\pi}{4}\epsilon^2\,\mathrm{CRLCD}^X_{L,u}(v)^2\right)\right),$$

as desired. □

We conclude this section with the following lemma, which shows that weighted sums of uniformly anti-concentrated random variables are not too close to being a constant.

Lemma 2.8. *Let $X := (X_1, \dots, X_n) \in \mathbb{C}^n$ be a complex random vector with independent coordinates such that $\Pr\left(b^{-1} \ge |\widetilde{X}_i| \ge b\right) \ge b$ for some $b \in (0,1)$. There exists a constant $c_{2.8} \in (0,1)$ depending only on b such that for all unit vectors $v := (v_1, \dots, v_n) \in \mathbb{S}^{n-1}_{\mathbb{C}}$,*

$$\rho_{c_{2.8},X}(v) \le 1 - c_{2.8}. \tag{2.1}$$

Proof: Let M be a sufficiently large constant depending only on b, to be determined during the course of the proof. We consider two cases, depending on $\|v\|_\infty$.

Case I: $\|v\|_\infty \ge M^{-1}$. Without loss of generality, suppose $|v_1| > M^{-1}$. Then, by conditioning on the variables $X_2, \dots, X_n$, we see that it suffices to prove that $\rho_{c,X_1}(v_1) \le 1 - c$, for some constant $c \in (0,1)$ depending only on b (and M). But this follows immediately since $\Pr(|\widetilde{X_1}| \ge b) \ge b$.

Case II: $\|v\|_\infty < M^{-1}$. In this case, it suffices to show that $\mathrm{CRLCD}^X_{L,u}(v) \ge Mb$, for $u = b^3$ and all L sufficiently large, for then, Equation (2.1) follows immediately from Proposition 2.7 by taking M to be sufficiently large depending on b. In order to show this, by definition, it suffices to show that for all $\theta \in \mathbb{C}$ such that $0 < |\theta| < Mb$,

$$\mathbb{E}\left[\mathrm{dist}^2(\theta v \star \widetilde{X}, (\mathbb{Z} + i\mathbb{Z})^n\right] \ge u|\theta|^2\|v\|_2^2.$$

For this, we begin by noting that for any such value of θ,

$$\begin{aligned}\operatorname{dist}^2(\theta v \star \widetilde{X}, (\mathbb{Z}+i\mathbb{Z})^n) &\geq \sum_{i=1}^n |\theta|^2 |v_i|^2 |\widetilde{X}_i|^2 \mathbf{1}\left[|\theta v_i \widetilde{X}_i| \leq \frac{1}{10}\right] \\ &\geq \sum_{i=1}^n |\theta|^2 |v_i|^2 |\widetilde{X}_i|^2 \mathbf{1}\left[|\widetilde{X}_i| \leq b^{-1}\right] \\ &\geq \sum_{i=1}^n |\theta|^2 |v_i|^2 |\widetilde{X}_i|^2 \mathbf{1}\left[b \leq |\widetilde{X}_i| \leq b^{-1}\right] \\ &\geq \sum_{i=1}^n b^2 |\theta|^2 |v_i|^2 \mathbf{1}\left[b \leq |\widetilde{X}_i| \leq b^{-1}\right].\end{aligned}$$

Therefore, taking the expectation on both sides, we see that

$$\begin{aligned}\mathbb{E}\left[\operatorname{dist}^2(\theta v \star \widetilde{X}, (\mathbb{Z}+i\mathbb{Z})^n)\right] &\geq \sum_{i=1}^n b^2 |\theta|^2 |v_i|^2 \mathbb{E}\left[\mathbf{1}\left[b \leq |\widetilde{X}_i| \leq b^{-1}\right]\right] \\ &\geq \sum_{i=1}^n b^2 |\theta|^2 |v_i|^2 \cdot b \\ &= b^3 |\theta|^2 \|v\|_2^2,\end{aligned}$$

which gives the desired conclusion. □

3. Proof of Theorem 1.1

In this section, we prove Theorem 1.1 following Livshyts (2021+); Livshyts et al. (2021). The only new ingredients are Lemma 3.1 and Proposition 3.8.

