J. Eur. Math. Soc. 23, 467–501 (2021) DOI 10.4171/JEMS/1015

Alexander E. Litvak · Anna Lytova · Konstantin Tikhomirov ·
Nicole Tomczak-Jaegermann · Pierre Youssef

Circular law for sparse random regular digraphs

Received January 16, 2018

Abstract. Fix a constant $C \geq 1$ and let $d = d(n)$ satisfy $d \leq \ln^C n$ for every large integer n. Denote by A_n the adjacency matrix of a uniform random directed d-regular graph on n vertices. We show that if $d \to \infty$ as $n \to \infty$, the empirical spectral distribution of the appropriately rescaled matrix A_n converges weakly in probability to the circular law. This result, together with an earlier work of Cook, completely settles the problem of weak convergence of the empirical distribution in a directed d-regular setting with the degree tending to infinity. As a crucial element of our proof, we develop a technique of bounding intermediate singular values of A_n based on studying random normals to rowspaces and on constructing a product structure to deal with the lack of independence between matrix entries.

Keywords. Circular law, logarithmic potential, random graphs, random matrices, regular graphs, sparse matrices, intermediate singular values

Contents

A. E. Litvak, N. Tomczak-Jaegermann: Department of Mathematical and Statistical Sciences, University of Alberta, Edmonton, AB, Canada, T6G 2G1;
e-mail: aelitvak@gmail.com, nicole.tomczak@ualberta.ca

A. Lytova: University of Opole, plac Kopernika 11A, 45-040 Opole, Poland;
e-mail: alytova@uni.opole.pl

K. Tikhomirov: Department of Mathematics, Princeton University,
Fine Hall, Washington Road, Princeton, NJ 08544, USA; e-mail: kt12@math.princeton.edu

P. Youssef: Université Paris Diderot, Laboratoire de probabilités et de modèles aléatoires,
75013 Paris, France; e-mail: youssef@math.univ-paris-diderot.fr

Mathematics Subject Classification (2020): Primary 60B20, 15B52, 46B06, 05C80; Secondary 46B09, 60C05

1. Introduction

Given an $n \times n$ random matrix B, its *empirical spectral distribution* (ESD) is the random probability measure on $\mathbb{C}$ given by

$$\mu_B := \frac{1}{n} \sum_{i=1}^{n} \delta_{\lambda_i},$$

where $(\lambda_i)_{i \le n}$ denote the eigenvalues of B (counting multiplicities). The study of the empirical spectral distribution is one of the major research directions in the theory of random matrices, with applications to other fields [29, 4, 6, 34, 14]. A fundamental fact in this area is the universality phenomenon which asserts that under very general conditions the empirical spectral distribution and some other characteristics of a random matrix asymptotically behave similarly to the empirical distribution (or corresponding characteristics) of the Gaussian random matrix of an appropriate symmetry type. This phenomenon has been confirmed for various models and in various senses (including limiting laws for the ESD, local eigenvalue statistics, distribution of eigenvectors). We refer to monographs [4, 6, 34, 14] for a (partial) exposition of the results.

For non-Hermitian random matrices with i.i.d. entries, the limit of the empirical spectral distribution is governed by the *circular law*. Compared to ESD's of the Wigner and sample covariance matrices, the study of the spectral distribution in the non-Hermitian setting is complicated due to its instability under small perturbations of the matrix entries and due to the fact that some of standard techniques, involving the moment method and truncation of the matrix entries, fail in the non-Hermitian case (we refer to [6, Section 11.1] for more information). As a specific example, while the bulk of the ESD of Hermitian matrices is stable under small-rank perturbations due to interlacing properties, the spectrum of random non-Hermitian matrices can be very sensitive even to a rank-one perturbation (see [6, Example 11.1] or [10, Example 1.2]).

Denote by μ_{circ} the uniform probability measure on the unit disk of the complex plane, that is,

$$d\mu_{\mathrm{circ}} = \pi^{-1} \mathbf{1}_{|z| \le 1}\, dx\, dy.$$

Convergence of the appropriately rescaled empirical spectral distribution of the standard Gaussian matrix with i.i.d. complex entries was derived in the first edition of [29, Chapter 15], and, much later, a corresponding result in the more delicate real case was obtained in [13]. Both results relied on the explicit formula for the joint distribution of eigenvalues, which is available in the Gaussian setting [15]. The circular law for non-Gaussian matrices with bounded densities of the entries was verified in [5] (following an earlier work [16], see also [18]); the density condition was removed in [17, 32, 19, 37, 39], with [39] establishing the circular law for the i.i.d. model under weakest moment assumptions. The sparse i.i.d. model was considered in [37, 38, 19, 8] (see also [26] for a non-i.i.d. sparse model). We refer to [10] for a detailed exposition and historical overview of the circular law in the i.i.d. setting, and for further references. For a review of other recent developments, including the limiting laws for inhomogeneous matrices and the local circular law, we refer to the introduction of [12].

In this paper, we are concerned with a sparse model of random matrices whose entries are not independent. In what follows, for any positive integers $d \le n$ we denote by $\mathcal{M}_{n,d}$ the set of all $n \times n$ matrices whose entries take values in $\{0, 1\}$ and the sum of elements in each row and in each column is equal to d. In other words, $\mathcal{M}_{n,d}$ is the set of adjacency matrices of d-regular directed graphs on n vertices, where we allow loops but do not allow multiple edges. We consider the random matrix A_n uniformly distributed on $\mathcal{M}_{n,d}$. Random directed d-regular graphs provide a basic model of a *typical* graph with predefined in- and out-degree sequences and in this connection are of interest in network analysis. In a more general setting, random (weighted) directed graphs are used to model connections between neurons and the eigenvalue distribution of their adjacency matrices (the *synaptic matrices* for neural networks) has been given considerable attention in the literature. We refer to the introduction of [12] for a discussion of those works.

In the directed d-regular setting, it was conjectured (see [10, Section 7]) that for any fixed $3 \le d \le n - 3$, μ_{A_n} converges to the probability measure with density

$$\frac{1}{\pi}\,\frac{d^2(d-1)}{(d^2 - |z|^2)^2}\mathbf{1}_{\{|z|<\sqrt{d}\}}$$

as $n \to \infty$. This measure is usually referred to as the *oriented Kesten–McKay distribution*, a non-symmetric version of the classical Kesten–McKay law for the limiting ESD of random undirected d-regular graphs [20, 28, 9]. Up to rescaling by $\sqrt{d}$, this measure tends to the circular law as $d \to \infty$. Proving the above conjecture remains a major challenge as of this writing.

In this paper we establish the circular law for sparse random directed d-regular graphs for any d going to infinity with n. We prove the following theorem.

Theorem 1.1 (The circular law). *Fix a constant $C \ge 1$ and for any $n > 1$ let $d = d(n)$ be a positive integer satisfying $d \le \ln^C n$. Assume that $d \to \infty$ as $n \to \infty$. Then the sequence of empirical spectral distributions $(\mu_{d^{-1/2}A_n})_n$, corresponding to the random matrices A_n uniformly distributed in $\mathcal{M}_{n,d}$, converges weakly in probability to the uniform distribution on the unit disk of the complex plane.*

The circular law for d-regular digraphs in the range $\ln^{96} n \le \min(d, n-d)$ was verified in earlier work [12] (see also [7]). Thus, our Theorem 1.1 closes the gap between the known limiting distribution for denser d-regular digraphs and the conjectured oriented Kesten–McKay limiting distribution for d-regular digraphs of constant degree. The proof of Theorem 1.1 combines some known methods, used previously in works on the circular law, with new crucial ingredients related to estimating the *intermediate* singular values of the shifted adjacency matrix.

