

# ON THE SMALLEST SINGULAR VALUE OF SYMMETRIC RANDOM MATRICES

VISHESH JAIN, ASHWIN SAH, AND MEHTAAB SAWHNEY

ABSTRACT. We show that for an  $n \times n$  random symmetric matrix  $A_n$ , whose entries on and above the diagonal are independent copies of a sub-Gaussian random variable  $\xi$  with mean 0 and variance 1,

$$\mathbb{P}[s_n(A_n) \leq \epsilon/\sqrt{n}] \leq O_\xi(\epsilon^{1/8} + \exp(-\Omega_\xi(n^{1/2}))) \quad \text{for all } \epsilon \geq 0.$$

This improves a result of Vershynin, who obtained such a bound with  $n^{1/2}$  replaced by  $n^c$  for a small constant  $c$ , and  $1/8$  replaced by  $(1/8) + \eta$  (with implicit constants also depending on  $\eta > 0$ ). Furthermore, when  $\xi$  is a Rademacher random variable, we prove that

$$\mathbb{P}[s_n(A_n) \leq \epsilon/\sqrt{n}] \leq O(\epsilon^{1/8} + \exp(-\Omega((\log n)^{1/4} n^{1/2}))) \quad \text{for all } \epsilon \geq 0.$$

The special case  $\epsilon = 0$  improves a recent result of Campos, Mattos, Morris, and Morrison, which showed that  $\mathbb{P}[s_n(A_n) = 0] \leq O(\exp(-\Omega(n^{1/2})))$ .

The main innovation in our work are new notions of arithmetic structure – the Median Regularized Least Common Denominator and the Median Threshold, which we believe should be more generally useful in contexts where one needs to combine anticoncentration information of different parts of a vector.

## 1. INTRODUCTION

Let  $M_n$  denote an  $n \times n$  random matrix, each of whose entries is an independent copy of a sub-Gaussian random variable  $\xi$  with mean 0 and variance 1. Prominent well-studied examples include the Ginibre ensemble (corresponding to  $\xi = \mathcal{N}(0, 1)$ ) and i.i.d. Rademacher matrices (corresponding to the Rademacher random variable  $\xi = \pm 1$  with probability  $1/2$  each).

A landmark result of Rudelson and Vershynin [21] shows that there are absolute constants  $C, c > 0$ , depending only on the sub-Gaussian norm of  $\xi$ , for which

$$\mathbb{P}[s_n(M_n) \leq \epsilon/\sqrt{n}] \leq C\epsilon + 2e^{-cn} \quad \text{for all } \epsilon \geq 0, \tag{1.1}$$

where  $s_n(M_n) = \inf_{v \in \mathbb{S}^{n-1}} \|Mv\|_2$  denotes the smallest singular value of  $M_n$ . Up to the constants  $C, c > 0$ , the above result is optimal, as can be seen by considering the two examples mentioned above. In particular, this result shows that the probability that an i.i.d. Rademacher matrix is singular is at most  $2 \exp(-cn)$  (for some  $c > 0$ ), thereby recovering (and substantially generalising) a well-known result of Kahn, Komlós, and Szemerédi [9]. We remark that after a series of intermediate works [2, 22, 23], a breakthrough result of Tikhomirov [24] established that the probability of singularity of an i.i.d. Rademacher matrix is at most  $(1/2 + o_n(1))^n$ , which is optimal up to the  $o_n(1)$  term.

In this paper, we will be concerned with  $n \times n$  symmetric random matrices  $A_n$  i.e.  $(A_n)_{ij} = (A_n)_{ji}$ , each of whose entries on and above the diagonal is an independent copy of a sub-Gaussian random variable  $\xi$  with mean 0 and variance 1. We note that the identical distribution assumption may be significantly relaxed (in particular, allowing for the diagonal entries to have a different distribution), although for the sake of simplicity, we do not deal with this modification here; the interested reader is referred to [25] and [12].

While symmetric matrices are especially convenient to work with linear algebraically, the lack of independence between the entries of  $A_n$  makes the non-asymptotic study of its smallest singular

value considerably more challenging than that of  $M_n$ . In the early 1990s, it was conjectured by Weiss that  $A_n(\text{Rad})$  (i.e.  $A_n$  where  $\xi$  is a Rademacher random variable) is invertible with probability  $1 - o_n(1)$ . This was only resolved in 2005 by Costello, Tao, and Vu [4], despite the corresponding statement for  $M_n$  (due to Komlós [10]) having been established almost 40 years prior.

Vershynin [25] showed that for any sub-Gaussian random variable  $\xi$  with mean 0 and variance 1, there are constants  $c, C_\eta$  depending only on the sub-Gaussian norm of  $\xi$  such that

$$\mathbb{P}[s_n(A_n) \leq \epsilon/\sqrt{n}] \leq C_\eta \epsilon^{1/8+\eta} + 2e^{-n^c}. \quad (1.2)$$

This improves (and generalizes) the nearly concurrent estimate of  $O_C(n^{-C})$  on the singularity probability of  $A_n(\text{Rad})$  obtained by Nguyen [17] using a novel quadratic variant of the inverse Littlewood–Offord theory. We note that in a subsequent work [18], Nguyen obtained estimates on the lower tail of  $s_n(A_n)$  for a large class of random variables  $\xi$ , including those not covered by [25], although the quantitative bounds in this work are much weaker than (1.2).

Recently, the upper bound on the singularity probability of  $A_n(\text{Rad})$  has been improved in a couple of works. Building on novel combinatorial techniques in [6], it was shown by Ferber and Jain [5] that this probability is at most  $\exp(-\Omega(n^{1/4}\sqrt{\log n}))$ . Subsequently, using a different combinatorial method inspired by the method of hypergraph containers [1], Campos, Mattos, Morris, and Morrison improved the bound to  $\exp(-\Omega(\sqrt{n}))$ . We note that both of these works deal only with  $A_n(\text{Rad})$ , and only with the singularity probability as opposed to quantitative estimates on  $s_n(A_n(\text{Rad}))$ .

The first main result of this paper is a strengthening of (1.2); the quantitative bounds are sufficiently powerful to generalize the aforementioned result of Campos et al. to all sub-Gaussian random variables.

**Theorem 1.1.** *Let  $A_n$  denote an  $n \times n$  random symmetric matrix, each of whose entries on and above the diagonal is an independent copy of a sub-Gaussian random variable  $\xi$  with mean 0 and variance 1. Then, there are constants  $C_{1.1}, c_{1.1}$  depending only on the sub-Gaussian norm of  $\xi$  such that, for all  $\epsilon \geq 0$ ,*

$$\mathbb{P}[s_n(A_n) \leq \epsilon/\sqrt{n}] \leq C_{1.1} \epsilon^{1/8} + 2e^{-c_{1.1} n^{1/2}}.$$

Next, we consider the particularly well studied case  $\xi = \text{Rad}$ ; setting  $\epsilon = 0$  in the theorem below improves the result of Campos et al. (see (3) in the Remark below).

**Theorem 1.2.** *Let  $A_n$  denote an  $n \times n$  random symmetric matrix, each of whose entries on and above the diagonal is an independent Rademacher random variable. Then, there are absolute constants  $C_{1.2}, c_{1.2}$  such that, for all  $\epsilon \geq 0$ ,*

$$\mathbb{P}[s_n(A_n) \leq \epsilon/\sqrt{n}] \leq C_{1.2} \epsilon^{1/8} + 2e^{-c_{1.2} n^{1/2} (\log n)^{1/4}}.$$

*Remark.* (1) We note that **Theorem 1.2** can be extended to the setting of discrete random variables covered in recent work of the authors [7]. We leave the details to an interested reader.

(2) The  $\epsilon^{1/8}$  term on the right hand side in **Theorem 1.1** improves on the  $\epsilon^{1/8+\eta}$  in [25]. It is believed that the correct dependence on  $\epsilon$  is  $O(\epsilon)$ , which would be optimal in light of the Gaussian example.

(3) The term  $\exp(-\Omega(n^{1/2}))$  on the right hand side in **Theorem 1.1** extends the result of Campos et al. to general sub-Gaussian random variables, whereas **Theorem 1.2** improves this result by a factor of  $(\log n)^{1/4}$  in the exponent, in the special case when  $\xi = \text{Rad}$ . A well-known conjecture is that one should be able to replace this with  $\exp(-\Omega(n))$ , although this will likely require significant new ideas. Indeed, as can be seen from our proof (see also the discussion in [3, Section 2.2]),  $\exp(-\tilde{\Omega}(n^{1/2}))$  is a natural barrier for techniques based on combining uniform anticoncentration estimates for a symmetric matrix row-vector product with tensorization.

The main innovation in our work are new notions of arithmetic structure of vectors, which we call the Median Regularized Least Common Denominator (MRLCD) (see [Section 3](#)) and the Median Threshold (see [Section 4](#)). Compared to the Regularized Least Common Denominator (RLCD) introduced in [\[25\]](#), and its natural threshold analogue, the MRLCD and median threshold are able to exploit the information that many different projections of a vector are arithmetically unstructured in a simple and transparent manner. Moreover, we are able to show that level sets of the MRLCD and median threshold admit sufficiently small nets at the appropriate scale – for the MRLCD, this follows by suitably adapting by-now standard bounds due to Rudelson and Vershynin [\[21\]](#), whereas for the median threshold, we adapt work of Tikhomirov [\[24\]](#) on the singularity of i.i.d Bernoulli random matrices. As the details are anyway short, we defer further discussion to [Sections 3](#) and [4](#).

We note that since its first appearance in [\[25\]](#), the RLCD has been used in many works (see, e.g., [\[13, 14, 16, 19, 26\]](#)); the MRLCD (and median threshold, for discrete distributions) can replace these applications in a black-box manner, and likely lead to improved quantitative estimates. We also note that a related use of combinatorially incorporating arithmetic unstructure of different projections of a vector appeared in recent work of the authors [\[8\]](#); however, the interaction with both the net and anticoncentration estimates is more delicate here.

**1.1. Notation.** We will drop the dimension in the subscript, henceforth denoting  $A_n$  by  $A$ , and denoting its rows by  $A_1, \dots, A_n$ . For an integer  $N$ ,  $\mathbb{S}^{N-1}$  denotes the set of unit vectors in  $\mathbb{R}^N$ , and  $\mathbb{B}_2^N$  denotes the unit ball in  $\mathbb{R}^N$  (i.e., the set of vectors of Euclidean norm at most 1).  $\|\cdot\|_2$  denotes the standard Euclidean norm of a vector, and for a matrix  $A = (a_{ij})$ ,  $\|A\|$  is its spectral norm (i.e.,  $\ell^2 \rightarrow \ell^2$  operator norm), and  $\|A\|_{\text{HS}}$  is its Hilbert-Schmidt norm, defined by  $\|A\|_{\text{HS}}^2 = \sum_{i,j} a_{ij}^2$ .