The first step in the proof of Theorem 1.1 is to decompose the sphere $\mathbb{S}_{\mathbb{C}}^{n-1}$. For some parameters $\delta, \rho \in (0,1)$ to be chosen later, following Rudelson and Vershynin (2008), we define the sets of sparse, compressible, and incompressible vectors as follows:

$$\begin{aligned}\operatorname{Sparse}(\delta) &:= \left\{u \in \mathbb{S}_{\mathbb{C}}^{n-1} : |\mathbf{supp}(u)| \leq \delta n\right\}, \\ \operatorname{Comp}(\delta,\rho) &:= \left\{u \in \mathbb{S}_{\mathbb{C}}^{n-1} : \operatorname{dist}(u, \operatorname{Sparse}(\delta)) \leq \rho\right\}, \\ \operatorname{Incomp}(\delta,\rho) &:= \mathbb{S}_{\mathbb{C}}^{n-1} \setminus \operatorname{Comp}(\delta,\rho).\end{aligned}$$

Here, for $u = (u_1, \dots, u_n)$, $\mathbf{supp}(u)$ denotes the set of coordinates $i \in [n]$ for which $u_i \neq 0$.

This results in

$$\mathbb{S}_{\mathbb{C}}^{n-1} = \operatorname{Comp}(\delta,\rho) \cup \operatorname{Incomp}(\delta,\rho).$$

By characterization (1.1) and the union bound, we have

$$\Pr(s_n(A_n) \leq \epsilon \cdot n^{-1/2}) \leq \Pr\left(\inf_{x \in \operatorname{Comp}(\delta,\rho)} \|A_n x\|_2 \leq \epsilon \cdot n^{-1/2}\right) + \Pr\left(\inf_{x \in \operatorname{Incomp}(\delta,\rho)} \|A_n x\|_2 \leq \epsilon \cdot n^{-1/2}\right).$$

We first deal with the compressible vectors. For this, as is standard, we begin with an estimate for 'invertibility with respect to a single vector', which in our case, follows directly by combining Lemma 2.8 with the so-called tensorization lemma (Rudelson and Vershynin, 2008, Lemma 2.2).

Lemma 3.1. *Let $A_{N,n}$ be an $N \times n$ complex random matrix whose entries $A_{i,j}$ are independent and satisfy* $\Pr\left(b^{-1} \geq |\widetilde{A_{i,j}}| \geq b\right) \geq b$ *for some $b \in (0,1)$. Then, for any fixed $v \in \mathbb{S}_{\mathbb{C}}^{n-1}$,*

$$\Pr\left(\|A_{N,n}v\|_2 \leq c_{3.1}\sqrt{N}\right) \leq (1 - c_{3.1})^N,$$

where $c_{3.1} \in (0,1)$ is a constant depending only on b.

Next, we need the following crucial theorem guaranteeing the existence of a suitable net on the sphere.

Theorem 3.2 (Modification of Livshyts, 2021+, Corollary 4)**.** *Fix $N, n \in \mathbb{N}$ and consider any subset $S \subset \mathbb{S}_{\mathbb{C}}^{n-1}$. For any $\mu \in (0,1)$, and for every $\epsilon \in (0, \mu^{c_0})$ (for some absolute constant $c_0 > 0$), let $\#N_\epsilon(S)$ denote the minimum number of Euclidean balls of radius ϵ needed to cover S. There exists a deterministic net $\mathcal{N} \subset \mathbb{C}^n$, with*

$$|\mathcal{N}| \leq \#N_\epsilon(S) \cdot (O(\epsilon))^{\mu n},$$

and there exist positive constants $C_1(\mu), C_2(\mu)$ such that the following holds. For every random $N \times n$ complex random matrix $A_{N,n}$ with independent columns, with probability at least $1 - e^{-C_1(\mu)n}$, for every $x \in S$, there exists $y \in \mathcal{N}$ so that

$$\|A_{N,n}(x - y)\|_2 \leq \frac{C_2(\mu)\epsilon}{\sqrt{n}}\sqrt{\mathbb{E}[\|A\|_{\mathrm{HS}}^2]}.$$

Remark 3.3. In Livshyts (2021+), this theorem is proved for nets of $\mathbb{S}_{\mathbb{R}}^{n-1}$; however, the same argument used there also works in the complex case.

Using Theorem 3.2 and the invertibility with respect to a single vector from Lemma 3.1, the following anti-concentration result for compressible vectors follows identically from Livshyts (2021+, Lemma 5.3).

