Annales de l'Institut Henri Poincaré - Probabilités et Statistiques
2020, Vol. 56, No. 4, 2307–2328
https://doi.org/10.1214/19-AIHP1032

ANNALES
DE L'INSTITUT
HENRI
POINCARÉ
PROBABILITÉS
ET STATISTIQUES

Sparse random matrices have simple spectrum

Kyle Luh[a,1] and Van Vu[b,2]

[a]*Department of Mathematics, University of Colorado Boulder, Boulder, CO, USA. E-mail: kyle.luh@colorado.edu*
[b]*Department of Mathematics, Yale University, New Haven, CT, USA . E-mail: van.vu@yale.edu*

Received 15 January 2019; revised 30 September 2019; accepted 8 October 2019

Abstract. Let M_n be a class of symmetric sparse random matrices, with independent entries $M_{ij} = \delta_{ij}\xi_{ij}$ for $i \le j$. δ_{ij} are i.i.d. Bernoulli random variables taking the value 1 with probability $p \ge n^{-1+\delta}$ for any constant $\delta > 0$ and ξ_{ij} are i.i.d. centered, subgaussian random variables. We show that with high probability this class of random matrices has simple spectrum (i.e. the eigenvalues appear with multiplicity one). We can slightly modify our proof to show that the adjacency matrix of a sparse Erdős–Rényi graph has simple spectrum for $n^{-1+\delta} \le p \le 1 - n^{-1+\delta}$. These results are optimal in the exponent. The result for graphs has connections to the notorious graph isomorphism problem.

Résumé. On définit une classe M_n de matrices symétriques clairsemées, à coefficients indépendants, en posant $M_{ij} = \delta_{ij}\xi_{ij}$ pour $i \le j$, où les δ_{ij} sont des variables aléatoires de Bernoulli i.i.d. prenant la valeur 1 avec probabilité $p \ge n^{-1+\delta}$ pour une constante $\delta > 0$ arbitraire, et les ξ_{ij} sont des variables aléatoires sous-gaussiennes i.i.d. centrées. Nous montrons qu'avec une grande probabilité, cette classe de matrices aléatoires a un spectre simple, c'est-à-dire que les valeurs propres sont de multiplicité 1. Une légère modification de la démonstration de ce résultat permet de montrer montrer que la matrice d'adjacence d'un graphe d'Erdős–Rényi clairsemé a un spectre simple pour $n^{-1+\delta} \le p \le 1 - n^{-1+\delta}$. Ces résultats sont optimaux en les exposants. Le résultat pour les graphes a des liens avec le célèbre problème de l'isomorphisme de graphe.

MSC2020 subject classifications: Primary 60B20; 15B52; secondary 05C80

Keywords: Random matrices; Sparse matrices; Eigenvalue degeneracy; Random graphs; Graph isomorphism

---

1. Introduction

The gaps between eigenvalues are natural objects to study in random matrix theory and are of central importance in the field. Since the introduction of the notion of a random matrix, there have been numerous inquiries into the spacings of consecutive eigenvalues of symmetric random matrices. For a matrix with eigenvalues λ_i, we denote the gaps by $\delta_i = \lambda_{i+1} - \lambda_i$. The limiting global gap distribution for Gaussian matrices (GOE and GUE) has been well understood for some time and can be deduced from Wigner's surmise [14, Chapter 6,7]. Recent progress on universality has extended these results to large classes of random variables [10,22]. At finer levels, meaning under proper normalization and for a particular gap, the limiting distribution for the GUE was only calculated in 2013 by Tao [19]. The four moment condition establishes that this distribution is universal for any random variable that matches the gaussian up to the first four moments. Using advanced dynamical techniques, Erdős and Yau removed this condition [9].

Although these results describe the behavior of a single gap, δ_i, bounds on the smallest gap, $\delta_{\min} = \min_i \delta_i$, for general matrices were still out of reach. Bourgade and Ben-Arous [5] showed that $\delta_{\min}$ is on the order of $n^{-5/6}$ for the GUE ensemble. Yet, currently, this issue does not fall into the scope of the four moment theorem. Although tail bounds for individual δ_i were known for more general matrices [21,22], they were too weak to survive the union bound over all i to conclude anything about $\delta_{\min}$. Under severe restrictions on the smoothness and decay of the entries, Erdős, Schlein and

[1]K. Luh is supported by the National Science Foundation under Award No. 1702533.
[2]V. Vu is supported by NSF grant DMS 1307797 and AFORS grant FA9550-12-1-0083.

Yau [8] proved that

$$\mathbb{P}\left(En^{1/2}-\frac{\varepsilon}{n^{1/2}}\le\lambda_i\le\lambda_k\le En^{1/2}+\frac{\varepsilon}{n^{1/2}}\text{ for some } i\right)=o\big(\varepsilon^{k^2}\big)$$

for any fixed $k \ge 1$, any $\varepsilon > 0$ and any bounded $E \in \mathbb{R}$. Applying a union bound to this result yields

$$\mathbb{P}\big(\delta_{\min}\le\delta n^{-1/2}\big)=o\big(n\delta^3\big)+\exp(-cn)$$

for any $\delta > 0$. Despite the strong bound, this result applies only to a small set of smooth random variables. Outside of this set, even whether $\delta_{\min}$ could equal zero could not be settled by these previous results and was only resolved in 2014. Phrased differently, the fact that a random matrix typically has simple spectrum (i.e. all eigenvalues have multiplicity one) is a recent result due to Tao and Vu [23]. They show that the probability that a random matrix has simple spectrum is bounded below by $1 - n^{-A}$ for any constant A. In [16], this qualitative statement was refined to quantitative tail bounds on the gaps between the eigenvalues and probability that a random matrix has simple spectrum was improved to $1 - \exp(-n^c)$ for a small unspecified constant c.

In the realm of graphs, whether or not a graph has simple spectrum (i.e. its adjancecy matrix has simple spectrum) has practical complexity implications. Although great strides have been made recently on the notorious graph isomorphism problem [1], the best running time guarantees are still *quasipolynomial*. However, Babai, Grigoryev and Mount [2] demonstrated ealier that the graph isomorphism problem restricted to the graphs with simple spectrum is in complexity class $\mathcal{P}$. A corollary of the random matrix result in [23] is that dense Erdős–Rényi random graphs have simple spectrum which answered a question of Babai's that had been open since the '80's.

In the past few years, there has been renewed interest in sparse random matrices due to their applications in data science, where they require less storage space and fewer operations to manipulate [6,7,15]. In other settings, sparse random matrices reflect the intuition that in many natural problems, each data point is dependent on only a few of the many parameters [11–13,25]. For random graphs, the more interesting behavior occurs for sparse graphs. Many real-world networks are sparse and applications often prefer graphs with fewer edges that maintain the necessary properties.

In this work, we establish that sparse random matrices have simple spectrum. Our result is nearly optimal in terms of the range of sparsity. In the dense range, our work improves the probability bound in [16] to $1 - \exp(-n^{1/128})$. The particular value of the constant $(1/128)$ is not meaningful and has not been optimized.

2. Main results

Let M_n be an $n \times n$ symmetric random matrix with entries $m_{ij} = \delta_{ij}\xi_{ij}$ for all $i \le j$, where δ_{ij} is a Bernoulli random variable that takes the value 1 with probability $p = p(n)$ and ξ_{ij} are iid random variables with mean zero, variance one, and subgaussian moment bounded by B. We remind the reader that the subgaussian moment of ξ_{ij} is the smallest t such that $\mathbb{E}\exp(\xi_{ij}^2/t^2) \le 2$. Our main result is the following.

Theorem 2.1. *For $0 < \delta \le 1$ a constant and $p \ge n^{-1+\delta}$, then with probability at least $1 - \exp(-(np)^{1/128})$, M_n has simple spectrum.*

Denote by $G(n, p)$, the random variable that takes values in the labeled graphs on $[n]$ vertices and distributed such that each edge appears independently with probability p.

Theorem 2.2. *Let A_n be the adjacency matrix of $G(n, p)$ and $0 < \delta \le 1$ a constant. For $n^{-1+\delta} \le p \le 1 - n^{-1+\delta}$, with probability at least $1 - \exp((np)^{-1/128})$, A_n has simple spectrum.*

Remark 2.3. Observe that for $p = o(\log n/n)$, there is likely to be at least two row of zeros in M_n and A_n. This yields a zero eigenvalue with multiplicity at least 2. Thus, our bound on p is near optimal. We record here that the upperbound on p does not appear in Theorem 2.1 as even with $p = 1$, there is additional randomness from the ξ_{ij}. However, for the adjacency matrix, for $p = 1$, we are left with the deterministic matrix $J_n - I_n$ which has eigenvalue -1 with multiplicity $n - 1$. By symmetry, the upperbound is also near optimal. In fact, we believe the true sparsity threshold is on the order of $p \ge \log n/n$, but our current method needs a technical refinement to achieve this bound and we postpone this matter for another occassion.

The remainder of the paper is organized as follows. In Section 3 we give a birds-eye view of the proof, avoiding any technical statements. In Section 4 we state several notational conveniences. In Sections 5 and 6 we develop the necessary tools to control the deviation of M_n acting on two different sets of vectors (compressible and incompressible respectively). Finally, in Section 7, we combine the results of the previous sections to obtain a short proof of Theorem 2.1. In the final section, we discuss the necessary modifications to handle adjacency matrices of sparse random graphs.

3. Proof strategy

The overall approach is analogous to that used in [23] and [16]. For M_n as in Theorem 2.1, we write

$$M_n = \begin{pmatrix} M_{n-1} & X \\ X^T & m_{nn} \end{pmatrix}, \tag{1}$$

where $X = (x_1, \dots, x_{n-1}) \in \mathbb{R}^{n-1}$. For a matrix X, let $\lambda_n(X) \le \cdots \le \lambda_1(X)$ be the eigenvalues of M_n. Let $v = (x, a)$ (where $x \in \mathbb{R}^{n-1}$ and $a \in \mathbb{R}$) be the unit eigenvector associated to $\lambda_i(M_n)$. By definition we have

$$\begin{pmatrix} M_{n-1} & X \\ X^T & m_{nn} \end{pmatrix} \begin{pmatrix} x \\ a \end{pmatrix} = \lambda_i(M_n) \begin{pmatrix} x \\ a \end{pmatrix}.$$

Restricting our attention to the top $n-1$ coordinates gives

$$\big(M_{n-1} - \lambda_i(M_n)\big)x + aX = 0.$$

Let w be the eigenvector of M_{n-1} corresponding to $\lambda_i(M_{n-1})$. After multiplying on the right by w^T, we deduce that

$$\big|aw^T X\big| = \big|w^T\big(M_{n-1} - \lambda_i(M_n)\big)x\big| = \big|\lambda_i(M_{n-1}) - \lambda_i(M_n)\big|\big|w^T x\big|.$$

By the Cauchy interlacing law, we must have $\lambda_i(M_n) \le \lambda_i(M_{n-1}) \le \lambda_{i-1}(M_n)$. Therefore, if we let $\mathcal{E}_i$ be the event that $\lambda_i(M_n) = \lambda_{i+1}(M_n)$, then assuming $a \neq 0$, on the event $\mathcal{E}_i$, this implies that $w^T X = 0$. A simple union bound over all choices of a in w removes our assumption that $a \neq 0$. Finally, if $\mathbb{P}(\mathcal{E}_i) = o(n^{-1})$ for all i, then a union bound yields the result.

Our task thus reduces to showing that an eigenvector w of M_{n-1} has the property that $\mathbb{P}(w^T X = 0)$ is small. Note that X and w are independent. By now, this is a well-studied phenomenon [17,18,20]. This small-ball probability is intimately related to the arithmetic structure of the vector w. The goal then is to prove that with high probability, an eigenvector of the submatrix M_{n-1} will not have much structure. For this intermediate objective, we make the simple observation that for v, a unit eigenvector of M_n, with eigenvalue λ,

$$(M_n - \lambda)v = 0.$$

For x close to v, $(M_n - \lambda)x \approx 0$. This is reminiscent of the least singular problem for a random matrix and the details of our argument draws heavily from the techniques of [3,18,24]. Choosing an appropriate net of the highly structured vectors in $\mathcal{S}^{n-1}$ and a net of potential eigenvalues, we show that these vectors are unlikely to be eigenvectors.