The rest of the introduction is divided into two parts. In the first part, we recall known techniques (such as Hermitization) and previously established facts about d-regular digraphs that will be needed for the proof. In the second part, we discuss limitations of existing tools (see remarks after Proposition 1.5) and describe our approach to bounding intermediate singular values of $A_n - z\,\mathrm{Id}$.

As in [16, 5, 19, 39] dealing with the i.i.d. setting, a key element in the proof of the circular law for d-regular digraphs is to transport the problem of the limiting ESD to the singular values distribution, which is much easier to study. This method—called the Hermitization technique—goes back to Girko [16] and exploits a close relation between the log-potential functions of the spectral and singular values distributions. Following Girko, this idea was used in various papers dealing with non-Hermitian random matrices, in particular in [5, 19, 39]. The Hermitization technique is presented in the literature in somewhat different forms; we follow the exposition in [10].

The *singular values distribution* of an $n\times n$ random matrix B is the random probability measure on $\mathbb{R}$ given by

$$\nu_B := \frac{1}{n}\sum_{i=1}^{n}\delta_{s_i},$$

where $(s_i)_{i\le n}$ are the singular values of B. Everywhere in this paper, we use non-increasing ordering for the singular values, that is, $s_1 = s_1(B)$ is the largest one and $s_n = s_n(B)$ is the smallest one.

The *logarithmic potential* $U_\mu : \mathbb{C} \to (-\infty, \infty]$ of a probability measure μ on $\mathbb{C}$ is defined for every $z \in \mathbb{C}$ by

$$U_\mu(z) := -\int_{\mathbb{C}} \ln|z-\lambda|\, d\mu(\lambda).$$

The logarithmic potential uniquely determines the underlying measure, that is, if $U_\mu = U_{\mu'}$ Lebesgue almost everywhere then $\mu = \mu'$ (see, in particular, [10, Lemma 4.1]).

Given an $n \times n$ matrix B, it is easy to check that

$$U_{\mu_B}(z) = -\frac{1}{n}\ln|\det(B - z\,\mathrm{Id})| = -\int_0^\infty \ln(t)\, d\nu_{B-z\,\mathrm{Id}}(t) = -\frac{1}{n}\sum_{i=1}^{n}\ln(s_i(B - z\,\mathrm{Id})).$$

Therefore, knowing $\nu_{B-z\,\mathrm{Id}}$ for almost all $z \in \mathbb{C}$, we can determine U_{μ_B}, hence μ_B itself. This observation lies at the heart of the method. We state its formalized version.

Lemma 1.2 (Hermitization, see [10, Lemma 4.3]). *For each n, let B_n be an $n \times n$ complex random matrix. Assume that for Lebesgue almost all $z \in \mathbb{C}$,*

(i) *there exists a probability measure ν_z on $\mathbb{R}_+$ such that $\nu_{B_n - z\,\mathrm{Id}}$ tends weakly to ν_z in probability;*
(ii) $\ln(\cdot)$ *is uniformly integrable for $\nu_{B_n-z\,\mathrm{Id}}$ in probability, i.e. for every $\varepsilon > 0$ there exists $T = T(z,\varepsilon) < \infty$ such that*

$$\sup_n \mathbb{P}\left\{\int_{\{|\ln(s)|>T\}} |\ln(s)|\, d\nu_{B_n-z\,\mathrm{Id}}(s) > \varepsilon\right\} \le \varepsilon.$$

Then μ_{B_n} converges weakly in probability to the unique probability measure μ on $\mathbb{C}$ whose logarithmic potential is given by

$$U_\mu(z) = -\int_0^\infty \ln(s)\, d\nu_z(s). \tag{1}$$

Thus in order to establish the circular law, it is enough to show the convergence of the empirical singular values distribution and the uniform integrability of the logarithm. For the first part, we will rely on a recent result of Cook [12], who uses the above strategy in order to establish the circular law for the uniform model on $\mathcal{M}_{n,d}$ for $d \geq \ln^{96} n$. The following is a version of [12, Proposition 7.2]. Note that its proof does not require that d is at least polylogarithmic in n, just $d \to \infty$ is enough (see Remark 5.2 below).

Proposition 1.3 (Weak convergence of singular values distributions, [12]). *Assume that $d = d(n) = o(\sqrt{n})$ and $d \to \infty$ as $n \to \infty$. Denote $B_z = d^{-1/2}A_n - z\,\mathrm{Id}$. Then for each $z \in \mathbb{C}$ there exists a probability measure ν_z on $\mathbb{R}_+$ such that ν_{B_z} converges weakly in probability to ν_z as $n \to \infty$. Moreover, the family $\{\nu_z\}_{z\in\mathbb{C}}$ satisfies (1) with $\mu = \mu_{\mathrm{circ}}$.*

In fact, this proposition was stated in [12] for the centralized matrix

$$X_n = A_n - \frac{d}{n}\mathbf{1}\mathbf{1}^t$$

instead of A_n, where $\mathbf{1}$ denotes the column vector with all components equal 1. However, since these two matrices differ by a rank-one matrix, using the interlacing of their singular values one can deduce that their empirical singular value distributions satisfy

$$\sup_{a>0} |\nu_{B_z}([0,a]) - \nu_{d^{-1/2}X_n - z\,\mathrm{Id}}([0,a])| \leq 1/n,$$

where B_z is as in Proposition 1.3. This has also been used in [12, (7.6)]. Therefore the two corresponding singular values distributions exhibit the same limiting behavior.

It is clear from the above discussion that the main obstacle to establishing Theorem 1.1 is in showing the uniform integrability of the logarithm. More precisely, one needs to prove that for any $\varepsilon \in (0,1)$ and any $z \in \mathbb{C}$ there exists $T = T(z,\varepsilon) > 0$ such that with probability going to 1 as $n \to \infty$,

$$\sum_{i:\,|\ln s_i(B_z)|\geq T} |\ln s_i(B_z)| \leq \varepsilon n. \tag{2}$$

A simple computation involving the Hilbert–Schmidt norm of B_z shows that the main contributors to the above sum are small singular values, namely those smaller than e^{-T}.

In [24], building upon ideas in [11] as well as on the authors' works [23, 22], a polynomial lower bound on the smallest singular value of B_z was obtained.

Theorem 1.4 ([24]). *There exists a universal constant $C \geq 1$ such that for all positive integers d, n satisfying $C \leq d \leq n/\ln^2 n$ and every $z \in \mathbb{C}$ with $|z| \leq d/6$ one has*

$$\mathbb{P}\{s_{\min}(A_n - z\,\mathrm{Id}) \geq n^{-6}\} \geq 1 - d^{-1/4}.$$

The above came as an improvement (in the sparse regime) of an earlier estimate of Cook [12], who derived his result under the additional assumption $d \geq \ln^C n$ for a universal constant C. Theorem 1.4 implies that the contribution of $o(n/\ln n)$ least singular values to the sum in (2) is negligible.

Together with the observation concerning largest singular values, this leaves the task of estimating the sum $\sum_J |\ln s_i(B_z)|$, where

$$J := \{i : i \leq n - o(n/\ln n) \text{ and } s_i(B_z) \leq e^{-T}\}.$$

Partially, the estimate comes from the following result of [12], obtained via comparison with Bernoulli random matrices.