We will let  $[N]$  denote the interval  $\{1, \dots, N\}$ ,  $\mathfrak{S}_{[N]}$  denote the set of permutations of  $[N]$ , and  $\binom{[N]}{k}$  denote the set of subsets of  $[N]$  of size exactly  $k$ . We will denote multisets by  $\{\{\}\}$ , so that  $\{\{a_1, \dots, a_n\}\}$ , with the  $a_i$ 's possibly repeated, is a multi-set of size  $n$ . For a vector  $v \in \mathbb{R}^N$  and  $T \subseteq [N]$ ,  $v|_T$  denotes the  $|T|$ -dimensional vector obtained by only retaining the coordinates of  $v$  in  $T$ . We write  $u \parallel v$  for  $u, v \in \mathbb{R}^N$  if there is  $t \in \mathbb{R}$  such that  $u = tv$  or  $tu = v$ .

We will also make use of asymptotic notation. For functions  $f, g$ ,  $f = O_\alpha(g)$  (or  $f \lesssim_\alpha g$ ) means that  $f \leq C_\alpha g$ , where  $C_\alpha$  is some constant depending on  $\alpha$ ;  $f = \Omega_\alpha(g)$  (or  $f \gtrsim_\alpha g$ ) means that  $f \geq c_\alpha g$ , where  $c_\alpha > 0$  is some constant depending on  $\alpha$ , and  $f = \Theta_\alpha(g)$  means that both  $f = O_\alpha(g)$  and  $f = \Omega_\alpha(g)$  hold.

All logarithms are natural, unless indicated otherwise, and floors and ceilings are omitted when they make no essential difference.

**1.2. Acknowledgements.** We thank Roman Vershynin for comments on the manuscript. The last two authors were supported by the National Science Foundation Graduate Research Fellowship under Grant No. 1745302.

## 2. PRELIMINARIES

We will need the decomposition of the unit sphere into compressible and incompressible vectors, as formalized by Rudelson and Vershynin [\[21\]](#).

**Definition 2.1** (Compressible and incompressible vectors). For  $c_0, c_1 \in (0, 1)$ ,  $\text{Comp}(c_0, c_1)$  consists of all vectors  $v \in \mathbb{S}^{n-1}$  which are within Euclidean distance  $c_1$  of some vector  $w \in \mathbb{R}^n$  satisfying  $|\text{Supp}(w)| \leq c_0 n$ . Moreover,  $\text{Incomp}(c_0, c_1) := \mathbb{S}^{n-1} \setminus \text{Comp}(c_0, c_1)$ .

In order to prove [Theorem 1.1](#), it suffices to analyze  $\inf_{x \in \text{Incomp}(c_0, c_1)} \|Ax\|_2$  due to the following.

**Lemma 2.2.** *There exist  $c_0, c_1, c \in (0, 1)$  depending only on the sub-Gaussian moment of  $\xi$  so that for any vector  $u \in \mathbb{R}^n$ , we have*

$$\mathbb{P} \left[ \inf_{v \in \text{Comp}(c_0, c_1)} \|Av - u\|_2 < c\sqrt{n} \right] \leq 2 \exp(-cn).$$

*Proof.* This follows immediately by combining [25, Proposition 4.2] with the concentration of the operator norm of random matrices with independent, uniformly sub-Gaussian centered entries (cf. [25, Lemma 2.3]).  $\square$

**Lemma 2.3** (Incompressible vectors are spread, cf. [25, Lemma 3.8]). *For every  $c_0, c_1 \in (0, 1)$ , we can choose  $c_{2.3} := c_{2.3}(c_0, c_1) \in (0, 1/5)$  depending only on  $c_0, c_1$  such that the following holds. For every  $v \in \text{Incomp}(c_0, c_1)$ , there are at least  $2\lceil c_{2.3}n \rceil$  indices  $k \in [n]$  such that*

$$\frac{c_1}{\sqrt{2n}} \leq |v_k| \leq \frac{1}{\sqrt{c_0n}}.$$

**Definition 2.4** (Spread set). For every  $c_0, c_1 \in (0, 1)$ , and for every  $v \in \text{Incomp}(c_0, c_1)$ , we assign a subset  $\text{Spread}(v) \subseteq [n]$  such that

$$|\text{Spread}(v)| = \lceil c_{2.3}n \rceil, \text{ and}$$

$$\frac{c_1}{\sqrt{2n}} \leq |v_k| \leq \frac{1}{\sqrt{c_0n}} \text{ for all } k \in \text{Spread}(v).$$

**Definition 2.5.** For every  $c_0, c_1 \in (0, 1)$  and for  $\lambda \in (0, c_{2.3}/2)$ , let  $c_{2.5}(\lambda)n$  be the largest multiple of  $\lceil \lambda n \rceil$  less than or equal to  $\lceil c_{2.3}n \rceil$ . Note that  $c_{2.5}(\lambda) \geq \frac{c_{2.3}}{2}$ .

To every  $v \in \text{Incomp}(c_0, c_1)$ , we assign  $\text{Spread}_\lambda(v) \subseteq \text{Spread}(v)$  such that

$$|\text{Spread}_\lambda(v)| = c_{2.5}(\lambda)n,$$

and choose a partition

$$\text{Spread}_\lambda(v) = \bigsqcup_{j=1}^k \text{Spread}_\lambda^j(v)$$

into  $k = c_{2.3}(\lambda)n / \lceil \lambda n \rceil$  disjoint subsets of size  $\lceil \lambda n \rceil$ . We further assume that the choice of  $\text{Spread}_\lambda(v)$  and  $\text{Spread}_\lambda^j(v)$  is uniform for a given choice of  $\lambda$  and  $\text{Spread}(v)$  (in particular, these choices do *not* depend directly on  $v$ ).

We recall the definition of the Lévy concentration function.

**Definition 2.6.** For a random variable  $X$  and  $\epsilon \geq 0$ , the Lévy concentration of  $X$  of width  $\epsilon$  is

$$\mathcal{L}(X, \epsilon) = \sup_{x \in \mathbb{R}} \mathbb{P}[|X - x| \leq \epsilon].$$

We will also need a slight variant of the standard tensorization lemma, whose proof follows from the usual argument (cf. [21, Lemma 2.2]). We include the details for completeness.

**Lemma 2.7** (Tensorization). *Let  $X = (X_1, \dots, X_N)$  be a random vector in  $\mathbb{R}^N$  with independent coordinates. Suppose that for all  $k \in [N]$ , there exist  $a_k, b_k \geq 0$  such that*

$$\sup_{X_1, \dots, X_{k-1}} \mathcal{L}(X_k | X_1, \dots, X_{k-1}, \epsilon) \leq a_k \epsilon + b_k \quad \text{for all } \epsilon \geq 0.$$

Then

$$\mathcal{L}(X, \epsilon\sqrt{N}) \leq e^N \prod_{k=1}^N (a_k \epsilon + b_k).$$

*Proof.* We have

$$\begin{aligned} \mathbb{P}[|X| \leq \epsilon\sqrt{N}] &= \mathbb{P}\left[\sum_{j=1}^N X_j^2 \leq \epsilon^2 N\right] = \mathbb{P}\left[N - \frac{1}{\epsilon^2} \sum_{j=1}^N X_j^2 \geq 0\right] \\ &\leq \mathbb{E} \exp\left(N - \frac{1}{\epsilon^2} \sum_{j=1}^N X_j^2\right) \\ &\leq e^N \prod_{j=1}^N \sup_{X_1, \dots, X_{j-1}} \mathbb{E} \exp(-X_j^2/\epsilon^2 | X_1, \dots, X_{j-1}). \end{aligned}$$

We finish by noting that for any realization of  $X_1, \dots, X_{j-1}$ ,

$$\begin{aligned} \mathbb{E} \exp(-X_j^2/\epsilon^2 | X_1, \dots, X_{j-1}) &= \int_0^\infty 2ue^{-u^2} \mathbb{P}[|X_k| < \epsilon u | X_1, \dots, X_{j-1}] du \\ &\leq \int_0^\infty 2ue^{-u^2} (a_j \epsilon u + b_j) du \leq a_j \epsilon + b_j. \quad \square \end{aligned}$$

We also recall the definition of essential least common denominator (LCD). We use a log-normalized version due to Rudelson (unpublished), which also appears in [25].

**Definition 2.8** (LCD). For  $L \geq 1$  and  $v \in \mathbb{S}^{N-1}$ , the *least common denominator (LCD)*  $D_L(v)$  is defined as

$$D_L(x) = \inf \left\{ \theta > 0 : \text{dist}(\theta v, \mathbb{Z}^N) < L \sqrt{\log_+(\theta/L)} \right\}.$$

Finally we will require the following anticoncentration inequality of Miroshnikov and Rogozin [15]; this generalizes a well-known inequality of Lévy-Kolmogorov-Rogozin [20].

**Lemma 2.9** ([15, Corollary 1]). *Let  $\xi_1, \dots, \xi_N$  be independent random variables. Then, for any real numbers  $r_1, \dots, r_N > 0$  and any real  $r \geq \max_{i \in [N]} r_N$ , we have*

$$\mathcal{L}\left(\sum_{i=1}^N \xi_i, r\right) \leq C_{2.9} r \left(\sum_{i=1}^N \frac{r_i^2 (1 - \mathcal{L}(\xi_i, r_i))}{\mathcal{L}(\xi_i, r_i)^2}\right)^{-1/2},$$

where  $C_{2.9} > 0$  is an absolute constant.

### 3. MEDIAN REGULARIZED LCD (MRLCD)

In this section, we introduce the median regularized LCD (MRLCD), which is the notion of arithmetic structure that we will use in the proof of [Theorem 1.1](#). As opposed to the regularized LCD (RLCD) (introduced in [25]) which guarantees only one arithmetically unstructured projection of the vector, a large MRLCD guarantees many arithmetically unstructured projections of the vector. This simple change allows the MRLCD to piece together various unstructured parts of the vector to obtain significantly better small-ball probability estimates ([Proposition 3.5](#)), while at the same time not significantly impacting the size of nets of level sets ([Proposition 3.3](#)).

**Definition 3.1** (Median Regularized LCD). For  $v \in \text{Incomp}(c_0, c_1)$ ,  $\lambda \in (0, c_{2.3})$ , and  $L \geq 1$ , the *median regularized LCD*, denoted  $\widehat{MD}_L(v, \lambda)$ , is defined as

$$\widehat{MD}_L(v, \lambda) = \text{median}\{\{D_L(v_I/\|v_I\|_2) : I = \text{Spread}_\lambda^j(v) \text{ for some } j\}\}.$$

Here, the median of an even number of elements is not an average, but instead the value of the upper half. We denote by  $I_M(v)$  the set  $\text{Spread}_\lambda^j(v)$  achieving the median (arbitrarily chosen from

among all such sets), and  $\mathcal{I}_M(v)$  the collection of sets attaining values at least that of the median. We let  $\mathcal{J}_M(v)$  be the collection of sets attaining values at most that of the median.

We will consider level sets obtained by dyadically chopping the range of the MRLCD.

