

# EIGENVECTOR DELOCALIZATION FOR NON-HERMITIAN RANDOM MATRICES AND APPLICATIONS

KYLE LUH AND SEAN O'ROURKE

ABSTRACT. Improving upon results of Rudelson and Vershynin, we establish delocalization bounds for eigenvectors of independent-entry random matrices. In particular, we show that with high probability every eigenvector is delocalized, meaning any subset of its coordinates carries an appropriate proportion of its mass. Our results hold for random matrices with genuinely complex as well as real entries. As an application of our methods, we also establish delocalization bounds for normal vectors to random hyperplanes. The proofs of our main results rely on a least singular value bound for genuinely complex rectangular random matrices, which generalizes a previous bound due to the first author, and may be of independent interest.

## 1. INTRODUCTION

Let  $G$  be an  $n \times n$  random matrix with independent and identically distributed (iid) entries whose real and imaginary parts are independent standard normal random variables. It is not difficult to see that the distribution of  $G$  is invariant under multiplication (either on the right or left) by unitary matrices. Among others, this implies that the unit eigenvectors of  $G$  are uniformly distributed on the complex unit sphere  $S_{\mathbb{C}}^{n-1}$ .

For an  $n \times n$  independent-entry matrix  $A$  with non-Gaussian entries no such invariance property exists, and the distribution of the eigenvectors is not easily described. In fact, if the entries of  $A$  are discrete random variables, then the eigenvectors cannot have continuous distribution. However, the universality phenomenon in random matrix theory asserts that, under some appropriate regularity conditions on the entries, the eigenvectors of  $A$  should be approximately uniform on the unit sphere for large enough dimension  $n$ . As such, we expect each eigenvector of  $A$  to have asymptotically the same properties as a vector uniformly distributed on the unit sphere.

The goal of this note is to quantify some of these properties for the eigenvectors of  $A$ . Let us begin by recalling some delocalization properties for random vectors uniformly distributed on the unit sphere. To fix some notation, for a vector  $v = (v_i)_{i=1}^n \in \mathbb{C}^n$ , we let  $\|v\|_{\infty}$  denote the  $\ell^{\infty}$ -norm of  $v$  and  $\|v\|_2$  denote the  $\ell^2$ -norm of  $v$ . In addition, for  $I \subset [n] := \{1, \dots, n\}$ , we let  $v_I$  denote the  $|I|$ -vector  $v_I = (v_i)_{i \in I}$ . Here,  $|I|$  denotes the cardinality of the finite set  $I$ .

**Proposition 1.1** (Largest coordinate of a uniformly distributed vector on the unit sphere). *Let  $v$  be uniformly distributed on the unit sphere in  $\mathbb{C}^n$  or  $\mathbb{R}^n$ . Then there exists an absolute constant  $C > 0$  such that*

$$(1.1) \quad \|v\|_{\infty} \leq C \sqrt{\frac{\log n}{n}}$$

---

*Date:* February 1, 2019.

K. Luh has been supported in part by the National Science Foundation under Award No. 1702533.

S. O'Rourke has been supported in part by NSF grants ECCS-1610003 and DMS-1810500.

with probability  $1 - o(1)$ <sup>1</sup>.

The bound on the  $\ell^\infty$ -norm in (1.1) rules out peaks in the distribution of mass of  $v$ . This bound is optimal, up to the choice of constant  $C$ . A similar bound was recently extended to eigenvectors of matrices with independent subgaussian entries [42].

**Definition 1.2** (Subgaussian random variable). A real random variable  $X$  is called *subgaussian* if there exists  $B > 0$  called the subgaussian moment of  $X$  such that

$$\mathbb{P}(|X| > t) \leq 2e^{-t^2/B^2}$$

for every  $t > 0$ .

**Theorem 1.3** (Theorem 1.1, [42]). *Let  $A$  be an  $n \times n$  matrix whose entries  $a_{ij}$  are independent real-valued random variables with mean zero, unit variance, and subgaussian moment bounded by  $B$ . Let  $t \geq 2$ . Then with probability at least  $1 - n^{1-t}$ , every eigenvector  $v$  of  $A$  satisfies*

$$\|v\|_\infty \leq \frac{Ct^{3/2} \log^{9/2} n}{\sqrt{n}} \|v\|_2.$$

Here  $C > 0$  depends only on  $B$ .

*Remark.* More generally, Theorem 1.3 holds in the case when the entries  $a_{ij}$  of  $A$  are complex-valued; see [42, Remark 1.2] for details.

In this note, we are interested in the smallest coordinates of the eigenvectors. For comparison, a random vector  $v$  uniformly distributed on the unit sphere has the following bounds for its smallest coordinates.

**Proposition 1.4** (Smallest coordinates of a vector uniformly distributed on the unit sphere). *Let  $v$  be uniformly distributed on the unit sphere in either  $\mathbb{R}^n$  or  $\mathbb{C}^n$ .*

- (Real case) *There exists constants  $C, c > 0$  such that if  $v$  is uniformly distributed on the unit sphere in  $\mathbb{R}^n$ , then for any integer  $1 \leq m \leq cn$*

$$(1.2) \quad \|v_I\|_2 \geq \frac{C}{\log^c n} \left(\frac{m}{n}\right)^{3/2} \quad \text{for all } I \subset [n], |I| \geq m$$

*with probability  $1 - o(1)$ .*

- (Complex case) *There exists constants  $C, c > 0$  such that if  $v$  is uniformly distributed on the unit sphere in  $\mathbb{C}^n$ , then for any integer  $1 \leq m \leq cn$*

$$(1.3) \quad \|v_I\|_2 \geq \frac{C}{\log^c n} \frac{m}{n} \quad \text{for all } I \subset [n], |I| \geq m$$

*with probability  $1 - o(1)$ .*

The bounds in (1.2) and (1.3) show that no set of  $m$  coordinates of  $v$  can have too little mass. This rules out “gaps” in how the mass of  $v$  is spread amongst the coordinates (or as described in [39], this shows that  $v$  lacks “almost empty zones”). This phenomenon was named *no-gaps delocalization* by Rudelson and Vershynin [43]. The bounds in (1.2) and (1.3) are conjectured to be optimal, modulo logarithmic corrections, for all values of  $m$ , and have been proven sharp for a number of regimes [18, 38].

Importantly, we emphasize the very different behavior displayed in Proposition 1.4 between a vector  $v$  uniformly distributed on the unit sphere in  $\mathbb{R}^n$  compared to the unit sphere in

---

<sup>1</sup>Here  $o(1)$  denotes a term which tends to zero as  $n \rightarrow \infty$ . See Section 2.5 for a complete description of our asymptotic notation.

$\mathbb{C}^n$ . This can be explained in a number of ways. In either case, the vector  $v$  has the same distribution as  $g/\|g\|_2$ , where  $g$  is the standard real or complex Gaussian vector in  $\mathbb{R}^n$  or  $\mathbb{C}^n$ . It follows that,  $\|g\|_2 = \Theta(\sqrt{n})$  with probability at least  $1 - Ce^{-cn}$ ; see for example, [31, Lemma 1]. Here  $C, c > 0$  are constants which may change from one occurrence to the next. In addition, for all  $\varepsilon > 0$ , the coordinates of  $g$  satisfy

$$\mathbb{P}(|g_i| \leq \varepsilon) \leq C\varepsilon$$

in the real case and

$$\mathbb{P}(|g_i| \leq \varepsilon) \leq C\varepsilon^2$$

in the complex case. These bounds imply the following bounds for the coordinates of a vector  $v$  uniformly distributed on the unit sphere:

$$(1.4) \quad \mathbb{P}\left(|v_i| \leq \frac{\varepsilon}{\sqrt{n}}\right) \leq C\varepsilon + Ce^{-cn}$$

in the real case and

$$(1.5) \quad \mathbb{P}\left(|v_i| \leq \frac{\varepsilon}{\sqrt{n}}\right) \leq C\varepsilon^2 + Ce^{-cn}$$

for the complex case. Importantly, the difference between  $\varepsilon$  appearing on the right-hand side of (1.4) and  $\varepsilon^2$  on the right-hand side of (1.5) leads to the differing behaviors seen in Proposition 1.4. Indeed, by the union bound, (1.4) and (1.5) can easily be used to deduce bounds for the smallest coordinate of  $v$ :

$$\mathbb{P}\left(\min_{1 \leq i \leq n} |v_i| \leq \frac{\varepsilon}{n^{3/2}}\right) \leq C\varepsilon + Cne^{-cn}$$

in the real case and

$$\mathbb{P}\left(\min_{1 \leq i \leq n} |v_i| \leq \frac{\varepsilon}{n}\right) \leq C\varepsilon^2 + Cne^{-cn}$$

for the complex case. These last two bounds agree with Proposition 1.4 for the case  $m = 1$ .

For eigenvectors of independent-entry matrices, Rudelson and Vershynin [43] proved the following analogue of (1.2).

**Theorem 1.5** (Theorem 1.5, [43]). *Let  $A$  be an  $n \times n$  random matrix whose entries are iid copies of the real-valued random variable  $\xi$ , which satisfies*

$$\sup_{u \in \mathbb{R}} \mathbb{P}(|\xi - u| \leq 1) \leq 1 - p, \quad \mathbb{P}(|\xi| > K) \leq p/2$$

for some  $K, p > 0$ . Choose  $M \geq 1$  such that the event  $\{\|A\| \leq M\sqrt{n}\}$  holds with probability at least  $1/2$ . Let  $\varepsilon \geq 1/n$  and  $s \geq c_1\varepsilon^{-7/6}n^{-1/6} + e^{-c_2/\sqrt{\varepsilon}}$ . Then, conditionally on the event  $\{\|A\| \leq M\sqrt{n}\}$ , the following holds with probability at least  $1 - (c_2s)^{\varepsilon n}$ . Every eigenvector  $v$  of  $A$  satisfies

$$\|v_I\|_2 \geq (\varepsilon s)^6 \|v\|_2 \quad \text{for all } I \subset [n], |I| \geq \varepsilon n.$$

Here  $c_1, c_2, c_3$  depend on  $p, k$ , and  $M$ .

*Remark.* We have stated Theorem 1.5 for real iid random matrices, but the results in [43] also extend to the case where the  $(i, j)$ -entry depends on the  $(j, i)$ -entry as well as the case when the entries are complex-valued. We refer the reader to [43, Section 1] for details.

In view of numerical simulations and heuristic arguments coming from (1.4) and (1.5), the bounds in Theorem 1.5 appear to be suboptimal. In this article, we improve the bounds in Theorem 1.5 for random matrices with genuinely complex entries.

**Definition 1.6.** Following [33], we say an  $N \times n$  random matrix  $A$  is *genuinely complex* if the entries of  $A$  are independent and<sup>2</sup>  $a_{ij} = \xi_{ij} + \sqrt{-1}\xi'_{ij}$  where  $\xi_{ij}$  and  $\xi'_{ij}$  are independent real random variables with mean zero, unit variance, and subgaussian moment bounded by  $B$ .

Eigenvectors of random matrices have been widely studied in the mathematics and physics literature. We refer the reader to [2, 15, 19, 22, 28, 29, 36, 47, 1, 3, 7, 5, 8, 6, 11, 14, 13, 20, 26, 30, 32, 44, 45, 46, 52, 55, 56, 49, 50, 10, 24, 23, 25, 27, 4, 9, 12, 16, 35, 38, 42, 43, 51, 57] and references therein for many results concerning various models of random matrices. The majority of these results apply to eigenvectors of Hermitian or real symmetric random matrices. Significantly less appears to be known for independent-entry matrices. In the case of the complex Ginibre ensemble, where the entries are iid standard complex Gaussians, a number of results have described the asymptotic correlations and overlaps between eigenvectors. Important contributions in this line of research were made by Chalker and Mehlig [15, 35] with significant improvements and generalizations being made recently by Fyodorov [28] as well as by Bourgade and Dubach [12]. Other recent results include [4, 9, 16] and references therein, and there still appears to be significant work to be done in this area.

Shortly after this paper appeared on the arXiv, an improved version of Theorem 1.5 was proved by Lytova and Tikhomirov [34]. For larger values of  $m$ , the results in [34] achieve the optimal bounds depicted in Proposition 1.4. In particular, the results from [34] improve upon our bounds when  $m \geq \log^C n$  for some constant  $C > 0$ . The techniques used by Lytova and Tikhomirov are significantly different than those employed in this paper. In [34], a geometric approach is taken, which utilizes test projections and involves studying random ellipsoids generated by projections of independent vectors. Compared to [34], the main results in the present paper hold with higher probability and include bounds for the cases when  $1 \leq m \leq \log^C n$ .

## 2. MAIN RESULTS

Our main results hold for random matrices with genuinely complex entries as well as random matrices with real entries. In this section, we also discuss an application of our methods to normal vectors of random hyperplanes. We continue to use the notation introduced above: for a vector  $v = (v_i)_{i=1}^n \in \mathbb{C}^n$ , we let  $\|v\|_2$  denote the  $\ell^2$ -norm of  $v$ . In addition, for  $I \subset [n] := \{1, \dots, n\}$ , we let  $v_I$  denote the  $|I|$ -vector  $v_I = (v_i)_{i \in I}$ , where,  $|I|$  denotes the cardinality of the finite set  $I$ . Recall that  $\sqrt{-1}$  denotes the imaginary unit.

**2.1. Results for eigenvectors of genuinely complex matrices.** Our first main result improves upon Theorem 1.5 for large enough values of  $m$ .

**Theorem 2.1.** *Assume  $A$  is an  $n \times n$  genuinely complex random matrix. Then there exist constants  $C, c, c' > 0$  such that for every  $t \geq e^{-\log^2 n}$  and  $\log^2 n \leq m \leq c'n$ , with probability at least  $1 - (Ct)^m - Ce^{-c'n}$ , every eigenvector  $v$  of  $A$  satisfies*

$$\|v_I\|_2 \geq c\sqrt{t} \left(\frac{m}{n}\right)^{3/2} \|v\|_2 \text{ for all } I \subset [n] \text{ with } |I| \geq m.$$

Here  $C, c, c'$  depend only on the subgaussian moment bound  $B$ .

For smaller values of  $m$ , we have the following bound.

---

<sup>2</sup>We use  $\sqrt{-1}$  to denote the imaginary unit and reserve  $i$  as an index. See Section 2.5 for a complete description of our notation.

**Theorem 2.2.** *Assume  $A$  is an  $n \times n$  genuinely complex random matrix. Then there exist constants  $C, c, c' > 0$  such that for every  $t \geq e^{-c'n}$  and  $1 \leq m \leq \log^2 n$ , with probability at least  $1 - (Ct)^m - Ce^{-c'n}$ , every eigenvector  $v$  of  $A$  satisfies*

$$\|v_I\|_2 \geq c \frac{\sqrt{t}}{\log^2 n} \left(\frac{m}{n}\right)^{3/2+1/m} \|v\|_2 \text{ for all } I \subset [n] \text{ with } |I| \geq m.$$

Here  $C, c, c'$  depend only on the subgaussian moment bound  $B$ .

As a corollary, we immediately obtain the following in the case that  $m = 1$ .

**Corollary 2.3.** *Assume  $A$  is an  $n \times n$  genuinely complex random matrix. Then there exist constants  $C, c, c' > 0$  such that for every  $t \geq e^{-c'n}$ , with probability at least  $1 - Ct$ , every eigenvector  $v$  of  $A$  satisfies*

$$|v_i| \geq c\sqrt{t} \frac{1}{n^{5/2} \log^2 n} \|v\|_2 \text{ for all } 1 \leq i \leq n.$$

Here  $C, c, c'$  depend only on the subgaussian moment bound  $B$ .

Corollary 2.3 implies that with probability at least  $1 - Ce^{-c'n}$ , every coordinate of every eigenvector is nonzero. In particular, this implies that, with the same probability, each eigenspace of  $A$  has dimension one. Indeed, if  $A$  has an eigenspace of dimension greater than one, then this eigenspace must have a non-trivial intersection with the orthogonal complement of the space spanned by  $e_i$ , where  $e_1, \dots, e_n$  are the standard basis elements in  $\mathbb{C}^n$ .

**2.2. Results for eigenvectors of matrices with real entries.** In this subsection, we consider eigenvectors of real matrices. Our first result is the analogue of Theorem 2.1 for the eigenvectors of  $A$  corresponding to real eigenvalues. When the entries of  $A$  are iid real standard normal random variables, the number of real eigenvalues was studied in [21]. The existence of real eigenvalues for random matrices with non-normal entries was established more recently in [53], under the assumption the entries match the first four moments of the standard Gaussian distribution.