Proposition 1.5. *There are absolute constants $C > 1 > c > 0$ such that the following holds. Let $C \leq d \leq n$ be positive integers and $z \in \mathbb{C}$. Assume that $d = d(n) = o(\sqrt{n})$ and $d \to \infty$ as $n \to \infty$. Denote $B_z = d^{-1/2} A_n - z\,\mathrm{Id}$. Then for large enough n, with probability at least $1 - \exp(-n/2)$, for every $k \leq n - Cnd^{-1/48}$ one has*

$$s_k(B_z) \geq c\,\frac{n-k}{n}.$$

This proposition is an immediate consequence of [12, Proposition 7.3] (see Section 5 below). In fact, [12, Proposition 7.3] was stated for d polylogarithmic in n. In Remark 5.2 below we indicate the changes to make in [12] in order to derive Proposition 1.5 without this restriction on d (these changes are actually implicitly mentioned in [12]).

Proposition 1.5 can be viewed as a (weak local) form of the *Marchenko–Pastur law* for the singular values distribution [27, 41]. When d is at least polylogarithmic in n (with an appropriate power of the logarithm), the proposition is enough to cover the entire range of singular values indexed by the set J and complete the proof. This is the approach taken in [12]. However, when d is smaller than a power of $\ln n$, the above result leaves untreated the range of smallish singular values s_k for $n - Cnd^{-1/48} \leq k \leq n - o(n/\ln n)$.

The idea of the proof of [12, Proposition 7.3] is to compare the uniform directed d-regular model with directed Erdős–Rényi graph, that is, to replace the matrix A_n by a matrix $\mathcal{B}_n$, whose entries are i.i.d. Bernoulli random variables with the parameter d/n. At this step one has to condition on the event that the Erdős–Renyi graph is d-regular, which is of very small probability, superexponential in n [30]. In this way satisfactory estimates for the intermediate singular values of the shifted adjacency matrix $A_n - z\,\mathrm{Id}$ can be obtained only if very strong estimates are available in the Bernoulli setting, which hold with probability at least $1 - \exp(-\omega(n))$, where $\omega(n)/n \to \infty$ as $n \to \infty$. Currently, no estimates of this type are available in the very sparse regime, and it is not clear whether such strong estimates can be obtained at all. This forces us to develop a completely different approach to bound the singular values s_k of $A_n - z\,\mathrm{Id}$ in the range $n - Cnd^{-c} \leq k \leq n - o(n/\ln n)$. We obtain the following bounds.

Theorem 1.6 (Intermediate singular values). *There exists a universal constant $C \geq 1$ with the following property. Let d, n be integers satisfying $C \leq d \leq \ln^{96} n$ and let $z \in \mathbb{C}$ be such that $|z| \leq \sqrt{d}\,\ln d$. Then for all k satisfying*

$$n - 2nd^{-3/2} \leq k \leq n - 3n/\ln^{144} n$$

one has

$$\mathbb{P}\Big\{A_n \in \mathcal{M}_{n,d} \,:\, s_k(A_n - z\,\mathrm{Id}) \geq \exp\Big(-C\Big(\frac{n}{n-k}\Big)^{1/144}\Big)\Big\} \geq 1 - C\,\frac{n-k}{n}.$$

In particular,

$$\mathbb{P}\{A_n \in \mathcal{M}_{n,d} \,:\, s_k(A_n - z\,\mathrm{Id}) \geq \exp(-C\,d^{1/96}) \textit{ for all } k \leq n - 2nd^{-3/2}\} \geq 1 - \frac{C}{d^{3/2}}.$$

In the above, we restricted our analysis to $d \leq \ln^{96} n$ as it complements what is covered by Proposition 1.5. Our approach can be extended to higher powers of $\ln n$ (even possibly to any $d \leq \exp(\sqrt{\ln n})$ as in [25]), but we prefer to prove the above statement as it is sufficient for our purposes and improves the exposition. Equipped with Theorem 1.4, Proposition 1.5, and Theorem 1.6, we have bounds on all singular values, which would allow us to show the uniform integrability of the logarithm and thus to establish the circular law. We note that the idea of splitting the singular values into different regimes is standard in this context (see [36, Chapter 2, Section 8] for more details) as one needs different levels of precision depending on the magnitude of the singular values. In our case, the sparsity adds a serious challenge and the comparison methods described above are ineffective. Moreover, due to the lack of independence, standard approaches to estimating the singular values are not applicable in our setting. For example, one cannot use Talagrand's concentration inequality [36, Theorem 2.1.13] in this context the same way as was previously done in the literature (see, in particular, [39]). The issues appear when one tries to follow the standard scheme which reduces estimates for the singular values to distance estimates for the matrix rows. Namely, the second moment identity [36] or the restricted invertibility principle (see, for example, [31, Theorem 9]) relates the intermediate singular values to quantities of the form

$$\mathrm{dist}\big(R_i(B_z), \mathrm{span}\,\{R_j(B_z)\}_{j\in I}\big)$$

for $I \subset [n]$ and $i \in [n]\setminus I$, where $R_i(B_z)$ denotes the i-th row of B_z. When these rows are independent, one can condition on a realization of $E := \mathrm{span}\,\{R_j(B_z)\}_{j\in I}$ and then use the randomness of the i-th row together with standard anti-concentration arguments to get a lower bound for $\|P_{E^\perp} R_i(B_z)\|_2 = \mathrm{dist}(R_i(B_z), E)$. On the other hand, the randomness of E is used to ensure that its normal vector is well spread for the anti-concentration argument to work. In our setting, i.e., for random d-regular graphs, the lack of product structure adds serious complications to the problem. Studying the distribution of a row conditioned on the realization of other rows involves careful application of the expansion properties of the underlying graph. In particular, such a direction was pursued by the third and last named authors [40] to establish, for denser d-regular graphs, a large deviation inequality for the inner product of a row with an arbitrary vector, conditioned on a realization of a block of rows. At the same time, the technical approach of [40] is not applicable here as we deal with very sparse random graphs and are interested in a small ball inequality instead of large deviations.

The key idea behind the argument developed in this paper is to inject additional randomness and create a sort of product structure, which would allow us to use the randomness of each of the (dependent) quantities involved. Similar ideas were used in asymptotic geometric analysis in the study of volume distribution in convex bodies [3, 33, 1, 2]. We provide a rough illustration of this idea. Fix $I \subset [n]$ and $i \in [n] \setminus I$, and observe that

$$\operatorname{dist}(R_i(B_z), E)^2 = \|P_{E^\perp} R_i(B_z)\|_2^2 = \mathbb{E}_G\, |\langle P_{E^\perp} G, R_i(B_z)\rangle|^2, \tag{3}$$

where G is a standard Gaussian vector in $\mathbb{C}^n$ and the expectation is taken with respect to G. Now standard Gaussian concentration allows us to remove the expectation above and to benefit from the randomness of G to study the quantity $\langle P_{E^\perp} G, R_i(B_z)\rangle$. The vector $P_{E^\perp} G$ plays the role of a *uniform random normal* to E. In other words, instead of working with a fixed vector normal to E, we choose a normal vector randomly according to the Gaussian distribution on $E^\perp$. As the key technical ingredient, we prove that the random normal is typically unstructured, i.e., has many levels of coordinates. In this sense, one of the most important inputs of this paper is a statement about the kernel of submatrices of $A_n - z\,\mathrm{Id}$ formed by removing a small proportion of rows (see Theorem 4.2). Once equipped with this statement, we switch back to the randomness of $R_i(B_z)$ in order to establish an anti-concentration inequality. Note that this also requires additional efforts as we deal with a sum of dependent random variables with non-trivial conditional distributions (conditioned on a realization of E) as opposed to the standard estimates in the independent case.