Proposition 3.4. *Let A be an $n \times n$ random matrix whose entries $A_{i,j}$ are independent and satisfy $\mathbb{E}[\|A\|_{\mathrm{HS}}^2] \leq Kn^2$ for some $K > 0$, and $\Pr\left(|\widetilde{A_{i,j}}| \geq b\right) \geq b$ for some $b \in (0,1)$. Then, there exist $\rho, \delta \in (0,1)$ and $C_{3.4}, c_{3.4} > 0$, depending only on K and b, such that*

$$\Pr\left(\inf_{x \in \mathrm{Comp}(\delta,\rho)} \|Ax\|_2 \leq C_{3.4}\sqrt{n}\right) \leq 2e^{-c_{3.4}n}.$$

For the incompressible vectors, we use an 'invertibility via distance' bound similar to Rudelson and Vershynin (2008). The precise version we use appears in Livshyts et al. (2021).

Lemma 3.5 (Invertibility via distance, Livshyts et al. (2021, Lemma 6.1))**.** *Fix a pair of parameters $\delta, \rho \in (0, 1/2)$ and assume that $n \geq 4/\delta$. Then, for any $\epsilon > 0$,*

$$\Pr\left(\inf_{x \in \mathrm{Incomp}(\delta,\rho)} \|Ax\|_2 \leq \epsilon\frac{\rho}{\sqrt{n}}\right) \leq \frac{4}{\delta n} \inf_{I \subset [n], |I| = n - \lfloor \delta n/2 \rfloor} \sum_{j \in I} \Pr\left(\mathrm{dist}(A_j, H_j) \leq \epsilon\right),$$

where A_j denotes the j^{th} column of A and H_j denotes the subspace spanned by all the columns of A except for A_j.

From the previous lemma, in order to control $\|Ax\|_2$ for incompressible vectors x, it suffices to understand the anti-concentration of $\mathrm{dist}(A_j, H_j)$. For this, we begin by noting that $\mathrm{dist}(A_j, H_j) \geq |\langle A_j, \nu_j \rangle|$, where ν_j denotes any unit vector normal to H_j, so that the anti-concentration of $\mathrm{dist}(A_j, H_j)$ reduces to studying the anti-concentration properties of a unit normal to a random hyperplane.

Before proceeding to the details, we will need the following lemma, which shows that incompressible vectors have sufficiently large CRLCD.

Lemma 3.6 (Incompressibles have large CRLCD, Livshyts et al., 2021, Lemma 2.10)**.** *For any $b, \delta, \rho \in (0,1)$ and $c > 0$, there are $n_0 = n_0(b, \delta, \rho, c)$, $h_{3.6} = h_{3.6}(b, \delta, \rho, c) \in (0,1)$ and $u_{3.6} = u_{3.6}(b, \delta, \rho, c) \in (0, 1/4)$ with the following property. Let $n \geq n_0$, let $v \in \mathrm{Incomp}_n(\delta, \rho)$, and assume that a random vector $X = (X_1, \ldots, X_n)$ with independent components satisfies $\Pr\left(|\widetilde{X}_i| \geq b\right) \geq b$ for all $1 \leq i \leq n$, and $\mathbb{E}[\|X\|^2] \leq T$, for some $T \geq cn$. Then, for any $L > 0$, we have*

$$\mathrm{CRLCD}^X_{L, u_{3.6}}(v) \geq h_{3.6} \cdot \frac{n}{\sqrt{T}}$$

Remark 3.7. In Livshyts et al. (2021), the above proposition is proved for the notion of RLCD defined there, but the same proof goes through for the CRLCD as well.

We can now prove the desired invertibility on incompressible vectors.