**Definition 3.2** (Level sets of MRLCD). For  $\lambda \in (0, c_{2.3})$ ,  $L \geq 1$ , and  $D \geq 1$ , we define the set

$$S_D = \{v \in \text{Incomp}(c_0, c_1) : \widehat{MD}_L(v, \lambda) \in [D, 2D]\}.$$

**3.1. Nets for level sets of MRLCD.** The main result of this subsection is the following.

**Proposition 3.3.** *Let  $c_0, c_1 \in (0, 1)$ . There exists  $C_{3.3} = C_{3.3}(c_0, c_1) > 0$  for which the following holds. Let  $\lambda \in (C_{3.3}/n, c_{2.3}/2)$  and  $L \geq 1$ . For every  $D \geq 1$ ,  $S_D$  has a  $\beta$ -net  $\mathcal{N}$  such that*

$$\beta = \frac{L\sqrt{\log(2D)}}{D}, \quad |\mathcal{N}| \leq D^{1/\lambda} \left( \frac{C_{3.3}D}{\sqrt{\log(2D)}} \right)^n \cdot \left( \frac{\sqrt{\log(2D)}}{\sqrt{\lambda n}} \right)^{c_{2.3}n/8}.$$

*Remark.* By changing  $C_{3.3}$  by a constant factor we can further assume that  $\mathcal{N} \subseteq S_D$ . Also, the  $L$  dependence here is not optimal – one can save a factor of  $L^{n(1-c_{2.3}/8)}$  by being more careful, although this does not affect the overall bounds if  $L$  is of constant order as in our application.

The proof of [Proposition 3.3](#) relies on a bound on the size of nets for level sets of the LCD.

**Lemma 3.4** (Corollary of Lemma 7.8 in [\[25\]](#)). *Let  $m \in \mathbb{N}$ ,  $D \geq 1$ , and  $c \in (0, 1)$  be such that  $D > c\sqrt{m} \geq 2$ . There exists a constant  $C$  depending only on  $c$  for which the following holds. Let  $\chi > 1$ ,  $L \geq 1$ , and  $\lambda > 0$ . Then the set*

$$\{x \in \sqrt{\chi\lambda}\mathbb{B}_2^m : c\sqrt{m} < D_L(x/\|x\|_2) \leq D\}$$

*has a  $\beta\sqrt{\chi\lambda}$ -net  $\mathcal{N}$  such that*

$$\beta = \frac{4L\sqrt{\log(2D)}}{D}, \quad |\mathcal{N}| \leq \left( \frac{CD}{\sqrt{m}} \right)^m D^2.$$

Now we conclude the result.

*Proof of [Proposition 3.3](#).* Let  $r = \lceil \lambda n \rceil$  and  $k = c_{2.5}(\lambda)n$ , so that  $r|k$  by definition.

Now, we pay a factor of  $2^n$  in a union bound over possible realizations of  $\text{Spread}(v)$ , which determines  $\text{Spread}_\lambda(v)$  and  $\text{Spread}_\lambda^j(v)$  for  $1 \leq j \leq k/r$ . We pay an additional factor of at most  $2^n$  to reveal which sets  $\text{Spread}_\lambda^j(v)$  are in  $\mathcal{J}_M(v)$ . Let  $J \subseteq [k/r]$  be the collection of corresponding indices  $j$ . We see  $|J| \geq k/(2r)$  by definition of median. Write  $J = \{j_1, \dots, j_t\}$  and let  $I_i = \text{Spread}_\lambda^{j_i}(v)$ .

Note that, given  $I_1, \dots, I_t$ , we know that  $D_L(v_{I_i}/\|v_{I_i}\|_2) \leq 2D$  for all  $1 \leq i \leq t$ . Moreover, since  $I_i \subseteq \text{Spread}_\lambda(v)$ , it follows that  $\|v_{I_i}\|_2 \leq \sqrt{\chi\lambda}$  for some  $\chi$  depending only on  $c_0$ . Further, by [\[25, Lemma 6.2\]](#), it follows (again, since  $I_i \subseteq \text{Spread}_\lambda(v)$ ) that  $D_L(v_{I_i}/\|v_{I_i}\|_2) \geq c\sqrt{\lceil \lambda n \rceil}$  for some  $c$  depending only on  $c_0, c_1$ . Hence, by [Lemma 3.4](#), we have a  $\beta\sqrt{\chi\lambda}$ -net for  $v_{I_i}$  where

$$\beta = \frac{2L\sqrt{\log(4D)}}{D}$$

of size at most

$$\left( \frac{CD}{\sqrt{\lceil \lambda n \rceil}} \right)^{\lceil \lambda n \rceil} D^2.$$

Finally, we take a product of these nets over  $1 \leq i \leq t$ , along with a standard  $\beta$ -net of  $\mathbb{B}_2^{I_0}$  (this net has size at most  $(1 + 3/\beta)^{|I_0|}$ ), where we let  $I_0 = [n] \setminus (I_1 \cup \dots \cup I_t)$ , to obtain the desired conclusion upon adjusting the value of  $\beta$  by standard arguments.  $\square$

**3.2. Anticoncentration via MRLCD.** We derive anticoncentration for a fixed vector with respect to MRLCD; the key idea is to patch together anticoncentration estimates on different segments of the vector through the use of [Lemma 2.9](#).

**Proposition 3.5** (Anticoncentration via the MRLCD). *Let  $\xi_1, \dots, \xi_n$  be i.i.d. random variables. Suppose that there exist  $\epsilon_0, p_0, M_0 > 0$  such that  $\mathcal{L}(\xi_k, \epsilon_0) \leq 1 - p_0$  and  $\mathbb{E}[|\xi_k|] \leq M_0$  for all  $k$ . Finally, let  $c_0, c_1 \in (0, 1)$ . Then, there exist  $C_{3.5}$ , depending only on  $\epsilon_0, p_0, M_0$  and  $C'_{3.5}$  depending on  $\epsilon_0, p_0, M_0, c_0, c_1$  such that the following holds.*

*Let  $L \geq p_0^{-1/2}$ ,  $\lambda \in (C'_{3.5}L^2/n, c_{2.3})$ ,  $v \in \text{Incomp}(c_0, c_1)$ ,  $J \subseteq \text{Spread}_\lambda(v)$ , and  $S_J = \sum_{k \in J} v_k \xi_k$ . Suppose that  $J$  is a union of sets in  $\mathcal{I}_M(v)$ . Then for every  $\epsilon \geq 0$ , we have for sufficiently large  $n$  (depending on  $\epsilon_0, p_0, M_0, c_0, c_1$ ) that*

$$\mathcal{L}(S_J, \epsilon) \leq C_{3.5}L \left( \frac{\epsilon}{\sqrt{|J|/n}} + \frac{\sqrt{\lambda n/|J|}}{\widehat{MD}_L(v, \lambda)} \right).$$

*Remark.* The above proposition should be compared with [\[25, Proposition 6.9\]](#), which bounds the Lévy concentration function in terms of the regularized LCD. The key difference is that the term  $\sqrt{|J|/n}$  in the denominator of our bound is replaced by  $\sqrt{\lambda}$ , which is always smaller. In fact, in the application considered here,  $\lambda$  must be chosen to be  $O(1/\sqrt{n})$ , which makes the above proposition significantly more efficient than the corresponding proposition in [\[25\]](#) for most of the matrix row-vector products (which satisfy  $|J|/n = \Theta(1)$ ).

*Proof.* Since  $v \in \text{Incomp}(c_0, c_1)$ , we have  $\widehat{MD}_L(v, \lambda) \geq c\sqrt{\lceil \lambda n \rceil}$  for some  $c$  depending only on  $c_0, c_1$  ([\[25, Lemma 6.2\]](#)).

First, assume that  $\epsilon \leq 1/(c\sqrt{n})$ . Let  $r = \lceil \lambda n \rceil$  and  $k = c_{2.5}(\lambda)n$ , so that  $r|k$  by definition. For  $i \in [k/r]$ , let

$$S_i = \sum_{k \in \text{Spread}_\lambda^i(v)} v_k \xi_k.$$

Let  $I$  be such that  $J = \cup_{i \in I} \text{Spread}_\lambda^i(v)$ . Since  $J$  is a union of sets in  $\mathcal{I}_M(v)$ , and  $D_L(v_I/\|v_I\|_2) \geq \widehat{MD}_L(v, \lambda)$  for each  $I \in \mathcal{I}_M(v)$ , it follows by standard anticoncentration estimates based on the LCD (see [\[25, Proposition 6.9\]](#) for the logarithmic version), that there exists an absolute constant  $C > 0$  such that

$$\mathcal{L}(S_i, \epsilon) \leq CL \left( \frac{\epsilon}{\sqrt{\lambda}} + \frac{1}{\widehat{MD}_L(v, \lambda)} \right) < \frac{1}{2}$$

for all  $i \in I$ , where the latter inequality follows from the assumption that  $\epsilon \leq 1/(c\sqrt{n})$ , along with the lower bound on  $\lambda$  (by taking  $C'_{3.5}$  sufficiently large depending on various parameters).

Now note that

$$S_J = \sum_{i \in I} S_i$$

and that the  $S_i$  are independent. Also, note that  $|I| = |J|/\lceil \lambda n \rceil$ . Therefore, by [Lemma 2.9](#), we have

$$\begin{aligned} \mathcal{L}(S_J, \epsilon) &\leq C_{2.9} \epsilon \left( \sum_{i \in I} \frac{\epsilon^2(1 - \mathcal{L}(S_i, \epsilon))}{\mathcal{L}(S_i, \epsilon)^2} \right)^{-1/2} \\ &\leq \frac{C_{2.9}\sqrt{2}}{\sqrt{|I|}} \max_{i \in I} \mathcal{L}(S_i, \epsilon) \leq \frac{2C_{2.9}}{\sqrt{|J|/n}} \cdot CL \left( \epsilon + \frac{\sqrt{\lambda}}{\widehat{MD}_L(v, \lambda)} \right), \end{aligned}$$

which proves the desired conclusion for  $\epsilon \leq 1/c\sqrt{n}$ .

Finally, for  $\epsilon > 1/(c\sqrt{n})$ , we note that any interval of length  $2\epsilon$  can be tiled by at most  $2\epsilon/\epsilon_0$  intervals of length  $2\epsilon_0$ , where  $\epsilon_0 = 1/(2c\sqrt{n})$ . Moreover, for such  $\epsilon_0$ , we have

$$\epsilon_0 + \frac{\sqrt{\lambda}}{\widehat{MD}_L(v, \lambda)} \leq 4\epsilon_0.$$

Hence, we have that for all  $\epsilon > 1/(c\sqrt{n})$ ,

$$\begin{aligned} \mathcal{L}(S_J, \epsilon) &\leq \frac{2\epsilon}{\epsilon_0} \cdot \mathcal{L}(S_J, \epsilon_0) \\ &\leq \frac{2\epsilon}{\epsilon_0} \cdot \frac{2C_{2.9}CL}{\sqrt{|J|/n}} \cdot 4\epsilon_0 \leq 16C_{2.9}CL \cdot \frac{\epsilon}{\sqrt{\lambda}}, \end{aligned}$$

as desired.  $\square$

Next, we derive a small-ball result for (symmetric) matrix-vector products.