The structure of normal vectors to subspaces spanned by the rows of random d-regular graphs was investigated by the authors in [25]. In particular, it was shown that if the subspace E is of large dimension, then any vector normal to it either is very steep (has a sudden drop at the beginning of the non-increasing rearrangement of absolute values of its coordinates) or has moderate coordinate decay and is unstructured (i.e., has many levels of coordinates). The latter property is essential for the anti-concentration argument to be effective. However, in general, a normal vector can be not sufficiently unstructured for our purposes. To improve this, we pass to a uniform random normal. Informally speaking, one of the advantages of introducing the additional randomness is that the random Gaussian vector picks the best normal vector and benefits from better structural properties. This vague observation will become more rigorous and clear from the proof of Theorem 4.2 (see also remarks following that theorem). We expect that some elements of our proof can be fruitful in the study of other matrix models with the lack of independence.

The paper is organized as follows. In Section 2, we derive the circular law assuming the estimates on the intermediate singular values. In Section 3, we introduce notations. In Section 4, we prove the structural theorem (Theorem 4.2) for uniform random normals after providing estimates for order statistics of projection of Gaussian vectors. In Section 5, we establish an anti-concentration estimate and combine it with the structural theorem in order to prove Theorem 1.6.

2. Proof of Theorem 1.1

In this section we prove Theorem 1.1—the circular law for the limiting spectral distribution—assuming the results mentioned in the introduction. As discussed before, we only need to verify uniform integrability of the logarithm, that is, item (ii) of Lemma 1.2.

Fix $z \in \mathbb{C}$, $\varepsilon > 0$ and, given n and d satisfying the assumptions of the theorem, set $B_z := d^{-1/2} A_n - z\,\mathrm{Id}$. We want to show that there exists $T = T(z, \varepsilon) > 0$ such that

$$\mathbb{P}\Big\{\sum_{i:\,|\ln s_i(B_z)|\geq T} |\ln s_i(B_z)| \geq \varepsilon n\Big\} \leq \varepsilon.$$

In the proof below, a sum over an empty set is always assumed to be 0.

For large singular numbers we will apply a deterministic bound which follows from d-regularity, namely we will use $\|A_n\|_{\mathrm{HS}}^2 = nd$, where $\|\cdot\|_{\mathrm{HS}}$ denotes the Hilbert–Schmidt norm. Choose a sufficiently large $T = T(z, \varepsilon) > 0$ to ensure that

$$\ln x \leq \frac{\varepsilon}{4(1+|z|^2)} x^2$$

whenever $x \geq e^T$. Then

$$\sum_{i:\,s_i(B_z)\geq e^T} \ln s_i(B_z) \leq \frac{\varepsilon}{4(1+|z|^2)} \sum_{i:\,s_i(B_z)\geq e^T} s_i^2(B_z) \leq \frac{\varepsilon}{4(1+|z|^2)} \|B_z\|_{\mathrm{HS}}^2$$
$$\leq \frac{\varepsilon}{2(1+|z|^2)} (\|d^{-1/2} A_n\|_{\mathrm{HS}}^2 + \|z\,\mathrm{Id}\,\|_{\mathrm{HS}}^2) = \frac{\varepsilon}{2} n.$$

Note that one could also use the spectral gap estimate for d-regular graphs (see [40] and references therein), which implies that with large probability all singular values of $d^{-1/2} A_n$ except for s_1 are bounded above by a universal constant.

Thus it is enough to show a bound for small singular values, more precisely, it is enough to show that

$$\mathbb{P}\Big\{\sum_{i\in I} |\ln s_i(B_z)| \geq \varepsilon n/2\Big\} \leq \varepsilon, \quad \text{where} \quad I = \{i : s_i(B_z) \leq e^{-T}\}.$$

We split the set I into four parts:

$$I_1 := I \cap \{i : i \leq n - Cnd^{-1/48}\}, \qquad I_2 := (I \cap \{i : i \leq n - 2n/d^{3/2}\}) \setminus I_1,$$
$$I_3 := (I \cap \{i : i \leq n - n/\ln^2 n\}) \setminus (I_1 \cup I_2), \qquad I_4 := I \cap \{i : i > n - n/\ln^2 n\},$$

where $C \geq 1$ is the absolute constant from Proposition 1.5. Proposition 1.5 implies that with probability at least $1 - \exp(-n/2)$, for all $i \leq n - Cd^{-1/48} n$ we have

$$s_i(B_z) \geq c(n-i)/n$$

for an absolute constant $c \in (0,1)$. Note that if $i \in I$ then this inequality implies $i \geq n(1 - 1/(ce^T))$. Thus $I_1 \neq \emptyset$ if and only if $d^{1/48} \geq ce^T$, in which case $n \gg ce^T$. Denoting

$$I_1' := \{i : n(1 - 1/(ce^T)) \leq i \leq (1 - Cd^{-1/48})n\}$$

and assuming $I_1 \neq \emptyset$ we obtain

$$\sum_{i\in I_1} |\ln s_i(B_z)| \leq \sum_{i\in I_1'} \ln \frac{n}{c(n-i)} \leq \sum_{k=1}^{n/(ce^T)} \ln \frac{n}{ck} \leq 2\int_1^{n/(ce^T)} \ln \frac{n}{ct}\, dt \leq \frac{2n(T+1)}{ce^T}.$$

For large enough T and for $n \geq 2\ln(4/\varepsilon)$, this implies

$$\mathbb{P}\Big\{\sum_{i\in I_1} |\ln s_i(B_z)| \geq \varepsilon n/8\Big\} \leq \varepsilon/4.$$

Further, by Theorem 1.6, for some universal constants C', C_0, with probability at least $1 - C'd^{-3/2}$ we have

$$\sum_{i\in I_2} |\ln s_i(B_z)| \leq |I_2|\,(C'd^{1/96}) \leq C_0 d^{-1/96} n \leq \varepsilon n/8$$

provided that $d \geq (8C_0/\varepsilon)^{96}$ and $d\ln^2 d \geq |z|^2$.

Next, by Theorem 1.4 applied to the matrix $A_n - z\sqrt{d}\,\mathrm{Id}$, with probability at least $1 - d^{-1/4}$ we have $s_n(B_z) \geq n^{-6}/\sqrt{d}$ and thus

$$\sum_{i\in I_4} |\ln s_i(B_z)| \leq \sum_{i>n-n/\ln^2 n} |\ln s_i(B_z)| \leq \frac{n}{\ln^2 n}\,|\ln s_n(B_z)| \leq \varepsilon n/8$$

provided that $d \geq 36|z|^2$ and $7/\ln n \leq \varepsilon/8$.

It remains to estimate the sum over I_3. Note that $I_3 \neq \emptyset$ only if $2n/d^{3/2} \geq n/\ln^2 n$. Consider the sequence of indices $i_0, i_1, \ldots$ defined by

$$i_u := \lfloor n - 2^{-u} d^{-3/2} n \rfloor$$

for $u \geq 0$ and let u_0 be the smallest integer such that $i_{u_0} \geq n - n/\ln^2 n$. Then

$$\sum_{i\in I_3} |\ln s_i(B_z)| \leq \sum_{u=0}^{u_0-1} (i_{u+1} - i_u)|\ln s_{i_{u+1}}(B_z)| \leq 4d^{-3/2} n \sum_{u=0}^{u_0-1} 2^{-(u+1)} |\ln s_{i_{u+1}}(B_z)|. \qquad (4)$$

Assuming that $d\ln^2 d \geq |z|^2$ and applying Theorem 1.6 again we find that for every $0 \leq u \leq u_0 - 1$,

$$\mathbb{P}\{s_{i_{u+1}}(B_z) \geq \exp(-C'd^{1/96} 2^{(u+1)/144})\} \geq 1 - \frac{C'}{d^{3/2} 2^{u+1}},$$

where $C' > 0$ is a universal constant. Taking the union bound we get, with probability at least $1 - C'd^{-3/2}$,

$$|\ln s_{i_{u+1}}(B_z)| \leq C'\, d^{1/96} 2^{(u+1)/144} \qquad \text{for all } 0 \leq u \leq u_0 - 1.$$

By (4) we obtain, with the same probability,

$$\sum_{i\in I_3} |\ln s_i(B_z)| \leq 2C'd^{-3/2} n \sum_{u=0}^{u_0-1} 2^{-(u+1)143/144} d^{1/96} \leq \varepsilon n/8$$

provided that $d \gg 1/\varepsilon$. Combining the estimates for the sums over $I_1, \ldots, I_4$ we obtain the result provided that $d \geq d_0 := \max\{36|z|^2, C_2/\varepsilon^{96}\}$ for a large universal constant $C_2 > 0$.