Proposition 3.8. *Let A be an $n \times n$ random matrix whose entries $A_{i,j}$ are independent and satisfy $\mathbb{E}[\|A\|^2_{\mathrm{HS}}] \leq Kn^2$ for some $K > 0$, and $\Pr\left(b^{-1} \geq |\widetilde{A_{i,j}}| \geq b\right) \geq b$ for some $b \in (0,1)$. Fix a pair of parameters $\delta, \rho \in (0, 1/2)$, and assume that $n \geq 4/\delta$. There exist absolute constants $C_{3.8}, c_{3.8}$ that only depend on δ, ρ, b, K such that for any $\epsilon \in (0,1)$,*

$$\Pr\left(\inf_{x \in \mathrm{Incomp}(\delta,\rho)} \|Ax\|_2 \leq \epsilon \frac{\rho}{\sqrt{n}}\right) \leq C_{3.8}\left(\epsilon + \exp(-c_{3.8}\epsilon^2 n)\right).$$

Proof: Let $\delta \in (0, 1/2)$ as in the statement of the proposition and let $A_1, \ldots, A_n$ denote the columns of A. Since $\mathbb{E}\|A\|^2_{\mathrm{HS}} \leq Kn^2$, there must be at least $(1 - \delta/4)n$ columns A_i of A which satisfy $\mathbb{E}[\|A_i\|^2_2] \leq 4Kn/\delta$. Let I denote the set of the first $n - \lfloor \delta n/2 \rfloor$ such indices. We will apply Lemma 3.5 with this choice of I.

For this, fix $i \in I$, and let H_i denote the span of all columns of the matrix except for A_i. Then, an identical argument to Proposition 3.4 shows that, except with probability at most $\exp(-c_{3.4}n/2)$, any unit vector ν which is orthogonal to H_i must belong to $\mathrm{Incomp}(\delta', \rho')$, where $\delta', \rho', c_{3.4}$ depend only on K, b. Henceforth, we restrict ourselves to this event, and let ν denote a unit normal vector to the (random) hyperplane H_i.

By Lemma 3.6, it follows that for any $L > 0$,

$$\mathrm{CRLCD}^{A_j}_{L, u_{3.6}}(\nu) \geq C(b, K, \delta)\sqrt{n},$$

where $u_{3.6}$ depends on K, b, δ. Therefore, by Proposition 2.7, it follows that

$$\begin{aligned}\Pr\left(\mathrm{dist}(A_j, H_j) \leq \epsilon\right) &\leq \rho_{2\epsilon, A_j}(\nu) + \exp(-c_{3.4}n/2) \\ &\leq C_{2.7}\left(2\epsilon u_{3.6}^{-1/2} + \exp\left(-\frac{1}{4}L^2\right) + \exp\left(-C'(b, K, \delta)\epsilon^2 n\right)\right) + \\ &\quad + \exp(-c_{3.4}n/2).\end{aligned}$$

Finally, taking $L > 2\sqrt{C'(b, K, \delta)n}$ and using Lemma 3.5 gives the desired conclusion. □

Proof of Theorem 1.1: The proof of Theorem 1.1 now follows from using characterization (1.1) and the union bound by combining Proposition 3.4 and Proposition 3.8. □

4. Proof of Theorem 1.3

By means of the so-called *replacement principle* (Tao and Vu, 2010a, Theorem 2.1), the following analogue of Tao and Vu (2010a, Proposition 2.2) suffices to prove Theorem 1.3. For a square matrix M, we denote its determinant by $\det(M)$.

Proposition 4.1. *Let $A_n(\boldsymbol{X})$ and $A_n(\boldsymbol{Y})$ be as in the statement of Theorem 1.3. Then, for every fixed $z \in \mathbb{C}$,*

$$\frac{1}{n}\log\left|\det\left(\frac{1}{\sqrt{n}}A_n(\boldsymbol{X}) - zI\right)\right| - \frac{1}{n}\log\left|\det\left(\frac{1}{\sqrt{n}}A_n(\boldsymbol{Y}) - zI\right)\right|$$

converges in probability to zero as $n \to \infty$.

By using Steps 2,3,4 in the proof of Tao and Vu (2010a, Theorem C.2) verbatim, the proof of Proposition 4.1 is reduced to proving the following.

Proposition 4.2. *Let $A_n(\boldsymbol{X})$ and $A_n(\boldsymbol{Y})$ be as in the statement of Theorem 1.3 and let $z \in \mathbb{C}$ be fixed. Let $X_1, \dots, X_n$ be the rows of $A_n(\boldsymbol{X}) - z\sqrt{n}I$ and, for each $1 \le i \le n$, let V_i be the $(i-1)$-dimensional space generated by $X_1, \dots, X_{i-1}$. Similarly, let $Y_1, \dots, Y_n$ be the rows of $A_n(\boldsymbol{Y}) - z\sqrt{n}I$ and, for each $1 \le i \le n$, let W_i be the $(i-1)$-dimensional space generated by $Y_1, \dots, Y_{i-1}$. Then,*

$$\frac{1}{n}\sum_{n-n^{0.99}\le i\le n}\left(\log \operatorname{dist}\left(\frac{1}{\sqrt{n}}X_i, V_i\right) - \log\operatorname{dist}\left(\frac{1}{\sqrt{n}}Y_i, W_i\right)\right)$$

converges in probability to zero as $n \to \infty$.