**Lemma 3.6.** *Fix  $K \geq 1$ ,  $c_0, c_1 \in (0, 1)$  and  $v \in \text{Incomp}(c_0, c_1)$ . There exists  $L$  depending only on the sub-Gaussian norm of  $\xi$ , and  $C_{3.6}, C_{3.6}$  depending on the sub-Gaussian norm of  $\xi$  and on  $c_0, c_1$  such that the following holds.*

*Let  $\lambda \in (C_{3.6}/n, C_{3.6})$  and suppose that  $v \in S_D$  (with MRLCD defined with respect to  $\lambda, L$ ). Then, for any  $u \in \mathbb{R}^n$ , we have*

$$\mathbb{P}[\|Av - u\|_2 \leq K\beta\sqrt{n}] \leq \left( \frac{C_{3.6}L^2\sqrt{\log(2D)}}{D} \right)^{n - \lceil \lambda n \rceil},$$

where

$$\beta = \frac{L\sqrt{\log(2D)}}{D}.$$

*Proof.* Fix  $u \in \mathbb{R}^n$  and  $v \in S_D$ . Note that for any permutation matrix  $P$ ,  $\|Av - u\|_2 \leq K\beta\sqrt{n}$  occurs if and only if

$$\|(PAP^{-1})Pv - Pu\| \leq K\beta\sqrt{n}.$$

Furthermore,  $PAP^{-1} = PAP^\top$  has the same distribution as  $A$ . Therefore, we will be able to permute the indices of  $[n]$  at our convenience (depending on  $v$ ).

In particular, we may assume that  $\text{Spread}_\lambda(v) = \lceil c_{2.5}(\lambda)n \rceil$ . Let  $A_t = \{k \in [n] : t\lceil \lambda n \rceil < k \leq (t+1)\lceil \lambda n \rceil\}$  (defined for  $0 \leq t \leq T-1$ ), where

$$T = \frac{c_{2.5}(\lambda)n}{\lceil \lambda n \rceil}.$$

We may also assume that  $A_t \in \mathcal{I}_M(v)$  for all  $t \leq \lceil T/2 \rceil - 1$ . Then, for all  $\lceil \lambda n \rceil \leq j \leq c_{2.5}(\lambda)n$ , the set  $J = [j]$  has a subset of at least half the size which satisfies the assumptions of [Proposition 3.5](#) (namely, the union of the first  $\lfloor j/\lceil \lambda n \rceil \rfloor$  sets  $A_t$ ).

Therefore [Proposition 3.5](#) implies that for all  $L$  sufficiently large depending on the sub-Gaussian norm of  $\xi$ , if  $j \geq \lceil \lambda n \rceil$  and  $\epsilon \geq 0$ , then

$$\mathcal{L}((Av - u)_j | (Av - u)_{j+1, \dots, n}, \epsilon) \leq CC_{3.5}L \left( \frac{\epsilon}{\sqrt{j/n}} + \frac{\sqrt{\lambda n/j}}{D} \right),$$

where  $C$  depends only on  $c_0, c_1$ . Here, we have used that the first  $j$  elements of the  $j^{\text{th}}$  row are independent of rows  $j+1, \dots, n$ , and that the Lévy concentration is monotone under removing independent random variables from a sum.

Let  $j' = \min(j, c_{2.5}(\lambda)n)$ . Then, by [Lemma 2.7](#), we deduce

$$\begin{aligned}
\mathbb{P}[\|Av - u\|_2 \leq K\beta\sqrt{n}] &\leq \mathbb{P}\left[\sum_{j=\lceil\lambda n\rceil}^n (Av - u)_j^2 \leq K^2\beta^2n\right] \\
&\leq (CL)^n \prod_{j=\lceil\lambda n\rceil}^n \left(\frac{K\beta\sqrt{n/(n - \lceil\lambda n\rceil)}}{\sqrt{j'/n}} + \frac{\sqrt{\lambda n/j'}}{D}\right) \\
&\leq (CL)^n \prod_{j=\lceil\lambda n\rceil}^n \left(\frac{KL\sqrt{n\log(2D)}}{D\sqrt{j'}}\right) \\
&\leq \left(\frac{CL^2\sqrt{\log(2D)}}{D}\right)^{n-\lceil\lambda n\rceil},
\end{aligned}$$

where the last inequality uses  $\prod_{j=1}^n (n/j) \leq e^n$ .  $\square$

**3.3. Structure theorem.** We will need the following structure theorem, which shows that, except with exponentially small probability, the preimage under  $A$  of any fixed vector is highly unstructured. This is our replacement for the key [\[25, Theorem 7.1\]](#). As usual,  $A$  denotes a random  $n \times n$  symmetric matrix with independent  $\xi$  entries on and above the diagonal.

**Theorem 3.7.** *Fix  $K \geq 1$ . Depending on the sub-Gaussian norm of  $\xi$ , we can choose  $L, c, C$  so that the following holds. For all  $\lambda \in (C/n, 1/\sqrt{n})$ , we have for any  $u \in \mathbb{R}^n$  that*

$$\mathbb{P}[\exists v \in \mathbb{S}^{n-1} : (Av \parallel u) \wedge (v \in \text{Comp}(c_0, c_1) \vee \widehat{MD}_L(v, \lambda) \leq 2^{\lambda n/C}) \wedge (\|A\| \leq K\sqrt{n})] \leq 2e^{-cn}.$$

*Proof.* This is an immediate consequence of [Proposition 3.3](#) and [Lemmas 2.2](#) and [3.6](#). Note that if  $Av = tu$  for  $t \in \mathbb{R}$ , then  $\|A\| \leq K\sqrt{n}$  implies  $\|tu\|_2 \leq K\sqrt{n}$ . For compressible vectors  $v$ , we use [Lemma 2.2](#) on a constant amount of target vectors parallel to  $u$  so as to cover the full range. For the rest, if the MRLCD is between  $D$  and  $2D$  (for some  $D \leq 2^{\lambda n/C}$ ), we take a net constructed in [Lemma 3.4](#) along with a  $1/D$ -net for  $\{tu : \|tu\|_2 \leq K\sqrt{n}\}$ , which adds an additional (unimportant) factor of  $KD\sqrt{n}$  to the size of our nets. Since

$$\begin{aligned}
KD\sqrt{n} \cdot D^{1/\lambda} &\left(\frac{C_{3.3}D}{\sqrt{\log(2D)}}\right)^n \cdot \left(\frac{\sqrt{\log(2D)}}{\sqrt{\lambda n}}\right)^{c_{2.3}n/8} \times \left(\frac{C_{3.6}L^2\sqrt{\log(2D)}}{D}\right)^{n-\lceil\lambda n\rceil} \\
&= KD\sqrt{n} \cdot D^{1/\lambda} \left(\frac{C'D}{\sqrt{\log(2D)}}\right)^{\lceil\lambda n\rceil} \left(\frac{\sqrt{\log(2D)}}{\sqrt{\lambda n}}\right)^{c_{2.3}n/8} \\
&\leq C'^n \cdot 2^{\lambda^2 n^2/C} \cdot (1/C)^{c_{2.3}n/8} \\
&\leq C'^n 2^{n/C} (1/C)^{c_{2.3}n/8},
\end{aligned}$$

the result follows by a union bound upon taking  $C$  sufficiently large. We omit the standard details, referring the reader to the proof of [\[25, Theorem 7.1\]](#) for a more detailed calculation.  $\square$

#### 4. MEDIAN THRESHOLD

We begin by defining an alternate notion of structure, based on the so-called threshold function ([Definition 4.1](#)), which will allow us to use results of Tikhomirov [\[24\]](#) to obtain a stronger bound for the probability of singularity Rademacher random symmetric matrices. We note that, although we have chosen to focus on the Rademacher case, our analysis can be extended to general real discrete distributions using recent results of the authors [\[7\]](#).

For a technical reason that will become clear later, we fix a sufficiently small absolute constant  $p \in (0, 1/2]$  throughout this section; for the case of Rademacher random variables, which is our focus, one can take  $p = 1/10$ .

**Definition 4.1.** For  $p \in (0, 1)$ ,  $L \geq 1$ , and  $v \in \mathbb{S}^{N-1}$ , the *threshold*  $\mathcal{T}_{p,L}(v)$  is defined as

$$\mathcal{T}_{p,L}(v) = \sup \left\{ t \in (0, 1) : \mathcal{L} \left( \sum_{i=1}^N b'_i v_i, t \right) > Lt \right\},$$

where the  $b'_i$  are i.i.d. random variables distributed as  $\text{Ber}(p) - \text{Ber}'(p)$ .

**Definition 4.2.** For  $p \in (0, 1)$ ,  $v \in \text{Incomp}(c_0, c_1)$ ,  $\lambda \in (0, c_{2.3})$ , and  $L \geq 1$ , the *median threshold*, denoted  $\widehat{\mathcal{T}}_{p,L}(v, \lambda)$ , is defined as

$$\widehat{\mathcal{T}}_{p,L}(v, \lambda) = \text{median} \{ \{ \mathcal{T}_{p,L}(v_I / \|v_I\|_2) : I = \text{Spread}_\lambda^j(v) \text{ for some } j \} \}.$$

**4.1. Threshold of random lattice points.** We next recall the key technical result of Tikhomirov [24], which upper bounds the number of vectors with “large” threshold within a lattice of appropriate size. The important fact is that the number of such vectors is *superexponentially* small compared to the size of the lattice, which is the key difference with the results coming from the MRLCD. First, we must establish some notation.

**Definition 4.3.** Choose  $N, n \geq 1$  and  $\delta \in (0, 1]$ , as well as  $K \geq 1$ . We say that  $\mathcal{A} \subseteq \mathbb{Z}^n$  is  $(N, n, K, \delta)$ -*admissible* if the following hold:

- $\mathcal{A} = A_1 \times \cdots \times A_n$ , where each  $A_i$  is an origin-symmetric subset of  $\mathbb{Z} \cap (-nN, nN)$ ,
- $A_i$  is an integer interval of size at least  $2N + 1$  for all  $i > \delta n$ ,
- $A_i$  is a union of two integer intervals of total size at least  $2N$  and  $A_i \cap [-N, N] = \emptyset$  for all  $i \leq \delta n$ , and
- $|\mathcal{A}| \leq (KN)^n$ .

**Theorem 4.4** (From [24, Corollary 4.3]). *Let  $\delta, \epsilon \in (0, 1]$ ,  $p \in (0, 1/2]$ , and  $K, M \geq 1$ . There exist  $n_{4.4} = n_{4.4}(\delta, \epsilon, K, M) \geq 1$  and  $L_{4.4} = L_{4.4}(\delta, \epsilon, K) > 0$  such that the following holds. If  $n \geq n_{4.4}$ ,  $1 \leq N \leq (1 - p + \epsilon)^{-n}$ , and  $\mathcal{A}$  is  $(N, n, K, \delta)$ -admissible, then*

$$\left| \left\{ x \in \mathcal{A} : \mathcal{L} \left( \sum_{i=1}^n b_i x_i, \sqrt{n} \right) \geq L_{4.4} N^{-1} \right\} \right| \leq \exp(-Mn) |\mathcal{A}|,$$

where the  $b_i$  are i.i.d.  $\text{Ber}(p)$  random variables. Furthermore,  $n_{4.4} = \exp(C_{4.4}(\delta, \epsilon, K)M^2)$  is allowable.