This aerial view of the argument obscures the technical obstacles that must be overcome when the matrices we deal with are sparse. As the random variables are zero with large probability, the small-ball probabilities that appear tend to be too large for direct union bounds to work. Delicate nets and careful balancing of probabilities is required to implement our overall strategy.

4. Notation

For a vector $v \in \mathbb{R}^n$ and an index set $I \subset [n]$, let $v_I \in \mathbb{R}^{|I|}$ be the restriction of v onto that index set and $P_I(v) \in \mathbb{R}^n$ be the vector v with all coordinates in I^c zeroed out.

We will also need finer control over index sets $I \subset [n]$. We let $ord(I)$ be the vector in $\mathbb{N}^{|I|}$ populated by elements of I in increasing order. Then, we define $I[k] := ord(I)_k$ and $I[k:k'] := ord(I)_{[k:k']}$, where $[k:k'] := \{i : k \le i \le k'\}$.

To avoid repition, we impose the assumption that, unless explicitly stated, any constants (usually numbered) in the statement of the Lemmas, Propositions and Theorems depend only on δ and the subgaussian moment B. Additionally, standard asymptotic notation (e.g. o, O) is stated with the assumption of n tending to infinity.

5. Compressible vectors

5.1. *Preliminaries*

We divide the unit sphere into two classes. The *compressible* vectors are those that are close to sparse vectors and the remaining vectors are called *incompressible*.

Definition 5.1. For $M \in \mathbb{N}$, a vector x is in $\mathrm{Sparse}(M)$ if $|\operatorname{Supp}(x)| \le M$. For a $\delta \in (0,1)$, we denote

$$\mathrm{Comp}(M,\delta) := \{x \in \mathcal{S}^{n-1} : \exists y \in \mathrm{Sparse}(M) \text{ such that } \|x-y\|_2 \le \delta\}.$$

The incompressible vectors are defined to be

$$\mathrm{Incomp}(M,\delta) := \{x \in \mathcal{S}^{n-1} : x \notin \mathrm{Comp}(M,\delta)\}.$$

We will often make use of the following bound on the operator norm of M_n.

Proposition 5.2. *There exist constants $K, c > 0$ such that*

$$\mathbb{P}(\|M_n\| \ge K\sqrt{pn}) \le \exp(-cpn).$$

Proof. The proof of Theorem 1.7 in [3] can easily be modified to handle symmetric random matrices. The details can be found in the proof of Theorem 1.14 [4]. □

5.2. *Compressible vectors*

Proposition 5.3. *There exist constants $C, \bar{C}, c, c' > 0$ depending only on B, such that for*

$$p \ge \frac{C\log n}{n}, \qquad \ell_0 := \left\lceil \frac{\log 1/(8p)}{\log\sqrt{pn}} \right\rceil \quad \text{and} \quad \lambda \in [-C\sqrt{pn}, C\sqrt{pn}]$$

we have

$$\mathbb{P}\big(\exists x \in \mathrm{Dom}\big(M,(K+C)^{-4}\big) \cup \mathrm{Comp}(M,\rho) \text{ s.t. } \big\|(M_n-\lambda)x\big\|_2 \le C\rho\sqrt{pn}\big) \le \exp\big(-c'pn\big),$$

where $\rho := \bar{C}^{-\ell_0-6}$ and $p^{-1} \le M \le cn$.

Remark 5.4. To gain some understanding of these parameters, observe that for $p = n^{-1+\delta}$ for some constant $\delta > 0$, then $\ell_0 = O(1)$. Near the threshold, when $p = \frac{\log n}{n}$, $\ell_0 = \Theta(\log n/\log\log n)$ so $\rho = \exp(-O(\log n/\log\log n))$.

Although this result is highly non-trivial, the proof follows from a straightforward adaptation of Proposition 3.1 in [3]. We include the proof with the necessary modifications in Appendix A.1.

From this result, we obtain a bound on the probability that an eigenvector is compressible.

Corollary 5.5. *For $p^{-1} \le M \le cn$, where c is the constant from Proposition 5.3, there exists a constant $c' > 0$ such that*

$$\mathbb{P}\big(\exists \text{ a unit eigenvector } \in \mathrm{Comp}(M,\rho)\big) \le \exp\big(-c'pn\big),$$

where $\rho := \bar{C}^{-\ell_0-6}$.

Proof. By Lemma 5.2, all eigenvalues of M_n are in the interval $I = [-K\sqrt{pn}, K\sqrt{pn}]$. Consider an n^{-1}-net of I which can be constructed to be of size at most $2Kn\sqrt{pn}$. For $\lambda \in I$ that is an eigenvalue of M_n with eigenvector $x \in \mathrm{Comp}(M,\rho)$, there exists an element of the net, λ_0, such that

$$\big\|(M_n-\lambda_0)x\big\|_2 = \big\|(\lambda-\lambda_0)x\big\|_2 \le n^{-1}.$$

However, by Proposition 5.3,

$$\mathbb{P}\big(\exists x \in \mathrm{Comp}(M,\rho) \text{ s.t. } \big\|(M_n-\lambda_0)x\big\|_2 \le n^{-1}\big) \le \exp\big(-c'pn\big).$$

Taking a union bound over the λ_0 and increasing C from Proposition 5.3 if necessary, yields the result. □

6. Incompressible vectors

For these vectors, we develop small-ball probability bounds that are dependent on a measure of arithmetic strucutre (Least Common Denominator) [18,24]. First, we introduce the following partition of the indices for $v \in \mathrm{Incomp}(M, \rho)$. Recall that $\rho := \bar{C}^{-\ell_0-6}$ with $\bar{C}$ from Proposition 5.3. Let

$$\sigma(v) := \left\{k : \frac{\rho}{\sqrt{2n}} \le |v_k| \le \frac{1}{\sqrt{M}}\right\}.$$

Due to the incompressibility of v, the cardinality of this set is large.

Lemma 6.1. *For $v \in \mathrm{Incomp}(M, \rho)$ where $\rho := \bar{C}^{-\ell_0-6}$,*

$$\left|\sigma(v)\right| \ge \frac{M\rho^2}{2}.$$

Proof. Define $\sigma_1(v) := \{k : |v_k| \le \frac{1}{\sqrt{M}}\}$. Since v is a unit vector, $|\sigma_1^c| \le M$. As $y = P_{\sigma_1^c} v$ is a sparse vector with support at most M, the definition of incompressible vectors implies $\|v - y\|_2 > \rho$ or $\|P_{\sigma_1}(v)\|_2^2 \ge \rho^2$. Define the following set to capture the lower bound.

$$\sigma_2(v) := \left\{k : |v_k| \ge \frac{\rho}{\sqrt{2n}}\right\}.$$

Clearly, $\|P_{\sigma_2}(v)\|_2^2 \le \rho^2/2$. Therefore,

$$\left\|P_\sigma(v)\right\|_2^2 \ge \left\|P_{\sigma_1}(v)\right\|_2^2 - \left\|P_{\sigma_2^c}(v)\right\|_2^2 \ge \rho^2/2.$$

By the upperbound on the coordinates in σ,

$$\frac{\rho^2}{2} \le \left\|P_\sigma(v)\right\|_2^2 \le \frac{|\sigma|}{M}.$$

Rearranging this inequality finishes the proof. □

For every $v \in \mathrm{Incomp}(M, \rho)$, we fix a set $\sigma(v)$ of size exactly $\lceil M\rho^2/2\rceil$. Let $\tau'(v)$ be the index set of the M coordinates with largest magnitude. If this set is not uniquely defined, choose one arbitrarily. Let $\tau(v) := \tau'(v) \setminus \sigma(v)$ and $\bar{\sigma} := [n] \setminus (\tau \cup \sigma)$. Now we divide $[n] \setminus \tau$ into disjoint sets $I_1, I_2, \dots, I_{k_0}$ and J, with $|I_k| = \lceil \alpha n\rceil \le M$ for $1 \le k \le k_0$ and $|J| \le \lceil \alpha n\rceil$ where $\alpha = o(1)$ is a parameter to be chosen later. For $1 \le k \le k_0$, we let

$$I_k := \sigma(v)_{[(k-1)\lceil \frac{M\rho^2}{2k_0}\rceil+1:k\lceil \frac{M\rho^2}{2k_0}\rceil]} \cup \bar{\sigma}(v)_{[(k-1)\lceil |\bar{\sigma}|/k_0\rceil+1:k\lceil |\bar{\sigma}|/k_0\rceil]}.$$

Finally, let $I_0 := J \cup \tau$ so $|I_0| \le 2M$ by our assumption on $\lceil \alpha n\rceil$. In words, I_0 is the index set of the large coordinates and the leftover smaller coordinates. Additionally, we have

$$\frac{1}{2\alpha} \le \frac{n - |\tau|}{\lceil \alpha n\rceil} - 1 \le k_0 \le \frac{n - |\tau|}{\lceil \alpha n\rceil} \le \frac{1}{\alpha}.$$

The purpose of this construction is to have substantial control over the ℓ_2 norm and the ℓ_∞ norm of each v_{I_k} for $1 \le k \le k_0$. In particular, we have

$$\|v_{I_k}\|_2 \ge \sqrt{\frac{M\rho^2\alpha}{8}\frac{\rho^2}{2n}} = \frac{\rho^2}{4}\sqrt{\frac{M\alpha}{n}} := \rho'. \tag{2}$$

Furthermore,

$$\|v_{I_k}\|_\infty \le \frac{1}{\sqrt{M}} \quad \text{and} \quad \|v_{I_k}\|_2 \le 2\sqrt{\frac{\alpha n}{M}}.$$

The I_k's are filled by drawing sequentially from σ and $\bar{\sigma}$ so that the entire partition is determined by τ' and σ. Thus, there are at most $\binom{n}{M}\binom{n}{M\rho^2/2}$ partitions for all the vectors in $\mathrm{Incomp}(M, \rho)$.

6.1. *Small-ball probability*

Recall from the proof strategy in Section 3 that we have reduced the problem to bounding the probability that an eigenvector of a random matrix is orthogonal to a random vector. As we will use various epsilon-net approximations, we need a more quantitative version of orthogonality. In particular, we need to bound the probability that the dot product of the eigenvector and the random vector are small. This leads naturally to the notion of small-ball probability.

Definition 6.2. The *Lévy concentration* of a random vector $Z \in \mathbb{R}^n$ is defined to be

$$\mathcal{L}(Z, \varepsilon) = \sup_{u \in \mathbb{R}^n} \mathbb{P}\big(\|Z - u\|_2 \le \varepsilon\big).$$

Intuitively, the structure of a vector, v, will highly affect the Lévy concentration of the random variable $v \cdot X$ where X is a random vector. To formalize this concept, we begin with a measure for the arithmetic structure of an entire unit vector.

Definition 6.3. We define the least common denominator (LCD) of $x \in \mathcal{S}^{n-1}$ as

$$D(x) = \inf\big\{\theta > 0 : \operatorname{dist}\big(\theta x, \mathbb{Z}^n\big) < (\delta_0 p)^{-1/2}\sqrt{\log_+(\sqrt{\delta_0 p}\theta)}\big\},$$

where δ_0 is an appropriate constant (see Remark 6.4 below). This particular form of the LCD was first used in [24].

Remark 6.4. There exists constants $\delta_0, \bar{\varepsilon}_0 \in (0,1)$ such that for any $\varepsilon \le \bar{\varepsilon}_0$, $\mathcal{L}(\delta\xi, \varepsilon) \le 1 - \delta_0 p$ where $\mathbb{P}(\delta = 1) = p$ and ξ is a subgaussian random variable with unit variance. We fix such a δ_0 in Definition 6.3.

The quantitative relationship between the arithmetic structure of a vector and small ball probability is captured in the following proposition.

Proposition 6.5 (Proposition 4.2, [3]). *Let $X \in \mathbb{R}^n$ be a random vector with i.i.d. coordinates of the form $\xi_j\delta_j$, where $\mathbb{P}(\delta_j = 1) = p$ and ξ_j's are random variables with unit variance and finite fourth moment. Then for any $v \in \mathcal{S}^{n-1}$,*

$$\mathcal{L}(X \cdot v, \sqrt{p}\varepsilon) \le C\left(\varepsilon + \frac{1}{\sqrt{p}D(v)}\right),$$

where C depends only on the fourth moment of ξ.