We can further reduce to proving the following high probability bounds on the extreme singular values of $A_n(\boldsymbol{X})$ and $A_n(\boldsymbol{Y})$.

Proposition 4.3. *Let $A_n(\boldsymbol{X})$ and $A_n(\boldsymbol{Y})$ be as in the statement of Theorem 1.3 and let $z \in \mathbb{C}$ be fixed. Then, there exists an absolute constant $C > 0$ such that*

(1) $\Pr\left(s_1(A_n(\boldsymbol{X}) - z\sqrt{n}I) \ge n^C\right) = o_n(1)$,
(2) $\Pr\left(s_n(A_n(\boldsymbol{X}) - z\sqrt{n}I) \le n^{-C}\right) = o_n(1)$,

and similarly for $A_n(\boldsymbol{Y})$.

Before proving Proposition 4.3, let us show how it implies Proposition 4.2. We will make use of the following linear algebraic fact.

Lemma 4.4 (Tao and Vu, 2010a, Lemma A.4)**.** *Let A be an invertible $n \times n$ matrix with singular values $s_1(A) \ge \dots \ge s_n(A) > 0$ and rows $X_1, \dots, X_n \in \mathbb{C}^n$. For each $1 \le i \le n$, let U_i be the hyperplane generated by the $n-1$ rows*

$$X_1, \dots, X_{i-1}, X_{i+1}, \dots, X_n.$$

Then,

$$\sum_{j=1}^n s_j(A)^{-2} = \sum_{j=1}^n \operatorname{dist}(X_j, U_j)^{-2}.$$

Proposition 4.3 implies Proposition 4.2: For $1 \le i \le n$, let U_i denote the hyperplane generated by the $n-1$ rows $X_1, \dots, X_{i-1}, X_{i+1}, \dots, X_n$ of $A_n(\boldsymbol{X}) - \sqrt{n}zI$. First, note that

$$\frac{1}{\sqrt{n}}\operatorname{dist}(X_i, U_i) = \operatorname{dist}\left(\frac{1}{\sqrt{n}}X_i, U_i\right) \le \operatorname{dist}\left(\frac{1}{\sqrt{n}}X_i, V_i\right) \le \frac{1}{\sqrt{n}}\|X_i\|_2,$$

and similarly for $A_n(\boldsymbol{Y}) - \sqrt{n}zI$. Next, by Lemma 4.4,

$$\operatorname{dist}\left(\frac{1}{\sqrt{n}}X_i, V_i\right) \ge \operatorname{dist}\left(\frac{1}{\sqrt{n}}X_i, U_i\right) \ge \frac{1}{n}s_n(A_n(\boldsymbol{X}) - \sqrt{n}zI),$$

and similarly for $A_n(\boldsymbol{Y}) - \sqrt{n}zI$.

Therefore, Proposition 4.2 follows if we can show that, except with probability $o_n(1)$, $\|X_i\|_2 \le n^{O(1)}$ (for all $1 \le i \le n$), $s_n(A_n(\boldsymbol{X}) - z\sqrt{n}I) \ge n^{-O(1)}$, and similarly for $A_n(\boldsymbol{Y}) - \sqrt{n}zI$. Indeed, in this case, except with probability $o_n(1)$, each summand of the sum appearing in Proposition 4.2 is bounded in absolute value by $O(\log n)$, so that the entire sum is bounded in absolute value by

$$\frac{1}{n} \cdot O(\log n) \cdot n^{0.99} \leq O\left(\frac{1}{n^{0.001}}\right).$$

Finally, note that for all $1 \leq i \leq n$, $\|X_i\|_2 \leq s_1(A_n(\boldsymbol{X}) - \sqrt{n} zI)$, so that the desired probability bounds on $\|X_i\|_2$ and $s_n(A_n(\boldsymbol{X}))$ (and similarly for $A_n(\boldsymbol{Y})$) follow from Proposition 4.3. □

Finally, we prove Proposition 4.3.