*Remark.* This is the same as [24, Corollary 4.3], except that we have claimed an explicit dependence between  $n$  and  $M$ , namely that one can take  $M$  growing as  $(\log n)^{1/2}$  (all other parameters fixed). This is an immediate consequence of unraveling the parameter dependencies in [24, Theorem 4.2]. We give a brief sketch, using the notation of [24, Theorem 4.2]. In the proof of [24, Theorem 4.2], one sets  $L = L_{4.5}(2M, p, \delta, \epsilon/2)$ , which can be checked to grow exponentially in  $M$  by examining the last line of the proof of [24, Proposition 4.5]. This shows that the parameter  $q$  in the proof of [24, Theorem 4.2] is chosen to be linear in  $M$ , and hence, the parameter  $\tilde{\epsilon}$  grows as  $M^{-1}$ . Next, it is required that  $n \geq n_{4.10}(p, \tilde{\epsilon}, \max(16\tilde{R}, L), \tilde{R}, 2M)$  and  $n \geq n_{4.5}(2M, p, \delta, \epsilon/2)$ . The more restrictive condition comes from [24, Proposition 4.10], and indeed, an examination of the first few lines of the proof of this proposition reveals that it suffices to have  $n$  growing as  $\exp(\Theta(M^2))$ . One also sees that  $\eta_{4.2} = \eta_{4.10}(p, \tilde{\epsilon}, \max(16\tilde{R}, L), \tilde{R}, 2M)$  decays as  $\exp(-\Theta(M^2))$ . Finally, the deduction of [24, Corollary 4.3] from [24, Theorem 4.2] requires  $n^{-1/2} \leq \eta$ , for which  $n$  growing as  $\exp(\Theta(M^2))$  is sufficient in light of the decay of  $\eta$  discussed above.

**4.2. Replacement.** In order to relate the anticoncentration of a vector with respect to Rademacher random variables to [Definitions 4.1 and 4.2](#), we will require the following inequality. This is closely related to the replacement trick employed by Kahn, Komlós, and Szemerédi [9] and later by Tao and Vu [23] (although the application here is substantially simpler).

**Lemma 4.5.** *There exists an absolute constant  $C_{4.5}$  for which the following holds. Let  $v \in \mathbb{R}^n$  and  $r > 0$ . Then, for any  $0 < p \leq (2 - \sqrt{2})/4$ ,*

$$\mathcal{L}\left(\sum_{i=1}^n b_i v_i, r\right) \leq C_{4.5} \mathcal{L}\left(\sum_{i=1}^n b'_i v_i, r\right),$$

where  $b_i$  are independent Rademacher random variables and  $b'_i$  are distributed as  $\text{Ber}(p) - \text{Ber}'(p)$ .

*Proof.* Note that by scaling  $v$ , we may assume without loss of generality that  $r = 1$ . Let  $X = \sum_{i=1}^n b'_i v_i$ . By Esseen's inequality and  $|\cos t| \leq (3 + \cos(2t))/4$ , we find

$$\begin{aligned} \mathcal{L}\left(\sum_{i=1}^n b_i v_i, 1\right) &\leq C \int_{-2}^2 \prod_{i=1}^n |\cos(v_i \theta)| d\theta \leq C \int_{-2}^2 \prod_{i=1}^n \left(\frac{3}{4} + \frac{1}{4} \cos(2v_i \theta)\right) d\theta \\ &\leq C \int_{-2}^2 \prod_{i=1}^n \mathbb{E} \exp\left(i\theta \cdot 2b'_i v_i\right) d\theta \leq 2C \int_{\mathbb{R}} \mathbb{1}_{[-2,2]} * \mathbb{1}_{[-2,2]}(\theta) \mathbb{E} \exp(i\theta(2X)) d\theta \\ &= 4C \mathbb{E} \left(\frac{\sin(4X)}{2X}\right)^2 \\ &\leq 4C \mathbb{E} \left(\frac{\sin(4X)}{2X} \cdot \mathbb{1}_{X \in [-1,1]}\right)^2 + \sum_{k=1}^{\infty} 4C \mathbb{E} \left(\frac{\sin(4X)}{2X} \cdot \mathbb{1}_{\pm X \in [2k-1, 2k+1]}\right)^2 \\ &\leq 16C \mathcal{L}(X, 1) + C \sum_{k=1}^{\infty} \frac{\mathcal{L}(X, 1)}{(2k-1)^2} \leq C' \mathcal{L}(X, 1). \end{aligned}$$

The third inequality uses  $p \leq (2 - \sqrt{2})/4$ , and the penultimate inequality uses  $\sin(4x)/(2x) \leq 2$  for  $x \in [-1, 1]$ .  $\square$

**4.3. Randomized rounding.** We will make use of a slight modification of [24, Lemma 5.3], proved using randomized rounding (cf. [11]). As the proof is identical we omit the details.

**Lemma 4.6.** *Let  $y = (y_1, \dots, y_n) \in \mathbb{R}^n$  be a vector,  $\Delta$  be a fixed distribution supported in  $[-1, 1]^n$ , and let  $\mu > 0$ ,  $\psi \in \mathbb{R}$  be fixed. There exist absolute constants  $c_{4.6}$  and  $C_{4.6}$  for which the following holds.*

*Suppose that for all  $t \geq \sqrt{n}$ ,*

$$\mathbb{P}\left[\left|\sum_{i=1}^n b_i y_i - \psi\right| \leq t\right] \leq \mu t,$$

where  $(b_1, \dots, b_n)$  are independent and distributed as  $\Delta$ . Then, there exists a vector  $y' \in \mathbb{Z}^n$  satisfying

- (R1)  $\|y - y'\|_{\infty} \leq 1$ ,
- (R2)  $\mathbb{P}[\|\sum_{i=1}^n b_i y'_i - \psi\| \leq t] \leq C_{4.6} \mu t$  for all  $t \geq \sqrt{n}$ , and
- (R3)  $\mathcal{L}(\sum_{i=1}^n b_i y'_i, \sqrt{n}) \geq c_{4.6} \mathcal{L}(\sum_{i=1}^n b_i y_i, \sqrt{n})$ .

Next, we prove a version of the above proposition for the case when  $\Delta = \text{Ber}(p) - \text{Ber}'(p)$  with  $p$  sufficiently small. The main difference is that the left hand side in (R2) above can be replaced by the Lévy concentration at width  $t$ ; this can be done since for a distribution with non-negative

characteristic function, the maximum concentration of given width is essentially obtained around  $\psi = 0$ .

**Lemma 4.7.** *Let  $y = (y_1, \dots, y_n) \in \mathbb{R}^n$  be a vector,  $p \in (0, 1)$ , and let  $\mu > 0$ ,  $\psi \in \mathbb{R}$  be fixed. There exist absolute constants  $c_{4.7}$  and  $C_{4.7}$  for which the following holds.*

*Suppose that for all  $t \geq \sqrt{n}$ ,*

$$\mathcal{L}\left(\sum_{i=1}^n b'_i y_i, t\right) \leq \mu t,$$

*where the  $b'_i$  are independent and distributed as  $\text{Ber}(p) - \text{Ber}'(p)$ . Then, there exists a vector  $y' = (y'_1, \dots, y'_n) \in \mathbb{Z}^n$  satisfying*

- (R1)  $\|y - y'\|_\infty \leq 1$ ,
- (R2)  $\mathcal{L}(\sum_{i=1}^n b'_i y'_i, t) \leq C_{4.7} \mu t$  for all  $t \geq \sqrt{n}$ , and
- (R3)  $\mathcal{L}(\sum_{i=1}^n b'_i y'_i, \sqrt{n}) \geq c_{4.7} \mathcal{L}(\sum_{i=1}^n b'_i y_i, \sqrt{n})$ .

*Proof.* We apply Lemma 4.6 to the distribution  $\Delta = \text{Ber}(p) - \text{Ber}'(p)$  and  $\psi = 0$ . From (R2),

$$\mathbb{P}\left[\left|\sum_{i=1}^n b'_i y_i\right| \leq t\right] \leq C_{4.6} \mu t$$

for all  $t \geq \sqrt{n}$ . Now let  $t \geq \sqrt{n}$  and  $X = (\sum_{i=1}^n b'_i y'_i)/t$ , and note that  $X$  has nonnegative characteristic function since  $b'_i$  does. Thus, for all  $\psi \in \mathbb{R}$ ,

$$\begin{aligned} \mathbb{P}[|X - \psi| \leq 1] &= \mathbb{E}[\mathbb{1}_{[-1,1]}(X - \psi)] \leq \mathbb{E}[\mathbb{1}_{[-1,1]} * \mathbb{1}_{[-1,1]}(X - \psi)] \\ &= \int_{\mathbb{R}} \left(\frac{2 \sin \theta}{\theta}\right)^2 \mathbb{E} \exp(i\theta(X - \psi)) d\theta \\ &\leq \int_{\mathbb{R}} \left(\frac{2 \sin \theta}{\theta}\right)^2 |\mathbb{E} \exp(i\theta X)| d\theta \\ &= \int_{\mathbb{R}} \left(\frac{2 \sin \theta}{\theta}\right)^2 \mathbb{E} \exp(i\theta X) d\theta \\ &= \mathbb{E}[\mathbb{1}_{[-1,1]} * \mathbb{1}_{[-1,1]}(X)] \leq 2\mathbb{P}[|X| \leq 2] \leq 4C_{4.6} \mu t. \quad \square \end{aligned}$$

**4.4. Threshold structure theorem.** We now prove the following improved version of Theorem 3.7.

**Theorem 4.8.** *Fix  $K \geq 1$  and  $0 < p \leq (2 - \sqrt{2})/4$ . We can choose  $L, c > 0$  and  $c' = c'(p)$  so that the following holds for sufficiently large  $n$ . For all  $\lambda \in (n^{-2/3}, c(\log n)^{1/4} n^{-1/2})$ , we have for any  $u \in \mathbb{R}^n$  that*

$$\mathbb{P}[\exists v \in \mathbb{S}^{n-1} : (Av \parallel u) \wedge (v \in \text{Comp}(c_0, c_1) \vee \widehat{\mathcal{T}}_{p,L}(v, \lambda) \geq 2^{-c'\lambda n}) \wedge (\|A\| \leq K\sqrt{n})] \leq 2e^{-cn},$$

*where  $A$  is a symmetric matrix with entries on and above the diagonal i.i.d. and distributed as the sum of a Rademacher random variable, and a Gaussian random variable of mean 0 and variance  $n^{-2n}$ .*

*Remark.* The Gaussian perturbation of the entries of  $A$  is not important here, and will only be used later, where it will be convenient to assume that various sub-matrices of  $A$  are invertible almost surely. Moreover, the variance of the Gaussian is chosen sufficiently small so that all anticoncentration claims that we need are essentially unaffected by this perturbation.

*Proof.* As in the proof of Theorem 3.7, we can deal with compressible vectors using Lemma 2.2. Therefore, it remains to deal with incompressible vectors with “large” median threshold.

By standard small-ball estimates for incompressible vectors (see [24, Lemma 5.1]), for  $v \in \text{Incomp}(c_0, c_1)$ , there is  $C_0 = C_0(p, c_0, c_1)$  such that  $\widehat{\mathcal{T}}_{p,L}(v, \lambda) \leq C_0(\lambda n)^{-1/2}$ . We let  $r = \lceil \lambda n \rceil$  and  $k = c_{2.5}(\lambda)n$ , so that  $r|k$  by definition, and let  $m = \lfloor k/(2r) \rfloor$ .