Following [24], we introduce a tool that can reveal the arithmetic structure in small segments of the vector x.

Definition 6.6 (Regularized LCD). Let $\alpha \in (0,1)$. We define the *regularized LCD* of a vector $v \in \operatorname{Incomp}(M, \rho)$ as

$$\widehat{D}(v, \alpha) = \max_{1 \le j \le k_0} D\big(x_{I_j}/\|x_{I_j}\|_2\big).$$

Recall that the α dependence stems from the constraint that $|I_j| = \lceil \alpha n \rceil$.

Combining Proposition 6.5 with the by now standard tensorization argument (see [18]), yields a bound on the Lévy concentration of $M_n x$.

Proposition 6.7 (Small ball probabilities of $M_n x$ via regularized LCD). *There exists a constant C such that for all $\varepsilon \ge 0$, and I is an index set of size $\lceil \alpha n \rceil$,*

$$\mathcal{L}\big(M_n x, \varepsilon\|v_I\|_2\sqrt{pm}\big) \le C^{n-\lceil \alpha n \rceil}\left(\varepsilon + \frac{1}{\sqrt{p}D(v_I/\|v_I\|_2)}\right)^{n-\lceil \alpha n \rceil}.$$

Therefore, by the bounds in (2) and the above proposition, we have

$$\mathcal{L}\big(M_n x, \varepsilon\rho'\sqrt{pn}\big) \le \left(C\varepsilon + \frac{C}{\sqrt{p}\hat{D}(v, \alpha)}\right)^{n-\alpha n}.$$

We also have the following simple lower bound for the LCD.

Proposition 6.8 (Lemma 6.2, [24]). *Let $x \in \mathcal{S}^{n-1}$. Then*

$$D(x) \geq \frac{1}{2\|x\|_\infty}.$$

We can deduce from this proposition and our bounds in (2), that

$$\hat{D}(v,\alpha) \geq \frac{1}{2}\rho'\sqrt{M}. \tag{3}$$

For the remainder of this section, we fix several parameters. For the readers' convenience, we have aggregated and highlighted several important variables below. Although we repeat these definitions, we urge the reader to refer to this section when verifying calculations later.

$$\boxed{M = \frac{n}{(np)^{1/16}}, \qquad \alpha = (np)^{-1/16}, \qquad \rho' = \frac{\rho^2}{4}\sqrt{\frac{M\alpha}{n}}}$$

Remark 6.9. Recall that $\rho := \bar{C}^{-\ell_0-6}$ with $\bar{C}$ from Proposition 5.3. As observed in Remark 5.4, due to the assumption that $p \geq n^{-1+\delta}$, we have that

$$c_\delta \leq \rho \leq c'_\delta$$

for two constants c_δ, c'_δ only depending on δ. We will often implicitly make use of the fact that $np \to \infty$.

6.2. *Vectors with mid-range and small LCD*

In this section, we show that matrices of the form $M_n - \lambda$ are unlikely to have vectors in their nullspace with mid-range or small LCD.

6.2.1. *Mid-range LCD*: $\frac{1}{\tilde{c}}\frac{n^{1/2}}{(pn)^{1/32}} \leq \hat{D} \leq \exp((np)^{1/32})$
One of the main technical contributions of this article is the following proposition.

Proposition 6.10 (Mid-range LCD). *For $\delta > 0$, $p \geq n^{-1+\delta}$ and $\lambda \in [-K\sqrt{pn}, K\sqrt{pn}]$. There exist constants*

$$c, c'', \tilde{c} > 0$$

such that for $M = \frac{n}{(np)^{1/16}}$,

$$\mathbb{P}\big(\exists v \in \hat{S}_D \text{ s.t. } \big\|(M_n - \lambda)v\big\|_2 \leq \tilde{c}\varepsilon_0(pn)^{7/16}\big) \leq \exp\big(-c''n\big),$$

where $\frac{1}{\tilde{c}}\frac{n^{1/2}}{(pn)^{1/32}} \leq D \leq \exp((np)^{1/32})$, $\varepsilon_0 = c\frac{n^{1/2}}{(np)^{1/32}D}$ and

$$\hat{S}_D := \big\{v \in \operatorname{Incomp}(M,\rho) : D \leq \widehat{D}(v) \leq 2D\big\}.$$

Recall that $\rho := \bar{C}^{-\ell_0-6}$, $\ell_0 := \lceil \frac{\log 1/(8p)}{\log\sqrt{pn}} \rceil$ where $\bar{C}$ is the constant from Proposition 5.3.

6.2.2. *Level sets for the usual LCD*
We remind the reader of some key terminology. Working in some metric space, a β-net of a set S is a subset $S' \subset S$ such that for every $s \in S$, there exists a $s' \in S'$ such that $\|s - s'\| \leq \beta$. We first construct level nets of the LCD (not regularized) for vectors of length $\lceil \alpha n \rceil$. We drop the ceiling function when such precision is not crucial. We keep the n dependence in this section as various parameters, e.g. $p(n)$, more conveniently depend on n rather than αn.

Lemma 6.11. *For $p \geq n^{-1+\delta}$,*

$$\beta = \frac{2\sqrt{\log(2\sqrt{\delta_0 p}D_0)}}{D_0\sqrt{\delta_0 p}}$$

and $D_0 > 0$, *the set* $\{v \in \mathcal{S}^{\alpha n-1} : D(v) \in (D_0, 2D_0]\}$ *has a* β*-net*, $\mathcal{N}$, *such that*

$$|\mathcal{N}| \leq \left(2 + \frac{\bar{c}D_0}{\sqrt{\alpha n}}\right)^{\alpha n}$$

for a universal constant $\bar{c}$.

Proof. For a v with $D(v) \in (D_0, 2D_0]$, by the definition of LCD, there exists a $\theta \in (D_0, 2D_0]$ and $z \in Z$ such that

$$\|\theta v - z\|_2 < \frac{\sqrt{\log(\sqrt{\delta_0 p}\theta)}}{\sqrt{\delta_0 p}}$$

which implies that

$$\left\| v - \frac{z}{\theta} \right\|_2 < \frac{\sqrt{\log(2\sqrt{\delta_0 p} D_0)}}{D_0\sqrt{\delta_0 p}}.$$

We also have

$$\left\| \left\| \frac{z}{\|z\|_2} \right\|_2 - \frac{\|z\|_2}{\theta} \right\|_2 = \left\| \|v\|_2 - \frac{\|z\|_2}{\theta} \right\|_2 < \left\| v - \frac{z}{\theta} \right\|_2.$$

Combining the above estimates gives

$$\begin{aligned}\left\| v - \frac{z}{\|z\|_2} \right\|_2 &\leq \left\| v - \frac{z}{\theta} \right\|_2 + \left\| \frac{z}{\theta} - \frac{z}{\|z\|_2} \right\|_2 \\ &\leq \left\| v - \frac{z}{\theta} \right\|_2 + \left\| \left\| \frac{z}{\|z\|_2} \right\|_2 - \frac{\|z\|_2}{\theta} \right\|_2 \\ &\leq \frac{2\sqrt{\log(\sqrt{\delta_0 p}\theta)}}{D_0\sqrt{\delta_0 p}}.\end{aligned}$$

Note that

$$\|z\|_2 \leq \|z - \theta v\|_2 + \|\theta v\|_2 \leq \frac{\sqrt{\log(\sqrt{\delta_0 p}\theta)}}{\sqrt{\delta_0 p}} + 2D_0 \leq 4D_0.$$

The last inequality follows from recalling that $D_0 \geq \rho'\sqrt{M} \geq \frac{c_\delta^2\sqrt{n}}{4(np)^{5/32}}$ so $D_0\sqrt{p} \geq \frac{c_\delta^2}{4}(np)^{11/32} \geq \sqrt{\log(\sqrt{\delta_0 p}\theta)}$. Let

$$Z := \left\{z \in \mathbb{Z}^m : \operatorname{supp}(z) \in I \text{ and } 0 < \|z\|_2 \leq 4D_0\right\}.$$

Define $\mathcal{N} := \{z/\|z\|_2 : z \in Z\}$. By the standard volumetric calculation,

$$|\mathcal{N}| \leq \left(2 + \frac{\bar{c}D_0}{\sqrt{m}}\right)^m$$

for some universal constant $\bar{c}$. $\mathcal{N}$ serves as an appropriate net. □

The above lemma can be modified so that β is a function of D rather than D_0.

Lemma 6.12. *For*

$$\beta = \frac{2\sqrt{\log(2\sqrt{\delta_0 p} D)}}{D\sqrt{\delta_0 p}}$$

and $D_0 > 0$, *the set* $\{v \in \mathcal{S}^{\alpha n-1} : D(v) \in (D_0, 2D_0]\}$ *has a* β*-net*, $\mathcal{N}$, *such that*

$$|\mathcal{N}| \leq \left(12 + \frac{\bar{c}D}{\sqrt{\alpha n}}\right)^{\alpha n}$$

for a universal constant $\bar{c}$.

Proof. By Lemma 6.11, the set is covered by at most $(2+\frac{\bar{c}D}{\sqrt{\alpha n}})^{\alpha n}$ balls of radius $\beta_0 = \frac{2\sqrt{\log(2\sqrt{\delta_0 p}D_0)}}{D_0\sqrt{\delta_0 p}}$. If $\beta \geq \beta_0$ then the result follows immediately. Assume $\beta < \beta_0$. A $\beta/2$ net of size $(4\beta_0/\beta)^{\alpha n} \leq (3D/D_0)^{\alpha n}$. Therefore, the number of small balls is at most

$$\left(2+\frac{\bar{c}D_0}{\sqrt{\alpha n}}\right)^{\alpha n}\left(\frac{3D}{D_0}\right)^{\alpha n} \leq \left(12+\frac{\bar{c}D_0}{\sqrt{\alpha n}}\right)^{\alpha n}.$$

□

Now we extend the net to cover all vectors with LCD less than $2D_0$.

Lemma 6.13. *For $D > f(n) = \omega(1)$, then the set*

$$\left\{v \in \mathcal{S}^{\alpha n-1} : f(n) \leq D(v) \leq D\right\}$$

has a β-net of size at most

$$\left(12+\frac{\bar{c}D_0}{\sqrt{\alpha n}}\right)^{\alpha n}\log(D).$$

Proof. Decompose the set

$$\left\{x \in \mathcal{S}^{m-1} : D(v) \leq D\right\} = \bigcup_k \left\{v \in \mathcal{S}^{m-1} : D(x) \in \left(2^{-k}D, 2^{-k+1}D\right]\right\},$$

where the union is over all k such that $(2^{-k}D, 2^{-k+1}D]$ has non-zero intersection with $[f(n), D]$. Each of these intervals has a β-net by Lemma 6.11. There are at most $\log D$ such k. □

6.2.3. *Nets for the level sets of the regularized LCD*
Let

$$S_{\hat{D}} := \left\{v \in \operatorname{Incomp}(M,\rho) : D < \widehat{D}(v) \leq 2D\right\}$$

and set

$$M = \frac{n}{(np)^{1/16}}, \qquad \alpha = (np)^{-1/16} \quad \text{and} \quad \varepsilon_0 = \frac{c\sqrt{\alpha n}}{D},$$

where c is the constant from Proposition 6.10. We record several useful bounds which are consequences of our choice of parameters,

$$\frac{c_\delta^2}{4}(np)^{-1/16} \leq \rho' \leq \frac{{c'_\delta}^2}{4}(np)^{-1/16}.$$

One can check that $p^{-1} \leq M$ since $np \to \infty$ so that M is in the range of Corollary 5.5.