Finally, we would like to comment on a purely technical aspect: why we can assume that $d \geq d_0$. Given $n \geq 1$, let $X_n \subset \mathbb{C}$ be the set of all eigenvalues of all d-regular $n \times n$

matrices divided by $\sqrt{d}$ (taken for all $d \le \ln^{96} n$). Since $X := \bigcup_n X_n$ has zero Lebesgue measure, it is enough to consider $z \notin X$. Now given a sequence $d(n) \to \infty$, $z \in \mathbb{C} \setminus X$, and $\varepsilon > 0$ choose $n_0 = n_0(z, \varepsilon)$ so that $d(n) \ge d_0$ whenever $n \ge n_0$. Set

$$\rho = \rho(z, \varepsilon) := \operatorname{dist}\Big(z, \bigcup_{n \le n_0} X_n\Big).$$

Then $\rho > 0$ and for every d-regular $n \times n$ matrix A_n with $n \le n_0$ the matrix B_z is invertible and the norm of its inverse can be estimated in terms of n, d, and ρ (e.g., via the formula for the inverse matrix, its Hilbert–Schmidt norm, and Hadamard's inequality). Since $n \le n_0$ and $s_n(B_z) = 1/\|B_z^{-1}\|$, we obtain a lower bound on $s_n(B_z)$ in terms of n_0 and ρ. Therefore, taking sufficiently large $T = T(z, \varepsilon)$, we find that for any $n \le n_0$ the set $\{i : |\ln s_i(B_z)| \ge T\}$ is empty.

3. Notation

Given positive integers $k \le \ell$, we denote $[k] = \{1, \dots, k\}$ and $[k, \ell] = \{k, k+1, \dots, \ell\}$. Given a sequence $(x_i)_{i=1}^n$, we denote by $(x_i^*)_{i=1}^n$ the non-increasing rearrangement of $(|x_i|)_{i=1}^n$. The vectors of the canonical basis of $\mathbb{C}^n$ are denoted by $e_1, \dots, e_n$, and the canonical Euclidean norm on $\mathbb{C}^n$ is denoted by $\|\cdot\|_2$. Given $E \subset \mathbb{C}^n$, the orthogonal projection on E is denoted by P_E. Given $J \subset [n]$, we denote by P_J the orthogonal projection on the space spanned by e_j, $j \in J$. Given an $n \times n$ matrix A we denote its rows by $R_i(A)$, $i \le n$. A set (or a subset of a certain set) of cardinality k is called a *k-set* (resp., *k-subset*). $\mathbf{1}$ denotes the column vector with all components 1.

As mentioned in the introduction, for every positive integer $d \le n$, we denote by $\mathcal{M}_{n,d}$ the set of all $n \times n$ matrices whose entries take values in $\{0, 1\}$ and the sum of elements of each row and of each column is equal to d. In other words, $\mathcal{M}_{n,d}$ is the set of adjacency matrices of directed d-regular graphs on n vertices. The random matrix uniformly distributed on $\mathcal{M}_{n,d}$ is denoted by A_n, and as before, we set $B_z := d^{-1/2} A_n - z\,\mathrm{Id}$, where $z \in \mathbb{C}$ and Id is the identity matrix. Below we often deal with a random subspace of $\mathbb{C}^n$ spanned by some rows of a random matrix. Given $I \subset [n]$, we denote by $E(A_n, I)$ (resp., $E(B_z, I)$) the random subspace spanned by the rows of A_n (resp., B_z) indexed by I.

A standard Gaussian variable in $\mathbb{C}$ is the variable $g = \xi_1 + i\xi_2$, where ξ_1 and ξ_2 are independent real Gaussians distributed according to $\mathcal{N}(0, 1/2)$. A standard Gaussian vector in $\mathbb{C}^n$ is a vector $G := (g_1, \dots, g_n)$, where g_i's are independent standard complex Gaussian variables. We always assume that G is independent of A_n. We use the fact that the distribution of G, denoted below by γ_n, is invariant under orthogonal transformations and that for every orthogonal projection P of rank $k \le n$ the vector PG is distributed as the standard Gaussian vector in $P\mathbb{C}^n \approx \mathbb{C}^k$. In particular, for every non-degenerate subspace E of $\mathbb{C}^n$ and every fixed $x \in \mathbb{C}^n \setminus \{0\}$ one has, for every $t > 0$,

$$\mathbb{P}\{|\langle x, P_E G\rangle| \le t\|P_E x\|_2\} = \mathbb{P}\left\{\left|\left\langle \frac{P_E x}{\|P_E x\|_2}, G\right\rangle\right| \le t\right\} = \mathbb{P}\{|g| \le t\} = 1 - \exp(-t^2). \tag{5}$$

In the next section we deal with *uniform random normals* which we define in the following way. Let $E \subset \mathbb{C}^n$ be a linear subspace and $E^\perp$ its orthogonal complement. The uniform random normal to E is a standard Gaussian vector in the orthogonal complement of E. Note that the uniform random normal to E is distributed as $P_{E^\perp}G$, which will often be denoted by Y.

4. Uniform random normals

The result of this section is based on the structural theorem proved in [25, Theorem 1.1]. We state a special case of this theorem, in which we fix several parameters and restrict the range of d and of the index subset $|I^c|$ according to our needs.

Theorem 4.1. *Let d, n be sufficiently large integers satisfying $d \le \ln^{96} n$ and $z \in \mathbb{C}$ be such that $|z| \le \sqrt{d}\ln d$. Let $a \in (d^{-1/2}, 1)$ and $\gamma = 1/288$. Fix a subset $I \subset [n]$ satisfying*

$$n/\ln^{1/\gamma} n \le |I^c| \le n/d^3.$$

Let $E = E(B_z, I)$ be the random subspace spanned by the rows of B_z indexed by I. Then with probability at least $1 - 1/n$ any non-zero vector $x \in E^\perp$ satisfies one of the two conditions:

- (Gradual with many levels) *For all $i \le a|I^c|$ one has $x_i^* \le 0.9(n/i)^3 x^*_{a|I^c|}$ and for all $\lambda \in \mathbb{C}$,*

$$\left|\left\{j \le n : |x_j - \lambda| \le \exp\bigl(-2(n/|I^c|)^\gamma\bigr)x^*_{a|I^c|}\right\}\right| \le \left(\frac{|I^c|}{n}\right)^{\gamma/2} n.$$

- (Very steep) *There exists $i \le a|I^c|$ such that $x_i^* > 0.9(n/i)^3 x^*_{a|I^c|}$.*

The idea, developed in this section, is that a normal vector picked uniformly at random in $E^\perp$ has better structural properties (in fact, is more "unstructured"). At the intuitive level, in the case of large codimensional $E \subset \mathbb{C}^n$, the vector $P_{E^\perp}G$ should be typically unstructured, i.e., should not have many coordinates of almost the same value. We will make this notion precise by combining Theorem 4.1 with some probabilistic arguments. The main result of this section is the following theorem.