**Step 1: Randomized rounding.** We consider the case  $\widehat{\mathcal{T}}_{p,L}(v, \lambda) \in [1/T, 2/T]$ , where  $T \in [C_0^{-1}\sqrt{\lambda n}, 2^{c'\lambda n}]$ . Then, by definition, there exist intervals  $I_1, \dots, I_m$  of the form  $\text{Spread}_\lambda^j(v)$  with

$$\mathcal{T}_{p,L}(v_{I_i}/\|v_{I_i}\|_2) \leq 2/T$$

for all  $i \in [m]$ .

Let  $D = C_1\sqrt{n}T$ , where  $C_1 = C_1(p, c_0, c_1) \geq 1$  will be an integer chosen later. Let  $y = Dv$ . By the definition of the threshold, for all  $t \geq \sqrt{\lceil \lambda n \rceil}$  we have

$$\begin{aligned} \mathcal{L}\left(\sum_{j \in I_i} b'_j y_j, t\right) &= \mathcal{L}\left(\sum_{j \in I_i} b'_j v_j, \frac{t}{D}\right) = \mathcal{L}\left(\sum_{j \in I_i} b'_j \frac{v_j}{\|v_{I_i}\|_2}, \frac{t}{D\|v_{I_i}\|_2}\right) \\ &\leq \mathcal{L}\left(\sum_{j \in I_i} b'_j \frac{v_j}{\|v_{I_i}\|_2}, \frac{\sqrt{2}t}{c_1 C_1 \sqrt{\lceil \lambda n \rceil} T}\right) \leq \frac{L}{T} \cdot \frac{2t}{\sqrt{\lceil \lambda n \rceil}}, \end{aligned}$$

as long as we chose  $C_1 > \sqrt{2}/c_1$ , where the  $b'_i$  are independent random variables distributed as  $\text{Ber}(p) - \text{Ber}'(p)$ . Applying Lemma 4.7 to the  $\lceil \lambda n \rceil$ -dimension vector  $y_{I_i}$ , we see that there is  $y'_{I_i} \in \mathbb{Z}^{\lceil \lambda n \rceil}$  satisfying the conclusions of Lemma 4.7 (with  $n$  replaced by  $\lceil \lambda n \rceil$ ). In particular, by (R3), we see that

$$\begin{aligned} \mathcal{L}\left(\sum_{j \in I_i} b'_j y'_{j_i}, \sqrt{\lceil \lambda n \rceil}\right) &\geq c_{4.7} \mathcal{L}\left(\sum_{j \in I_i} b'_j v_j, \sqrt{\lambda}/(C_1 T)\right) \\ &\geq c_{4.7} \mathcal{L}\left(\sum_{j \in I_i} b'_j \frac{v_j}{\|v_{I_i}\|_2}, \frac{2\sqrt{\lambda}}{C_1 T \cdot \sqrt{\lambda/c_0}}\right) \\ &\geq C_1^{-1} \sqrt{c_0} \cdot c_{4.7} \mathcal{L}\left(\sum_{j \in I_i} b'_j \frac{v_j}{\|v_{I_i}\|_2}, 2/T\right) \\ &\geq C_1^{-1} \sqrt{c_0} \cdot c_{4.7} \cdot 2LT^{-1}. \end{aligned} \tag{4.1}$$

Let  $I_0 = [n] \setminus (I_1 \cup \dots \cup I_m)$ . Then, by approximating each coordinate of  $y_{I_0}$  by the nearest integer, and combining with the above integer approximations of  $y_{I_1}, \dots, y_{I_m}$ , we obtain an integer vector  $y' \in \mathbb{Z}^n$ .

**Step 2: Size of nets of level sets.** We now estimate the number of possible realizations  $y'$ . This is the analogue of Proposition 3.3 in the present context. By paying an overall factor of at most  $6^n$ , we may fix  $\text{Spread}(v)$  (hence all the  $\text{Spread}_\lambda^j(v)$ ), as well as which  $\text{Spread}_\lambda^j(v)$  are in  $\mathcal{I}_M(v)$  and  $\mathcal{J}_M(v)$ . As above, let us denote the intervals  $\text{Spread}_\lambda^j(v)$  in  $\mathcal{I}_M(v)$  by  $I_1, \dots, I_m$ , and let  $I_0 = [n] \setminus (I_1 \cup \dots \cup I_m)$ .

First, note that the number of choices for  $y'_{I_0}$  is at most  $(CD/\sqrt{n})^{|I_0|}$  for an absolute constant  $C$  – this follows since  $y'_{I_0}$  is an integer point in a ball of radius  $D \geq \sqrt{n} \geq \sqrt{|I_0|}$  (provided that  $C_1$  is chosen sufficiently large), at which point, we can use a standard volumetric estimate for the number of integer points in  $\mathbb{R}^{|I_0|}$  in a ball of radius  $R \geq \sqrt{|I_0|}$ , together with the bound  $|I_0| \geq n/2$ .

Next, we fix  $i \in [m]$ , and bound the number of choices for  $y'_{I_i}$ . Note that for any  $r \geq 0$ ,

$$\mathcal{L}\left(\sum_{j \in I_i} b_j y'_j, r\right) \geq \mathcal{L}\left(\sum_{j \in I_i} (b_j - \tilde{b}_j) y'_j, r\right),$$

where  $b_i, \tilde{b}_i$  are independent copies of  $\text{Ber}(p)$ . Since  $b'_j$  is distributed as  $b_j - \tilde{b}_j$ , it follows from (4.1) that

$$\mathcal{L}\left(\sum_{j \in I_i} b_j y'_j, \sqrt{\lceil \lambda n \rceil}\right) \geq c_2 L N^{-1},$$

where  $b_j$  are i.i.d.  $\text{Ber}(p)$  random variables. From the definition of  $\text{Spread}_\lambda(v)$  and (R1), we see that  $y'_{I_i}$  lies within a  $(D/(C_2\sqrt{n}), \lceil \lambda n \rceil, K', 1)$ -admissible set  $\mathcal{A}$  for  $C_2$  and  $K'$  sufficiently large depending on  $c_0, c_1$ . Then, for  $L$  sufficiently large depending on  $c_0, c_1, p$ , by [Theorem 4.4](#) (noting that  $D$  is bounded by  $n2^{c'\lambda n}$  for all sufficiently large  $n$ , and that we can take  $c'$  to be sufficiently small depending on  $p$ ), we deduce that the number of potential  $y'_{I_i} \in \mathbb{Z}^{I_i}$  is bounded by

$$\exp(-M|I_i|)(CD/\sqrt{n})^{|I_i|},$$

where  $C$  depends on  $c_0, c_1$ , and  $M$  grows as  $\sqrt{\log \lceil \lambda n \rceil}$ , hence as  $\sqrt{\log n}$ . Explicitly, we can pick  $M \geq c_3\sqrt{\log n}$  for some small  $c_3 > 0$  depending only on  $c_0, c_1, p$ . Multiplying the total number of possibilities for  $y'_0, y'_1, \dots, y'_m$ , we see that the total number of possibilities for  $y' \in \mathbb{Z}^n$  is at most

$$\exp(-c_{2.3}Mn/8)(CD/\sqrt{n})^n,$$

for  $C$  depending on  $c_0, c_1$  and  $M \geq c_3\sqrt{\log n}$  with  $c_3$  depending on  $c_0, c_1, p$ .

**Step 3: Small-ball probability for net points.** Fix  $y' \in \mathbb{Z}^n$  resulting from the randomized rounding process and  $u \in \mathbb{R}^n$ . Our goal is to bound  $\mathbb{P}[\|Ay' - u\|_2 \leq Kn]$ . As in the proof of [Lemma 3.6](#), we can without loss of generality permute the coordinates of  $y'$  so that  $I_1, \dots, I_m$  are the first  $m$  blocks of size  $\lceil \lambda n \rceil$  within  $[n]$ . Then, for all  $\lceil \lambda n \rceil \leq j \leq c_{2.5}(\lambda)n$ , we have for all  $\epsilon \geq 0$  that

$$\mathcal{L}((Ay' - u)_j | (Ay' - u)_{j+1, \dots, n}, D\epsilon) \leq CL \left( \frac{\epsilon}{\sqrt{j/n}} + \frac{\sqrt{\lambda n/j}}{T} \right),$$

where  $C$  is an absolute constant. To deduce this, we use that the first  $j$  elements of row  $j$  are independent of rows  $j+1, \dots, n$ , then use [Lemma 4.5](#) to replace the Rademacher entries of  $A$  (plus the small Gaussian perturbation, which has variance so small that it can be disregarded) by  $\text{Ber}(p) - \text{Ber}'(p)$ , and finally use [Lemma 2.9](#) (as in the proof of [Proposition 3.5](#)) to stitch together the Lévy concentration properties of each  $y'_{I_i}$  (guaranteed by (R3) of [Lemma 4.7](#)). Combining this with [Lemma 2.7](#), we see that for  $y', u$  as above,

$$\mathbb{P}[\|Ay' - u\|_2 \leq Kn] \leq \left( \frac{C''L\sqrt{n}}{D} \right)^{n - \lceil \lambda n \rceil},$$

where  $C''$  depends only on  $c_0, c_1, p$ .

**Step 4: Union bound.** On the event  $\|A\| \leq K\sqrt{n}$ ,  $Av = tu$  with  $\|tu\|_2 \leq K\sqrt{n}$ . By splitting the range  $\{tu\}$  into  $(4D/\sqrt{n})^2$  intervals, and rounding  $v$  as in Step 2, we see that the probability of the event in question is bounded above by

$$\exp(-c_{2.3}Mn/8) \left( \frac{CD}{\sqrt{n}} \right)^{n+2} \sup_{y', u} \mathbb{P}[\|Ay' - u\|_2 \leq Kn],$$

where the supremum is over  $u \in \mathbb{R}^n$  and  $y' \in \mathbb{Z}^n$  such that each  $y'_{I_i}$  for  $i \in [m]$  satisfies the conclusions of [Lemma 4.7](#) (with  $n$  replaced by  $\lceil \lambda n \rceil$ ). Controlling the final factor by Step 3, we see that the probability is bounded above by

$$\exp(-c_{2.3}Mn/8) C^n D^{\lceil \lambda n \rceil + 2},$$

where  $C$  depends on  $c_0, c_1, p$ . Finally, since  $D \leq 2^{2c'\lambda n}$  for all  $n$  sufficiently large (depending on  $c_0, c_1, p$ ), we obtain an overall upper bound of

$$\exp(-c_{2.3}Mn/8 + n \log C + 2c'\lambda^2 n^2 + 6c'\lambda n).$$

Since  $M \geq c_3 \sqrt{\log n}$ , for  $c_3$  depending on  $c_0, c_1, p$ , the desired result follows by choosing  $\lambda < c(\log n)^{1/4} n^{-1/2}$  for  $c$  sufficiently small depending on  $c_0, c_1, p$ , so that the quantity above is bounded by  $\exp(-\Omega(n(\log n)^{1/2}))$ .  $\square$

## 5. PROOF OF THEOREM 1.1

In this section (along with [Appendix A](#)), we complete the proof of [Theorem 1.1](#) by closely following [\[25\]](#) with appropriate modifications. Since the smallest singular value is a continuous function of the entries of the matrix, by perturbing each entry of the random matrix by a Gaussian variable with arbitrarily small variance, one may assume that  $\xi$  is absolutely continuous with respect to the Lebesgue measure; in particular, one may freely assume that various square matrices whose entries are independent copies of  $\xi$  are invertible.