Let c^* be a constant less than $1/2C$ with C the constant from Proposition 6.7. We first create an $c^*\rho'\varepsilon_0/10K$-net for the coordinates in I_0. For this set, we use a trivial-net $\mathcal{N}_0$ of size at most

$$\left(\frac{10K}{c^*\varepsilon_0\rho'}\right)^{2M}.$$

Recall that $\|v_{I_k}\|_\infty \leq \frac{1}{\sqrt{M}}$, so by Porposition 6.8, $\hat{D}(v) \geq \rho'\sqrt{M}$ for any $v \in \operatorname{Incomp}(M,\rho)$. For each I_k with $1 \leq k \leq k_0$, by Lemma 6.13 and the fact that $\operatorname{LCD}(v_{I_k}/\|v_{I_k}\|_2) \leq 2D$, we can create a $\beta = \frac{2\sqrt{\log(2\sqrt{\delta_0 p}D)}}{D\sqrt{\delta_0 p}}$-net of size at most

$$\left(12+\frac{\bar{c}D_0}{\sqrt{\alpha n}}\right)^{\alpha n}\log(D).$$

Let $\bar{\mathcal{N}}_I$ be an $\frac{c^*\rho'\varepsilon_0}{10Kk_0}$-net of $[\rho', 1]$ of size at most $\frac{30Kk_0}{c^*\varepsilon_0\rho'}$.

Define

$$\mathcal{M} := \left\{ x + \sum_k t_k y_k : x \in \mathcal{N}_0, y \in \mathcal{N}_k, t \in \bar{\mathcal{N}_k} \right\}.$$

We note that

$$|\mathcal{M}| \le \left(\frac{10K}{c^*\varepsilon_0\rho'}\right)^{2M} \prod_k \left(12 + \frac{\bar{c}D_0}{\sqrt{\lambda n}}\right)^{\lambda n} \log(D)\left(\frac{30Kk_0}{c^*\varepsilon_0\rho'}\right).$$

For any $v \in S_D$, there exists a $m = x + \sum_k t_k y_k \in \mathcal{M}$ such that

$$\|x - v_{I_0}\|_2 \le \frac{c^*\rho'\varepsilon_0}{10K}, \qquad \left\| y_k - \frac{v_{I_k}}{\|v_{I_k}\|_2} \right\|_2 \le \beta, \quad \text{and} \quad \left| t_k - \|v_{I_k}\|_2 \right| \le \frac{c^*\rho'\varepsilon_0}{10Kk_0}.$$

Therefore,

$$\begin{aligned}
\|v - m\|_2 &\le \frac{c^*\rho'\varepsilon_0}{10K} + \sum_k \left(\left\| v_{I_k} - \|v_{I_k}\|_2 y_k \right\|_2 + \left\| \|v_{I_k}\|_2 y_k - t_k y_k \right\|_2 \right) \\
&\le \frac{c^*\rho'\varepsilon_0}{10K} + \sum_k \left(\left\| \frac{v_{I_k}}{\|v_{I_k}\|_2} - y_k \right\|_2 \|v_{I_k}\|_2 + \left\| \|v_{I_k}\|_2 y_k - t_k y_k \right\|_2 \right) \\
&\le \frac{c^*\rho'\varepsilon_0}{10K} + k_0\left(2\beta + \frac{c\rho'\varepsilon_0}{10Kk_0}\right) \\
&\le \frac{c^*\rho'\varepsilon_0}{5K} + \frac{1}{\alpha} \frac{2\sqrt{\log(2\sqrt{\delta_0 p}D)}}{D\sqrt{\delta_0 p}} \\
&= \frac{c^*\rho'\varepsilon_0}{5K} + \frac{4}{\sqrt{M}\rho^2} \frac{1}{\alpha} \frac{2\sqrt{\log(2\sqrt{\delta_0 p}D)}}{c\sqrt{\delta_0 p}} \rho'\varepsilon_0.
\end{aligned}$$

Using the upper bound on $D \le \exp((np)^{1/32}) \le \exp(\frac{c^2 M\rho^4\alpha^2 p}{K^2})$ where c is the constant from Proposition 6.10, we deduce that

$$\|v - m\|_2 \le \frac{c^*\rho'\varepsilon_0}{5K} + o(\rho'\varepsilon_0) \le \frac{\rho'\varepsilon_0}{4K}.$$

At this point, there is no guarantee that the elements of $\mathcal{M}$ lie in $\hat{S}_D$. We rectify this issue by slightly adjusting $\mathcal{M}$. For every $m \in \mathcal{M}$, if there exists a $v \in \hat{S}_D$ such that $\|v - m\|_2 \le \frac{c^*\rho'\varepsilon_0}{4K}$ then replace m by v. Otherwise, simply discard m. We call this new set $\mathcal{M}'$ and note that $|\mathcal{M}'| \le |\mathcal{M}|$. By the triangle inequality, $\mathcal{M}'$ is a $\frac{c^*\rho'\varepsilon_0}{2K}$-net of $\hat{S}_D$.

6.2.4. *Proof of Proposition* 6.10

Proof. Fix a $\lambda \in [-K, K]$. In the last section, we showed that for all the vectors with the same σ, τ, $\mathcal{M}'$ is an $c\rho'\varepsilon_0/2K$-net of $\hat{S}_D$. Let $\mathcal{E}_{\mathcal{M}'}$ be the event that there exists a $m \in \mathcal{M}'$ such that $\|(M_n - \lambda)m\|_2 > c^*\rho'\varepsilon_0\sqrt{pn}$. As we fixed $c < 1/2C$ with C from Proposition 6.7, one can verify that

$$\varepsilon_0 \ge \frac{1}{\sqrt{p}D},$$

by Lemma 6.7,

$$\mathbb{P}(\mathcal{E}_{\mathcal{M}}) \le |\mathcal{M}|\varepsilon_0^{n - \lceil \alpha n \rceil}.$$

By our lower bound on D, we have that $\bar{c}D/\sqrt{\alpha n} \ge 1$.

$$\mathbb{P}(\mathcal{E}_{\mathcal{M}'}) \le \binom{n}{2M}\binom{n}{M\rho^2}\left(\frac{10K}{c^*\varepsilon_0\rho'}\right)^{2M}\left(\frac{2\bar{c}D}{\sqrt{\alpha n}}\right)^n \log^{\alpha^{-1}}(2D)\left(\frac{30Kk_0}{c^*\varepsilon_0\rho'}\right)^{\alpha^{-1}} \varepsilon_0^{n - \lceil \alpha n \rceil}$$

$$
\begin{aligned}
&\leq \binom{n}{2M}\binom{n}{M\rho^2}\left(\frac{10K}{c^*\rho'}\right)^{2M}\left(\frac{2\bar{c}D}{\sqrt{\alpha n}}\right)^{n}\log^{\alpha^{-1}}(2D)\left(\frac{30Kk_0}{c^*\rho'}\right)^{\alpha^{-1}}\left(\frac{c\sqrt{\alpha n}}{D}\right)^{n-\lceil\alpha n\rceil-2M-\alpha^{-1}}\\
&\leq \exp\biggl(-n\biggl(-\frac{2M}{n}\log n-\frac{c'^2_\delta M}{n}\log n-\frac{2M}{3n}\log(pn)+\frac{n-\alpha n-2M-\alpha^{-1}}{n}\log(1/c)\\
&\quad-\log(2\bar{c})-\frac{\alpha^{-1}}{10n}\log(pn)-\frac{\alpha^{-1}}{n}\log\left(\frac{120Kk_0}{c^*c_\delta^2}(pn)^{1/32}\right)-\frac{\alpha n+2M+\alpha^{-1}}{n}\log(D/\sqrt{\alpha n})\biggr)\biggr)\\
&\leq \exp\left(-n\left(-\log(2\bar{c})+\frac{1}{2}\log\left(\frac{1}{c}\right)-\frac{\alpha n+2M+\alpha^{-1}}{n}(np)^{1/32}+o(1)\right)\right)\\
&\leq \exp\left(-n\left(-\log(2\bar{c})+\frac{1}{2}\log\left(\frac{1}{c}\right)+o(1)\right)\right)\\
&\leq \exp(-c''n)
\end{aligned}
$$

for small enough c from Proposition 6.10. On the event $\overline{\mathcal{E}_{\mathcal{M}'}}$, for any $v\in\hat{S}_D$, we can find a $m\in\mathcal{M}'$ such that $\|v-m\|_2\leq\rho'\varepsilon_0/2K$. Therefore,

$$
\left\|(M_n-\lambda)v\right\|_2\geq\left\|(M_n-\lambda)m\right\|_2-\|M_n-\lambda\|\|v-m\|_2\geq\frac{c\rho'\varepsilon_0\sqrt{pn}}{2}.
$$

The proof is complete upon setting $\tilde{c}=c^*c_\delta^2c$. □

Remark 6.14. Note that a trivial ε_0 net of the unit sphere is of size $(3D/c\sqrt{\alpha n})^n$ which is of the same order as our more involved construction. However, the key gain of our design is that it is $\bar{c}$ that appears in the dominant term of our net size and c from Proposition 6.10 can be defined independently.

6.2.5. *Small LCD*
For this range of regularized LCD, a nearly identical argument as Proposition 6.10 applies. As the choice of parameters is different, we show the necessary computations below.

Proposition 6.15 (Small LCD). *For $\delta>0$, $p\geq n^{-1+\delta}$ and $\lambda\in[-K\sqrt{pn},K\sqrt{pn}]$. There exist constants*

$$
c'',\tilde{c}>0
$$

such that for $M=\frac{n}{(np)^{1/8}}$

$$
\mathbb{P}\left(\exists v\in\hat{S}_D \text{ s.t. } \left\|(M_n-\lambda)v\right\|_2\leq\tilde{c}\varepsilon_0(pn)^{7/16}\right)\leq\exp(-c''n),
$$

where $\rho'\sqrt{M}\leq D\leq\frac{1}{\bar{c}}\sqrt{\alpha n}$, $\varepsilon_0'=(\rho'\sqrt{pM})^{-1/2}$ and

$$
\hat{S}_D:=\left\{v\in\mathrm{Incomp}(M,\rho):D\leq\widehat{D}(v)\leq 2D\right\}.
$$

Recall that $\rho:=\bar{C}^{-\ell_0-6}$, $\ell_0:=\frac{\log 1/(8p)}{\log\sqrt{pn}}$.

Using this new ε_0', we have

$$
\begin{aligned}
\|v-m\|_2&\leq\frac{c\rho'\varepsilon_0}{10K}+\sum_k\left(\left\|v_{I_k}-\|v_{I_k}\|_2y_k\right\|_2+\left\|\|v_{I_k}\|_2y_k-t_ky_k\right\|_2\right)\\
&\leq\frac{c\rho'\varepsilon_0'}{10K}+\sum_k\left(\left\|\frac{v_{I_k}}{\|v_{I_k}\|_2}-y_k\right\|_2\|v_{I_k}\|_2+\left\|\|v_{I_k}\|_2y_k-t_ky_k\right\|_2\right)\\
&\leq\frac{c\rho'\varepsilon_0'}{10K}+k_0\left(\beta+\frac{c\rho'\varepsilon_0'}{10Kk_0}\right)
\end{aligned}
$$

$$\le \frac{c\rho'\varepsilon_0'}{5K} + k_0 \frac{2\sqrt{\log(2\sqrt{\delta_0 p}\rho'\sqrt{M})}}{\rho'\sqrt{M}\sqrt{\delta_0 p}}$$
$$\le \frac{c\rho'\varepsilon_0'}{5K} + k_0 \frac{2\sqrt{\log(2\sqrt{\delta_0 p}\rho'\sqrt{M})}}{(\rho')\rho'^{1/2}(Mp)^{1/4}\sqrt{\delta_0}}\rho'\varepsilon_0'$$
$$\le \frac{c\rho'\varepsilon_0'}{10K}.$$

The third to last inequality follows from the observation that the function $x \to \frac{\sqrt{\log(c_1 x)}}{c_2 x}$ is a decreasing function for large values of x, $\sqrt{\delta_0 p}\rho'\sqrt{M} \ge (np)^{3/8}$ and $np \to \infty$. The last inequality follows from the simple calculation $\rho'\rho'^{1/2}(Mp)^{1/4} \to \infty$.