**Theorem 2.4.** *Assume  $A$  is an  $n \times n$  real random matrix whose entries are independent copies of a mean zero subgaussian random variable with unit variance. Then there exist constants  $C, c, c' > 0$  such that for every  $t \geq e^{-c'n}$  and  $\log^2 n \leq m \leq c'n$ , with probability at least  $1 - (Ct)^m - Ce^{-c'n}$ , every eigenvector  $v \in \mathbb{R}^n$  of  $A$  corresponding to a real eigenvalue satisfies*

$$\|v_I\|_2 \geq ct \left(\frac{m}{n}\right)^2 \|v\|_2 \text{ for all } I \subset [n] \text{ with } |I| \geq m.$$

Here  $C, c, c'$  depend only on the subgaussian moment of the entries.

For smaller values of  $m$ , we have the following bound.

**Theorem 2.5.** *Assume  $A$  is an  $n \times n$  real random matrix whose entries are independent copies of a mean zero subgaussian random variable with unit variance. Then there exist constants  $C, c, c' > 0$  such that for every  $t \geq e^{-c'n}$  and  $1 \leq m \leq \log^2 n$ , with probability at least  $1 - (Ct)^m - Ce^{-c'n}$ , every eigenvector  $v \in \mathbb{R}^n$  of  $A$  corresponding to a real eigenvalue satisfies*

$$\|v_I\|_2 \geq ct \left(\frac{m}{n}\right)^{2+1/m} \|v\|_2 \text{ for all } I \subset [n] \text{ with } |I| \geq m.$$

Here  $C, c, c'$  depend only on the subgaussian moment of the entries.

**2.3. Normal vectors to random hyperplanes.** As an application of our methods, we now consider delocalization bounds for normal vectors to random hyperplanes. Let  $A$  be an  $(n-1) \times n$  independent-entry random matrix. As this matrix is ill-conditioned, there exists at least one unit vector  $v$  so that  $Av = 0$ . Stated another way, this means that there is at least one unit vector that is orthogonal to the rows of  $A$ . In fact, under very general conditions on the entries,  $A$  has rank  $n-1$ . In this case,  $v$  is uniquely determined up to a phase. Nguyen and Vu studied the normal vector  $v$  when the entries of  $A$  are centered iid subgaussian random variables [37].

In this section, we extend the results in [37] to include some additional delocalization properties for the normal vector  $v$ . Intuitively, one expects  $v$  to behave like a random vector uniformly distributed on the unit sphere. In fact, in the case when  $A$  has standardized Gaussian entries this is precisely the distribution of  $v$ .

We begin by considering the case when  $m$  is proportional to  $n$ . Let us introduce the following notation. Let  $F$  be the cumulative distribution function of the  $\chi^2$ -distribution with two degrees of freedom. Following the notation in [17], let  $Q$  denote the quantile function of  $F$ . That is,

$$(2.1) \quad Q(s) := \inf\{x \in \mathbb{R} : F(x) \geq s\}, \quad 0 < s \leq 1, \quad Q(0) := \lim_{s \searrow 0} Q(s).$$

Define

$$(2.2) \quad H(s) := -Q(1-s), \quad 0 \leq s < 1.$$

**Theorem 2.6.** *Suppose  $\xi$  and  $\xi'$  are iid real subgaussian random variables with mean zero and unit variance. Let  $A$  be an  $(n-1) \times n$  iid matrix whose entries are iid copies of  $\xi + \sqrt{-1}\xi'$ , and let  $v$  be any unit vector which satisfies  $Av = 0$ . Then, for any fixed  $1 > \delta > 0$ ,*

$$\max_{I \subset [n]: |I| = \lfloor \delta n \rfloor} \|v_I\|_2^2 \longrightarrow - \int_0^\delta H(u) du$$

and

$$\min_{I \subset [n]: |I| = \lfloor \delta n \rfloor} \|v_I\|_2^2 \longrightarrow - \int_{1-\delta}^1 H(u) du$$

in probability as  $n \rightarrow \infty$ , where  $H$  is defined in (2.2).

*Remark.* Following [38], one can show that, as  $\delta$  tends to zero,

$$- \int_0^\delta H(u) du = \Theta(\delta \log(\delta^{-1}))$$

and

$$- \int_{1-\delta}^1 H(u) du = \Theta(\delta^2).$$

In other words, Theorem 2.6 implies that the smallest  $\delta n$  coordinates of an eigenvector contribute only  $\Theta(\delta^2)$  fraction of the mass, which matches the bounds from Proposition 1.4.

Our next results are the analogues of Theorems 2.1 and 2.2 for the normal vector.

**Theorem 2.7.** *Assume  $A$  is an  $(n-1) \times n$  genuinely complex matrix. Then there exist constants  $C, c, c' > 0$  such that for every  $t \geq e^{-\log^2 n}$  and  $\log^2 n \leq m \leq c'n$ , with probability at least  $1 - (Ct)^m - Ce^{-c'n}$ , every nonzero vector  $v$  orthogonal to the rows of  $A$  satisfies*

$$\|v_I\|_2 \geq c\sqrt{t} \left(\frac{m}{n}\right)^{3/2} \|v\|_2 \text{ for all } I \subset [n] \text{ with } |I| \geq m.$$

Here  $C, c, c'$  depend only on the subgaussian moment bound  $B$ .

**Theorem 2.8.** *Assume  $A$  is an  $(n-1) \times n$  genuinely complex matrix. Then there exist constants  $C, c, c' > 0$  such that for every  $t > 0$  and  $1 \leq m \leq \log^2 n$ , with probability at least  $1 - (Ct)^m - Ce^{-c'n}$ , every nonzero vector  $v$  orthogonal to the rows of  $A$  satisfies*

$$\|v_I\|_2 \geq c \frac{\sqrt{t}}{\log n} \left(\frac{m}{n}\right)^{3/2} \|v\|_2 \text{ for all } I \subset [n] \text{ with } |I| \geq m.$$

Here  $C, c, c'$  depend only on the subgaussian moment bound  $B$ .

*Remark.* More generally, Theorems 2.7 and 2.8 can be extended to cases where  $A$  is an  $(n-k) \times n$  matrix and  $m \geq k$  using the same methods, but the lower bound for  $\|v_I\|_2$  in these cases is substantially more cumbersome to notate.

**2.4. Outline of the paper.** The paper is organized as follows. In Section 3, we give an overview of our argument by showing how delocalization properties for the eigenvectors of the square matrix  $A$  can be reduced to questions concerning the least singular value of rectangular sub-matrices of  $A$ . Similar reductions have been utilized before, and our arguments in this section follow closely those in [43]. We establish a bound for the least singular value of genuinely complex rectangular random matrices in Section 4. This bound is based on a similar bound for genuinely complex square random matrices established by the first author [33]. The main results in Section 2 are proven in Section 5. The proofs of Propositions 1.1 and 1.4 are presented in Appendix A.

**2.5. Notation.** We use asymptotic notation (such as  $O, o$ ) under the assumption that  $n \rightarrow \infty$ . We use  $X = O(Y)$ ,  $Y = \Omega(X)$ ,  $X \ll Y$ , or  $Y \gg X$  to denote the estimate  $|X| \leq CY$  for some constant  $C > 0$  independent of  $n$  and all  $n \geq C$ . If  $C$  depends on another parameter, e.g.,  $C = C_k$ , we will indicate this by subscripts, e.g.,  $X = O_k(Y)$  or  $X \ll_k Y$ . We write  $X = \Theta(Y)$  if  $X \ll Y \ll X$ . We write  $X = o(Y)$  if  $|X| \leq c_n Y$  for some  $c_n$  that tends to zero as  $n \rightarrow \infty$ .

$|S|$  denotes the cardinality of the finite set  $S$ . We use  $\sqrt{-1}$  to denote the imaginary unit and reserve  $i$  as an index.  $[n]$  denotes the discrete interval  $\{1, \dots, n\}$ .

We use  $S_{\mathbb{R}}^{n-1}$  and  $S_{\mathbb{C}}^{n-1}$  to denote the unit spheres in  $\mathbb{R}^n$  and  $\mathbb{C}^n$ , respectively. For a vector  $x = (x_i)_{i=1}^n \in \mathbb{C}^n$ , we let  $\|x\|_2$  denote the  $\ell^2$ -norm of  $x$ . In addition, for  $I \subset [n] := \{1, \dots, n\}$ , we let  $x_I$  denote the  $|I|$ -vector  $x_I = (x_i)_{i \in I}$ . Similarly, for an  $N \times n$  matrix  $A = (a_{ij})_{i \in [N], j \in [n]}$  and a subset  $J \subset [n]$ , we let  $A_J$  denote the  $N \times |J|$  matrix  $A_J = (a_{ij})_{i \in [N], j \in J}$ . We let  $\mathcal{M}_{N \times n}^{\mathbb{F}}$  denote the set of  $N \times n$  matrices over the field  $\mathbb{F}$  (here,  $\mathbb{F}$  is either  $\mathbb{R}$  or  $\mathbb{C}$ ).

Recall that the singular values of a matrix  $A$  are the square roots of the eigenvalues of  $A^*A$ . For an  $N \times n$  matrix  $A$ , we let  $s_1(A) \geq \dots \geq s_n(A)$  denote the ordered singular values of  $A$ . Of particular importance are the largest and smallest singular values:

$$s_1(A) = \max_{\|x\|_2=1} \|Ax\|_2, \quad s_n(A) = \min_{\|x\|_2=1} \|Ax\|_2.$$

We will let  $\|A\|$  denote the spectral norm of  $A$ , i.e.,  $\|A\| = s_1(A)$ . For convenience, we will often let  $s_{\min}(A)$  denote the smallest singular value of  $A$ .

### 3. OUTLINE OF THE ARGUMENT

**3.1. Reduction of delocalization to invertibility.** For an  $n \times n$  matrix  $A$ , introduce the localization event

$$\text{Loc}(A, m, \delta) = \{\exists \text{ eigenvector } v \in S_{\mathbb{C}}^{n-1} \text{ of } A, \exists I \subset [n], |I| = m, \|v_I\|_2 < \delta\}.$$

Intuitively,  $\text{Loc}(A, m, \delta)$  captures the event that  $A$  has an eigenvector which has a subset of  $m$  coordinates which carry a disproportionately small proportion of the mass.

We will also extend this notion to rectangular matrices, but first we fix some notation. If  $A$  is a square matrix and  $\lambda \in \mathbb{C}$ , then  $A - \lambda$  denotes the matrix  $A - \lambda I$ , where  $I$  is the identity matrix. Similarly, if  $A$  is a rectangular matrix, we define  $A - \lambda$  to be the  $N \times n$  matrix with entries  $A_{ij} - \lambda \delta_{ij}$ , where  $\delta_{ij}$  is the Kronecker delta.

For an  $N \times n$  matrix  $A$  and  $\lambda_0 \in \mathbb{C}$ , we define the localization event

$$\text{Loc}_{\lambda_0}(A, M, m, \delta) = \{\exists v \in S_{\mathbb{C}}^{n-1}, \|(A - \lambda_0)v\|_2 \leq \delta M \sqrt{n}, \exists I \subset [n], |I| = m, \|v_I\|_2 < \delta\}.$$

In the case when  $A$  is square,  $\text{Loc}_{\lambda_0}(A, M, m, \delta)$  is the event that an approximate eigenvector  $v$  (with approximate eigenvalue  $\lambda_0$ ) is localized. For Theorems 2.7 and 2.8 it is important that we allow this event to also apply to rectangular matrices.

The following three propositions are based on [43, Proposition 4.1] and show that the study of the localization events defined above can be reduced to a question involving the least singular value of the random matrix  $A$ .

**Proposition 3.1** (Reduction of delocalization to invertibility for approximate eigenvectors).

*Let  $A$  be an  $N \times n$  random matrix with arbitrary distribution. Let  $M \geq 1$ ,  $\delta \in (0, 1/2)$ ,  $p_0 \in (0, 1)$ ,  $m \in [n]$ , and  $\lambda_0 \in \mathbb{C}$  with  $|\lambda_0| \leq M\sqrt{n}$ . Assume that for any set  $I \subset [n]$  with  $|I| = m$ , we have*

$$(3.1) \quad \mathbb{P}(s_{\min}((A - \lambda_0)_{I^c}) \leq 6\delta M \sqrt{n} \text{ and } \|A\| \leq M\sqrt{n}) \leq p_0.$$

*Then*

$$\mathbb{P}(\text{Loc}_{\lambda_0}(A, M, m, \delta) \text{ and } \|A\| \leq M\sqrt{n}) \leq \left(\frac{ne}{m}\right)^m p_0.$$

*Proof.* Assume  $\|A\| \leq M\sqrt{n}$  and the localization event  $\text{Loc}_{\lambda_0}(A, M, m, \delta)$  holds. Then there exists a unit vector  $v$  and an index set  $I \subset [n]$  with  $|I| = m$  such that

$$(3.2) \quad \|(A - \lambda_0)v\|_2 \leq \delta M \sqrt{n} \text{ and } \|v_I\|_2 < \delta.$$

We decompose the vector  $v$  as  $v = v_I + v_{I^c}$  to obtain

$$(A - \lambda_0)v = (A - \lambda_0)_I v_I + (A - \lambda_0)_{I^c} v_{I^c}.$$

Using (3.2), we find

$$\|(A - \lambda_0)_{I^c} v_{I^c}\|_2 \leq 3\delta M \sqrt{n}$$

and so

$$s_{\min}((A - \lambda_0)_{I^c}) \|v_{I^c}\|_2 \leq 3\delta M \sqrt{n}.$$

Since  $\|v_I\|_2 < \delta \leq 1/2$ , we obtain  $\|v_{I^c}\|_2 \geq 1/2$ , and hence

$$(3.3) \quad s_{\min}((A - \lambda_0)_{I^c}) \leq 6\delta M \sqrt{n}.$$

In other words, we have shown that the events  $\|A\| \leq M\sqrt{n}$  and  $\text{Loc}_{\lambda_0}(A, M, m, \delta)$  imply the existence of a subset  $I \subset [n]$  with  $|I| = m$  such that (3.3) holds. Applying the union bound and (3.1), we conclude that

$$\mathbb{P}(\text{Loc}_{\lambda_0}(A, m, \delta) \text{ and } \|A\| \leq M\sqrt{n}) \leq \binom{n}{m} p_0 \leq \left(\frac{ne}{m}\right)^m p_0,$$

as desired.  $\square$



**Proposition 3.2** (Reduction of delocalization to invertibility for eigenvectors). *Let  $A$  be an  $n \times n$  random matrix with arbitrary distribution. Let  $M \geq 1$ ,  $\delta \in (0, 1/2)$ ,  $p_0 \in (0, 1)$ , and  $m \in [n]$ . Assume that for any set  $I \subset [n]$  with  $|I| = m$  and any  $\lambda_0 \in \mathbb{C}$  with  $|\lambda_0| \leq M\sqrt{n}$ , we have*

$$(3.4) \quad \mathbb{P}(s_{\min}((A - \lambda_0)_{I^c}) \leq 6\delta M\sqrt{n} \text{ and } \|A\| \leq M\sqrt{n}) \leq p_0.$$

Then

$$\mathbb{P}(\text{Loc}(A, m, \delta) \text{ and } \|A\| \leq M\sqrt{n}) \leq \frac{9}{\delta^2} \left(\frac{ne}{m}\right)^m p_0.$$

*Proof.* Suppose  $\|A\| \leq M\sqrt{n}$  and the localization event  $\text{Loc}(A, m, \delta)$  holds. Then there exists an eigenvector  $v \in S_{\mathbb{C}}^{n-1}$  of  $A$  and an index set  $I \subset [n]$  with  $|I| = m$  such that  $\|v_I\|_2 < \delta$ . Let  $\lambda$  be the eigenvalue of  $A$  corresponding to the eigenvector  $v$ . Then  $|\lambda| \leq \|A\| \leq M\sqrt{n}$ . Let  $\mathcal{N}$  be a  $(\delta M\sqrt{n})$ -net of the disc  $\{z \in \mathbb{C} : |z| \leq M\sqrt{n}\}$ . A simple volume argument shows that one can construct the net  $\mathcal{N}$  to have cardinality

$$(3.5) \quad |\mathcal{N}| \leq \frac{9}{\delta^2}.$$

Choose  $\lambda_0 \in \mathcal{N}$  such that  $|\lambda - \lambda_0| \leq \delta M\sqrt{n}$ . Then the eigenvalue-eigenvector equation  $Av = \lambda v$  implies

$$(A - \lambda_0)v = (\lambda - \lambda_0)v,$$

and hence

$$\|(A - \lambda_0)v\|_2 \leq |\lambda - \lambda_0| \leq \delta M\sqrt{n}.$$

To summarize, we have shown that the events  $\|A\| \leq M\sqrt{n}$  and  $\text{Loc}(A, m, \delta)$  imply the existence of  $\lambda_0 \in \mathcal{N}$  such that  $\text{Loc}_{\lambda_0}(A, M, m, \delta)$  holds. We conclude from Proposition 3.1 and the union bound that

$$\mathbb{P}(\text{Loc}(A, m, \delta) \text{ and } \|A\| \leq M\sqrt{n}) \leq |\mathcal{N}| \left(\frac{ne}{m}\right)^m p_0.$$

Combining this bound with (3.5) completes the proof.  $\square$

To work with the eigenvectors of real matrices, we also require the following event:

$$\text{Loc}_{\mathbb{R}}(A, m, \delta) := \{\exists \text{ eigenvector } v \in S_{\mathbb{R}}^{n-1} \text{ of } A, \exists I \subset [n], |I| = m, \|v_I\|_2 < \delta\}.$$

In this case, we have the following analogue of Proposition 3.2.