Proof of Proposition 4.3: **Bound on s_1:** By the triangle inequality for $s_1(= \|\cdot\|)$, it suffices to show that there is an absolute constant $C > 0$ such that

$$\Pr\left(s_1(A_n(\boldsymbol{X})) \geq n^C\right) = o_n(1).$$

Note that by assumptions (i), (ii) and 1. in the statement of Theorem 1.3, we have $\mathbb{E}[\|A_n(\boldsymbol{X})\|_{\mathrm{HS}}^2] = O(n^{C'})$, so that $\mathbb{E}[s_1^2(A_n(\boldsymbol{X}))] = O(n^{C'})$. The desired conclusion now follows from Markov's inequality.

Bound on s_n: We begin by verifying that $P := A_n(\boldsymbol{X}) - z\sqrt{n}I$ satisfies the assumptions of Theorem 1.1. An identical argument works for $A_n(\boldsymbol{Y}) - z\sqrt{n}I$ as well.

Assumptions (i), (ii), and 1. of Theorem 1.3 show that $\mathbb{E}\sum_{i,j} |P_{i,j}|^2 \leq Kn^2$ for some $K > 0$.

Moreover, assumptions 1. and 2. of Theorem 1.3 show that there exists some $b' \in (0,1)$ such that $\Pr(b'^{-1} \geq |\widetilde{P_{i,j}}| \geq b') \geq b'$ for all i, j – indeed, assumption 2. of Theorem 1.3 shows that $\Pr(|\widetilde{P_{i,j}}| \geq b/\beta) \geq b$, and assumption 1. shows that $\mathbb{E}[|\widetilde{P_{i,j}}|^2] \leq \beta^2$ (therefore, by Markov's inequality, $\Pr(|\widetilde{P_{i,j}}| \geq \beta \cdot \sqrt{2/b}) \leq b/2$) so that we can conclude using the union bound and taking b' to be sufficiently small.

Finally, we can apply Theorem 1.1 to P with $\epsilon = n^{-1/4}$ (say) to obtain the desired conclusion. □

Acknowledgements

We are grateful to an anonymous referee for their careful reading of the manuscript and for their valuable comments.

References

Edelman, A. Eigenvalues and condition numbers of random matrices. *SIAM J. Matrix Anal. Appl.*, **9** (4), 543–560 (1988). MR964668.

Jain, V. Quantitative invertibility of random matrices: a combinatorial perspective. *ArXiv Mathematics e-prints* (2019). arXiv: 1908.11255.

Jain, V. The strong circular law: A combinatorial view. *Random Matrices: Theory and Applications*, p. 2150031 (2020). DOI: 10.1142/S2010326321500313.

Livshyts, G. V. The smallest singular value of heavy-tailed not necessarily i.i.d. random matrices via random rounding (2021+). To appear in *Journal d'Analyse Mathematique.*

Livshyts, G. V., Tikhomirov, K., and Vershynin, R. The smallest singular value of inhomogeneous square random matrices. *Ann. Probab.*, **49** (3), – (2021). MR4255145.

Rebrova, E. and Tikhomirov, K. Coverings of random ellipsoids, and invertibility of matrices with i.i.d. heavy-tailed entries. *Israel J. Math.*, **227** (2), 507–544 (2018). MR3846333.

Rudelson, M. Invertibility of random matrices: norm of the inverse. *Ann. of Math. (2)*, **168** (2), 575–600 (2008). MR2434885.

Rudelson, M. and Vershynin, R. The Littlewood-Offord problem and invertibility of random matrices. *Adv. Math.*, **218** (2), 600–633 (2008). MR2407948.

Tao, T. and Vu, V. Random matrices: the circular law. *Commun. Contemp. Math.*, **10** (2), 261–307 (2008). MR2409368.

Tao, T. and Vu, V. Random matrices: universality of ESDs and the circular law. *Ann. Probab.*, **38** (5), 2023–2065 (2010a). MR2722794.

Tao, T. and Vu, V. Smooth analysis of the condition number and the least singular value. *Math. Comp.*, **79** (272), 2333–2352 (2010b). MR2684367.