Again, it is easy to check that $\varepsilon_0 \ge \frac{1}{\sqrt{pD}}$, so by Lemma 6.7,

$$\mathbb{P}(\mathcal{E}_{\mathcal{M}}) \le |\mathcal{M}|\varepsilon_0^{n-\lceil \alpha n\rceil}.$$

As, $\bar{c}D/\sqrt{\alpha n} < 1$,

$$\mathbb{P}(\mathcal{E}_{\mathcal{M}}) \le \binom{n}{2M}\binom{n}{M\rho^2}\left(\frac{10K}{c\varepsilon_0\rho'}\right)^{2M}(13)^n \log^{\alpha^{-1}}(2D)\left(\frac{30Kk_0}{c\varepsilon_0\rho'}\right)^{\alpha^{-1}}\varepsilon_0^{n-\lceil \alpha n\rceil}$$
$$\le \binom{n}{2M}\binom{n}{M\rho^2}\left(\frac{10K}{c\rho'}\right)^{2M}(13)^n \log^{\alpha^{-1}}(2D)\left(\frac{30Kk_0}{c\rho'}\right)^{\alpha^{-1}}\left(\frac{1}{\sqrt{\rho'}(pM)^{1/4}}\right)^{n-\lceil \alpha n\rceil-2M-\alpha^{-1}}$$
$$\le \exp(-n).$$

We extend this to all the vectors in $\hat{S}_D$ by the same approximation argument.

7. Proof of Theorem 2.1

Proof. By Corollary 5.5, with probability $1-\exp(-cn)$ with c from Corollary 5.5, the eigenvectors of M_{n-1} are not compressible. We now show that the eigenvectors of M_{n-1} do not have mid-range or small regularized LCD. We begin with the mid-range vectors. Let $\frac{1}{\bar{c}}\frac{n^{1/2}}{(pn)^{1/32}} \le D \le \exp((np)^{1/32})$. We demonstrate that an eigenvector is unlikely to be in $\hat{S}_D$. Let $\mathcal{P}$ be a $\tilde{c}\varepsilon_0(pn)^{7/16}$-net of $[-K\sqrt{pn}, K\sqrt{pn}]$ with $\tilde{c}$ from Proposition 6.10 and

$$|\mathcal{P}| \le \frac{2K\sqrt{pn}}{\tilde{c}\varepsilon_0(pn)^{7/16}} \le \exp\bigl((np)^{1/16}\bigr).$$

If $v \in \hat{S}_D$ and is an eigenvector with eigenvalue λ, then there is a $\lambda_0 \in \mathcal{P}$ with $|\lambda-\lambda_0| \le \tilde{c}\varepsilon_0(pn)^{7/16}$. Therefore,

$$\bigl\|(M_n-\lambda_0)v\bigr\|_2 \le |\lambda-\lambda_0| \le \tilde{c}\varepsilon_0(pn)^{7/16}.$$

Thus, by Proposition 6.10 and a union bound, with probability greater than $1-\exp(-c''n/2)$, an eigenvector of M_{n-1} will not lie in $\hat{S}_D$. Consider the following decomposition.

$$\left\{v \in \mathrm{Incomp}(M,\rho) : \frac{1}{\bar{c}}\frac{n^{1/2}}{(pn)^{1/32}} \le \hat{D}(v) \le \exp\bigl((np)^{1/32}\bigr)\right\}$$
$$= \bigcup_k \bigl\{v \in \mathrm{Incomp}(M,\rho) : D(v) \in \bigl(2^{-k}\exp\bigl((np)^{1/32}\bigr), 2^{-k+1}\exp\bigl((np)^{1/32}\bigr)\bigr]\bigr\},$$

where k takes values such that $(2^{-k}\exp((np)^{1/32}), 2^{-k+1}\exp((np)^{1/32})]$ has a non-zero intersection with $[\frac{1}{\bar{c}}\frac{n^{1/2}}{(pn)^{1/32}}, \exp((np)^{1/32})]$. There are at most $(np)^{1/16}$ such k, so by a simple union bound, we can guarantee that the event, $\mathcal{E}_{\mathrm{mid}}$, that M_{n-1} does not have eigenvectors with regularized LCD in $[\frac{1}{\bar{c}}\frac{n^{1/2}}{(pn)^{1/32}}, \exp((np)^{1/32})]$ occurs with probability at least

$1 - \exp(-c''/3)$. By an identical argument, replacing Proposition 6.10 with Proposition 6.15, we have that the event, $\mathcal{E}_{\mathrm{small}}$, that the eigenvectors of M_{n-1} have regularized LCD that lie outside of the interval $[\rho'\sqrt{M}, \frac{1}{\tilde{c}}\frac{n^{1/2}}{(pn)^{1/32}}]$ occurs with probability at least $1 - \exp(-cn/3)$ with c from Proposition 6.15. On the event $\mathcal{E}_{\mathrm{mid}} \cap \mathcal{E}_{\mathrm{small}}$, by Proposition 6.5,

$$\mathbb{P}(\mathcal{E}_i | \mathcal{E}_{\mathrm{mid}} \cap \mathcal{E}_{\mathrm{small}}) \leq \mathbb{P}(X \cdot v = 0 | \mathcal{E}_{\mathrm{mid}} \cap \mathcal{E}_{\mathrm{small}}) \leq \frac{C}{\sqrt{p}\exp((np)^{1/32})} \leq \exp\bigl(-(np)^{1/64}\bigr),$$

where X is from the decomposition of M_n in (1). Taking a union bound over all i, we conclude that M_n has simple spectrum with probability at least $1 - \exp(-(np)^{1/64}/2)$. □

8. Erdős–Rényi random graphs

Let G_n be a random variable that takes values in the simple graphs on n vertices with vertex set $[n]$. G_n is distributed such that an edge appears between two vertices independently with probability p. Let A_n denote the adjacency matrix of G_n. Note that the entries of A_n have mean p. Thus, Theorem 2.1 does not immediately apply. However, the expected adjancecy matrix is $p(J - I)$ where J is the $n \times n$ all ones matrix. However, J is a rank one matrix and we can exploit this fact to adjust our proof to handle this case. As the proof is only slightly modified, we do not repeat the argument and only highlight the necessary changes. These adjustments follow those in [3, Section 7].

In preparation for the proof of Theorem 2.2, we need analogues of Proposition 5.3, Proposition 6.10 and Proposition 6.15. The necessary changes for Proposition 5.3 are discussed in Appendix B. For Propositions 6.10 and Proposition 6.15, the first step was to obtain estimates on the Lévy concentration. As this function is insensitive to shifts in the mean, the first part of the proof holds without change. For the net arguments to hold, we simply make the observation that $A_n - p(J_n - I_n)$ is a mean zero random matrix so the standard arguments (e.g. those in Proposition 5.2) yield

$$\mathbb{P}\bigl(\|A_n - p(J_n - I_n)\|_2 \geq K'\sqrt{pn}\bigr) \leq \exp\bigl(c'pn\bigr).$$

Therefore we have the following two propositions.

Proposition 8.1 (Mid-range LCD). *For $\delta > 0$, $n^{-1+\delta} \leq p \leq 1/2$ and $\lambda \in [-K\sqrt{pn}, K\sqrt{pn}]$. There exist constants*

$$c, c'', \tilde{c} > 0$$

such that for $M = \frac{n}{(np)^{1/16}}$,

$$\mathbb{P}\bigl(\exists v \in \hat{S}_D \text{ s.t. } \|(A_n - p(J_n - I_n) - \lambda)v\|_2 \leq \tilde{c}\varepsilon_0(pn)^{7/16}\bigr) \leq \exp\bigl(-c''n\bigr),$$

where $\frac{1}{\tilde{c}}\frac{n^{1/2}}{(np)^{1/32}} \leq D \leq \exp((pn)^{1/32})$, $\varepsilon_0 = cn^{1/2-\delta/12}/D$ and

$$\hat{S}_D := \bigl\{v \in \mathrm{Incomp}(M, \rho) : D \leq \widehat{D}(v) \leq 2D\bigr\}.$$

Proposition 8.2 (Small LCD). *For $\delta > 0$, $n^{-1+\delta} \leq p \leq 1/2$ and $\lambda \in [-K\sqrt{pn}, K\sqrt{pn}]$. There exist constants*

$$c, c'', \tilde{c} > 0$$

such that for $M = \frac{n}{(np)^{1/8}}$

$$\mathbb{P}\left(\exists v \in \hat{S}_D \text{ s.t. } \|(M_n - p(J_n - I_n) - \lambda)v\|_2 \leq \frac{\tilde{c}}{2}\varepsilon_0(pn)^{7/16}\right) \leq \exp\bigl(-c''n\bigr),$$

where $\rho'\sqrt{M} \leq D \leq \frac{1}{\tilde{c}}\sqrt{\alpha n}$, $\varepsilon_0' = (\rho'\sqrt{pM})^{-1/2}$ and

$$\hat{S}_D := \bigl\{v \in \mathrm{Incomp}(M, \rho) : D \leq \widehat{D}(v) \leq 2D\bigr\}.$$

8.1. *Proof of Theorem 2.2*

Proof. (Sketch) We first handle the case where $p \le 1/2$. Observe that the set $\{J_n x : x \in \mathcal{S}^{n-1}\} = \{\theta \cdot \mathbf{1} : \theta \in [-n, n]\}$ where $\mathbf{1}_n$ is the vector of all ones. Let $\mathcal{X}_n = \{\kappa \cdot \mathbf{1} : \kappa \in [-pn, pn]\}$. As this is a one-dimensional set, we can create a net with small cardinality. Let $\mathcal{B}$ be a $\tilde{c}\varepsilon_0(pn)^{7/16}$-net of $\mathcal{X}_n$ with $\tilde{c}$ from Proposition 8.1 and

$$|\mathcal{B}| \le \frac{2pn}{\tilde{c}\varepsilon_0(pn)^{7/16}} \le \exp\bigl((np)^{1/16}\bigr).$$

By the triangle inequality, for $x, x' \in \mathcal{X}_n$,

$$\bigl|\inf\bigl\|\bigl(A_n - p(J_n - I_n) - \lambda\bigr)v - x\bigr\|_2 - \bigl\|\bigl(A_n - p(J_n - I_n) - \lambda\bigr)v - x'\bigr\|_2\bigr| \le \bigl|x - x'\bigr|.$$

Therefore, the standard union bound and triangle inequality argument shows that for D in the mid-range LCD,

$$\mathbb{P}\Bigl(\inf_{x\in\mathcal{X}_n}\inf_{v\in S_D}\bigl\|\bigl(A_n - p(J_n - I_n) - \lambda\bigr)v - y\bigr\|_2 \le \tilde{c}\varepsilon_0(pn)^{7/16}\Bigr) \le \exp(-cn).$$

The same applies for the low-range. Finally observing that,

$$\inf_{x\in\mathcal{X}_n}\inf_{v\in S_D}\bigl\|\bigl(A_n - p(J_n - I_n) - \lambda\bigr)v - y\bigr\|_2 \le \inf_{v\in S_D}\bigl\|\bigl(A_n - (\lambda - p)\bigr)v\bigr\|_2,$$

summing over the level sets as before and using the same net argument on λ, we can conclude that any eigenvector of A_n has large LCD. The rest of the argument proceeds as in the proof of Theorem 2.1.

For the remaining $p > 1/2$ case, we observe that the adjacency matrix of $G(n, p)$, $A_n(p)$, has the same distribution as $J_n - I_n - A_n(1-p)$. Therefore, to control $\|(A_n(p) - p(J_n - I_n)\lambda)v\|_2$ it suffices to manage $\|(A_n(1-p) - (1-p)(J_n - I_n))v\|_2$, for which our previous argument applies. □

9. Concluding remarks

As mentioned before, we believe the threshold for a random matrix to have simple spectrum should be $p \sim \log n/n$ rather than $p \sim n^{-1+\delta}$. The calculations near the threshold are more involved and will appear elsewhere. Additionally, our arguments naturally offer a quantitative bound on the size of the gaps between eigenvalues and the smallest absolute value of an eigenvalue (which is needed to bound the condition number of the matrix). We have made no attempt to optimize these bounds so we pursue this line of work in a separate article.