**Proposition 3.3** (Reduction of delocalization to invertibility for real eigenvectors). *Let  $A$  be an  $n \times n$  real random matrix with arbitrary distribution. Let  $M \geq 1$ ,  $\delta \in (0, 1/2)$ ,  $p_0 \in (0, 1)$ , and  $m \in [n]$ . Assume that for any set  $I \subset [n]$  with  $|I| = m$  and any  $\lambda_0 \in \mathbb{R}$  with  $|\lambda_0| \leq M\sqrt{n}$ , we have*

$$(3.6) \quad \mathbb{P}(s_{\min}((A - \lambda_0)_{I^c}) \leq 6\delta M\sqrt{n} \text{ and } \|A\| \leq M\sqrt{n}) \leq p_0.$$

Then

$$\mathbb{P}(\text{Loc}_{\mathbb{R}}(A, m, \delta) \text{ and } \|A\| \leq M\sqrt{n}) \leq \frac{3}{\delta} \left(\frac{ne}{m}\right)^m p_0.$$

*Proof.* The proof follows a similar argument as the proof of Proposition 3.2. Suppose  $\|A\| \leq M\sqrt{n}$  and the localization event  $\text{Loc}_{\mathbb{R}}(A, m, \delta)$  holds. Then there exists an eigenvector  $v \in S_{\mathbb{R}}^{n-1}$  of  $A$  and an index set  $I \subset [n]$  with  $|I| = m$  such that  $\|v_I\|_2 < \delta$ . Let  $\lambda$  be the eigenvalue of  $A$  corresponding to the eigenvector  $v$ . Since the matrix  $A$  has real entries, the eigenvalue  $\lambda$  must be real. In addition,  $|\lambda| \leq \|A\| \leq M\sqrt{n}$ . Let  $\mathcal{N}$  be a  $(\delta M\sqrt{n})$ -net of the

real interval  $[-M\sqrt{n}, M\sqrt{n}]$ . A simple volume argument shows that one can construct the net  $\mathcal{N}$  to have cardinality

$$(3.7) \quad |\mathcal{N}| \leq \frac{3}{\delta}.$$

Choose  $\lambda_0 \in \mathcal{N}$  such that  $|\lambda - \lambda_0| \leq \delta M\sqrt{n}$ . Then the eigenvalue-eigenvector equation  $Av = \lambda v$  implies

$$(A - \lambda_0)v = (\lambda - \lambda_0)v,$$

and hence

$$\|(A - \lambda_0)v\|_2 \leq |\lambda - \lambda_0| \leq \delta M\sqrt{n}.$$

To summarize, we have shown that the events  $\|A\| \leq M\sqrt{n}$  and  $\text{Loc}_{\mathbb{R}}(A, m, \delta)$  imply the existence of  $\lambda_0 \in \mathcal{N}$  such that  $\text{Loc}_{\lambda_0}(A, M, m, \delta)$  holds. We conclude from Proposition 3.1 and the union bound that

$$\mathbb{P}(\text{Loc}(A, m, \delta) \text{ and } \|A\| \leq M\sqrt{n}) \leq |\mathcal{N}| \left(\frac{ne}{m}\right)^m p_0.$$

Combining this bound with (3.7) completes the proof.  $\square$

**3.2. Least singular value of rectangular matrices.** In order to apply Propositions 3.1, 3.2, and 3.3, we will need bounds on the least singular value of genuinely complex random matrices. These bounds are the key technical achievement of this paper. Indeed, the results below provide an analogue of the main result in [41] for genuinely complex random matrices.

**Theorem 3.4.** *Let  $A$  be an  $N \times n$  random genuinely complex matrix,  $N \geq n$ , and  $\lambda \in \mathbb{C}$  with  $|\lambda| \leq M\sqrt{N}$  for  $M \geq 1$ . Then, for every  $\varepsilon > 0$ , we have*

$$(3.8) \quad \mathbb{P}\left(s_n(A - \lambda) \leq \varepsilon(\sqrt{N} - \sqrt{n-1})\right) \leq (C\varepsilon)^{2(N-n+1)-1} + e^{-cN}$$

where  $C, c > 0$  depend (polynomially) only on the subgaussian moment  $B$  and  $M$ .

*Remark.* Note that in [41], the upperbound in (3.8) for real random matrices and  $\lambda = 0$  is of the form  $(C\varepsilon)^{N-n+1} + e^{-cN}$ . Essentially, we have replaced this  $\varepsilon$  in the real case with  $\varepsilon^2$  in the genuinely complex case. The right hand side is near optimal up to a factor of  $\varepsilon$ .

By slightly altering the proof of Theorem 3.4, we are able to prove a bound that is more effective in the regime where  $N - n$  is small.

**Theorem 3.5.** *Let  $A$  be an  $N \times n$  random genuinely complex matrix,  $N \geq n \geq N - T$ , for some integer  $T$ . Consider  $\lambda \in \mathbb{C}$  with  $|\lambda| \leq M\sqrt{N}$  for  $M \geq 1$ . Then, for every  $\varepsilon > 0$ , we have*

$$(3.9) \quad \mathbb{P}\left(s_n(A - \lambda) \leq \varepsilon(\sqrt{N} - \sqrt{n-1})\right) \leq (C\sqrt{T}\varepsilon)^{2(N-n+1)} + e^{-cN}$$

where  $C, c > 0$  depend (polynomially) only on the subgaussian moment  $B$  and  $M$ .

*Remark.* Note that when  $T = O(1)$ , we recover the optimal bound.

## 4. PROOF OF THEOREM 3.4

Our proof follows [41] closely and also combines several ideas from [33]. We mirror the notation from [41, 33] for ease of comparison.

### 4.1. Preliminaries.

4.1.1. *Nets.* Consider a subset  $D$  of  $\mathbb{C}^n$ , and let  $\varepsilon > 0$ . Recall that an  $\varepsilon$ -net of  $D$  is a subset  $\mathcal{N} \subseteq D$  such that for every  $x \in D$  one has  $\text{dist}(x, \mathcal{N}) \leq \varepsilon$ .

The following lemma is the complex analogue of Proposition 2.1 in [41]. The proof is identical to the real case if one identifies  $\mathbb{C}^n$  with  $\mathbb{R}^{2n}$ .

**Proposition 4.1** (Nets). *Let  $S$  be a subset of  $S_{\mathbb{C}}^{n-1}$ , and let  $\varepsilon > 0$ . Then there exists an  $\varepsilon$ -net of  $S$  of cardinality at most*

$$4n \left(1 + \frac{2}{\varepsilon}\right)^{2n-1}.$$

Using the standard net argument, one can show the following bound on the operator norm of rectangular matrices.

**Proposition 4.2.** *Let  $A$  be an  $N \times n$  genuinely complex random matrix, with  $N \geq n$  and  $\lambda \in \mathbb{C}$  with  $|\lambda| \leq M\sqrt{N}$  for  $M \geq 1$ . Then*

$$\mathbb{P}(\|A - \lambda\| > t\sqrt{N}) \leq e^{-c_0 t^2 N} \quad \text{for } t \geq C_0,$$

where  $C_0, c_0 > 0$  depend only on the subgaussian moment  $B$  and  $M$ .

4.1.2. *Converting between  $\mathbb{R}$  and  $\mathbb{C}$ .* Following [33], for a vector  $v \in \mathbb{C}^n$ , we denote by  $\underline{v}$  its associated real vector defined to be

$$\underline{v} := \begin{pmatrix} \Re(v) \\ \Im(v) \end{pmatrix}$$

and  $[v]$  denote its associated  $2n \times 2$  real matrix defined to be

$$[v] := \begin{pmatrix} \Re(v) & -\Im(v) \\ \Im(v) & \Re(v) \end{pmatrix}$$

We generalize this notion from [33] to include matrices. For a  $M \in \mathcal{M}_{n \times m}^{\mathbb{C}}$  matrix with

$$M = A + iB,$$

where  $A, B \in \mathcal{M}_{n \times m}^{\mathbb{R}}$ , we define  $[M] \in \mathcal{M}_{2n \times 2m}^{\mathbb{R}}$  to be

$$[M] := \begin{pmatrix} A & -B \\ B & A \end{pmatrix}.$$

We record some useful consequences of these definitions below.

**Lemma 4.3.** *For  $a \in \mathbb{C}$ ,  $x, y \in \mathbb{C}^n$  and  $M \in \mathcal{M}_{m \times n}^{\mathbb{C}}$ ,*

$$\|x - y\|_2 = \|\underline{x} - \underline{y}\|_2,$$

$$\underline{ax} = [x]\underline{a},$$

$$\underline{Mx} = [M]\underline{x}.$$

4.1.3. *Decomposition of the unit sphere.* In our proof of Theorem 3.4, we utilize a partition of the unit sphere due to Rudelson and Vershynin [40].

**Definition 4.4.** Let  $\delta, \rho \in (0, 1)$ . A vector  $x \in \mathbb{C}^n$  is *sparse* if  $|\text{supp}(x)| \leq \delta n$ . A vector  $x \in \mathbb{C}^n$  is *compressible* if there exists a sparse vector  $y$  such that  $\|x - y\|_2 \leq \rho$ . A vector  $x \in \mathbb{C}^n$  is *incompressible* if it is not compressible. We denote the sets of compressible and incompressible vectors by  $\text{Comp}(\delta, \rho)$  and  $\text{Incomp}(\delta, \rho)$  respectively.

We now recall two simple results.

**Lemma 4.5** (Lemma 5.3, [33]). *Let  $x \in \text{Incomp}(\delta, \rho)$ . Then there exists a set  $\sigma \subseteq [n]$  of cardinality  $|\sigma| \geq \nu_1 n$  and such that*

$$\frac{\nu_2}{\sqrt{n}} \leq |x_k| \leq \frac{\nu_3}{\sqrt{n}} \quad \text{for all } k \in \sigma$$

where  $0 < \nu_1, \nu_2, \nu_3$  are constants depending only on  $\delta$  and  $\rho$ .

The next lemma controls the norm of the images of compressible vectors. We omit the proof which is a straightforward adaptation of Section 2.2 in [40].

**Lemma 4.6.** *Let  $A$  be a  $N \times n$  genuinely complex random matrix,  $N \geq n/2$  and  $\lambda \in \mathbb{C}$  with  $|\lambda| \leq M\sqrt{N}$  for  $M \geq 1$ . There exist  $\delta, \rho, c_3$  depending only on the subgaussian moment  $B$  and  $M$  such that*

$$\mathbb{P}\left(\inf_{x \in \text{Comp}(\delta, \rho)} \|(A - \lambda)x\|_2 \leq c_3\sqrt{N}\right) \leq e^{-c_3 N}.$$

**4.2. Small ball probability and arithmetic structure in  $\mathbb{R}$ .** At several points in the proof of Theorem 3.4, we will need quantitative control on the spread of a random variable.

**Definition 4.7.** The *Lévy concentration function* of a random vector  $S \in \mathbb{R}^m$  (or  $\mathbb{C}^m$ ) is defined for  $\varepsilon > 0$  as

$$\mathcal{L}(S, \varepsilon) = \sup_{v \in \mathbb{R}^m \text{ (or } \mathbb{C}^m)} \mathbb{P}(\|S - v\|_2 \leq \varepsilon).$$

Below we recount several results for real random variables.

**Lemma 4.8** (Lemma 2.6, [41]). *Let  $\xi$  be a real random variable with mean zero, unit variance, and finite fourth moment. Then for every  $\varepsilon \in (0, 1)$ , there exists  $p \in (0, 1)$  which depends only on  $\varepsilon$  and on the fourth moment, and such that*

$$\mathcal{L}(\xi, \varepsilon) \leq p.$$

This rather crude bound can be significantly improved when more is known about the random variable. In particular, a well-developed theory exists when  $S = \sum_{k=1}^N a_k \xi_k$  where  $a_k$  are fixed vectors and  $\xi_k$  are independent random variables. This question is the basis of Littlewood-Offord theory and the situation when  $a_k$  are scalars has a long history in random matrix theory [54, 40]. The fundamental observation in Littlewood-Offord theory is that the Lévy concentration function is dependent on the additive structure of the coefficients,  $a_k$ . For the scalar case, Rudelson and Vershynin [40] defined the essential least common denominator for the vector of coefficients,  $a = (a_1, \dots, a_N)$ , to be

$$\text{LCD}_{\alpha, \gamma}(a) := \inf \{ \theta > 0 : \text{dist}(\theta a, \mathbb{Z}^N) < \min(\gamma \|\theta a\|_2, \alpha) \}$$

which roughly captures the length of the shortest arithmetic progression in which  $a$  can be embedded.

In [41], Rudelson and Vershynin generalized this notion to higher dimensions. If we now allow  $a = (a_1, \dots, a_N)$  to be a sequence of vectors  $a_k \in \mathbb{R}^m$ , then we define the product of such a multi-vector  $a$  and a vector  $\theta \in \mathbb{R}^m$  as

$$\theta \cdot a = (\langle \theta, a_1 \rangle, \dots, \langle \theta, a_N \rangle) \in \mathbb{R}^N.$$

Then we define, for  $\alpha > 0$  and  $\gamma \in (0, 1)$ ,

$$\text{LCD}_{\alpha, \gamma}(a) := \inf \{ \|\theta\|_2 : \theta \in \mathbb{R}^m, \text{dist}(\theta \cdot a, \mathbb{Z}^N) < \min(\gamma \|\theta \cdot a\|_2, \alpha) \}$$

The following theorem provides a bound on the small ball probability in terms of this generalized essential least common denominator.

**Theorem 4.9** (Theorem 3.3, [41]). *Let  $a \in (a_1, \dots, a_N)$  be a sequence of vectors  $a_k \in \mathbb{R}^m$  which satisfy*

$$\sum_{k=1}^N \langle a_k, x \rangle^2 \geq \|x\|_2^2 \quad \text{for every } x \in \mathbb{R}^m.$$

*Let  $\xi_1, \dots, \xi_N$  be independent real random variables, such that  $\mathcal{L}(\xi_k, 1) \leq 1 - b$  for some  $b > 0$ . Consider the random sum  $S = \sum_{k=1}^N a_k \xi_k$ . Then, for every  $\alpha > 0$  and  $\gamma \in (0, 1)$ , and for*

$$\varepsilon \geq \frac{\sqrt{m}}{\text{LCD}_{\alpha, \gamma}(a)},$$

*we have*

$$\mathcal{L}(S, \varepsilon \sqrt{m}) \leq \left( \frac{C\varepsilon}{\gamma \sqrt{b}} \right)^m + C^m e^{-2b\alpha^2}.$$

*Remark.* In [41], the statement of the theorem requires identically distributed, mean zero random variables, but the proof (which begins with symmetrization anyways) can be easily altered to handle random variables with arbitrary and possibly different means. The identical distribution requirement can also be relaxed as long as the random variables have unit variance and a uniform bound on the subgaussian moment.

**4.3. Arithmetic structure in  $\mathbb{C}$ .** In [33], the first author generalized the notion of essential least common denominator to the complex setting.

**Definition 4.10.** If we let  $a = (a_1, \dots, a_N)$  be a vector of complex numbers, we define the essential least common denominator of  $a$  to be

$$\text{LCD}_{\alpha, \gamma}(a) := \inf \{ \|\theta\|_2 : \theta \in \mathbb{R}^2, \text{dist}([a]\theta, \mathbb{Z}^{2N}) < \min(\gamma \|a\theta\|, \alpha) \}$$

By Lemma 4.3, an equivalent definition is

$$\text{LCD}_{\alpha, \gamma}(a) := \inf \{ |\theta| : \theta \in \mathbb{C}, \text{dist}(\underline{\theta}a, \mathbb{Z}^{2N}) < \min(\gamma |a\theta|, \alpha) \}$$

We extend this definition to higher dimensions below.