Theorem 4.2. *Let d, n be sufficiently large integers satisfying $d \le \ln^{96} n$ and let $z \in \mathbb{C}$ be such that $|z| \le \sqrt{d}\ln d$. Let $\gamma = 1/288$ and fix a subset $I \subset [n]$ satisfying*

$$n/\ln^{1/\gamma} n \le |I^c| \le n/d^3.$$

Let $E = E(B_z, I)$ be the random subspace spanned by the rows of B_z indexed by I. Then

$$\mathbb{P}\Bigl\{\text{for every } J \subset [n] \text{ with } |J| \le 2(|I^c|/n)^{\gamma/2} n \text{ there is } \lambda \in \mathbb{C} \text{ such that}$$
$$\bigl|\{j \in [n] \setminus J : |\langle P_{E^\perp}G, e_j\rangle - \lambda| \le \exp(-C(n/|I^c|)^\gamma)\}\bigr| > |I^c|\Bigr\} \le |I^c|/n,$$

where we take the product probability measure on $\mathcal{M}_{n,d} \times (\mathbb{C}^n, \gamma_n)$, i.e. assume that G and A_n are independent, and C is a universal positive constant.

We would like to note that using a better version of the structural theorem, namely [25, Theorem 4.1], one could prove a more general statement covering a wider range of d and $|I^c|$. Since the above statement is sufficient for our purposes, we prefer to avoid additional technicalities.

Theorem 4.1 states that any normal vector to E which is not very steep (in the above sense) necessarily has at least $(n/|I^c|)^{\gamma/2}$ levels of coordinates. Theorem 4.2 improves this by asserting that the uniform normal has as many as $n/|I^c|$ levels of coordinates. Also, as noticed in (3), there is a straightforward connection between the distance from a vector x to E and the inner product of x with $P_{E^\perp}G$. This connection together with Theorem 4.2 and the anti-concentration machinery developed in Section 5 allows us to get bounds on the intermediate singular values.

4.1. *Order statistics of uniform random normals*

Given $E \subset \mathbb{C}^n$, let $Y = Y(E) = (Y_1, \ldots, Y_n) = P_{E^\perp}G$. We also deal with linear combinations of vectors distributed as Y. Given $p \geq 1$ and $x \in \mathbb{C}^p$, denote

$$Y(x) = Y(x, p) := \sum_{j=1}^{p} x_j Y^{(j)}, \tag{6}$$

where $Y^{(j)}$, $j \leq p$, are independent copies of Y. In this subsection, we derive bounds on the order statistics of Y and $Y(x)$. We start with the following lemma.

Lemma 4.3 (Small ball probability for order statistics). *There exist absolute positive constants c and C such that the following holds. Let $E \subset \mathbb{C}^n$ be a fixed subspace of $\mathbb{C}^n$ with $m := \dim E^\perp \geq C$. Then*

$$\mathbb{P}\{Y^*_{cm} \leq cm/n\} \leq \exp(-cm).$$

Proof. Note that for every $i \leq n$ we have

$$Y^*_i \geq \min\{\|P_J Y\|_2/\sqrt{n} : J \subset [n],\ |J^c| = i\}.$$

Therefore,

$$\mathbb{P}\{Y^*_i \leq \tau\} \leq \binom{n}{i} \max_{|J^c|=i} \mathbb{P}\{\|P_J Y\|_2 \leq \tau\sqrt{n}\}.$$

Denoting $W = P_J P_{E^\perp}$, and applying a small ball probability estimate for Gaussian vectors ([21, Proposition 2.6], see also Remark 4.6 below), we have

$$\mathbb{P}\{\|P_J Y\|_2 \leq \tau\sqrt{n}\} \leq \left(\frac{\tau\sqrt{n}}{\|W\|_{\mathrm{HS}}}\right)^{c'\|W\|^2_{\mathrm{HS}}/\|W\|^2} \qquad \text{for } \tau \leq c_0\|W\|_{\mathrm{HS}}/\sqrt{n},$$

where $c_0, c' \in (0, 1)$ are universal constants. Note that $\|W\| \leq 1$ and

$$\|W\|^2_{\mathrm{HS}} = \mathrm{Tr}(P_J P_{E^\perp}) \geq m - i.$$

Therefore for $\tau < c_0\|W\|_{\mathrm{HS}}/\sqrt{n}$ and $i = c'm/4$ we have

$$\mathbb{P}\{Y_i^* \le \tau\} \le \left(\frac{en}{i}\right)^i \left(\tau\sqrt{\frac{n}{m-i}}\right)^{c'(m-i)} \le \left(\frac{4en}{c'm}\right)^{c'm/4} \left(\tau\sqrt{\frac{2n}{m}}\right)^{c'm/2}$$
$$\le \left(\frac{8n\tau}{\sqrt{c'}\,m}\right)^{c'm/2}.$$

The choice of $\tau = \sqrt{c'}\,m/(8en)$ and $c = \min\{\sqrt{c'}/(8e), c'/4\}$ completes the proof. □

As a consequence of Lemma 4.3, we obtain a bound for linear combinations.

Proposition 4.4 (Small ball for linear combinations). *Let n be a large enough integer, and $E \subset \mathbb{C}^n$ be a fixed subspace of $\mathbb{C}^n$ with $m := \dim E^\perp \ge n^{1/2}$. Given $p \le \sqrt{n}/\ln^2 n$ and $x \in \mathbb{C}^p$, let $Y(x) = Y(x, p)$ be defined as in (6). Then*

$$\mathbb{P}\Big\{\inf_{\|x\|_2=1} (Y(x))^*_{c_{4.4}m} \le c_{4.4}\, m/n\Big\} \le \exp(-c_{4.4}m),$$

where $c_{4.4} > 0$ is a universal constant.

Proof. Let $\mathcal{N}$ be a $c/(pn^2)$-net on the set of unit vectors in $\mathbb{C}^p$ with cardinality $|\mathcal{N}| \le (3pn^2/c)^{2p}$, where c is the constant from Lemma 4.3. Since for every unit vector x the vector $Y(x)$ has the same distribution as Y, Lemma 4.3 together with the union bound implies

$$\mathbb{P}\Big\{\inf_{x\in\mathcal{N}} (Y(x))^*_{cm} \le cm/n\Big\} \le |\mathcal{N}|\exp(-cm) \le \exp(-cm + 2p\ln(3pn^2/c)).$$

By the definition of $\mathcal{N}$, for any unit vector $x \in \mathbb{C}^p$ there is $y = y(x) \in \mathcal{N}$ such that $\|x - y\|_2 \le c/(pn^2)$, hence

$$\|Y(x) - Y(y)\|_2 = \Big\|\sum_{j=1}^p x_j Y^{(j)} - \sum_{j=1}^p y_j Y^{(j)}\Big\|_2 \le \sum_{j=1}^p |x_j - y_j|\,\|Y^{(j)}\|_2$$
$$\le \frac{c}{n^2}\max_{j\le p}\|Y^{(j)}\|_2.$$

This immediately implies that

$$(Y(x))^*_{cm} \ge (Y(y))^*_{cm} - \frac{c}{n^2}\max_{j\le p}\|Y^{(j)}\|_2.$$

Thus, we obtain a deterministic relation

$$\inf_{\|x\|_2=1} (Y(x))^*_{cm} \ge \inf_{x\in\mathcal{N}} (Y(x))^*_{cm} - \frac{c}{n^2}\max_{j\le p}\|Y^{(j)}\|_2.$$

This together with the rough bound $\mathbb{P}\{\max_{j\le p}\|Y^{(j)}\|\ge n\}<e^{-n}$ yields

$$\mathbb{P}\Big\{\inf_{\|x\|_2=1}(Y(x))^*_{cm}\le\frac{cm}{2n}\Big\}\le\mathbb{P}\Big\{\inf_{x\in\mathcal{N}}(Y(x))^*_{cm}\le\frac{cm}{2n}+\frac{c}{n}\Big\}+e^{-n}$$
$$\le\mathbb{P}\Big\{\inf_{x\in\mathcal{N}}(Y(x))^*_{cm}\le cm/n\Big\}+e^{-n}$$
$$\le\exp(-cm+2p\ln(3pn^2/c))+e^{-n}.$$

Since $m\ge\sqrt{n}\ge p\ln^2 n$, this completes the proof. □

We now pass to upper bounds.