The proof of our result for adjacency matrices applies almost without change to matrices of the form $R_n + M_n$ where R_n is a deterministic low-rank matrix. However, to generalize this result to arbitrary non-zero mean matrices requires several new tools which we are currently developing. The ε-net arguments that lie at the core of our current work fail in this setting as we no longer have the necessary control on the operator norm of the matrix and the image of the matrix may not be a perturbation of a low-dimensional space as for the adjacency matrix. To address these new concerns, it will be necessary to use sparse versions of the Inverse Littlewood–Offord theorems of the second author and Nguyen.

Appendix A: Proof of Proposition 5.3

A.1. *Matrix lemma*

The following observation was first utilized in [3]. If we fix the submatrix corresponding to the support of a sparse vector, it is likely that many of these rows will contain exactly one non-zero entry. In this case, in the product of the matrix with the sparse vector, there is no cancellation in these coordinates. As we are dealing with deviations of a matrix from a fixed vector u, we simply modify the lemma to show that there are many rows with exactly one non-zero coordinate with a convenient sign.

Lemma A.1. *Let M_n be a $n \times n$ matrix with independent entries $m_{ij} = \delta_{ij}\xi_{ij}$ where δ_{ij} are Bernoulli random variables with $\mathbb{P}(\delta_{ij} = 1) = p$, where $p \ge C\log n/n$ and ξ_{ij} are iid random variables with* $\max\{\mathbb{P}(\xi_{ij} \ge 1), \mathbb{P}(\xi_{ij} \le -1)\} \ge c_0$. *For $\kappa \in \mathbb{N}$, we define $\mathcal{E}_c^{JJ'}$ to be the event that for any vector of signs $\{\varepsilon_j\}_{j=1}^n$ there are at least $c\kappa pn$ rows of M_n for which* *there is exactly one non-zero entry m_{ij} with $m_{ij}\varepsilon_j \ge 1$ and $i \ne j$ in the columns corresponding to J, and all zero entries in the columns corresponding to J'. Let*

$$\mathfrak{m} = \kappa := \kappa\sqrt{pn} \wedge \frac{1}{8p}.$$

Then, there exists constants $0 < c_{A.1}, c'_{A.1}$, depending only on c_0 such that

$$\mathbb{P}\Bigg(\bigcap_{\kappa \le (8p\sqrt{pn})^{-1}\vee 1}\ \bigcap_{J\in\binom{[n]}{\kappa}}\ \bigcap_{J'\in\binom{[n]}{\mathfrak{m}}, J\cap J'=\varnothing} \mathcal{E}^{J,J'}_{c'_{A.1}}\Bigg) \ge 1-\exp(-cpn).$$

Proof. The same proof as in [3, Proof of Lemma 3.2] yields the result when applied to the upper $\lfloor n/2\rfloor \times \lfloor n/2\rfloor$ submatrix of M_n. The entries in this submatrix are independent. □

A.2. Very sparse vectors

Definition A.2. For any $x \in S^{n-1}$, let $\pi_x : [n] \to [n]$ be a permutation which arranges the absolute values of the coordinates of x in an non-increasing order. For $1 \le m \le m' \le n$, denote by $x_{[m:m']} \in \mathbb{R}^n$ the vector with coordinates

$$x_{[m:m']}(j) = x(j)\cdot \mathbf{1}_{[m:m']}\big(\pi_x(j)\big).$$

In other words, we include in $x_{[m:m']}$ the coordinates of x which take places from m to m' in the non-increasing rearrangement.

For $\alpha < 1$ and $m \le n$ define the set of vectors with dominated tail as follows:

$$\mathrm{Dom}(m,\alpha) := \big\{x \in S^{n-1} \mid \|x_{[m+1:n]}\|_2 \le \alpha\sqrt{m}\|x_{[m+1:n]}\|_\infty\big\}.$$

Lemma A.3. *Denote*

$$\ell_0 := \left\lceil \frac{\log(1/8p)}{\log\sqrt{pn}} \right\rceil,$$

$$\begin{aligned}&\mathbb{P}\big(\exists x \in \mathrm{Dom}\big(1/8p, (CK)^{-1}\big)\\ &\qquad \text{such that } \big\|(M_n-\lambda)x\big\|_2 \le \big(C'K\big)^{-\ell_0}\sqrt{pn}\\ &\qquad \text{and } \|M_n - \lambda I\| \le K\sqrt{pn}\big)\\ &\quad \le \exp(-cpn).\end{aligned}$$

Proof. We begin by diving $[n]$ into two roughly equal sets. Let $n_0 = \lceil n/2\rceil$. We denote this decomposition by

$$M_n - \lambda I = \begin{pmatrix} A & B\\ B^T & C\end{pmatrix}, \qquad x = \begin{pmatrix} y\\ z\end{pmatrix}.$$

Thus, we have the following equivalence.

$$\big\|(M_n-\lambda)x\big\|_2^2 = \|Ay+Bz\|_2^2 + \big\|B^Ty + Cz\big\|_2^2.$$

We condition on a realization of A and C.

$$\mathbb{P}\big(\exists x \in \mathrm{Sparse}(m)\cap \mathcal{S}^{n-1} \text{ such that } \big\|(M_n-\lambda)x\big\|_2 \le \sqrt{cnp}\big).$$

Let us begin with the assumption that $p \ge (1/4)n^{-1/3}$. In this regime, $\ell_0 = 1$. For $k \in [n]$ let $J_k = \{k\}$ and $J'_k = \mathrm{supp}(x)\setminus J_k$. Define the following vectors of signs. $\{\varepsilon_j\}_{j=1}^n = \{\mathrm{sgn}(z_j)\cdot \mathrm{sgn}((Ay)_j)\}_{j=1}^n$.

$$\big\|(M_n-\lambda)x\big\|_2^2 \ge \sum_{k\in\mathrm{supp}(x)\cap[1,n_0]}\ \sum_{i\in I_k}\big((Bz)_i + (Ay)_i)\big)^2 + \sum_{k\in\mathrm{supp}(x)\cap[n_0+1,n]}\ \sum_{i\in I_k}\big(\big(B^Ty\big)_i + (Cz)_i\big)^2$$

$$\geq \sum_{k\in \text{supp}(x)\cap[1,n_0]} \sum_{i\in I_k} (Bz)_i^2 + \sum_{k\in \text{supp}(x)\cap[n_0+1,n]} \sum_{i\in I_k} \big(B^T y\big)_i^2$$

$$\geq \sum_{k\in \text{supp}(x)\cap[1,n_0]} c_{A.1} pnz_i^2 + \sum_{k\in \text{supp}(x)\cap[n_0+1,n]} cpny_i^2 = cpn,$$

where in the final inequality we have invoked Lemma A.1 with the necessary signs. Now we extend this estimate to $\text{Dom}(1/8p, (CK)^{-1})$. Let $m = (8p)^{-1}$. Assume that

$$\big\|(M_n - \lambda)x\big\| < \frac{1}{2}\sqrt{cpn}.$$

Since $x \in \mathcal{S}^{n-1}$, we have $\|x_{[m+1:n]}\|_\infty \leq m^{-1/2}$. Therefore,

$$\|x_{[m+1:n]}\|_2 \leq (CK)^{-1}\sqrt{m}\|x_{[m+1:n]}\|_\infty \leq (CK)^{-1}.$$

Therefore we have

$$\big\|(M_n - \lambda)x_{[1:m]}\big\|_2 \leq \big\|(M_n - \lambda)x\big\|_2 + (K\sqrt{pn})(CK)^{-1} < \frac{3}{4}\sqrt{cpn}$$

for $C \geq \frac{4}{\sqrt{c}}$.

Furthermore,

$$\Big|\big\|(M_n - \lambda)(x_{[1:m]}/\|x_{[1:m]})\big\|_2 - \big\|(M_n - \lambda)x_{[1:m]}\big\|_2\Big| \leq K\big|1 - \|x_{[1:m]}\|_2\big| \leq \frac{\sqrt{cpn}}{2}.$$

Now we address the remaining $\frac{C\log n}{n} \leq p \leq (1/4)n^{-1/3}$. Note that

$$\frac{1}{8p\sqrt{pn}} > 1.$$

Let $x \in \text{Dom}(1/8p, (CK)^{-1})$. We rearrange the coordinates of x by decreasing magnitude and group them into blocks of size $(pn)^{\ell/2}$ with $\ell = 1, \dots, \ell_0$. From here on, for simplicity, we assume that $(pn)^{\ell_0/2} = 1/8p$. In other words, set

$$z_\ell = x_{[(pn)^{(\ell-1)/2}+1:(pn)^{\ell/2}]},$$

and

$$z_{\ell_0+1} = x_{[(pn)^{\ell_0/2}+1:n]}.$$

For ease of notation, let $m = (pn)^{\ell_0/2}$. We now find a block of substantial ℓ_2 norm. Observe that

$$\|z_{\ell_0+1}\|_2 \leq (CK)^{-1}\sqrt{m}\|z_{\ell_0+1}\|_\infty \leq \sqrt{2}(CK)^{-1}\|z_{\ell_0}\|_2. \tag{4}$$

Since $x \in \mathcal{S}^{n-1}$ implies $\sum_{\ell+1}^{\ell_0+1} \|z_\ell\|_2^2 = 1$, we have

$$\sum_{\ell=1}^{\ell_0} \|z_\ell\|_2^2 \geq 1 - 2(CK)^{-2}.$$

On the other hand, for $K \geq 1$, if $C > 2$, then $3\sum_{\ell=1}^{\infty}(CK)^{-\ell} < 1$. Therefore,

$$\sum_{\ell=1}^{\ell_0}(CK)^{-2\ell} < \sum_{\ell=1}^{\ell_0} \|z_\ell\|_2^2,$$

from which one can deduce that there exists $\ell \leq \ell_0$ such that $\|z_\ell\|_2 \geq (CK)^{-\ell}$. Let ℓ_* be the largest index with this property and define $u = \sum_{\ell=1}^{\ell_*} z_\ell$, $v = \sum_{\ell=\ell_*+1}^{\ell_0+1} z_\ell$. We begin with the case $\ell_* < \ell_0$. By the triangle inequality and (4), we have

$$\|v\|_2 \le \sum_{\ell=\ell_*+1}^{\ell_0+1} \|z_\ell\|_2 \le 2\sqrt{2}(CK)^{-\ell_*+1}.$$

Let $\kappa = (pn)^{(\ell_*-1)/2}$. Note that

$$\kappa \le (np)^{(\ell_0-1)/2} \le \frac{1}{8p\sqrt{pn}}.$$

We apply Lemma A.1 with this choice fo κ. Divide the support of u into $\sqrt{pn}$ blocks of size κ. Define $L_{\ell_*} := \pi_x^{-1}([1,(np)^{\ell_*/2}])$, where π_x is the permutation arranging the coordinates of x in decreasing order with respect to magnitude. For $s \in [(pn)^{1/2}]$, define $J_s := \pi_x^{-1}([(s-1)\kappa+1, s\kappa])$, and set $J'_s = L_{\ell_*} \setminus J_s$. Since $|J'_s| \le |L_{\ell_*}| = \kappa\sqrt{pn}$, we apply Lemma A.1 to get a set $\mathcal{A}$ with large probability, such that on $\mathcal{A}$, there exists subset of rows I_s with $|I_s| \ge c\kappa pn$ for all $s \in [\sqrt{pn}]$, such that for every $i \in I_s$, we have $|a_{i,j_0}| \ge 1$ for only one index $j_0 \in J_s$ and $a_{i,j} = 0$ for all $j \in J_s \cup J'_s \setminus \{j_0\}$. It can further be checked that $I_1, I_2, \dots, I_{\sqrt{pn}}$ are disjoint subsets. Therefore, on set $\mathcal{A}$ for any $i \in I_s$,

$$\left|\big((M_n-\lambda)u\big)_i\right| = \left|(M_n)_{i,j_0}u(j_0)\right| = \left|(M_n)_{i,j_0}\right|\cdot\left|u(j_0)\right| \ge \left|x\big(\pi_x^{-1}(s\kappa)\big)\right|.$$

Here we used that π_x is a non-increasing rearrangement. Now note that for $i \notin \operatorname{supp}(u)$,

$$\big((M_n-\lambda)u\big)_i = (M_n u)_i, \quad \text{and} \quad \operatorname{supp}(u) = \kappa\sqrt{np} \ll c\kappa np,$$

as long as $np \to \infty$. Therefore,

$$\begin{aligned}
\left\|(M_n-\lambda)u\right\|_2^2 &\ge \sum_{s=1}^{(pn)^{1/2}} \sum_{i\in I_s\setminus \operatorname{supp}(u)} \big((M_n u)_i\big)^2 \\
&\ge \frac{cpn}{2}\sum_{s=1}^{(pn)^{1/2}} \kappa\big(x\big(\pi_x^{-1}(s\kappa)\big)\big)^2 \\
&\ge \frac{cpn}{2}\sum_{k=(pn)^{(\ell_*-1)/2}}^{(pn)^{\ell_*/2}} \big(x\big(\pi_x^{-1}(k)\big)\big)^2 \\
&= \frac{cpn}{2}\|z_{\ell_*}\|_2^2 \ge \frac{cpn}{2}\cdot (C_{A.3}K)^{-2\ell_*},
\end{aligned} \tag{5}$$

where the third inequality uses the monotonicity of the sequence $\{|x(\pi_x^{-1}(k))|\}_{k=1}^n$. Combining the above with the bound on $\|v\|_2$, on the set $\mathcal{A}$, we get that

$$\begin{aligned}
\left\|(M_n-\lambda)x\right\|_2 &\ge \left\|(M_n-\lambda)u\right\|_2 - \|M_n-\lambda\|\cdot\|v\|_2 \\
&\ge \sqrt{\frac{cpn}{2}}(CK)^{-\ell_*} - (K+R)\sqrt{pn}\cdot 2\sqrt{2}(CK)^{-(\ell_*+1)} \ge \sqrt{pn}\big(C'K\big)^{-\ell_*}\sqrt{pn},
\end{aligned}$$

where the last inequality follows if the constants C, C' are chosen large enough independently of ℓ_*.