**Definition 4.11.** Let  $a = (a_1, \dots, a_N)$  be a sequence of vectors  $a_k \in \mathbb{C}^m$ . Then we define, for  $\alpha > 0$  and  $\gamma \in (0, 1)$ ,

$$\text{LCD}_{\alpha, \gamma}(a) := \inf \{ \|\theta\|_2 : \theta \in \mathbb{R}^{2m}, \text{dist}([A]^T \theta, \mathbb{Z}^{2N}) < \min(\gamma \|\theta \cdot a\|_2, \alpha) \}$$

where  $A$  is the matrix with columns  $a_k$ . An equivalent, more geometric, definition is the following:

$$\text{LCD}_{\alpha, \gamma}(a) := \inf \{ \|\theta\|_2 : \theta \in \mathbb{C}^m, \text{dist}(\underline{\theta} \cdot a, \mathbb{Z}^{2N}) < \min(\gamma \|\theta \cdot a\|_2, \alpha) \}$$

where we define the product of such a multi-vector  $a$  and a vector  $\theta \in \mathbb{C}^m$  as

$$\theta \cdot a = (\langle \theta, a_1 \rangle, \dots, \langle \theta, a_N \rangle) \in \mathbb{C}^N.$$

*Remark.* Note that the first definition makes it clear that the LCD of  $N$  complex vectors can be related to the LCD of  $2N$  real vectors (the  $2N$  columns of  $[A]$ ). This allows us to use Theorem 4.9 in the complex setting.

**4.4. Least common denominator of incompressible vectors.** We recall a lemma from [33] which provides a lower bound on the LCD of incompressible vectors.

**Lemma 4.12** (Lemma 5.12, [33]). *There exist constants  $\gamma, \lambda > 0$  only depending on  $\delta$  and  $\rho$  such that for any incompressible vector  $x \in \text{Incomp}(\delta, \rho)$  one has for every  $\alpha > 0$ ,*

$$\text{LCD}_{\alpha, \gamma}(x) \geq \lambda \sqrt{n}.$$

**4.5. Distance to subspaces and arithmetic structure.** In this section, we utilize the arithmetic structure of subspaces to control the distances of random vectors to random subspaces. It is in this section that we exploit having a genuinely complex random matrix and we gain the extra factor of  $\varepsilon$  to  $\varepsilon^2$  as compared to the real case. In particular, we show the following optimal bound on the distance of a random vector to a random subspace.

**Theorem 4.13** (Distance to random subspace). *Let  $X$  be a vector in  $\mathbb{C}^N$  whose coordinates are genuinely complex (but not necessarily centered) and independent. Let  $H$  be a random subspace in  $\mathbb{C}^N$  spanned by  $N - m$  genuinely complex random vectors (not necessarily centered) independent of  $X$ , with  $0 < m < \tilde{c}N$ . Then, for every  $v \in \mathbb{C}^N$  and every  $\varepsilon > 0$ , we have*

$$\mathbb{P}(\text{dist}(X, H + v) < \varepsilon\sqrt{m}) \leq (C\varepsilon)^{2m} + e^{-cN},$$

where  $C, c, \tilde{c} > 0$  depend only on the subgaussian moment  $B$ .

We deduce Theorem 4.13 via a covering argument that first requires a bound that holds for a fixed subspace and depends on the arithmetic structure of that subspace. For  $\alpha > 0$  and  $\gamma \in (0, 1)$ , we define the *essential least common denominator of a subspace  $E$*  in  $\mathbb{C}^N$  to be

$$\text{LCD}_{\alpha, \gamma} := \inf \{ \text{LCD}_{\alpha, \gamma}(a) : a \in S(E) \}$$

where  $S(E)$  denotes the intersection of the unit sphere with  $E$ . One can see that this is equivalent to

$$\text{LCD}_{\alpha, \gamma} = \inf \{ \|\theta\|_2 : \theta \in E, \text{dist}(\underline{\theta}, \mathbb{Z}^{2N}) < \min(\gamma\|\theta\|_2, \alpha) \}.$$

We now combine this notion with Theorem 4.9 to yield the following bound on the distance.

**Theorem 4.14** (Distance to a general subspace). *Let  $X$  be a genuinely complex random vector (not necessarily centered) in  $\mathbb{C}^N$ . Let  $H$  be a subspace in  $\mathbb{C}^N$  of dimension  $N - m > 0$ . Then for every  $v \in \mathbb{C}^N$ ,  $\alpha > 0$ ,  $\gamma \in (0, 1)$ , and for*

$$\varepsilon \geq \frac{\sqrt{m}}{\text{LCD}_{\alpha, \gamma} H^\perp},$$

we have

$$\mathbb{P}(\text{dist}(X, H + v) < \varepsilon\sqrt{m}) \leq \left( \frac{C\varepsilon}{\gamma} \right)^{2m} + C^m e^{-c\alpha^2}$$

where  $C, c > 0$  depend only on the subgaussian moment  $B$ .

*Proof.* We write  $\underline{X}$  in coordinates,  $\underline{X} = (\xi_1, \dots, \xi_N, \xi_{N+1}, \dots, \xi_{2N})$ . By Lemma 4.8, each coordinate of  $\underline{X}$  satisfies  $\mathcal{L}(\xi_k, 1/2) \leq 1 - b$  for some  $b > 0$  that only depends on the subgaussian moment  $B$ . Thus, the random variables  $\xi_k/2$  satisfy the assumptions of Theorem 4.9.

Now, we convert the distance problem into a small ball probability calculation for a sum of independent vectors. Let  $P_H$  signify the orthogonal projection onto a subspace  $H$ .

$$(4.1) \quad \text{dist}(X, H + v) = \|P_{H^\perp}(X - v)\|_2 = \left\| \sum_{k=1}^{2N} a_k \xi_k - w \right\|_2,$$

where

$$a_k = \begin{pmatrix} \Re(P_{H^\perp} e_k) \\ \Im(P_{H^\perp} e_k) \end{pmatrix}, \quad a_{N+k} = \begin{pmatrix} -\Im(P_{H^\perp} e_k) \\ \Re(P_{H^\perp} e_k) \end{pmatrix}, \quad w = \underline{P_{H^\perp} v}$$

for  $1 \leq k \leq N$ . For this sequence of vectors  $a = (a_1, \dots, a_{2N})$ , we have

$$\sum_{k=1}^{2N} \langle a_k, \underline{x} \rangle^2 = \sum_{k=1}^N |\langle P_{H^\perp} e_k, x \rangle|^2 = \sum_{k=1}^N |\langle e_k, x \rangle|^2 = \|x\|_2^2 \quad \text{for any } x \in H^\perp$$

so we can apply Theorem 4.9 in the space  $H^\perp$  (which can be identified with  $\mathbb{C}^m$  under a suitable isometry).

For any  $\theta \in H^\perp$  we have  $\langle \theta, P_{H^\perp} e_k \rangle = \langle P_{H^\perp} \theta, e_k \rangle = \langle \theta, e_k \rangle$  so by Lemma 4.3,

$$\begin{pmatrix} a_k^T \\ a_{N+k}^T \end{pmatrix} \bar{\theta} = [(P_{H^\perp} e_k)^T] \bar{\theta} = \underline{(P_{H^\perp} e_k)^T \bar{\theta}} = \underline{(P_{H^\perp} e_k, \theta)} = \overline{\langle \theta, e_k \rangle}.$$

As conjugation will not alter the norm, we have

$$\text{LCD}_{\alpha, \gamma}(H^\perp) = \text{LCD}_{\alpha, \gamma}(a).$$

The result now follows from a direct application of Theorem 4.9.  $\square$

To prove the distance bound we carry out a covering argument to exclude those possible  $H^\perp$  with small LCD of a random subspace  $H^\perp$ . In fact, we show that the LCD of such a subspace is typically exponentially large.

**Theorem 4.15** (Structure of a random subspace). *Let  $H$  be a random subspace in  $\mathbb{C}^N$  spanned by  $N - m$  genuinely complex random vectors,  $1 \leq m \leq \tilde{c}N$ . Then, for  $\alpha = c\sqrt{N}$ , we have*

$$\mathbb{P}(\text{LCD}_{\alpha, c}(H^\perp) < c\sqrt{N}e^{cN/m}) \leq e^{-cN},$$

where  $c \in (0, 1)$  and  $\tilde{c} \in (0, 1/2)$  depend only on the subgaussian moment  $B$ .

For now, if we assume this result, we can complete the proof of Theorem 4.13.

*Proof of Theorem 4.13.* Consider the event

$$\mathcal{E} := \{\text{LCD}_{\alpha, c}(H^\perp) \geq c\sqrt{N}e^{cN/m}\}.$$

By Theorem 4.15,  $\mathbb{P}(\mathcal{E}^c) \leq e^{-cN}$ . We now condition on a realization of  $H$  in  $\mathcal{E}$ . By the independence of  $H$  and  $X$ , Theorem 4.14 applied with  $\alpha = c\sqrt{N}$  and  $\gamma = c$  yields

$$\mathbb{P}(\text{dist}(X, H) < \varepsilon\sqrt{m} | \mathcal{E}) \leq (C_1\varepsilon)^{2m} + C^m e^{-c_1 N}$$

for any

$$\varepsilon > C_2 \sqrt{\frac{m}{N}} e^{-cN/m}.$$

Since  $m \leq \tilde{c}N$ , for an appropriate choice of  $\tilde{c}$  we have

$$C_2 \sqrt{\frac{m}{N}} e^{cN/m} \leq \frac{1}{C_1} e^{-c_3 N/m} \quad \text{and} \quad C^m e^{-c_1 N} \leq e^{-c_3 N}.$$

Thus, for every  $\varepsilon > 0$ ,

$$\mathbb{P}(\text{dist}(X, H) < \varepsilon\sqrt{m} | \mathcal{E}) \leq (C_1\varepsilon)^{2m} + 2e^{-c_3 N} \leq (C_1\varepsilon)^{2m} + e^{-c_4 N}$$

$\square$

4.5.1. *Proof of Structure Theorem 4.15.* Throughout the proof we assume that  $N > N_0$  for some suitably large number  $N_0$  which only depends on the subgaussian moment  $B$ . Indeed, the assumption on  $m$  implies that  $N > 1/\tilde{c}$ . Thus, choosing  $\tilde{c}$  small enough, we can make  $N_0$  suitably large.

Let  $X_1, \dots, X_{N-m}$  denote the independent random vectors that span the subspace  $H$ . Consider the  $(N-m) \times N$  random matrix  $B$  with rows  $\bar{X}_k$ . Then

$$H^\perp \subseteq \ker(B).$$

Therefore, for every set  $S$  in  $\mathbb{C}^N$  we have:

$$(4.2) \quad \inf_{x \in S} \|Bx\|_2 > 0 \text{ implies } H^\perp \cap S = \emptyset.$$

This observation reduces the intersection problem to bounding the infimum of the image of  $S$  under  $B$ .

We now show that a typical subspace is entirely contained in  $Incomp(\delta, \rho)$ .

**Lemma 4.16.** *There exist  $\delta, \rho \in (0, 1)$  such that*

$$\mathbb{P}(H^\perp \cap S^{n-1} \subseteq Incomp(\delta, \rho)) \geq 1 - e^{-cN}.$$

*Proof.* Since  $N - m > (1 - \tilde{c})N$  and  $\tilde{c} < 1/2$ , we can apply Lemma 4.6 for the matrix  $B$ . Therefore, there exist  $\delta, \rho \in (0, 1)$  such that

$$\mathbb{P}\left(\inf_{x \in Comp(\delta, \rho)} \|Bx\|_2 \geq c_3\sqrt{N}\right) \geq 1 - e^{-c_3N}.$$

Thus, by (4.2),  $H^\perp \cap Comp(\delta, \rho) = \emptyset$  with probability at least  $1 - e^{-c_3N}$ .  $\square$

Fix the values of  $\delta$  and  $\rho$  for the rest of this section. We decompose the incompressible vectors into level sets,  $S_D$  by the value of the essential least common denominator. For each level set except those where  $D$  is exponentially large, we show that  $\inf_{x \in S_D} \|Bx\|_2 > 0$ .

Let  $\alpha = \mu\sqrt{N}$ , where  $\mu > 0$  is a small number to be chosen later, which depends only on the subgaussian moment  $B$ . By Lemma 4.12,

$$LCD_{\alpha, c} \geq c_0\sqrt{N} \quad \text{for every } x \in Incomp(\delta, \rho).$$

**Definition 4.17** (Level Sets). Let  $D \geq c_0\sqrt{N}$ . Define  $S_D \subseteq S_{\mathbb{C}}^{n-1}$  as

$$S_D := \{x \in Incomp : D \leq LCD_{\alpha, c}(x) < 2D\}.$$

We first derive a lower bound for  $\|Bx\|_2$  for a fixed vector  $x$ .

**Lemma 4.18.** *Let  $x \in S_D$ . Then for every  $t > 0$  we have*

$$(4.3) \quad \mathbb{P}(\|Bx\|_2 < t\sqrt{N}) \leq \left(Ct^2 + \frac{C}{D} + Ce^{-c\alpha^2}\right)^{N-m}$$

*Proof.* We examine the coordinates of  $Bx$ .

$$\underline{(Bx)}_j = [x]^T \underline{B}_j^T = \sum_{k=1}^{2N} \xi_k [x]_j.$$

Since  $x \in S_{\mathbb{C}}^{N-1}$ , we have  $\sum_{k=1}^{2N} \langle [x]_j, y \rangle^2 = \|y\|_2^2$  for every  $y \in \mathbb{R}^2$ . We can apply Theorem 4.9 with  $m = 2$ .

$$\mathbb{P}(|(Bx)_j| < t) \leq Ct^2 + \frac{C}{D} + Ce^{-c\alpha^2}.$$



Since the rows of  $B$  are independent, we can use the Tensorization Lemma 2.2 of [40] to conclude that

$$\mathbb{P}\left(\sum_{j=1}^{N-m} |(Bx)_j|^2 \leq t^2(N-m)\right) \leq \left(C''t^2 + \frac{C''}{D} + C''e^{-c\alpha^2}\right)^{N-m}.$$

This completes the proof since  $\sum_{j=1}^{N-m} |(Bx)_j|^2 = \|Bx\|_2^2$ .  $\square$

We recall the following bound from [33] on the size of an  $\varepsilon$ -net of a level set.

**Lemma 4.19** (Lemma 5.14, [33]). *There exists a  $(2\alpha/D)$ -net of  $S_D$  of cardinality at most  $D^2(C_0D/\sqrt{N})^{2N}$ .*

Using this bound on the net size and our anti-concentration for a single vector, we can generate a lower bound for an entire level set.

**Lemma 4.20** (Lower bound for a level set). *There exist  $c_1, c_2, \mu \in (0, 1)$  such that the following holds. Let  $\alpha = \mu\sqrt{N} \geq 1$  and  $D \leq c_1\sqrt{N}e^{c_1N/m}$ . Then*

$$\mathbb{P}\left(\inf_{x \in S_D} \|Bx\|_2 < c_2N/D\right) \leq 2e^{-N}.$$

*Proof.* By Proposition 4.2, there exists  $K \geq 1$  such that

$$\mathbb{P}(\|B\| > K\sqrt{N}) \leq e^{-N}.$$

To complete the proof, it suffices to find  $\nu > 0$  which depends only on  $B$  such that the event

$$\mathcal{E} := \left\{ \inf_{x \in S_D} \|Bx\|_2 < \frac{\nu N}{2D} \text{ and } \|B\| \leq K\sqrt{N} \right\}$$

has probability at most  $e^{-N}$ .