**5.1. Quadratic small-ball probabilities.** To prove [Theorems 1.1](#) and [1.2](#), we need the following small-ball inequalities for quadratic forms. The derivation is almost identical to the approach in [\[25, Theorem 8.1\]](#), with improvements coming from [Theorem 3.7](#) and [Proposition 3.5](#) (respectively [Theorem 4.8](#)). We include details in the appendix for the reader's convenience.

**Theorem 5.1.** *Let  $A$  be an  $n \times n$  symmetric random matrix whose independent entries are identical copies of a sub-Gaussian random variable  $\xi$  with variance 1. Suppose  $X$  is a random vector (independent of  $A$ ) whose entries are independent copies of  $\xi$ . Then, for every  $\epsilon \geq 0$  and  $u \in \mathbb{R}$ , we have*

$$\mathbb{P} \left[ \frac{|\langle A^{-1}X, X \rangle - u|}{\sqrt{1 + \|A^{-1}X\|_2^2}} \leq \epsilon \wedge \|A\| \leq K\sqrt{n} \right] \leq C_{5.1} \epsilon^{1/8} + 2 \exp(-c_{5.1} n^{1/2}).$$

We similarly derive the following strengthening for Rademacher entries.

**Theorem 5.2.** *Let  $A$  be an  $n \times n$  symmetric random matrix whose independent entries are distributed as the sum of a Rademacher random variable and a centered Gaussian with variance  $n^{-2n}$ . Suppose  $X$  is a random vector (independent of  $A$ ) whose entries are independent Rademachers. Then, for all sufficiently large  $n$ , and for every  $\epsilon \geq 0$  and  $u \in \mathbb{R}$ , we have*

$$\mathbb{P} \left[ \frac{|\langle A^{-1}X, X \rangle - u|}{\sqrt{1 + \|A^{-1}X\|_2^2}} \leq \epsilon \wedge \|A\| \leq K\sqrt{n} \right] \leq C_{5.1} \epsilon^{1/8} + 2 \exp(-c_{5.1} n^{1/2} (\log n)^{1/4}).$$

**5.2. Putting it together.** Given the above, the proofs of [Theorems 1.1](#) and [1.2](#) follows from a modification (due to Vershynin) of the invertibility-via-distance paradigm due to Rudelson and Vershynin. We reproduce the details from [\[25\]](#) for the reader's convenience

*Proof of [Theorems 1.1](#) and [1.2](#).* Fix  $c_0, c_1, c \in (0, 1)$ , as guaranteed by [Lemma 2.2](#). We can clearly assume that  $\epsilon \leq c$ . Then, by the union bound and [Lemma 2.2](#), we have

$$\begin{aligned} \mathbb{P}[s_n(A) \leq \epsilon/\sqrt{n}] &\leq \mathbb{P}[\exists v \in \text{Comp}(c_0, c_1) : \|Av\|_2 \leq c\sqrt{n}] + \mathbb{P}[\exists v \in \text{Incomp}(c_0, c_1) : \|Av\|_2 \leq \epsilon/\sqrt{n}] \\ &\leq 2 \exp(-cn) + \mathbb{P}[\exists v \in \text{Incomp}(c_0, c_1) : \|Av\|_2 \leq \epsilon/\sqrt{n}]. \end{aligned}$$

Let  $A_1, \dots, A_n$  denote the rows of  $A$ , and note that, by symmetry,

$$Av = \sum_{i=1}^n v_i A_i^T.$$

In particular,

$$\|Av\|_2 \geq |v_i| \text{dist}(A_i, H_i),$$

where  $H_i$  is the span of the rows  $A_j$  for  $j \neq i$ . Since  $|v_i| \geq c_1/2\sqrt{n}$  for all  $i \in \text{Spread}(v)$ , it follows that if  $\|Av\|_2 \leq \epsilon/\sqrt{n}$  for some  $v \in \text{Incomp}(c_0, c_1)$ , then we must necessarily have

$$\text{dist}(A_i, H_i) \leq \frac{\epsilon\sqrt{2}}{c_1}$$

for at least  $c_{2.3}n$  indices  $i \in [n]$ . Thus, we see that the probability that  $s_n(A) \leq \epsilon/\sqrt{n}$  is at most

$$2 \exp(-cn) + \frac{1}{c_{2.3}n} \sum_{i=1}^n \mathbb{P} \left[ \text{dist}(A_i, H_i) \leq \frac{\epsilon\sqrt{2}}{c_1} \right].$$

Therefore, for [Theorem 1.1](#) it suffices to show that

$$\mathbb{P}[\text{dist}(A_1, H_1) \leq \epsilon] \leq C\epsilon^{1/8} + 2 \exp(-cn^{1/2}).$$

A direct computation ([\[25, Proposition 5.1\]](#)) shows that

$$\text{dist}(A_1, H_1) = \frac{|\langle (A')^{-1}X, X \rangle - a_{11}|}{\sqrt{1 + \|(A')^{-1}X\|_2^2}},$$

where  $A'$  is the bottom right  $(n-1) \times (n-1)$  block of  $A$ , and  $X$  is the first column of  $A$  with the top element removed. At this point, we can apply [Theorem 5.1](#) to conclude. If  $A$  has Rademacher entries, by continuity we can transfer the singular value estimate to the model where the distribution is perturbed by a centered Gaussian with sufficiently small variance, at which point, an application of [Theorem 5.2](#) allows us to conclude.  $\square$

## REFERENCES

- [1] József Balogh, Robert Morris, and Wojciech Samotij, *The method of hypergraph containers*, Proceedings of the International Congress of Mathematicians—Rio de Janeiro 2018. Vol. IV. Invited lectures, World Sci. Publ., Hackensack, NJ, 2018, pp. 3059–3092.
- [2] Jean Bourgain, Van H. Vu, and Philip Matchett Wood, *On the singularity probability of discrete random matrices*, Journal of Functional Analysis **258** (2010), 559–603.
- [3] Marcelo Campos, Letícia Mattos, Robert Morris, and Natasha Morrison, *On the singularity of random symmetric matrices*, Duke Mathematical Journal (2020), to appear.
- [4] Kevin P. Costello, Terence Tao, and Van Vu, *Random symmetric matrices are almost surely nonsingular*, Duke Math. J. **135** (2006), 395–413.
- [5] Asaf Ferber and Vishesh Jain, *Singularity of random symmetric matrices—a combinatorial approach to improved bounds*, Forum of Mathematics, Sigma, vol. 7, Cambridge University Press, 2019.
- [6] Asaf Ferber, Vishesh Jain, Kyle Luh, and Wojciech Samotij, *On the counting problem in inverse Littlewood–Offord theory*, arXiv:1904.10425.
- [7] Vishesh Jain, Ashwin Sah, and Mehtaab Sawhney, *Singularity of discrete random matrices II*, arXiv:2010.06554.
- [8] Vishesh Jain, Ashwin Sah, and Mehtaab Sawhney, *The smallest singular value of dense random regular digraphs*, arXiv:2008.04755.
- [9] Jeff Kahn, János Komlós, and Endre Szemerédi, *On the probability that a random  $\pm 1$ -matrix is singular*, Journal of the American Mathematical Society **8** (1995), 223–240.
- [10] János Komlós, *On determinant of  $(0, 1)$  matrices*, Studia Science Mathematica Hungarica **2** (1967), 7–21.
- [11] Galyna V Livshyts, *The smallest singular value of heavy-tailed not necessarily iid random matrices via random rounding*, arXiv:1811.07038.
- [12] Galyna V Livshyts, Konstantin Tikhomirov, and Roman Vershynin, *The smallest singular value of inhomogeneous square random matrices*, arXiv:1909.04219.
- [13] Patrick Lopatto and Kyle Luh, *Tail bounds for gaps between eigenvalues of sparse random matrices*, arXiv:1901.05948.
- [14] Kyle Luh and Van Vu, *Sparse random matrices have simple spectrum*, arXiv:1802.03662.
- [15] A. L. Mirošnikov and B. A. Rogozin, *Inequalities for concentration functions*, Teor. Veroyatnost. i Primenen. **25** (1980), 178–183.
- [16] Hoi Nguyen, Terence Tao, and Van Vu, *Random matrices: tail bounds for gaps between eigenvalues*, Probability Theory and Related Fields **167** (2017), 777–816.

- [17] Hoi H. Nguyen, *Inverse Littlewood–Offord problems and the singularity of random symmetric matrices*, Duke Math. J. **161** (2012), 545–586.
- [18] Hoi H. Nguyen, *On the least singular value of random symmetric matrices*, Electron. J. Probab. **17** (2012), no. 53, 19.
- [19] Sean O’Rourke and Behrouz Touri, *On a conjecture of Godsil concerning controllable random graphs*, SIAM Journal on Control and Optimization **54** (2016), 3347–3378.
- [20] B. A. Rogozin, *On the increase of dispersion of sums of independent random variables*, Teor. Veroyatnost. i Primenen **6** (1961), 106–108.
- [21] Mark Rudelson and Roman Vershynin, *The Littlewood–Offord problem and invertibility of random matrices*, Advances in Mathematics **218** (2008), 600–633.
- [22] Terence Tao and Van Vu, *On random  $\pm 1$  matrices: singularity and determinant*, Random Structures & Algorithms **28** (2006), 1–23.
- [23] Terence Tao and Van H. Vu, *On the singularity probability of random Bernoulli matrices*, Journal of the American Mathematical Society **20** (2007), 603–628.
- [24] Konstantin Tikhomirov, *Singularity of random Bernoulli matrices*, Ann. of Math. (2) **191** (2020), 593–634.
- [25] Roman Vershynin, *Invertibility of symmetric random matrices*, Random Structures Algorithms **44** (2014), 135–182.
- [26] Feng Wei, *Investigate invertibility of sparse symmetric matrix*, arXiv:1712.04341.

## APPENDIX A. QUADRATIC SMALL-BALL PROBABILITIES

The purpose of this appendix is to prove [Theorem 5.1](#) for completeness. We will also briefly note the necessary modifications to deduce [Theorem 5.2](#).