Now we consider the case when $\ell_* = \ell_0$. Note that in this setting, using (5), we have that

$$\|(M_n-\lambda u\|_2 \ge \sqrt{\frac{cpn}{2}}\|z_{\ell_0}\|_2,$$

and from (4), we have $\|v\|_2 = \|z_{\ell_0+1}\|_2 \le \sqrt{2}(CK)^{-1}\|z_{\ell_0}\|_2$. Now proceeding similarly as before, on $\mathcal{A}$, we obtain that

$$\left\|(M_n-\lambda)x\right\|_2 \ge \sqrt{pn}\big(C(K+R)\big)^{-\ell_0}\sqrt{pn}.$$

Since by Lemma A.1, $\mathbb{P}(\mathcal{A}) \ge 1-\exp(-cpn)$, the proof is complete. □

A.3. Moderately sparse vectors

A.3.1. Small ball probability

Lemma A.4. *Let δ_i be independent Bernoulli random variables taking value* 1 *with probability p and ξ_i be independent random variables with mean* 0, *variance* 1, *and subgaussian moment bounded by B. For a vector $x = (x_1, \dots, x_n) \in \mathbb{R}^n$, we have*

$$\mathcal{L}\left(\sum_{i=1}^n \delta_i \xi_i x_i, \frac{1}{4}\|x\|_2\sqrt{p}\right) \le 1 - \frac{cp}{(\|x\|_\infty/\|x\|_2)^2 + p}.$$

Proof. We can assume that x is a unit vector. We use the symmetrization technique to reduce the Levy function to a bound on the small ball probability around the origin. Let $\delta_1', \dots, \delta_n'$ and $\xi_1', \dots, \xi_n'$ be independent copies of $\delta_1, \dots, \delta_n$ and $\xi_1, \dots, \xi_n$. For any $r \in \mathbb{R}$,

$$\mathbb{P}^2\left(\left|\sum_{i=1}^n \delta_i \xi_i x_i - r\right| \le t\right) = \mathbb{P}\left(\left|\sum_{i=1}^n \delta_i \xi_i x_i - r\right| \le t\right)\mathbb{P}\left(\left|\sum_{i=1}^n \delta_i' \xi_i' x_i - r\right| \le t\right) \tag{6}$$

$$\le \mathbb{P}\left(\left|\sum_{i=1}^n (\delta_i \xi_i - \delta_i' \xi_i') x_i\right| \le 2t\right). \tag{7}$$

Let $\zeta_i := \delta_i \xi_i - \delta_i' \xi_i'$ and $S := \sum_{i=1}^n \zeta_i x_i$. Observe $\mathbb{E}\zeta_i = 0$, $\mathbb{E}\zeta_i^2 = 2p$, $\mathbb{E}\zeta_i^3 = 0$ and $\mathbb{E}\zeta^4 = 2p\mathbb{E}\xi^4 + 6p^2(\mathbb{E}\xi^2)^2 \le Cp$ for some constant C depending only on the subgaussian moment B. $\mathbb{E}S^2 = \sum_{i=1}^n \mathbb{E}\zeta_i^2 \cdot x_i^2 = 2p$ and

$$\mathbb{E}S^4 = \sum_{i=1}^n \mathbb{E}\zeta_i^4 \cdot x_i^4 + 3\sum_{j \ne k} \mathbb{E}\zeta_j^2 x_j^2 \cdot \mathbb{E}\zeta_k^2 x_k^2 \le C\|x\|_\infty^2 p + 12p^2$$

for some constant C' depending only on B. Thus by the Paley–Zygmund inequality, for $2t \le \sqrt{2p}$,

$$\mathbb{P}\big(|S| \le 2t\big) \le 1 - \frac{(\mathbb{E}S^2 - 4t^2)^2}{\mathbb{E}S^4}.$$

Therefore,

$$\mathbb{P}\left(|S| \le \frac{1}{2}\sqrt{p}\right) \le 1 - \frac{cp}{C\|x\|_\infty^2 + p}.$$

Combining this with inequality (6) yields

$$\mathbb{P}\left(\left|\sum_{i=1}^n \delta_i \xi_i x_i - r\right| \le \frac{1}{4}\sqrt{p}\right) \le \sqrt{1 - \frac{c'p}{\|x\|_\infty^2 + p}}$$

and setting $c = c'/2$ yields the result. □

Lemma A.5. *For a random variable, X, with subgaussian moment bounded by B. Then for any $k \in \mathbb{N}^+$, we have*

$$\mathbb{E}\big(|X|^k\big)^{1/k} \le 2B\sqrt{k}.$$

Lemma A.6. *Let $V_1, \dots, V_n$ be non-negative independent random variables such that $\mathbb{P}(V_i > 1) \ge q$, for all $i \in [n]$, and for some $q \in (0, 1/2)$. Then there exist constants $0 < c, c' < \infty$, such that*

$$\mathbb{P}\left(\sum_{j=1}^n V_j \le \frac{cqn}{\log(1/q)}\right) \le \exp\big(-c'_{A.6} n\big).$$

Corollary A.7. *Let M_n be a symmetric random matrix. Then for any $\alpha > 1$, there exist $\beta, \gamma > 0$ such that for $x \in \mathbb{R}^n$ satisfying*

$$\|x\|_\infty / \|x\|_2 \le \alpha\sqrt{p},$$

we have

$$\mathbb{P}\big(\big\|(M_n - \lambda)x\big\|_2, \beta\sqrt{pn}\|x\|_2\big) \le \exp(-\gamma n).$$

Proof. Without loss of generality, we can assume that the coordinates of x are organized by their magnitudes in decreasing order. Let $n_0 = \lceil n/2 \rceil$.

$$M_n - \lambda I = \begin{pmatrix} A & B \\ B^T & E \end{pmatrix}, \qquad x = \begin{pmatrix} y \\ z \end{pmatrix}.$$

Then,

$$\|y\|_2 \ge \frac{1}{2}\|x\|_2$$

so then $\|y\|_\infty / \|y\|_2 \le 2\alpha\sqrt{p}$. Fix a $w \in \mathbb{R}^n$ and let $V_j = \frac{16}{p\|y\|_2^2}((B^T y + Ez)_j - w_j)^2$. Also, by our assumptions,

$$\frac{c}{(\|y\|_\infty / \|y\|_2)^2 + p} \ge \frac{c}{4\alpha^2 + 1}$$

with c from Lemma A.4. By Lemmas A.4 and A.6 we have

$$\mathbb{P}\big(\big\|(M_n - \lambda)x\big\|_2, \beta\sqrt{pn}\|x\|_2\big) \le \mathbb{P}\big(\big\|B^T y + Ez\big\|_2, 2\beta\sqrt{pn}\|y\|_2\big).$$

□

A.4. Compressible vectors

Proof of Proposition 5.3. We begin by diving $[n]$ into two roughly equal sets. Let $n_0 = \lceil n/2 \rceil$. We denote this decomposition by

$$M_n - \lambda I = \begin{pmatrix} A & B \\ B^T & E \end{pmatrix}, \qquad x = \begin{pmatrix} y \\ z \end{pmatrix}, \qquad u = \begin{pmatrix} v \\ w \end{pmatrix},$$

where A is $n_0 \times n_0$ and C is $n - n_0 \times n - n_0$. Thus, we have the following equivalence.

$$\big\|(M_n - \lambda I)x\big\|_2^2 = \|Ay + Bz\|_2^2 + \big\|B^T y + Ez\big\|_2^2.$$

We condition on a realization of A and E. Let

$$W := \mathrm{Sparse}(M) \setminus \big(\mathrm{Comp}\big((8p)^{-1}, \rho\big) \cup \mathrm{Dom}\big((8p)^{-1}, (CK)^{-1}\big).$$

Denote $m = (8p)^{-1}$ so $m < M/2$.

Case I: Let's begin by assuming $p \ge \frac{1}{4}n^{-1/3}$. In this regime, $\ell_0 = 1$ and so $\rho = (C'K)^{-2}$ for C' from Lemma A.3. Observe that for $x \in V$,

$$\|x_{[m+1:M]}\|_\infty / \|x_{[m+1:M]}\|_2 \le CK\sqrt{8p}$$

for C from Lemma A.3. Since, $x \notin \mathrm{Comp}(m, \rho)$, $\|x_{[m+1:M]}\|_2 \ge \rho$. Thus, by Corollary A.7,

$$\mathbb{P}\big(\big\|(M_n - \lambda)x\big\|_2 \le \big(C'K\big)^{-3}\sqrt{pn}\|x_{[m+1:M]}\|_2\big) \le \exp\big(-c'n\big).$$

Now we extend this bound to all vectors in V. Define $\varepsilon = (C'K)^{-4}\rho$. There exists an ε-net $\mathcal{N} \subset V$ of cardinality less than

$$\binom{n}{M}\left(\frac{3}{\varepsilon}\right)^M \le \exp\left(cn\log\left(\frac{3e}{c_{5.3}\varepsilon}\right)\right).$$

Since $\lim_{x\to 0} x\log(1/x) = 0$ there exists a constant c so that $c\log(\frac{3e}{c\varepsilon}) \le c'/2$. Therefore, by the union bound,

$$\mathbb{P}\big(\exists x \in \mathcal{N} : \big\|(M_n - \lambda I)x\big\|_2 \le \big(C'K\big)^{-3}\sqrt{pn}\|x_{[m+1:M]}\|_2\big) \le \exp\big(-\big(c'/2\big)n\big).$$

We now extend this result to all of W. Assume for all $x \in \mathcal{N}$,

$$\big\|(M_n - \lambda I)x\big\|_2 \ge \big(C'K\big)^{-3}\sqrt{pn}\|x_{[m+1:M]}\|_2.$$

Let $x' \in V$. There exists a $x \in \mathcal{N}$ such that $\|x' - x\|_2 \le \varepsilon$. We have

$$\begin{aligned}\big\|(M_n - \lambda I)x'\big\|_2 &\ge \big\|(M_n - \lambda I)x\big\|_2 - \|M_n - \lambda I\|\big\|x - x'\big\|_2\\ &\ge \big(C'K\big)^{-3}\sqrt{pn}\|x_{[m+1:M]}\|_2 - K\varepsilon\\ &\ge \frac{1}{2}\big(C'K\big)^{-3}\sqrt{pn}\rho.\end{aligned}$$