We verify that this holds with the following choice of parameters:

$$\nu = \frac{1}{(3CC_0)^3e}, \quad \mu = \frac{\nu}{9K}, \quad c_1 = c\mu^2 \leq \nu.$$

Choosing  $\tilde{c}$  in the statement of Theorem 4.15 to be sufficiently small, we can assume that  $N > \nu^{-2}$ . We apply Lemma 4.18 with  $t = \nu\sqrt{N}/D$ . By our choice of parameters, the  $Ct$  term dominates in the right hand side of (4.3). Therefore, for  $x_0 \in S_D$ ,

$$\mathbb{P}\left(\|Bx_0\|_2 < \frac{\nu N}{D}\right) \leq \left(\frac{3C\nu\sqrt{N}}{D}\right)^{2(N-m)}.$$

By Lemma 4.19, there exists a  $(2\alpha/D)$ -net,  $\mathcal{N}$ , of size at most  $D^2(C_0D/\sqrt{N})^{2N}$

$$p := \mathbb{P}\left(\inf_{x_0 \in \mathcal{N}} \|Bx_0\|_2 < \frac{\nu N}{D}\right) \leq D^2(C_0D/\sqrt{N})^{2N} \left(\frac{3C\nu\sqrt{N}}{D}\right)^{2(N-m)}.$$

Denote  $C_1 := 3CC_0$ .

$$p \leq C_1^{2N} D^2 \left(\frac{D}{\sqrt{N}}\right)^{2m} \nu^{2(N-m)} \leq C_1^{2N} D^2 (\nu e^{\nu N/m})^{2m} \nu^{N-m} \leq C_1^{3N} \nu^N = e^{-N}.$$

We assume that  $\mathcal{E}$  occurs. Fix a  $x \in S_D$  for which  $\|Bx\|_2 < \frac{\nu N}{2D}$ . There exists an element  $x_0 \in \mathcal{N}$  such that  $\|x - x_0\|_2 \leq \frac{2\mu\sqrt{N}}{D}$ . Therefore, by the triangle inequality,

$$\|Bx_0\|_2 \leq \|Bx\|_2 + \|B\|\|x - x_0\|_2 \leq \frac{\nu N}{2D} + K\sqrt{N} \frac{2\mu\sqrt{N}}{D} < \frac{\nu N}{D}.$$

□

*Proof of Theorem 4.15.* Consider  $x \in S_{\mathbb{C}}^{N-1}$  such that

$$\text{LCD}_{\alpha,c}(x) < c_1 \sqrt{N} e^{c_1 N/m},$$

where  $c_1$  is the constant from Lemma 4.20. Either  $x$  is compressible or  $x \in S_D$  for some  $D \in \mathcal{D}$ , where

$$\mathcal{D} := \{D : c_0 \sqrt{N} \leq D < c_1 \sqrt{N} e^{c_1 N/m}, D = 2^k, k \in \mathbb{N}\}.$$

We can now decompose the desired probability as

$$\begin{aligned} p &:= \mathbb{P}(\text{LCD}_{\alpha,c}(H^\perp) < c_1 \sqrt{N} e^{c_1 N/m}) \\ &\leq \mathbb{P}(H^\perp \cap \text{Comp} \neq \emptyset) + \sum_{D \in \mathcal{D}} \mathbb{P}(H^\perp \cap S_D \neq \emptyset). \end{aligned}$$

By Lemma 4.16, the first term on the right is bounded by  $e^{-cN}$ . By Lemma 4.20 each term in the summation on the right can be bounded by  $2e^{-N}$ . Since  $|\mathcal{D}| \leq C'N$ , we have

$$p \leq e^{-cN} + C'N e^{-N} \leq e^{-c'N}.$$

□

**4.6. Invertibility via uniform distance bounds.** The remainder of the proof is identical to [41] and is included with the obvious modifications for the reader's convenience. We first make several reductions. Without loss of generality, we may assume that our random variables have an absolutely continuous distribution. Indeed, we can add to each entry an independent complex gaussian random variable with small variance  $\sigma$  and later let  $\sigma$  tend to zero.

Let  $N = n - 1 + d$  for some  $d \geq 1$ . We can assume that

$$(4.4) \quad 1 \leq d \leq c_0 n,$$

as when  $d$  is above a constant proportion of  $n$ , our matrix is sufficiently rectangular for a simple epsilon argument (cf. Introduction of [41]). Note that

$$\sqrt{N} - \sqrt{n-1} \leq \frac{d}{\sqrt{n}}.$$

Therefore,

$$\begin{aligned} (4.5) \quad \mathbb{P}\left(s_n(A - \lambda) \leq \varepsilon(\sqrt{N} - \sqrt{n-1})\right) &\leq \mathbb{P}\left(s_n(A - \lambda) \leq \varepsilon \frac{d}{\sqrt{n}}\right) \\ &\leq \mathbb{P}\left(\inf_{x \in \text{Comp}(\delta, \rho)} \|(A - \lambda)x\|_2 \leq \varepsilon \frac{d}{\sqrt{n}}\right) + \mathbb{P}\left(\inf_{x \in \text{Incomp}(\delta, \rho)} \|(A - \lambda)x\|_2 \leq \varepsilon \frac{d}{\sqrt{n}}\right). \end{aligned}$$

We can conclude from Lemma 4.6 that

$$(4.6) \quad \mathbb{P}\left(\inf_{x \in \text{Comp}(\delta, \rho)} \|(A - \lambda)x\|_2 \leq \varepsilon \frac{d}{\sqrt{n}}\right) \leq e^{-c_3 N}.$$

Therefore, in this section, we focus on a lower bound for incompressible vectors.

Let  $X_1, \dots, X_n \in \mathbb{C}^N$  be the columns of the matrix  $A$ . Given a subset  $J \subseteq [n]$  of cardinality  $d$ , we consider the subspace

$$H_J := \text{span}(X_k)_{k \in J} \subseteq \mathbb{C}^N.$$

For levels  $K_1, K_2 > 0$  that only depend on  $\delta, \rho$ , we define the set of totally spread vectors

$$(4.7) \quad \text{Spread}_J := \left\{ y \in S(\mathbb{C}^J) : \frac{K_1}{\sqrt{d}} \leq |y_k| \leq \frac{K_2}{\sqrt{d}} \text{ for all } k \in J \right\}.$$

In the following lemma, we let  $J$  be a random subset uniformly distributed over all subsets of  $[n]$  of cardinality  $d$ . To avoid confusion, we often denote the probability and expectation over the random set  $J$  by  $\mathbb{P}_J$  and  $\mathbb{E}_J$ , and with respect to the random matrix  $A$  by  $\mathbb{P}_A$  and  $\mathbb{E}_A$ .

**Lemma 4.21** (Total spread). *For every  $\delta, \rho \in (0, 1)$ , there exist  $K_1, K_2, c_0 > 0$  which depend only on  $\delta, \rho$ , and such that the following holds. For every  $x \in \text{Incomp}(\delta, \rho)$ , the event*

$$\mathcal{E}(x) := \left\{ \frac{P_J x}{\|P_J x\|_2} \in \text{Spread}_J \quad \text{and} \quad \frac{\rho\sqrt{d}}{\sqrt{2n}} \leq \|P_J x\|_2 \leq \frac{\sqrt{d}}{\sqrt{\delta n}} \right\}$$

satisfies  $\mathbb{P}_J(\mathcal{E}(x)) > c_0^d$ .

*Proof.* Let  $\sigma \subset [n]$  be the subset from Lemma 4.5. By choosing  $c_0$  sufficiently small in (4.4), we may assume that  $d \leq |\sigma|/2$ . By Stirling's approximation,

$$\mathbb{P}_J(J \subset \sigma) = \binom{|\sigma|}{d} / \binom{n}{d} > \left(\frac{\nu_1}{e}\right)^d = c_0^d.$$

Lemma 4.5 also provides the two-sided bound on  $\|P_J x\|_2$ . Thus, we can set  $K_1 = \nu_2/\nu_3$  and  $K_2 = 1/K_1$ .  $\square$

We recall the following lemma from [41]. Although the lemma in [41] is stated for real vector spaces, the same proof carries over for complex vector spaces.

**Lemma 4.22** (Lemma 6.2, [41]). *There exist  $C_1, c_1 > 0$  which depend only on  $\delta, \rho$ , and such that the following holds. Let  $J$  be any  $d$ -element subset of  $[n]$ . Then for every  $\varepsilon > 0$*

$$(4.8) \quad \mathbb{P}\left(\inf_{x \in \text{Incomp}(\delta, \rho)} \|(A - \lambda)x\|_2 < c_1 \varepsilon \sqrt{\frac{d}{n}}\right) \leq C_1^d \cdot \mathbb{P}\left(\inf_{z \in \text{Spread}_J} \text{dist}((A - \lambda)z, H_{J^c}) < \varepsilon\right).$$

**4.7. Uniform distance bound.** In this section we bound the probability in the right hand side of (4.8) following [41].

**Theorem 4.23** (Uniform distance bound). *For every  $t > 0$ ,*

$$\mathbb{P}\left(\inf_{z \in \text{Spread}_J} \text{dist}((A - \lambda)z, H_{J^c}) < t\sqrt{d}\right) \leq (Ct)^{2d-1} + e^{-cN}.$$

Since  $H_{J^c}$  is the span of  $n - d$  independent random vectors and the distribution of the vectors is uniformly continuous, we can assume that

$$\dim(H_{J^c}) = n - d.$$

Without loss of generality, in the proof of Theorem 4.23, we can assume that

$$(4.9) \quad t \geq t_0 = e^{-\bar{c}N/d}.$$

Let us now represent the distance problem in matrix notation. Let  $P$  be the orthogonal projection in  $\mathbb{C}^N$  onto  $(H_{J^c})^\perp$ , and let

$$(4.10) \quad W := PA|_{\mathbb{C}^J}.$$

Then for every  $v \in \mathbb{C}^N$ , the following identity holds:

$$(4.11) \quad \text{dist}((A - \lambda)z, H_{J^c} + v) = \|Wz - w\|_2, \quad \text{where } w = P(v + \lambda z).$$

We omit the standard proof to the following proposition.

**Proposition 4.24** (Proposition 7.3, [40]).

$$\mathbb{P}(\|W\| > t\sqrt{d}) \leq e^{-c_0 t^2 d} \quad \text{for } t \geq C_0.$$

Having controlled the operator norm of  $W$ , we can run through the standard approximation argument to uniformly control the distance.

**Proposition 4.25.** *For every  $t$  that satisfies (4.9) we have*

$$(4.12) \quad \mathbb{P}\left(\inf_{z \in \text{Spread}_J} \|Wz - w\|_2 < t\sqrt{d} \text{ and } \|W\| \leq K_0\sqrt{d}\right) \leq (C_2 t)^{2d-1}.$$

*Proof.* Let  $\varepsilon = t/K_0$ . By Proposition 4.1, there exists an  $\varepsilon$ -net  $\mathcal{N}$  of  $\text{Spread}_J \subseteq S(\mathbb{C}^J)$  of cardinality

$$|\mathcal{N}| \leq 4d \left(a + \frac{2}{\varepsilon}\right)^{2d-1} \leq 4d \left(\frac{3K_0}{t}\right)^{2d-1}.$$

Consider the event

$$\mathcal{E} := \left\{ \inf_{z \in \mathcal{N}} \|Wz - w\|_2 < 2t\sqrt{d} \right\}.$$

Taking a union bound, we obtain

$$\mathbb{P}(\mathcal{E}) \leq |\mathcal{N}| \max_{z \in \mathcal{N}} \mathbb{P}(\|Wz - w\|_2 \leq 2t\sqrt{d}) \leq 4d \left(\frac{3K_0}{t}\right)^{2d-1} (2C_1 t)^{4d-2} \leq (C_2 t)^{2d-1}.$$

Now, suppose the event in (4.12) holds, i.e. there exists  $z' \in \text{Spread}_J$  such that

$$\|Wz' - w\|_2 < t\sqrt{d} \text{ and } \|W\| \leq K_0\sqrt{d}.$$

Choose  $z \in \mathcal{N}$  such that  $\|z - z'\|_2 \leq \varepsilon$ . Then by the triangle inequality

$$\|Wz - w\|_2 \leq \|Wz' - w\|_2 + \|W\| \|z - z'\|_2 < t\sqrt{d} + K_0\sqrt{d}\varepsilon \leq 2t\sqrt{d}.$$

□

We now invoke a proposition from [41] which allows us to decouple the behavior of  $\|W\|$  and  $\|Wz\|_2$ . The proof is a simple translation of the real version.

**Proposition 4.26** (Decoupling, Proposition 7.5, [41]). *Let  $W$  be an  $N \times d$  matrix whose columns are independent random vectors. Let  $\beta > 0$  and let  $z \in S^{d-1}$  be a vector satisfying  $|z_k| \geq \frac{\beta}{\sqrt{d}}$  for all  $k \in \{1, \dots, d\}$ . Then for every  $0 < a < b$ , we have*

$$\mathbb{P}(\|Wz\|_2 < a, \|W\| > b) \leq 2 \sup_{x \in S_c^{d-1}, w \in \mathbb{C}^N} \mathbb{P}\left(\|Wx - w\|_2 < \frac{\sqrt{2}}{\beta} a\right) \mathbb{P}\left(\|W\| > \frac{b}{\sqrt{2}}\right).$$

We apply this proposition to prove the following lemma.

**Lemma 4.27.** *Let  $W$  be a random matrix as in (4.10), where  $P$  is the orthogonal projection of  $\mathbb{C}^N$  onto the random subspace  $(H_{Jc})^\perp$ , defined as in Theorem 4.23. Then for every  $s \geq 1$  and every  $t$  that satisfies (4.9), we have*

$$(4.13) \quad \begin{aligned} & \mathbb{P}\left(\inf_{z \in \text{Spread}_J} \|Wz\|_2 < t\sqrt{d} \text{ and } sK_0\sqrt{d} < \|W\| \leq 2sK_0\sqrt{d}\right) \\ & \leq (C_3 t e^{-c_3 s^2})^{2d-1} + e^{-cN}. \end{aligned}$$

*Proof.* Let  $\varepsilon = t/2sK_0$ . By Proposition 4.1, there exists an  $\varepsilon$ -net  $\mathcal{N}$  of  $\text{Spread}_J \subset S(\mathbb{R}^J)$  of cardinality

$$|\mathcal{N}| \leq 2d \left(1 + \frac{2}{\varepsilon}\right)^{2d-1} \leq 2d \left(\frac{6sK_0}{t}\right)^{2d-1}.$$

Consider the event

$$\mathcal{E} := \left\{ \inf_{z \in \mathcal{N}} \|Wz\|_2 < 2t\sqrt{d} \text{ and } \|W\| > sK_0\sqrt{d} \right\}.$$

We condition on a realization of the subspace  $H_{J^c}$  which allows us to consider the columns of  $W$  as independent. By the definition of  $\text{Spread}_J$ , we can apply the decoupling proposition 4.26 with  $\beta = K_1$ . Applying a union bound, we have that

$$\begin{aligned} \mathbb{P}(\mathcal{E} \mid H_{J^c}) &\leq |\mathcal{N}| \cdot \max_{z \in \mathcal{N}} \mathbb{P}(\|Wz\|_2 \leq 2t\sqrt{d} \text{ and } \|W\| > sK_0\sqrt{d} \mid H_{J^c}) \\ &\leq |\mathcal{N}| \cdot 2 \max_{z \in S(\mathbb{R}^J), w \in \mathbb{R}^N} \mathbb{P}\left(\|Wz - w\|_2 < \frac{\sqrt{2}}{K_1} \cdot 2t\sqrt{d} \mid H_{J^c}\right) \\ &\quad \cdot \mathbb{P}\left(\|W\| > \frac{sK_0\sqrt{d}}{\sqrt{2}} \mid H_{J^c}\right). \end{aligned}$$

Assuming that  $\text{LCD}_{\alpha,c}(H_{J^c}^\perp) \geq c\sqrt{N}e^{cN/m}$ , where  $\alpha$  and  $c$  are as in Theorem 4.15, then by Proposition 4.24 and representation (4.11), we can conclude as in the proof of Theorem 4.13 that

$$\mathbb{P}(\mathcal{E} \mid H_{J^c}) \leq 8d \left(\frac{6sK_0}{t}\right)^{2d-1} \cdot (C't)^{4d-2} \cdot e^{-c's^2d}$$

for any  $t$  satisfying (4.9). Since  $s \geq 1$  and  $d \geq 1$ , we can use the following upperbound

$$\mathbb{P}(\mathcal{E} \mid H_{J^c}) \leq (C_3te^{-c_3s^2/2})^{2d-1}.$$

Additionally, by Theorem 4.15,

$$\begin{aligned} \mathbb{P}(\mathcal{E}) &\leq \mathbb{P}(\mathcal{E} \mid \text{LCD}_{\alpha,c}(H_{J^c}^\perp) \geq c\sqrt{N}e^{cN/m}) + \mathbb{P}(\text{LCD}_{\alpha,c}(H_{J^c}^\perp) < c\sqrt{N}e^{cN/m}) \\ &\leq (C_3te^{-c_3s^2})^{2d-1} + e^{-cN}. \end{aligned}$$

Now, suppose the event in (4.13) holds. There exists  $z' \in \text{Spread}_J$  such that

$$\|Wz'\|_2 < t\sqrt{d} \text{ and } sK_0\sqrt{d} < \|W\| \leq 2sK_0\sqrt{d}.$$

Choose  $z \in \mathcal{N}$  such that  $\|z - z'\|_2 \leq \varepsilon$ . Then by the triangle inequality

$$\|Wz\|_2 \leq \|Wz'\|_2 + \|W\| \cdot \|z - z'\|_2 < t\sqrt{d} + 2sK_0\sqrt{d} \cdot \varepsilon \leq 2t\sqrt{d}.$$

Thus,  $\mathcal{E}$  holds. The conclusion follows from the bound on the probability of  $\mathcal{E}$ .  $\square$

*Proof of Theorem 4.23.* Recall that we can safely assume (4.9) holds. Let  $k_1$  be the smallest natural number such that

$$(4.14) \quad 2^{k_1} \cdot K_0\sqrt{d} > C_0\sqrt{N},$$

where  $C_0$  and  $K_0$  are constants from Proposition 4.2 and Lemma 4.27 respectively. Summing the probability bounds from Proposition 4.25 and Lemma 4.27 for  $s = 2^k$ ,  $k = 1, \dots, k_1$ , we

find that

$$\begin{aligned} & \mathbb{P}\left(\inf_{z \in \text{Spread}_J} \|Wz\|_2 < t\sqrt{d}\right) \\ & \leq (C_2 t)^{2d-1} + \sum_{s=2^k, k=1, \dots, k_1} \left( (C_3 t e^{-c_3 s^2})^{2d-1} + e^{-cN} \right) + \mathbb{P}(\|W\| > C_0 \sqrt{N}) \\ & \leq (C_4 t)^{2d-1} + k_1 e^{-c'N} + \mathbb{P}(\|W\| > C_0 \sqrt{N}). \end{aligned}$$

By (4.14) and Proposition 4.2, the last expression is upperbounded by  $(Ct)^{2d-1} + e^{-c'N}$ .  $\square$

#### 4.8. Proofs of Theorems 3.4 and 3.5.

*Proof.* By Lemma 4.22 and Theorem 4.23, we can conclude that

$$\mathbb{P}\left(\inf_{x \in \text{Incomp}(\delta, \rho)} \|(A - \lambda)x\|_2 \leq \varepsilon \frac{d}{\sqrt{n}}\right) \leq (C\varepsilon)^{2d-1} + e^{-cN}.$$

By (4.5) this concludes the proof.  $\square$

A more direct approach suffices for the proof of Theorem 3.5. The proof is essentially identical to the square case (c.f. [40, 33]).