Lemma 4.5 (Large deviations of order statistics). *Let E be as in Lemma 4.3. Then for every $i\le n/2$ and $\tau>0$ one has*

$$\mathbb{P}\{Y_i^*\ge C\sqrt{\ln(n/i)}\}\le(i/n)^i,$$

where $C>0$ is a universal constant.

Proof. Note that for a fixed $i\le n$ we have

$$Y_i^*\le\max\{\|P_JY\|_2/\sqrt{i}\,:\,J\subset[n],\ |J|=i\}.$$

Thus,

$$\mathbb{P}\{Y_i^*\ge\tau\}\le\binom{n}{i}\max_{|J|=i}\mathbb{P}\{\|WG\|_2\ge\tau\sqrt{i}\},$$

where $W=P_JP_{E^\perp}$. Using $\mathbb{E}\|WG\|_2^2=\mathrm{Tr}(W)\le i$, we get

$$\mathbb{P}\{\|WG\|_2\ge\tau\sqrt{i}\}\le\mathbb{P}\{\|WG\|_2^2\ge\mathbb{E}\,\|WG\|_2^2+(\tau^2-1)i\}.$$

Applying the Hanson–Wright inequality (see for example [35, Theorem 1.1, Remark 3.3] and Remark 4.6), we deduce that for any $\tau\ge\sqrt{2}$,

$$\mathbb{P}\{\|WG\|_2\ge\tau\sqrt{i}\}\le\exp(-c\tau^2 i)$$

for some absolute positive constant c. Taking $\tau=C\sqrt{\ln(n/i)}$ for a sufficiently large constant C completes the proof. □

Remark 4.6. The results of both [21] and [35] used in this section are formulated for real matrices and real random vectors. However, this is easily overcome by noticing that with any $n\times n$ complex matrix W and $x\in\mathbb{C}^n$, one may associate the $2n\times 2n$ matrix

$$\widetilde{W}=\begin{bmatrix}\mathrm{Re}(W)&-\mathrm{Im}(W)\\ \mathrm{Im}(W)&\mathrm{Re}(W)\end{bmatrix}\quad\text{and}\quad\widetilde{x}=\begin{bmatrix}\mathrm{Re}(x)\\ \mathrm{Im}(x)\end{bmatrix},$$

where Re and Im denote the real and imaginary parts. Now notice that $\|\widetilde{W}\widetilde{x}\|_2=\|Wx\|_2$ and thus $\|\widetilde{W}\|=\|W\|$. Moreover, one can check that $\|\widetilde{W}|_{\mathrm{HS}}^2=2\|W\|_{\mathrm{HS}}^2$. Therefore, one could apply the results of [21] and [35] to $\widetilde{W}$ and deduce the analogous results for the complex case.

As a consequence of Lemma 4.5 we obtain a bound for linear combinations.

Proposition 4.7. *Let n be a large enough integer, and E a fixed subspace of $\mathbb{C}^n$ with $m := \dim E^{\perp} \geq n^{1/2}$. Given $p \leq 2n^{1/4}$ and $x \in \mathbb{C}^p$, let $Y(x) = Y(x,p)$ be defined as in (6). Then*

$$\mathbb{P}\Big\{\sup_{\|x\|_2=1} (Y(x))_i^* \geq C_{4.7}\, p\sqrt{\ln(np/i)} \textit{ for some } i \leq m/4\Big\} \leq 8/\sqrt{n},$$

where $C_{4.7}$ is a universal positive constant.

Proof. Fix $i \leq n$ and a collection $\{z^1, \dots, z^p\}$ of n-dimensional vectors. Observe that for every subset $J \subset [n]$ of cardinality i, one has

$$\min_{j\in J} |(z^1 + \cdots + z^p)_j| \leq \min_{j\in J} \sum_{\ell=1}^{p} |z_j^\ell| \leq p \min_{j\in J} \max_{\ell\leq p} |z_j^\ell| =: pa.$$

For any $j \in J$ there is $\ell = \ell(j) \leq p$ such that $|z_j^\ell| \geq a$. Hence, by the pigeonhole principle, there is $\ell_0 \leq p$ such that $|z_j^{\ell_0}| \geq a$ for at least $|J|/p = i/p$ indices from J. Thus,

$$\min_{j\in J} |(z^1 + \cdots + z^p)_j| \leq p \max_{\ell\leq p} (z^\ell)^*_{\lceil i/p\rceil}.$$

Note that the right hand side does not depend on the choice of J, therefore

$$(z^1 + \cdots + z^p)_i^* \leq p \max_{\ell\leq p} (z^\ell)^*_{\lceil i/p\rceil}.$$

Returning to vectors $Y^{(1)}, \dots, Y^{(p)}$ we get, for any unit complex vector x,

$$(Y(x))_i^* \leq p \max_{\ell\leq p} (x_\ell Y^{(\ell)})^*_{\lceil i/p\rceil} \leq p \max_{\ell\leq p} (Y^{(\ell)})^*_{\lceil i/p\rceil}.$$

Recall that $m = \dim E^{\perp}$. Applying Lemma 4.5, and using the fact that the function $f(t) = (t/n)^t$ is decreasing on $(0, n/e)$, we obtain, for an appropriate absolute constant $C > 0$,

$$\begin{aligned}
&\mathbb{P}\Big\{\sup_{\|x\|_2=1} (Y(x))_i^* \geq Cp\sqrt{\ln(np/i)} \text{ for some } i \leq m/4\Big\}\\
&\leq \mathbb{P}\{(Y^{(\ell)})^*_{\lceil i/p\rceil} \geq C\sqrt{\ln(np/i)} \text{ for some } i \leq m/4 \text{ and } \ell \leq p\}\\
&\leq p\sum_{i=1}^{m/4}\left(\frac{\lceil i/p\rceil}{n}\right)^{\lceil i/p\rceil} \leq p^2 \sum_{j=1}^{m/(4p)} \left(\frac{j}{n}\right)^j \leq p^2\left(\frac{1}{n} + \frac{m}{4p}\,\frac{4}{n^2}\right) \leq \frac{2p^2}{n} \leq \frac{8}{\sqrt{n}}
\end{aligned}$$

provided that n is large enough. This completes the proof. □

4.2. Strongly correlated indices

Let E be a fixed subspace of $\mathbb{C}^n$ and let $Y = P_{E^\perp} G$ as before. Let $\alpha, \beta > 0$ be parameters. We say that a pair (i, j) of indices is (α, β)-*strongly correlated* (with respect to E) if

$$\mathbb{P}\{|Y_i - Y_j| \geq \alpha\} \leq \beta.$$

Next, we construct inductively a sequence $(U_\ell)_{\ell\geq 1} = (U_\ell(\alpha, \beta))_{\ell\geq 1}$ of (non-random) sets satisfying $\bigcup_{\ell\geq 1} U_\ell = [n]$, in the following way. At the first step, choose U_1 as the largest subset of $[n]$ such that there is $u_1 \in U_1$ such that (u_1, u) is (α, β)-strongly correlated for all $u \in U_1$. At the ℓ-th step, we define