Case II: We now tackle the remaining case where $\frac{C\log n}{n} \le p \le \frac{1}{4}n^{-1/3}$. Let $I, J \subset [n]$ be disjoint sets such that $|I| = m$, $|J| = M - m$. Let ε, τ be positive numbers to be chosen later. The sets

$$B_I := \big\{u \in B_2^n : \operatorname{supp}(u) \subset I\big\},$$

and

$$R_J := \big\{u \in \mathcal{S}^{n-1} : \operatorname{supp}(u) \subset J \text{ and } \|u\|_\infty \le 4CK\sqrt{p}\big\}$$

admit an ε-net $\mathcal{N}_I \subset B_I$ and a τ-net $\mathcal{N}_J \subset R_J$ of sizes

$$|\mathcal{N}_I| \le \left(\frac{3}{\varepsilon}\right)^{|I|}$$

and

$$|\mathcal{N}_J| \le \left(\frac{3}{\tau}\right)^{|J|}.$$

Let $\mathcal{N}_0$ be an ε-net in $[\rho/\sqrt{2}, 1] \subset \mathbb{R}$, and let

$$\mathcal{M}_{IJ} := \{u + lw : u \in \mathcal{N}_I, w \in \mathcal{N}_J, l \in \mathcal{N}_0\}$$

and

$$\mathcal{M} := \Bigg(\bigcup_{\substack{I\subset[n],\\ |I|=m}} \bigcup_{\substack{J\subset[n],\\ |J|=M-m, I\cap J=\varnothing}} \mathcal{M}_{IJ}\Bigg).$$

We now verify that this is an appropriate net for W. Let $x \in W$ be decomposed as $x = u_x + v_x$ where $u_x = x_{[1:m]}$ and $v_x = x_{[m+1:M]}$. Since $x \notin \operatorname{Comp}(m, \rho) \cup \operatorname{Dom}(m, (CK)^{-1})$, this implies that

$$\|v_x\|_2 \ge \rho \quad \text{and} \quad \|v_x\|_\infty < CK\sqrt{8p}\|v_x\|_2. \tag{8}$$

Choose $\bar{u} \in \mathcal{N}_I$, $\bar{v} \in \mathcal{N}_J$ and $l \in \mathcal{N}_0$ such that

$$\|\bar{u} - u\| \le \varepsilon, \qquad \big\|v_x/\|v_x\|_2 - \bar{v}\big\|_2 \le \tau \quad \text{and} \quad \big|\|v_x\|_2 - l\big| \le \varepsilon.$$

We can easily modify the net $\mathcal{M}$ so that $\mathcal{M} \subset W$ at the cost of adjusting ε and τ by a factor of 2. Thus, by (8) we have for a fixed $\bar{x} \in \mathcal{M}$

$$\mathbb{P}\big((M_n - \lambda I)\bar{x} \le \big(C'K\big)^{-3}\sqrt{pn}\|v_{\bar{x}}\|_2\big) \le e^{-c'n}.$$

Now for $x \in W$,

$$\big\|(M_n-\lambda)x\big\|_2 \ge \big\|(M_n-\lambda)\bar{x}\big\|_2 - \|M_n-\lambda\|\big(\|u_x-u_{\bar{x}}\|_2+\|v_x-v_{\bar{x}}\|_2\big).$$

We observe that

$$\|v_x-v_{\bar{x}}\|_2 \le \left\|\frac{v_x}{\|v_x\|_2}-\frac{v_{\bar{x}}}{\|v_{\bar{x}}\|_2}\right\|_2 \|v_{\bar{x}}\|_2 + \|v_x\|_2\left|1-\frac{\|v_{\bar{x}}\|_x}{\|v_x\|_2}\right| \le 2\tau\|v_{\bar{x}}\|_2+2\varepsilon.$$

Therefore, letting $\mu' := (CK)^{-3}$,

$$\big\|(M_n-\lambda)x\big\|_2 \ge \mu'\|v_{\bar{x}}\|_2\sqrt{pn} - K\sqrt{pn}\big(3\varepsilon+2\tau\|v_{\bar{x}}\|_2\big).$$

Setting $\varepsilon := \frac{\mu'\rho}{12K}$ and $\tau = \frac{\mu'}{8K}$ implies

$$\big\|(M_n-\lambda)x\big\|_2 \ge \frac{1}{2}\mu'\rho\sqrt{pn}.$$

To take the union bound over all points in the net, we must obtain an upperbound on the size of the cardinality of the net.

$$|\mathcal{M}| \le \binom{n}{m}\binom{n-m}{M-m}\left(\frac{3}{\varepsilon}\right)^m\left(\frac{3}{\tau}\right)^{M-m}\frac{1}{\varepsilon}.$$

We first bound

$$\binom{n}{m}\binom{n-m}{M-m} \le \binom{n}{m}\binom{n}{M} \le \left(\frac{en}{m}\right)^m\left(\frac{en}{M}\right)^M \le (8epn)^{(8p)^{-1}}\left(\frac{e}{c}\right)^{cn}.$$

Thus,

$$|\mathcal{M}| \le \left(\frac{288eKpn}{\mu'\rho}\right)^{(8p)^{-1}}\left(\frac{24eK}{c\mu'}\right)^{cn}.$$

We claim that

$$\left(\frac{288eKpn}{\mu'\rho}\right)^{(8p)^{-1}} \le \left(\frac{24eK}{c\mu'}\right)^{cn}.$$

This reduces to the assertion that

$$p^{-1}\log\left(\frac{pn}{\rho}\right) = o(n)$$

which is obvious by our assumption that $np \to \infty$ and $\ell_0 = o(np)$. Finally, we conclude that

$$|\mathcal{M}| \le \exp\big(-c'n/2\big)$$

if we choose c small enough since $\lim_{x\to 0} x\log(1/x) = 0$. Therefore, a union bound concludes the proof. □

Appendix B: Non-centered version of Proposition 5.3

To derive an analogue of Proposition 5.3. We begin by diving $[n]$ into two roughly equal sets. Let $n_0 = \lceil n/2\rceil$. We denote this decomposition by

$$A_n - p(J_n - I_n) = \begin{pmatrix} E & B \\ B^T & C \end{pmatrix}, \qquad x = \begin{pmatrix} y \\ z \end{pmatrix}.$$

To lowerbound $\|(A_n - p(J_n - I_n))x\|_2^2$, it suffices to lower bound $\|Ey + Bz\|_2^2$. For very sparse vectors, we can use the sign-matching argument from Section A.3 after conditioning on a realization of E. For moderately sparse vectors, the Lévy concentration argument is insensitive to shifts and for the net argument, we add an extra net over the low-dimensional image as in Section 8. We omit the details.

References

[1] L. Babai. Graph isomorphism in quasipolynomial time. In *Proceedings of the 48th Annual ACM SIGACT Symposium on Theory of Computing* 684–697. ACM, 2016. MR3536606

[2] L. Babai, D. Grigoryev and D. Mount. Isomorphism of graphs with bounded eigenvalue multiplicity. In *Proceedings of the Fourteenth Annual ACM Symposium on Theory of Computing* 310–324. ACM, 1982.

[3] A. Basak and M. Rudelson. Invertibility of sparse non-Hermitian matrices. *Adv. Math.* **310** (2017) 426–483. MR3620692 https://doi.org/10.1016/j.aim.2017.02.009

[4] A. Basak and M. Rudelson. Sharp transition of the invertibility of the adjacency matrices of sparse random graphs, ArXiv preprint, 2018. Available at arXiv:1809.08454.

[5] G. Ben Arous, P. Bourgade et al. Extreme gaps between eigenvalues of random matrices. *The Annals of Probability* **41** (4) (2013) 2648–2681. MR3112927 https://doi.org/10.1214/11-AOP710

[6] K. Clarkson and D. Woodruff. Low-rank approximation and regression in input sparsity time. *J. ACM* **63(6):Art. 54** (2017) 45. MR3614862 https://doi.org/10.1145/3019134

[7] A. Dasgupta, R. Kumar and T. Sarlós. A sparse Johnson–Lindenstrauss transform. In *STOC'10—Proceedings of the 2010 ACM International Symposium on Theory of Computing* 341–350. ACM, New York, 2010. MR2743282

[8] L. Erdős, B. Schlein and H. T. Yau. Wegner estimate and level repulsion for Wigner random matrices. *International Mathematics Research Notices* **2010** (3) (2009) 436–479. MR2587574 https://doi.org/10.1093/imrn/rnp136

[9] L. Erdos and H. T. Yau. Gap universality of generalized Wigner and beta-ensembles. ArXiv preprint, 2012. Available at arXiv:1211.3786.

[10] L. Erdős and H. T. Yau. Gap universality of generalized Wigner and β-ensembles. *J. Eur. Math. Soc. (JEMS)* **17** (8) (2015) 1927–2036. MR3372074 https://doi.org/10.4171/JEMS/548

[11] S. Foucart and H. Rauhut. *A Mathematical Introduction to Compressive Sensing. Applied and Numerical Harmonic Analysis*. Birkhäuser/Springer, New York, 2013. MR3100033 https://doi.org/10.1007/978-0-8176-4948-7

[12] A. Ganesh, A. Wagner, Z. Zhou, A. Yang, Y. Ma and J. Wright. Face recognition by sparse representation. In *Compressed Sensing* 515–539. Cambridge Univ. Press, Cambridge, 2012. MR2963577

[13] K. Luh and V. Vu. Dictionary learning with few samples and matrix concentration. *IEEE Trans. Inform. Theory* **62** (3) (2016) 1516–1527. MR3472263 https://doi.org/10.1109/TIT.2016.2517011

[14] M. Mehta. *Random Matrices*, **142**. Academic press, 2004. MR2129906

[15] J. Nelson and H. Nguyen. OSNAP: Faster numerical linear algebra algorithms via sparser subspace embeddings. In *2013 IEEE 54th Annual Symposium on Foundations of Computer Science—FOCS 2013* 117–126. IEEE Computer Soc., Los Alamitos, CA, 2013. MR3246213 https://doi.org/10.1109/FOCS.2013.21

[16] H. Nguyen, T. Tao and V. Vu. Random matrices: Tail bounds for gaps between eigenvalues. *Probab. Theory Related Fields* **167** (3–4) (2017) 777–816. MR3627428 https://doi.org/10.1007/s00440-016-0693-5

[17] H. Nguyen and V. Vu. Optimal inverse Littlewood–Offord theorems. *Adv. Math.* **226** (6) (2011) 5298–5319. MR2775902 https://doi.org/10.1016/j.aim.2011.01.005

[18] M. Rudelson and R. Vershynin. The Littlewood–Offord problem and invertibility of random matrices. *Adv. Math.* **218** (2) (2008) 600–633. MR2407948 https://doi.org/10.1016/j.aim.2008.01.010

[19] T. Tao. The asymptotic distribution of a single eigenvalue gap of a Wigner matrix. *Probability Theory and Related Fields* **157** (1–2) (2013) 81–106. MR3101841 https://doi.org/10.1007/s00440-012-0450-3

[20] T. Tao and V. Vu. Inverse Littlewood–Offord theorems and the condition number of random discrete matrices. *Ann. of Math. (2)* **169** (2) (2009) 595–632. MR2480613 https://doi.org/10.4007/annals.2009.169.595

[21] T. Tao and V. Vu. Random matrices: Universality of local eigenvalue statistics up to the edge. *Communications in Mathematical Physics* **298** (2) (2010) 549–572. MR2669449 https://doi.org/10.1007/s00220-010-1044-5

[22] T. Tao and V. Vu. Random matrices: Universality of local eigenvalue statistics. *Acta Math.* **206** (1) (2011) 127–204. MR2784665 https://doi.org/10.1007/s11511-011-0061-3

[23] T. Tao and V. Vu. Random matrices have simple spectrum. *Combinatorica* **37** (3) (2017) 539–553. MR3666791 https://doi.org/10.1007/s00493-016-3363-4

[24] R. Vershynin. Invertibility of symmetric random matrices. *Random Structures Algorithms* **44** (2) (2014) 135–182. MR3158627 https://doi.org/10.1002/rsa.20429

[25] J. Yang, J. Wright, T. Huang and Y. Ma. Image super-resolution via sparse representation. *IEEE Trans. Image Process.* **19** (11) (2010) 2861–2873. MR2814193 https://doi.org/10.1109/TIP.2010.2050625