In this setting, we can use a more straightforward reduction to the distance problem.

**Lemma 4.28** (Lemma 3.4, [40]). *For  $\lambda \in \mathbb{C}^n$  and  $|\lambda| \leq M\sqrt{N}$ ,*

$$\mathbb{P}\left(\inf_{x \in \text{Incomp}(\delta, \rho)} \|(A - \lambda)x\|_2 < \varepsilon \rho n^{-1/2}\right) \leq \frac{1}{\delta n} \sum_{k=1}^n \mathbb{P}(\text{dist}(X_k, H_k) < \varepsilon)$$

where  $X_k$  denotes the  $k$ -th column of  $A - \lambda$  and  $H_k$  is the span of all the columns excluding the  $k$ -th.

*Remark.* The proof in [40] applies equally well in the rectangular setting.

*Proof of Theorem 3.5.* By (4.5) and Lemma 4.28, our task reduces to bounding

$$\mathbb{P}(\text{dist}(X_k, H_k) < \varepsilon).$$

By Theorem 4.13,

$$\mathbb{P}(\text{dist}(X_k, H_k) < \varepsilon) \leq (C\sqrt{T}\varepsilon)^{2(N-n+1)} + e^{-cN}.$$

$\square$

## 5. PROOF OF MAIN RESULTS

This section is dedicated to the proof of our main results in Section 2. We record the following standard bound for the spectral norm of a random matrix with independent subgaussian entries.

**Lemma 5.1.** *Let  $A$  be an  $N \times n$  genuinely complex random matrix. There exists constants  $M \geq 1$  and  $C, c > 0$  such that*

$$\mathbb{P}(\|A\| \geq M\sqrt{\max\{N, n\}}) \leq C \exp(-c \max\{N, n\}).$$

Here  $M, C, c$  depend only on the uniform subgaussian moment bound  $B$ .

*Proof.* The result essentially follows immediately from [48, Exercise 2.33], which applies only to square matrices. One can easily obtain the bound for rectangular matrices by padding the matrix with zeros to create a square matrix. Alternatively, one can apply the same net argument as in Proposition 4.2.  $\square$

We begin with the proofs of Theorems 2.1 and 2.2.

*Proof of Theorem 2.1.* Without loss of generality, assume  $1 \geq t \geq e^{-\log^2 n}$  (as the bound is trivial when  $t \geq 1$ ). Let  $M \geq 1$  be the constant from Lemma 5.1. Let  $\varepsilon, \delta$  be positive values to be chosen later, and take

$$q := \mathbb{P}(\text{Loc}(A, m, \delta) \text{ and } \|A\| \leq M\sqrt{n}).$$

Proposition 3.2 implies that

$$q \leq \frac{9}{\delta^2} \left(\frac{ne}{m}\right)^m p_0,$$

where  $p_0$  satisfies (3.4). Choose  $\delta$  in terms of  $\varepsilon$  via the following identity:

$$(5.1) \quad 6\delta M\sqrt{n} = \varepsilon(\sqrt{n} - \sqrt{n-m-1}).$$

In other words, once we specify  $\varepsilon, \delta$  will also be determined. Using Theorem 3.4, we find

$$p_0 \leq (C\varepsilon)^{2m+1} + e^{-cn},$$

and hence

$$q \ll \frac{\varepsilon}{\delta^2} \left(\frac{ne}{m} C^2 \varepsilon^2\right)^m + \frac{1}{\delta^2} \left(\frac{ne}{m}\right)^m e^{-cn}.$$

Returning to (5.1), we see

$$\delta \geq \frac{\varepsilon m}{12Mn}.$$

This implies that

$$q \ll \left(\left(\frac{n}{m}\right)^{1+2/m} e C^2 \varepsilon^{2-1/m}\right)^m + \frac{1}{\varepsilon^2} \left(\frac{n}{m}\right)^2 \left(\frac{ne}{m}\right)^m e^{-cn}.$$

We now choose  $\varepsilon$ . Indeed, take

$$(5.2) \quad \varepsilon := t^{m/(2m-1)} \left(\frac{m}{n}\right)^{(m+2)/(2m-1)},$$

and recall that this choice of  $\varepsilon$  also determines  $\delta$  by (5.1). In addition, this choice implies that

$$\left(\frac{n}{m}\right)^{1+2/m} \varepsilon^{2-1/m} = t,$$

which means

$$(5.3) \quad q \ll (C^2 et)^m + \frac{1}{\varepsilon^2} \left(\frac{n}{m}\right)^2 \left(\frac{ne}{m}\right)^m e^{-cn}.$$

We now simplify the expression for  $\varepsilon$  given in (5.2) using the fact that  $m \geq \log^2 n$ . Indeed, in this case it follows that

$$\left(\frac{m}{n}\right)^{(m+2)/(2m-1)} = \Theta\left(\left(\frac{m}{n}\right)^{1/2}\right)$$

and, using the fact that  $1 \geq t \geq e^{-\log^2 n}$ ,

$$t^{m/(2m-1)} = \Theta(\sqrt{t}).$$

We conclude that

$$(5.4) \quad \varepsilon = \Theta \left( \sqrt{t} \left( \frac{m}{n} \right)^{1/2} \right),$$

and hence

$$(5.5) \quad \delta \gg \sqrt{t} \left( \frac{m}{n} \right)^{3/2}.$$

Returning to (5.3), we use (5.4) and  $t \geq e^{-\log^2 n}$  to see that

$$q \ll (C^2 et)^m + e^{-c'n}$$

for  $m \leq c'n$ , where  $c' > 0$  is a sufficiently small constant. In conclusion, we have now shown that

$$\mathbb{P}(\text{Loc}(A, m, \delta) \text{ and } \|A\| \leq M\sqrt{n}) \ll (C^2 et)^m + e^{-c'n}$$

for some  $\delta > 0$  which satisfies (5.5). In view of Lemma 5.1, the proof is complete.  $\square$

*Proof of Theorem 2.2.* The proof is similar to the proof of Theorem 2.1. Without loss of generality assume  $1 \geq t \geq e^{-c'n}$  (as the bound is trivial when  $t \geq 1$ ) for a sufficiently small constant  $c' > 0$  to be chosen later. Let  $M \geq 1$  be the constant from Lemma 5.1. Let  $\varepsilon, \delta$  be positive values to be chosen later, and take

$$q := \mathbb{P}(\text{Loc}(A, m, \delta) \text{ and } \|A\| \leq M\sqrt{n}).$$

Proposition 3.2 implies that

$$q \leq \frac{9}{\delta^2} \left( \frac{ne}{m} \right)^m p_0,$$

where  $p_0$  satisfies (3.4). Set  $\delta$  in terms of  $\varepsilon$  again via (5.1), so that  $\delta$  is determined completely once we select  $\varepsilon$ . Using Theorem 3.5 and the bound  $m \leq \log^2 n$ , we find that

$$p_0 \leq (C\varepsilon \log n)^{2(m+1)} + e^{-cn}.$$

Thus, we have

$$q \ll (\log n)^2 \frac{\varepsilon^2}{\delta^2} \left( \frac{neC^2 \log^2 n \varepsilon^2}{m} \right)^m + \frac{1}{\delta^2} \left( \frac{ne}{m} \right)^m e^{-cn}.$$

From (5.1), we see that

$$(5.6) \quad \delta \geq \frac{\varepsilon}{12M} \frac{m}{n},$$

and so

$$q \ll \left( \left( \frac{n}{m} \right)^{1+2/m} eC^2 (\log n)^{2+2/m} \varepsilon^2 \right)^m + \frac{1}{\delta^2} \left( \frac{ne}{m} \right)^m e^{-cn}.$$

Define  $\varepsilon$  by the following identity:

$$\left( \frac{n}{m} \right)^{1+2/m} (\log n)^{2+2/m} \varepsilon^2 = t.$$

This implies that

$$\varepsilon = \frac{\sqrt{t}}{(\log n)^{1+1/m}} \left( \frac{m}{n} \right)^{(m+2)/2m} \geq \frac{\sqrt{t}}{\log^2 n} \left( \frac{m}{n} \right)^{(m+2)/2m}.$$

In view of (5.6) we see that

$$(5.7) \quad \delta \gg \left( \frac{m}{n} \right)^{3/2+1/m} \frac{\sqrt{t}}{\log^2 n}.$$



In addition, we obtain

$$q \ll (C^2 et)^m + \frac{n^5}{t} (\log n)^4 (ne)^{\log^2 n} e^{-cn}.$$

Using the assumption that  $t \geq e^{-c'n}$  and taking  $c'$  sufficiently small, we deduce that

$$q \ll (C^2 et)^m + e^{-c''n}$$

for some constant  $c'' > 0$ .

In conclusion, we have now shown that

$$\mathbb{P}(\text{Loc}(A, m, \delta) \text{ and } \|A\| \leq M\sqrt{n}) \ll (C^2 et)^m + e^{-c''n}$$

for some  $\delta > 0$  which satisfies (5.7). In view of Lemma 5.1, the proof is complete.  $\square$

We now turn to the proofs of Theorem 2.4 and 2.5. We will need the following least singular value bound for real iid matrices, adopted from [41].

**Theorem 5.2.** *Let  $A$  be an  $N \times n$  real random matrix,  $N \geq n$ , whose elements are independent copies of a mean zero subgaussian random variable with unit variance. Then for every  $\varepsilon > 0$  and  $\lambda \in \mathbb{R}$  with  $|\lambda| \leq M\sqrt{N}$  for some  $M \geq 1$ , we have*

$$\mathbb{P}\left(s_n(A - \lambda) \leq \varepsilon(\sqrt{N} - \sqrt{n-1})\right) \leq (C\varepsilon)^{N-n+1} + e^{-cN}$$

where  $C, c > 0$  depend (polynomially) only on the subgaussian moment of the entries and  $M$ .

The  $\lambda = 0$  case of this theorem appears as [41, Theorem 1.1]. However, a close inspection of their proof confirms that their argument can be adapted to the shifted case, in the same way that we have explicitly done in the proof of Theorem 3.4.

*Proof of Theorem 2.4.* The proof is similar to the proof of Theorem 2.1. Without loss of generality, assume  $1 \geq t \geq e^{-c'n}$  for some constant  $c' > 0$  to be chosen later (as the bound is trivial when  $t \geq 1$ ). By [41, Proposition 2.3], there exists  $M \geq 1$  such that

$$(5.8) \quad \mathbb{P}(\|A\| \leq M\sqrt{n}) \geq 1 - C_0 e^{-c_0 n},$$

where  $M, C_0, c_0 > 0$  depend only on the subgaussian moment of the entries. Let  $\varepsilon, \delta$  be positive values to be chosen later, and take

$$q := \mathbb{P}(\text{Loc}_{\mathbb{R}}(A, m, \delta) \text{ and } \|A\| \leq M\sqrt{n}).$$

Proposition 3.3 implies that

$$q \leq \frac{3}{\delta} \left(\frac{ne}{m}\right)^m p_0,$$

where  $p_0$  satisfies (3.6). Choose  $\delta$  in terms of  $\varepsilon$  via (5.1), and again note that once we specify  $\varepsilon$ ,  $\delta$  will also be determined. Using Theorem 5.2, we find

$$p_0 \leq (C\varepsilon)^{m+1} + e^{-cn},$$

and hence

$$q \ll \frac{\varepsilon}{\delta} \left(\frac{ne}{m} C\varepsilon\right)^m + \frac{1}{\delta} \left(\frac{ne}{m}\right)^m e^{-cn}.$$

Returning to (5.1), we see

$$\delta \geq \frac{\varepsilon m}{12Mn}.$$

This implies that

$$q \ll \left(\left(\frac{n}{m}\right)^{1+1/m} eC\varepsilon\right)^m + \frac{1}{\varepsilon} \left(\frac{ne}{m}\right)^{m+1} e^{-cn}.$$

We now choose  $\varepsilon$ . Indeed, take

$$(5.9) \quad \varepsilon := t \left( \frac{m}{n} \right)^{(m+1)/m},$$

and recall that this choice of  $\varepsilon$  also determines  $\delta$  by (5.1). In addition, this choice implies that

$$(5.10) \quad q \ll (Cet)^m + \frac{1}{\varepsilon} \left( \frac{ne}{m} \right)^{m+1} e^{-cn}.$$

We now simplify the expression for  $\varepsilon$  given in (5.9) using the fact that  $m \geq \log^2 n$ . Indeed, in this case it follows that

$$\left( \frac{m}{n} \right)^{(m+1)/m} = \Theta \left( \frac{m}{n} \right).$$

We conclude that

$$(5.11) \quad \varepsilon = \Theta \left( t \frac{m}{n} \right),$$

and hence

$$(5.12) \quad \delta \gg t \left( \frac{m}{n} \right)^2.$$

Returning to (5.10), we use (5.11) and  $t \geq e^{-c'n}$  to see that

$$q \ll (Cet)^m + e^{-c'n}$$

for  $m \leq c'n$ , where  $c' > 0$  is a sufficiently small constant. In conclusion, we have now shown that

$$\mathbb{P}(\text{Loc}_{\mathbb{R}}(A, m, \delta) \text{ and } \|A\| \leq M\sqrt{n}) \ll (Cet)^m + e^{-c'n}$$

for some  $\delta > 0$  which satisfies (5.12). In view of (5.8), the proof is complete.  $\square$

Theorem 2.5 follows from similar arguments as those presented in the proofs of Theorems 2.2 and 2.4; we omit the details.

We now turn to the proofs of results from Section 2.3.

For the proof of Theorem 2.6, we first recall a result from [37].

**Theorem 5.3** (Theorem 1.4, [37]). *For a  $n - 1 \times n$  genuinely complex random matrix  $A$ , let  $x$  be a vector normal to all the rows. There exists a positive constants  $c$  and  $c'$  such that for any  $d$ -tuple  $(i_1, \dots, i_d)$  with  $d = n^c$  and  $\Omega \in \mathbb{C}^d$ ,*

$$|\mathbb{P}((\sqrt{n}x_{i_1}, \dots, \sqrt{n}x_{i_d}) \in \Omega) - \mathbb{P}(\mathbf{g}_{\mathbb{C},1}, \dots, \mathbf{g}_{\mathbb{C},d}) \in \Omega)| \leq d^{-c'}.$$

We model our proof of Theorem 2.6 after the proof of Theorem 5.1 in [38].