We essentially replicate the argument in [25, Section 8] with the obvious modifications. In brief, our improved anticoncentration estimate [Proposition 3.5](#) will allow us to replace the  $\epsilon^{1/8+\eta}$  dependence in [25] with  $\epsilon^{1/8}$ , and the improved range of arithmetic structure derived in [Theorem 3.7](#) will allow us to achieve an error term of  $\exp(-\Omega(\sqrt{n}))$ . For the sake of simplicity, we define the event

$$\mathcal{E}_K := \{\|A\| \leq K\sqrt{n}\}.$$

**Proposition A.1** (Analogue of [25, Proposition 8.2]). *Let  $A$  be a symmetric random matrix whose independent entries are identical copies of a sub-Gaussian random variable with mean 0 and variance 1. Let  $X$  be a random vector (independent of  $A$ ) whose coordinates are i.i.d. copies of  $\xi$ . There exist constants  $C, c > 0$  depending only on the sub-Gaussian moment of  $\xi$  for which the following holds.*

*For  $\lambda \in (C/n, 1/\sqrt{n})$ ,  $A$  satisfies the following with probability at least  $1 - 2e^{-cn}$ : if  $\mathcal{E}_K$  holds, then for every  $\epsilon > 0$ :*

- $\|A^{-1}X\|_2 \geq c$  with probability at least  $1 - e^{-cn}$  in the randomness of  $X$ .
- $\|A^{-1}X\|_2 \leq \epsilon^{-1/2}\|A^{-1}\|_{\text{HS}}$  with probability at least  $1 - \epsilon$  in the randomness of  $X$ .
- $\|A^{-1}X\|_2 \geq \epsilon\|A^{-1}\|_{\text{HS}}$  with probability at least  $1 - C\epsilon - 2e^{-c\lambda n}$  in the randomness of  $X$ .

*Proof.* The first two parts have the same proof as in [25, Proposition 8.2]. The last part also has essentially the same proof, except that we use [Proposition 3.5](#) in place of [25, Proposition 6.9] and use [Theorem 3.7](#) in place of [25, Theorem 7.1].  $\square$

*Remark.* For [Theorem 5.2](#), we note that if  $A$  has entries which are Rademacher plus a centered Gaussian with sufficiently small variance, we can prove the same statement with  $n^{-2/3} < \lambda < c(\log n)^{1/4}n^{-1/2}$ . We use [Theorem 4.8](#) instead of [Theorem 3.7](#) and the analogue of [Proposition 3.5](#) for the threshold. The remaining part of the proof is exactly the same.

Next, we require the following decoupling lemma from [25]; this use of decoupling to establish singularity for symmetric random matrices originates in work of Costello, Tao, and Vu [4], and has been used in essentially all follow-up works.

**Lemma A.2** ([25, Lemma 8.4]). *Let  $G$  be an arbitrary symmetric  $n \times n$  matrix, and let  $X, X'$  be independent samples of a random vector in  $\mathbb{R}^n$  with independent coordinates. Let  $J \subseteq [n]$ . Then,*

for every  $\epsilon \geq 0$  we have

$$\mathcal{L}(\langle GX, X \rangle, \epsilon)^2 \leq \mathbb{P}_{X, X'}[|\langle GP_{J^c}(X - X'), P_J X \rangle - v| \leq \epsilon]$$

for some random variable  $v$  determined by  $G|_{J^c \times J^c}$  and  $P_{J^c} X, P_{J^c} X'$ .

We can now prove [Theorem 5.1](#); we refer the reader to [\[25\]](#) for a more detailed exposition.

*Proof of [Theorem 5.1](#).* We randomly choose  $J \subseteq [n]$  by sampling elements independently with probability  $1 - c_{2.3}/2$ . We trivially see by the Chernoff bound that if  $\mathcal{E}_J = \{|J^c| \leq c_{2.3}n\}$ , then

$$\mathbb{P}[\mathcal{E}_J] \geq 1 - 2e^{-cn}.$$

For  $J$  satisfying  $\mathcal{E}_J$ , let us assign the set  $\text{Spread}(v)$  for  $v \in \text{Incomp}(c_0, c_1)$  in a way such that  $\text{Spread}(x) \subseteq |J|$ . We can do this since, in [Lemma 2.3](#), we chose  $c_{2.3}$  so as to have at least  $2c_{2.3}n$  spread coordinates. We will then use this assignment to obtain the median regularized LCDs that are used.

Next, consider the event  $\mathcal{E}_D$  given by

$$\epsilon_0^{1/2} \sqrt{1 + \|A^{-1}X\|_2^2} \leq \|A^{-1}\|_{\text{HS}} \leq \frac{1}{\epsilon_0} \|A^{-1}P_{J^c}(X - X')\|_2.$$

Applying [Proposition A.1](#) to  $X$  and  $Y_i = \delta_i(X_i - X'_i)$ , where  $\delta_i$  is the indicator of  $i \in J^c$ , and adjusting constants appropriately, we find that

$$\mathbb{P}_{J, A, X, X'}[\mathcal{E}_D \vee \mathcal{E}_K^c] \geq 1 - C\epsilon_0 - 2e^{-c\lambda n} - 2e^{-cn},$$

where the constants depend only on the sub-Gaussian norm of  $\xi$ . Now define

$$x_0 = \frac{A^{-1}P_{J^c}(X - X')}{\|A^{-1}P_{J^c}(X - X')\|_2},$$

which is a random vector. If the denominator is 0, we can use an arbitrary fixed vector. Let  $\mathcal{E}_U$  be the event (analogue of [\[25, Equation 8.11\]](#)) that

$$x_0 \in \text{Incomp}(c_0, c_1), \quad \widehat{MD}_L(x_0, \lambda) \geq 2^{\lambda n/C},$$

where we choose  $L$  as in [Theorem 3.7](#).

Now condition on some  $J$  satisfying  $\mathcal{E}_J$  and some  $X, X'$ . By [Theorem 3.7](#), we deduce that

$$\mathbb{P}_A[\mathcal{E}_U \vee \mathcal{E}_K^c | X, X', J] \geq 1 - 2e^{-cn}.$$

Thus (analogue of [\[25, Equation 8.12\]](#))

$$\mathbb{P}_{J, A, X, X'}[(\mathcal{E}_J \wedge \mathcal{E}_D \wedge \mathcal{E}_U) \vee \mathcal{E}_K^c] \geq 1 - p_0,$$

where  $p_0 = C \max(\epsilon_0, 2^{-\lambda n/C})$ . Hence, there is a realization of  $J$  such that  $\mathcal{E}_J$  holds and

$$\mathbb{P}_{A, X, X'}[(\mathcal{E}_D \wedge \mathcal{E}_U) \vee \mathcal{E}_K^c] \geq 1 - p_0.$$

We fix this choice of  $J$  for the remainder of the proof. Now let  $\mathcal{E}_A$  be the event, dependent only on  $A$ , that simultaneously  $\mathcal{E}_K$  and

$$\mathbb{P}_{X, X'}[\mathcal{E}_D \wedge \mathcal{E}_U | A] \geq 1 - p_0^{1/2}.$$

By Fubini's theorem, Markov's inequality, and the fact that  $\mathcal{E}_K$  depends only on  $A$ , we see from the above that (analogue of [\[25, Equation 8.13\]](#))

$$\mathbb{P}_A[\mathcal{E}_A \vee \mathcal{E}_K^c] \geq 1 - p_0^{1/2}.$$

Now, if  $\mathcal{E}$  is the desired event

$$\frac{|\langle A^{-1}X, X \rangle - u|}{\sqrt{1 + \|A^{-1}X\|_2^2}} \leq \epsilon,$$

then

$$\mathbb{P}_{A,X}[\mathcal{E}] \leq \mathbb{P}[\mathcal{E}_K^c] + p_0^{1/2} + \sup_{A \in \mathcal{E}_A} \mathbb{P}_X[\mathcal{E}|A].$$

Fix some  $A \in \mathcal{E}_A$  for the remainder of the proof. We need to bound

$$\mathbb{P}_X[\mathcal{E}|A] \leq \mathbb{P}_{X,X'}[\mathcal{E} \wedge \mathcal{E}_D|A] + p_0^{1/2}.$$

Using  $\mathcal{E}_D$  along with  $\mathcal{E}$ , we see that

$$\mathbb{P}_{X,X'}[\mathcal{E} \wedge \mathcal{E}_D|A] \leq \mathbb{P}_{X,X'}[|\langle A^{-1}X, X \rangle - u| \leq \epsilon \epsilon_0^{-1/2} \|A^{-1}\|_{\text{HS}}|A] =: p_1.$$

Then by [Lemma A.2](#), we find that this satisfies

$$p_1^2 \leq \mathbb{P}_{X,X'}[|\langle A^{-1}P_{J^c}(X - X'), P_JX \rangle - v| \leq \epsilon \epsilon_0^{-1/2} \|A^{-1}\|_{\text{HS}}|A],$$

where  $v = v(A^{-1}, P_{J^c}X, P_{J^c}X')$  is some random variable depending on these parameters only. This last probability is at most

$$p_0^{1/2} + \mathbb{P}_{X,X'}[|\langle A^{-1}P_{J^c}(X - X'), P_JX \rangle - v| \leq \epsilon \epsilon_0^{-1/2} \|A^{-1}\|_{\text{HS}} \wedge \mathcal{E}_D \wedge \mathcal{E}_U|A].$$

Now by using  $\mathcal{E}_D$  again, and dividing the inequality in question by  $\|A^{-1}P_{J^c}(X - X')\|_2$ , we see that (analogue of [\[25, Equation 8.15\]](#))

$$p_1^2 \leq p_0^{1/2} + \mathbb{P}_{X,X'}[|\langle x_0, P_JX \rangle - w| \leq \epsilon_0^{-3/2} \epsilon \wedge \mathcal{E}_U|A].$$

Finally, we can apply [Proposition 3.5](#) to this random variable. Note that  $x_0, w$  do not depend on  $P_JX$ . Also, we know from  $\mathcal{E}_U$  that  $\widehat{MD}_L(x_0, \lambda) \geq 2^{\lambda n/C}$ . It suffices to check that  $\text{Spread}_\lambda(x_0) \subseteq J$  by the definitions chosen at the beginning; hence, we can drop the randomness in  $P_JX$  of all coordinates except for those in the spread set and apply the result. We again technically need to check that  $\text{Spread}_\lambda(x_0)$  satisfies the conditions on [Proposition 3.5](#), which we have already implicitly verified before. Overall, we deduce

$$\mathbb{P}_{X,X'}[|\langle x_0, P_JX \rangle - w| \leq \epsilon_0^{-3/2} \epsilon \wedge \mathcal{E}_U|A] \leq C \epsilon_0^{-3/2} \epsilon + 2^{-\lambda n/C}.$$

Finally, tracing it all back, we have

$$\mathbb{P}[\mathcal{E}] \leq \mathbb{P}[\mathcal{E}_K^c] + 2p_0^{1/2} + p_1 \leq 2e^{-cn} + C\epsilon_0^{1/2} + C\epsilon_0^{1/4} + C\epsilon_0^{-3/4} \epsilon^{1/2} + 2^{-\lambda n/C}$$

by sub-Gaussian concentration of the operator norm of  $A$  ([\[25, Lemma 2.3\]](#)). Now choosing  $\epsilon_0 = \epsilon^{1/2}$  and  $\lambda = 1/\sqrt{n}$ , which is the biggest permitted by [Theorem 3.7](#), we are done.  $\square$

SIMONS INSTITUTE FOR THE THEORY OF COMPUTING, BERKELEY, CA 94720, USA  
*Email address:* visheshj@stanford.edu

DEPARTMENT OF MATHEMATICS, MASSACHUSETTS INSTITUTE OF TECHNOLOGY, CAMBRIDGE, MA 02139, USA  
*Email address:* {asah,msawhney}@mit.edu