*Proof of Theorem 2.6.* Let  $Z$  and  $Z'$  be standard normal distributions with cumulative distribution function  $\Phi(x)$ . Recall that  $F(x)$  is the cumulative distribution function of  $Z^2/2 + Z'^2/2$ .

$$\begin{aligned} F(x) &= \frac{1}{2\pi} \int \int_{z^2 + z'^2 \leq 2x} e^{-z^2/2} e^{-z'^2/2} dz dz' \\ &= \int_0^{\sqrt{2x}} e^{-r^2/2} r dr \\ &= 1 - e^{-x} \end{aligned}$$

For convenience, we introduce the function

$$G(x) := F(x^2) = 1 - e^{-x^2}.$$

By direct calculation,

$$\begin{aligned} - \int_{1-\delta}^1 H(u) du &= \int_0^\delta F^{-1}(u) du \\ &= \int_0^{\sqrt{F^{-1}(\delta)}} 2x^3 F'(x) dx \\ &= \int_0^{G^{-1}(\delta)} x^2 G'(x) dx \\ &= - \int_0^\delta \log(1-u) du \end{aligned}$$

and

$$- \int_0^\delta H(u) du = \int_{1-\delta}^1 F^{-1}(u) du = - \int_{1-\delta}^1 \log(1-u) du.$$

Thus, it suffices to show

$$(5.13) \quad \left| \max_{S \subset [n]: |S| = \lfloor \delta n \rfloor} \|x_S\|_2^2 + \int_0^\delta \log(1-u) du \right| \leq \varepsilon$$

and

$$(5.14) \quad \left| \min_{S \subset [n]: |S| = \lfloor \delta n \rfloor} \|x_S\|_2^2 + \int_{1-\delta}^1 \log(1-u) du \right| \leq \varepsilon$$

In fact, we can simply focus on (5.14) as (5.13) follows from the identity

$$\max_{S \subset [n]: |S| = \lfloor \delta n \rfloor} \|x_S\|_2^2 + \min_{S \subset [n]: |S| = \lfloor \delta n \rfloor} \|x_S^c\|_2^2 = 1$$

and  $\int_0^1 \log(1-u) du = 1$ .

Define

$$N(c, k) := \sum_{j=1}^n \mathbf{1}_{\{c(k-1) \leq \sqrt{n}|v(j)| < ck\}}$$

Let  $Z$  be a complex gaussian and define

$$f(c, k) := n \mathbb{P}(c(k-1) \leq |Z| < ck).$$

Note that

$$f(c, k) = n(G(ck) - G(c(k-1))).$$

By Theorem 5.3, we have that

$$\mathbb{P}(c(k-1) \leq \sqrt{n}|x(j)| < ck) = \mathbb{P}(c(k-1) \leq |Z| < ck)(1 + o(1))$$

uniformly for all  $1 \leq j \leq n$ . Thus,

$$\mathbb{E}N(c, k) = (1 + o(1))f(c, k).$$

Similarly, we can verify that

$$\text{Var}(N(c, k)) = o(n^2).$$

By Chebyshev's inequality, we can conclude that

$$(5.15) \quad N(c, k) = (1 + o(1))f(c, k)$$

with probability  $1 - o(1)$ .

We choose  $c > 0$  and  $k_0 \in \mathbb{N}$  so that

$$(5.16) \quad c \left[ G^{-1}(\delta) \right]^2 < \frac{\varepsilon}{2}$$

and  $ck_0 = G^{-1}(\delta)$ . This definition ensures that

$$(5.17) \quad \sum_{k=1}^{k_0} G(ck) - G(c(k-1)) = G(ck_0) - G(0) = \delta.$$

Additionally, we have

$$(5.18) \quad \begin{aligned} & \left| \sum_{k=1}^{k_0} c^2(k-1)^2(G(ck) - G(c(k-1))) - \int_0^{G^{-1}(\delta)} x^2 G'(x) dx \right| \\ &= \left| \sum_{k=1}^{k_0} c^2(k-1)^2(G(ck) - G(c(k-1))) - \int_{c(k-1)}^{ck} x^2 G'(x) dx \right| \\ &= \left| \sum_{k=1}^{k_0} \left[ (c^2(k-1)^2 - c^2 k^2) G(ck) + 2 \int_{c(k-1)}^{ck} x G(x) dx \right] \right| \\ &= 2 \left| \sum_{k=1}^{k_0} \int_{c(k-1)}^{ck} x(G(x) - G(ck)) dx \right| \\ &\leq c \left[ G^{-1}(\delta) \right]^2 \\ &< \frac{\varepsilon}{2} \end{aligned}$$

by integration by parts and (5.16). The first inequality follows from the mean value theorem, the identity  $G'(x) = 2xe^{-x^2}$  and the bound  $|G'(x)| \leq \sqrt{2}e^{-2} \leq 1$ . By an identical argument, we can show that

$$(5.19) \quad \left| \sum_{k=1}^{k_0} c^2 k^2 (G(ck) - G(c(k-1))) - \int_0^{G^{-1}(\delta)} x^2 G'(x) dx \right| < \frac{\varepsilon}{2}.$$

By (5.15), for any  $1 \leq k \leq k_0 + 1$ ,

$$(5.20) \quad N(c, k) = (1 + o(1))f(c, k) = (1 + o(1))2n(G(ck) - G(c(k-1)))$$

with probability  $1 - o(1)$ . (5.17) implies that

$$\sum_{k=1}^{k_0} f(c, k) = \delta n.$$

Therefore, by a union bound, with probability  $1 - o(1)$ ,

$$\sum_{k=1}^{k_0} N(c, k) = (1 + o(1))\delta n = \lceil \delta n \rceil + o(n).$$

We have the two-sided bound

$$\sum_{k=1}^{k_0} c^2(k-1)^2 N(c, k) \leq n \min_{S \subset [n]: |S| = \sum_{k=1}^{k_0} N(c, k)} \|x_S\|_2^2 \leq \sum_{k=1}^{k_0} c^2 k^2 N(c, k).$$

With probability  $1 - o(1)$ , there exists a sequence  $\tau_n$  with  $\tau_n \rightarrow 0$  such that

$$\begin{aligned} & \sum_{k=1}^{k_0} c^2(k-1)^2 N(c, k) - \tau_n c^2 k_0^2 N(c, k_0) \\ & \leq \min_{S \subset [n]: |S| = \lceil \delta n \rceil} \|x_S\|_2^2 \leq \sum_{k=1}^{k_0} c^2 k^2 N(c, k) + \tau_n c^2 (k_0 + 1)^2 N(c, k_0 + 1). \end{aligned}$$

By (5.20), we also have that

$$\begin{aligned} & \sum_{k=1}^{k_0} c^2(k-1)^2 (G(ck) - G(c(k-1)))(1 + o(1)) \\ & \leq \min_{S \subset [n]: |S| = \lceil \delta n \rceil} \|x_S\|_2^2 \leq \sum_{k=1}^{k_0} k_0 c^2 k^2 (G(ck) - G(c(k-1)))(1 + o(1)) \end{aligned}$$

with probability  $1 - o(1)$ . Finally, combining (5.18) and (5.19), we can conclude that

$$\left| \min_{S \subset [n]: |S| = \lceil \delta n \rceil} \|x_S\|_2^2 - \int_0^{G^{-1}(\delta)} x^2 G'(x) dx \right| \leq \varepsilon$$

with probability  $1 - o(1)$ .  $\square$

*Proof of Theorem 2.7.* The proof closely mirrors the proof of Theorem 2.1. Without loss of generality, assume  $1 \geq t \geq e^{-\log^2 n}$  (as the bound is trivial when  $t \geq 1$ ). Let  $M \geq 1$  be the constant from Lemma 5.1. Let  $\varepsilon, \delta$  be positive values to be chosen later, and take

$$q := \mathbb{P}(\text{Loc}_0(A, M, m, \delta) \text{ and } \|A\| \leq M\sqrt{n}).$$

(Here, we have set  $\lambda_0 = 0$  in the definition of  $\text{Loc}_{\lambda_0}(A, M, m, \delta)$ .) Proposition 3.1 implies that

$$q \leq \left(\frac{ne}{m}\right)^m p_0,$$

where  $p_0$  satisfies (3.1). Choose  $\delta$  in terms of  $\varepsilon$  via the following identity:

$$(5.21) \quad 6\delta M\sqrt{n} = \varepsilon(\sqrt{n-1} - \sqrt{n-m-1}).$$

In other words, once we specify  $\varepsilon, \delta$  will also be determined. Using Theorem 3.4, we find

$$p_0 \leq (C\varepsilon)^{2m-1} + Ce^{-cn},$$

and so

$$q \ll \left(\frac{neC^2}{m}\varepsilon^{(2m-1)/m}\right)^m + \left(\frac{ne}{m}\right)^m e^{-cn}.$$

Choose  $\varepsilon > 0$  such that

$$\frac{n}{m}\varepsilon^{(2m-1)/m} = t.$$

This choice implies that

$$q \ll (eC^2t)^m + \left(\frac{ne}{m}\right)^m e^{-cn}$$

and

$$\varepsilon = \left(\frac{m}{t}\right)^{m/(2m-1)}.$$

We now use the assumption that  $m \geq \log^2 n$  to simplify this expression for  $\varepsilon$ . Indeed, in this case it follows that

$$\left(\frac{m}{n}\right)^{m/(2m-1)} = \Theta\left(\left(\frac{m}{n}\right)^{1/2}\right).$$

Similarly, since  $1 \geq t \geq e^{-\log^2 n}$ , we have

$$t^{m/(2m-1)} = \Theta(\sqrt{t}).$$

Thus, we conclude that

$$\varepsilon = \Theta\left(\sqrt{t} \left(\frac{m}{n}\right)^{1/2}\right).$$

Combining this with (5.21), we see that

$$(5.22) \quad \delta \gg \frac{m}{n} \varepsilon \gg \sqrt{t} \left(\frac{m}{n}\right)^{3/2}.$$

In addition, there exists a sufficiently small constant  $c' > 0$  such that

$$q \ll (eC^2t)^m + \left(\frac{ne}{m}\right)^m e^{-cn} \ll (eC^2t)^m + e^{-c'n}$$

for  $m \leq c'n$ . To conclude, we have shown that

$$\mathbb{P}(\text{Loc}_0(A, M, m, \delta) \text{ and } \|A\| \leq M\sqrt{n}) \ll (eC^2t)^m + e^{-c'n}$$

for some  $\delta$  which satisfies (5.22). In view of Lemma 5.1, the proof is complete.  $\square$

*Proof of Theorem 2.8.* The proof follows closely the proofs of Theorems 2.2 and 2.7. Without loss of generality, assume  $1 \geq t > 0$  (as the bound is trivial when  $t \geq 1$ ). Let  $M \geq 1$  be the constant from Lemma 5.1. Let  $\varepsilon, \delta$  be positive values to be chosen later, and take

$$q := \mathbb{P}(\text{Loc}_0(A, M, m, \delta) \text{ and } \|A\| \leq M\sqrt{n}).$$

Proposition 3.1 implies that

$$q \leq \left(\frac{ne}{m}\right)^m p_0,$$

where  $p_0$  satisfies (3.1). Again take  $\delta$  in terms of  $\varepsilon$  by (5.21), so that  $\delta$  is completely determined once we specify  $\varepsilon$ . Applying Theorem 3.5 and using the fact that  $1 \leq m \leq \log^2 n$ , we find

$$p_0 \leq (C\varepsilon \log n)^{2m} + Ce^{-cn}.$$

This gives

$$q \ll \left(\frac{n}{m} eC^2 (\log n)^2 \varepsilon^2\right)^m + \left(\frac{ne}{m}\right)^m e^{-cn}.$$

Define

$$\varepsilon := \frac{\sqrt{t}}{\log n} \sqrt{\frac{m}{n}},$$

so that

$$\varepsilon^2 \frac{n}{m} \log^2 n = t.$$

From (5.21), we see that this choice of  $\varepsilon$  gives

$$(5.23) \quad \delta \gg \frac{m}{n} \varepsilon \geq \frac{\sqrt{t}}{\log n} \left(\frac{m}{n}\right)^{3/2}.$$

In addition, it follows that

$$q \ll (eC^2t)^m + \left(\frac{ne}{m}\right)^m e^{-cn}.$$

Since  $1 \leq m \leq \log^2 n$ ,

$$\left(\frac{ne}{m}\right)^m e^{-cn} \ll e^{-c'n}$$

for some sufficiently small constant  $c' > 0$ . We conclude that

$$\mathbb{P}(\text{Loc}_0(A, M, m, \delta) \text{ and } \|A\| \leq M\sqrt{n}) \ll (eC^2t)^m + e^{-c'n}$$

for some  $\delta$  which satisfies (5.23). In view of Lemma 5.1, the proof is complete.  $\square$

#### APPENDIX A. PROOF OF PROPOSITIONS 1.1 AND 1.4

*Proof of Proposition 1.1.* We prove the result for  $\mathbb{R}^n$ , but an analogous argument applies in  $\mathbb{C}^n$ . We model the uniform distribution on the unit sphere by sampling a gaussian vector  $g \sim N(0, I_n)$  and normalizing by  $\|g\|_2^{-1}$ . Let  $\mathcal{E}$  denote the event that  $\|g\|_2 \geq \frac{\sqrt{n}}{10}$ . By standard concentration bounds, we have that

$$\mathbb{P}(\mathcal{E}^c) \leq \exp(-cn).$$

Similarly, for  $1 \leq i \leq n$ ,

$$\mathbb{P}(|g_i| \geq t) \leq 2 \exp(-t^2/8).$$

Therefore,

$$\begin{aligned} \mathbb{P}\left(|v_i| \geq C\sqrt{\frac{\log n}{n}}\right) &= \mathbb{P}\left(\frac{|g_i|}{\|g\|_2} \geq C\sqrt{\frac{\log n}{n}}\right) \\ &\leq \mathbb{P}(|g_i| \geq C\sqrt{\log n}) + \mathbb{P}(\mathcal{E}^c) \\ &\leq 1/n^2 \end{aligned}$$

for large enough  $C$ . Therefore, applying the union bound,

$$\|v\|_\infty \leq C\sqrt{\frac{\log n}{n}}$$

with probability  $1 - o(1)$ .  $\square$

*Proof of Proposition 1.4.* We address the complex case first. As we are not trying to optimize the constant in the exponent of the logarithm, we can conveniently assume that  $C \log n \leq m \leq n/\log n$  for any constant  $C$ . We follow the convention that  $C, c$  denote absolute constants that may change from line to line. Again, we model the uniform distribution on the unit sphere in  $\mathbb{C}^n$  by considering a random variable  $g \sim N_{\mathbb{C}}(0, 1)$  that is normalized by  $\|g\|_2^{-1}$ . Let  $\mathcal{E}$  denote the event that  $\|g\|_2 \leq 10\sqrt{n}$ . We have that

$$\mathbb{P}(\mathcal{E}^c) \leq \exp(-cn).$$

We let  $Y_1 < \dots < Y_n$  denote the order statistics of the magnitudes of  $|g_1|, \dots, |g_n|$ . Therefore, (A.1)

$$\begin{aligned} \mathbb{P}\left(\|v_I\|_2 \leq \frac{C}{\log^c n} \frac{m}{n} \text{ for all } I \subset [n], |I| = m\right) &\leq \mathbb{P}\left(\|v_I\|_2 \leq \frac{C}{\log^c n} \frac{m}{n} \text{ for all } I \text{ and } \mathcal{E}\right) + \mathbb{P}(\mathcal{E}^c) \\ &\leq \mathbb{P}\left(\sum_{i=1}^m Y_i^2 \leq \frac{C}{\log^c n} \frac{m^2}{n}\right) + \exp(-cn). \end{aligned}$$

We use a simple counting and grouping argument to control the latter probability. We define the following random variables that count the number of coordinates with magnitude in a fixed range.

$$\eta_k := \sum_{i=1}^n \mathbb{1}_{\{\frac{\delta}{n} 2^{k-1} \leq |g_i|^2 \leq \frac{\delta}{n} 2^k\}}$$

for  $1 \leq k \leq L$  where  $\delta := 1/\log n$  and  $L = \lfloor \log_2(m/2\delta) \rfloor$ . Additionally, we denote the probability of a coordinate falling in this range by

$$p_k := \mathbb{P}\left(\frac{\delta}{n} 2^{k-1} \leq |g_i|^2 \leq \frac{\delta}{n} 2^k\right).$$

As  $\eta_k$  is the sum of independent random variables, we have that

$$\mathbb{E}\eta_k = np_k$$

and

$$\text{Var}(\eta_k) = np_k(1 - p_k) \leq np_k.$$

By Chebyshev's inequality, for  $t > 0$ ,

$$(A.2) \quad \mathbb{P}(|\eta_k - \mathbb{E}\eta_k| > t) \leq \frac{np_k}{t^2}.$$

As  $|g_1|^2$  is a chi-squared distribution with two degrees of freedom and  $\frac{\delta}{n}2^k \leq 1$  for all  $k$ , by the bounded density of the chi-squared distribution, we deduce that

$$(A.3) \quad p_k = \Theta\left(\frac{\delta}{n}2^{k-1}\right).$$

Let  $\mathcal{E}'$  denote the event that

$$(A.4) \quad \eta_k \geq c\delta 2^{k-1} - 2^{(2/3)k}$$

for all  $1 \leq k \leq L$ . Combining (A.2) and (A.3), we can conclude that  $P(\mathcal{E}') \geq 1 - O(\delta)$ . In particular, observe that for  $k \geq \log \log n$ , say, (A.4) implies that  $\eta_k \geq \delta 2^{k-2}$  for large enough  $n$ .

Recall that the cumulative distribution function of a chi-squared distribution with two degrees of freedom is  $F(x) = 1 - e^{-x/2}$  for  $x \geq 0$ . Therefore, from our choice of  $L = \lceil \log_2(m/2\delta) \rceil$  we find that

$$\frac{\delta 2^L}{n} \leq \frac{m}{2n} \leq -2 \log(1 - m/2n)$$

Using the cumulative distribution function, we find that

$$\mathbb{P}\left(|g_1|^2 \leq \frac{\delta 2^L}{n}\right) \leq \mathbb{P}(|g_1|^2 \leq -2 \log(1 - m/2n)) = m/2n.$$

Thus, by Chernoff's bound,

$$P(Y_m \geq \delta 2^L) \leq 2e^{-cm} = O(\delta),$$

where the last equality follows from the assumptions that  $m \geq C \log n$  and  $\delta > 1/n$ . Finally, we have that with probability at least  $1 - O(\delta)$ ,

$$\sum_{i=1}^m Y_i^2 \geq \sum_{k=\log \log n}^L \delta 2^{k-1} \frac{\delta}{n} 2^{k-1} = \Omega\left(\frac{\delta^2 2^{2L}}{n}\right) = \Omega\left(\frac{m^2}{n \log^2 n}\right).$$

We have shown that

$$\mathbb{P}\left(\sum_{i=1}^m Y_i^2 \leq \frac{C}{\log^c n} \frac{m^2}{n}\right) = O(\delta).$$

From A.1, we can infer that

$$\mathbb{P}\left(\|v_I\|_2 \leq \frac{C}{\log^c n} \frac{m}{n} \text{ for all } I \subset [n], |I| = m\right) = O(\delta) = o(1),$$

which concludes the proof of the complex case.

The real case follows the same outline. The choice of parameters is slightly different as the density of the chi squared distribution with one degree of freedom no longer has bounded density but grows as  $x^{-1/2}$  near zero. We use the same notation as in the complex case.



Again, we can assume that  $C \log n \leq m \leq n/\log n$  for any large constant  $C$ . We model the uniform distribution on the sphere in  $\mathbb{R}^n$  by considering a random variable  $g \sim N_{\mathbb{R}}(0, 1)$  that is normalized by  $\|g\|_2^{-1}$ . Let  $\mathcal{E}$  denote the event that  $\|g\|_2 \leq 10\sqrt{n}$ . We have that

$$\mathbb{P}(\mathcal{E}^c) \leq \exp(-cn).$$

We let  $Y_1 < \dots < Y_n$  denote the order statistics of the magnitudes of  $|g_1|, \dots, |g_n|$ . Therefore,

$$\begin{aligned} \text{(A.5)} \quad \mathbb{P}\left(\|v_I\|_2 \leq \frac{C}{\log^c n} \left(\frac{m}{n}\right)^{3/2} \text{ for all } I \subset [n], |I| = m\right) \\ \leq \mathbb{P}\left(\|v_I\|_2 \leq \frac{C}{\log^c n} \left(\frac{m}{n}\right)^{3/2} \text{ for all } I \text{ and } \mathcal{E}\right) + \mathbb{P}(\mathcal{E}^c) \\ \leq \mathbb{P}\left(\sum_{i=1}^m Y_i^2 \leq \frac{C}{\log^c n} \frac{m^3}{n^2}\right) + \exp(-cn). \end{aligned}$$

To control the latter probability, we define the following random variables that count the number of coordinates with magnitude in a fixed range.

$$\eta_k := \sum_{i=1}^n \mathbb{1}_{\left\{\frac{\delta^2}{n^2} 2^{k-1} \leq |g_i|^2 \leq \frac{\delta^2}{n^2} 2^k\right\}}$$

for  $1 \leq k \leq L$  where  $\delta := 1/\log n$  and  $L = \lfloor 2 \log_2(m/C^* \delta) \rfloor$  where  $C^*$  is a constant to be fixed later. We denote the probability of a coordinate falling in this range by

$$p_k := \mathbb{P}\left(\frac{\delta^2}{n^2} 2^{k-1} \leq |g_i|^2 \leq \frac{\delta^2}{n^2} 2^k\right).$$

As  $\eta_k$  is the sum of independent random variables, we have that

$$\mathbb{E}\eta_k = np_k$$

and

$$\text{Var}(\eta_k) = np_k(1 - p_k) \leq np_k.$$

As a chi-squared distributed random variable with one degree of freedom has probability density function  $\Theta(x^{-1/2})$  near zero and  $\frac{\delta^2}{n^2} 2^k \leq 1$  for all  $k$ , we deduce that

$$\text{(A.6)} \quad p_k = \Theta\left(\frac{\delta}{n} 2^{k/2}\right).$$

Let  $\mathcal{E}'$  denote the event that

$$\text{(A.7)} \quad c\delta 2^{k/2} - 2^{k/3} \leq \eta_k \leq C\delta 2^{k/2} + 2^{k/3}$$

for all  $1 \leq k \leq L$  and  $C, c$  are the implied constants in (A.6). By Chebyshev's inequality and (A.6), we can conclude that the probability of (A.7) is larger than  $1 - O(\delta)$ . In particular, observe that for  $k \geq \log \log n$ , say, (A.4) implies that  $\eta_k \geq c\delta 2^{k/2}$  for large enough  $n$ .

By our choice of  $L$ , the probability that  $\sum_{i=0}^L \eta_k \geq m$  is at most  $O(\delta)$  for large enough  $C^*$ . By a simple calculation, we also have that  $P(Y_1 < \delta^2/n^2) = O(\delta)$ . Therefore, we can

have shown that

$$\begin{aligned} \mathbb{P}\left(\sum_{i=1}^m Y_i^2 \geq \sum_{k=\log \log n}^L \delta 2^{k/2} \frac{\delta^2 2^k}{n^2}\right) &\geq \mathbb{P}\left(\sum_{i=1}^m Y_i^2 \geq \frac{C\delta^3 2^{3L/2}}{n^2}\right) \\ &\geq \mathbb{P}\left(\sum_{i=1}^m Y_i^2 \geq \frac{Cm^3}{n^2 \log^c n}\right) \\ &= 1 - O(\delta). \end{aligned}$$

From A.5, we can conclude that

$$\mathbb{P}\left(\|v_I\|_2 \leq \frac{C}{\log^c n} \left(\frac{m}{n}\right)^{3/2} \text{ for all } I \subset [n], |I| = m\right) = O(\delta) = o(1),$$

which finishes the proof of the real case.  $\square$

## REFERENCES

- [1] R. Allez and J.-P. Bouchaud. Eigenvector dynamics under free addition. *Random Matrices Theory Appl.*, 3(3):1450010, 17, 2014.
- [2] A. Athreya, C. E. Priebe, M. Tang, V. Lyzinski, D. J. Marchette, and D. L. Sussman. A limit theorem for scaled eigenvectors of random dot product graphs. *Sankhya A*, 78(1):1–18, 2016.
- [3] Z. D. Bai, B. Q. Miao, and G. M. Pan. On asymptotics of eigenvectors of large sample covariance matrix. *Ann. Probab.*, 35(4):1532–1572, 2007.
- [4] S. Belinschi, M. A. Nowak, R. Speicher, and W. Tarnowski. Squared eigenvalue condition numbers and eigenvector correlations from the single ring theorem. *J. Phys. A*, 50(10):105204, 11, 2017.
- [5] F. Benaych-Georges. A universality result for the global fluctuations of the eigenvectors of Wigner matrices. *Random Matrices Theory Appl.*, 1(4):1250011, 23, 2012.
- [6] F. Benaych-Georges and A. Guionnet. Central limit theorem for eigenvectors of heavy tailed matrices. *Electron. J. Probab.*, 19:no. 54, 27, 2014.
- [7] F. Benaych-Georges and R. R. Nadakuditi. The eigenvalues and eigenvectors of finite, low rank perturbations of large random matrices. *Adv. Math.*, 227(1):494–521, 2011.
- [8] F. Benaych-Georges and S. Péché. Largest eigenvalues and eigenvectors of band or sparse random matrices. *Electron. Commun. Probab.*, 19:no. 4, 9, 2014.
- [9] F. Benaych-Georges and O. Zeitouni. Eigenvectors of non normal random matrices. Available at arXiv:1806.06806, 2018.
- [10] A. Bloemendal, L. Erdős, A. Knowles, H.-T. Yau, and J. Yin. Isotropic local laws for sample covariance and generalized Wigner matrices. *Electron. J. Probab.*, 19:no. 33, 53, 2014.
- [11] C. Bordenave and A. Guionnet. Localization and delocalization of eigenvectors for heavy-tailed random matrices. *Probab. Theory Related Fields*, 157(3-4):885–953, 2013.
- [12] P. Bourgade and G. Dubach. The distribution of overlaps between eigenvectors of ginibre matrices. Available at arXiv:1801.01219, 2018.
- [13] P. Bourgade, J. Huang, and H.-T. Yau. Eigenvector statistics of sparse random matrices. *Electron. J. Probab.*, 22:Paper No. 64, 38, 2017.
- [14] P. Bourgade and H.-T. Yau. The eigenvector moment flow and local quantum unique ergodicity. *Comm. Math. Phys.*, 350(1):231–278, 2017.
- [15] J. T. Chalker and B. Mehlis. Eigenvector statistics in non-hermitian random matrix ensembles. *Phys. Rev. Lett.*, 81:3367–3370, Oct 1998.
- [16] N. Crawford and R. Rosenthal. Eigenvector correlations in the complex ginibre ensemble. Available at arXiv:1805.08993, 2018.
- [17] S. Csorgo, E. Haeusler, and D. M. Mason. The asymptotic distribution of extreme sums. *Ann. Probab.*, 19(2):783–811, 04 1991.
- [18] H. A. David and H. N. Nagaraja. *Order statistics*. Wiley Series in Probability and Statistics. Wiley-Interscience [John Wiley & Sons], Hoboken, NJ, third edition, 2003.
- [19] Y. Dekel, J. R. Lee, and N. Linial. Eigenvectors of random graphs: nodal domains. *Random Structures Algorithms*, 39(1):39–58, 2011.

- [20] I. Dumitriu and S. Pal. Sparse regular random graphs: spectral density and eigenvectors. *Ann. Probab.*, 40(5):2197–2235, 2012.
- [21] A. Edelman, E. Kostlan, and M. Shub. How many eigenvalues of a random matrix are real? *J. Amer. Math. Soc.*, 7(1):247–267, 1994.
- [22] R. Eldan, M. Z. Rácz, and T. Schramm. Braess’s paradox for the spectral gap in random graphs and delocalization of eigenvectors. *Random Structures Algorithms*, 50(4):584–611, 2017.
- [23] L. Erdős and A. Knowles. Quantum diffusion and delocalization for band matrices with general distribution. *Ann. Henri Poincaré*, 12(7):1227–1319, 2011.
- [24] L. Erdős and A. Knowles. Quantum diffusion and eigenfunction delocalization in a random band matrix model. *Comm. Math. Phys.*, 303(2):509–554, 2011.
- [25] L. Erdős, A. Knowles, H.-T. Yau, and J. Yin. Delocalization and diffusion profile for random band matrices. *Comm. Math. Phys.*, 323(1):367–416, 2013.
- [26] L. Erdős, B. Schlein, and H.-T. Yau. Semicircle law on short scales and delocalization of eigenvectors for Wigner random matrices. *Ann. Probab.*, 37(3):815–852, 2009.
- [27] L. Erdős, H.-T. Yau, and J. Yin. Bulk universality for generalized Wigner matrices. *Probab. Theory Related Fields*, 154(1-2):341–407, 2012.
- [28] Y. V. Fyodorov. On statistics of bi-orthogonal eigenvectors in real and complex Ginibre ensembles: combining partial Schur decomposition with supersymmetry. *Comm. Math. Phys.*, 363(2):579–603, 2018.
- [29] L. Geisinger. Convergence of the density of states and delocalization of eigenvectors on random regular graphs. *J. Spectr. Theory*, 5(4):783–827, 2015.
- [30] A. Knowles and J. Yin. Eigenvector distribution of Wigner matrices. *Probab. Theory Related Fields*, 155(3-4):543–582, 2013.
- [31] B. Laurent and P. Massart. Adaptive estimation of a quadratic functional by model selection. *Ann. Statist.*, 28(5):1302–1338, 2000.
- [32] J. O. Lee and K. Schnelli. Extremal eigenvalues and eigenvectors of deformed Wigner matrices. *Probab. Theory Related Fields*, 164(1-2):165–241, 2016.
- [33] K. Luh. Complex random matrices have no real eigenvalues. *Random Matrices: Theory and Applications*, 7(01):1750014, 2018.
- [34] A. Lytova and K. Tikhomirov. On delocalization of eigenvectors of random non-hermitian matrices. Available at arXiv:1810.01590, 2018.
- [35] B. Mehlhlig and J. T. Chalker. Statistical properties of eigenvectors in non-Hermitian Gaussian random matrix ensembles. *J. Math. Phys.*, 41(5):3233–3256, 2000.
- [36] P. Mitra. Entrywise bounds for eigenvectors of random graphs. *Electron. J. Combin.*, 16(1):Research Paper 131, 18, 2009.
- [37] H. H. Nguyen and V. H. Vu. Normal vector of a random hyperplane. *International Mathematics Research Notices*, 2018(6):1754–1778, 2018.
- [38] S. O’Rourke, V. Vu, and K. Wang. Eigenvectors of random matrices: a survey. *J. Combin. Theory Ser. A*, 144:361–442, 2016.
- [39] M. Rudelson. Delocalization of eigenvectors of random matrices. lecture notes. Available at arXiv:1707.08461, 2017.
- [40] M. Rudelson and R. Vershynin. The Littlewood-Offord problem and invertibility of random matrices. *Adv. Math.*, 218(2):600–633, 2008.
- [41] M. Rudelson and R. Vershynin. Smallest singular value of a random rectangular matrix. *Comm. Pure Appl. Math.*, 62(12):1707–1739, 2009.
- [42] M. Rudelson and R. Vershynin. Delocalization of eigenvectors of random matrices with independent entries. *Duke Math. J.*, 164(13):2507–2538, 2015.
- [43] M. Rudelson and R. Vershynin. No-gaps delocalization for general random matrices. *Geom. Funct. Anal.*, 26(6):1716–1776, 2016.
- [44] J. Schenker. Eigenvector localization for random band matrices with power law band width. *Comm. Math. Phys.*, 290(3):1065–1097, 2009.
- [45] J. W. Silverstein. On the eigenvectors of large-dimensional sample covariance matrices. *J. Multivariate Anal.*, 30(1):1–16, 1989.
- [46] J. W. Silverstein. Weak convergence of random functions defined by the eigenvectors of sample covariance matrices. *Ann. Probab.*, 18(3):1174–1194, 1990.
- [47] F. Slanina. Localization of eigenvectors in random graphs. *Eur. Phys. J. B*, 85(11):Art. 361, 12, 2012.
- [48] T. Tao. *Topics in random matrix theory*, volume 132 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2012.

- [49] T. Tao and V. Vu. Random matrices: universality of local eigenvalue statistics up to the edge. *Comm. Math. Phys.*, 298(2):549–572, 2010.
- [50] T. Tao and V. Vu. Random matrices: universality of local eigenvalue statistics. *Acta Math.*, 206(1):127–204, 2011.
- [51] T. Tao and V. Vu. Random covariance matrices: universality of local statistics of eigenvalues. *Ann. Probab.*, 40(3):1285–1315, 2012.
- [52] T. Tao and V. Vu. Random matrices: universal properties of eigenvectors. *Random Matrices Theory Appl.*, 1(1):1150001, 27, 2012.
- [53] T. Tao and V. Vu. Random matrices: universality of local spectral statistics of non-Hermitian matrices. *Ann. Probab.*, 43(2):782–874, 2015.
- [54] T. Tao and V. H. Vu. Inverse Littlewood-Offord theorems and the condition number of random discrete matrices. *Ann. of Math. (2)*, 169(2):595–632, 2009.
- [55] K. Truong and A. Ossipov. Statistics of eigenvectors in the deformed Gaussian unitary ensemble of random matrices. *J. Phys. A*, 49(14):145005, 11, 2016.
- [56] K. Truong and A. Ossipov. Statistical properties of eigenvectors and eigenvalues of structured random matrices. *J. Phys. A*, 51(6):065001, 12, 2018.
- [57] V. Vu and K. Wang. Random weighted projections, random quadratic forms and random eigenvectors. *Random Structures Algorithms*, 47(4):792–821, 2015.

CENTER OF MATHEMATICAL SCIENCES AND APPLICATIONS, HARVARD UNIVERSITY

*E-mail address:* `kluh@cmsa.fas.harvard.edu`

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF COLORADO AT BOULDER, BOULDER, CO 80309

*E-mail address:* `sean.d.orourke@colorado.edu`