$$U_\ell \subset \bar{U}_\ell := [n] \setminus (U_1 \cup \cdots \cup U_{\ell-1})$$

as the largest subset of $\bar{U}_\ell$ such that there is an index $u_\ell \in U_\ell$ such that (u_ℓ, u) is (α, β)-strongly correlated for all $u \in U_\ell$ (if $\bar{U}_\ell = \emptyset$ then we set $U_\ell = \emptyset$ as well). Further, it will be convenient for us to assume that the sequence $(U_\ell)_{\ell\geq 1}$ is uniquely defined. This can be achieved, for example, by defining a total order respecting cardinality on the set of all subsets of $[n]$ and, at each step above, choosing the largest admissible set with respect to that order. Observe that by the construction of U_ℓ's, the sequence $(|U_\ell|)_{\ell\geq 1}$ of cardinalities is non-increasing, and for every ℓ and all $i, j \in U_\ell$, the pair (i, j) is $(2\alpha, 2\beta)$-strongly correlated. Note that together with $(U_\ell)_{\ell\geq 1}$ we have also constructed a sequence $(u_\ell)_{\ell\geq 1}$ of indices, which can also be defined in a unique way.

Lemma 4.8. *Assume that a pair (i, j) is not (α, β)-strongly correlated for some $\alpha > 0$ and $\beta \in (0, 1/2]$. Then for every $s > 0$,*

$$\mathbb{P}\{|Y_i - Y_j| \leq \alpha s/\sqrt{\ln(1/\beta)}\} \leq s^2.$$

Proof. Set $\xi := Y_i - Y_j$. Observe that ξ is a centered complex Gaussian variable and denote its variance by σ^2. By the assumption of the lemma and by (5), we have

$$\beta \leq \mathbb{P}\{|\xi| > \alpha\} = e^{-\alpha^2/\sigma^2},$$

which implies that $\sigma \geq \alpha/\sqrt{\ln(1/\beta)}$. Since for every $s > 0$,

$$\mathbb{P}\{|\xi| \leq s\sigma\} = 1 - e^{-s^2} \leq s^2,$$

the desired result follows. □

The last lemma, combined with averaging arguments, implies the following lemma.

Lemma 4.9. *Let $\alpha > 0$ and $\beta \in (0, 1/2]$, and let $(U_\ell)_{\ell\geq 1}$ be defined as above. Let $k \geq 1$ and $b > 0$ be such that $|U_k| \leq b$. Then for every $s > 0$,*

$$\mathbb{P}\Big\{\exists \lambda \in \mathbb{C} : \Big|\Big\{j \in \bigcup_{\ell\geq k} U_\ell : |Y_j - \lambda| \leq \alpha s/\sqrt{4\ln(1/\beta)}\Big\}\Big| \geq 2b\Big\} \leq \frac{(sn)^2}{2b^2}.$$

Proof. Let $U = \bigcup_{\ell \geq k} U_\ell$ and for every $i \in U$ set

$$K_i = \{j \in U : (i, j) \text{ are } (\alpha, \beta)\text{-strongly correlated}\}.$$

By the construction of $(U_\ell)_{\ell \geq 1}$, for every $i \in U$ we have $|K_i| \leq b$. Applying Lemma 4.8 we obtain, for all $s > 0$ and $j \in U \setminus K_i$,

$$\mathbb{P}\{|Y_i - Y_j| \leq \alpha s/\sqrt{\ln(1/\beta)}\} \leq s^2.$$

Fix now $s > 0$ and for every $i \in U$ define the event

$$\mathcal{E}_i := \Big\{\big|\{j \in U : |Y_i - Y_j| \leq \alpha s/\sqrt{\ln(1/\beta)}\}\big| \geq 2b\Big\}.$$

As $|K_i| \leq b$, $\mathcal{E}_i$ is contained in the event $\{|\{j \in U \setminus K_i : |Y_i - Y_j| \leq \alpha s/\sqrt{\ln(1/\beta)}\}| \geq b\}$. Hence, applying Markov's inequality, we get

$$\mathbb{P}(\mathcal{E}_i) \leq \frac{1}{b} \sum_{j \in U \setminus K_i} \mathbb{P}\{|Y_i - Y_j| \leq \alpha s/\sqrt{\ln(1/\beta)}\} \leq s^2 n/b.$$

Next, given $\lambda \in \mathbb{C}$, denote

$$J_\lambda = \{j \in U : |Y_j - \lambda| \leq \alpha s/\sqrt{4\ln(1/\beta)}\}.$$

By the triangle inequality, for every $i, j \in J_\lambda$ we have $|Y_i - Y_j| \leq \alpha s/\sqrt{\ln(1/\beta)}$. Therefore, using Markov's inequality again, we observe

$$\begin{aligned}
&\mathbb{P}\{\exists \lambda \in \mathbb{C} : |J_\lambda| \geq 2b\} \\
&\qquad \leq \mathbb{P}\Big\{\big|\big\{i \in U : |\{j \in U : |Y_i - Y_j| \leq \alpha s/\sqrt{\ln(1/\beta)}\}| \geq 2b\big\}\big| \geq 2b\Big\} \\
&\qquad = \mathbb{P}\Big\{\sum_{i \in U} \chi_{\mathcal{E}_i} \geq 2b\Big\} \leq \frac{1}{2b} \sum_{i \in U} \mathbb{P}(\mathcal{E}_i) \leq \frac{n}{2b} \cdot \frac{s^2 n}{b}.
\end{aligned}$$

This completes the proof. □

We will use all properties of Gaussian vectors established previously to show that if the number of strongly correlated pairs associated to E is large, then we can construct an orthogonal vector to E satisfying none of the assumptions of Theorem 4.1, i.e., a normal vector to E which is neither very steep nor gradual with many levels.

Lemma 4.10. *There exist absolute positive constants C and $c_{4.10}$ such that the following holds. Let $\gamma > 0$ and $C_\gamma = 2 + \max(C, 2/\gamma)$. Let E be a fixed subspace of $\mathbb{C}^n$ with $m := \dim E^\perp \geq n^{3/4}$. Denote*

$$\alpha := \exp\left(-C_\gamma \left(\frac{n}{m}\right)^\gamma\right), \quad \beta := \frac{1}{4}\left(\frac{m}{4n}\right)^3, \quad V = 2n\left(\frac{m}{n}\right)^{\gamma/2}.$$

Let $(U_\ell)_{\ell \geq 1}$ be defined as above and $p \leq 2n/m$. Suppose that $|\bigcup_{\ell=1}^p U_\ell| > V$. Then there exists a vector $w \in \mathbb{C}^n$ orthogonal to E such that

$$\forall i \leq c_{4.10} m : \quad w_i^* \leq 0.9(n/i)^3 w^*_{c_{4.10} m}$$

and for some $\lambda \in \mathbb{C}$*,*

$$\left|\{i \le n : |w_i - \lambda| \le \exp(-2(n/m)^\gamma) w^*_{c_{4.10} m}\}\right| > n\left(\frac{m}{n}\right)^{\gamma/2}.$$

In other words, there exists a vector $z \in E^\perp$ *which is neither very steep nor gradual with many levels in the sense of Theorem 4.1.*

Proof. Let as before $Y^{(1)}, \dots, Y^{(p)}$ be independent copies of the vector Y and let $(u_\ell)_{\ell\ge1}$ be the sequence of indices which was defined together with $(U_\ell)_{\ell\ge1}$ at the beginning of Subsection 4.2. For any realization of $Y^{(1)}, \dots, Y^{(p)}$, let $\xi \in \mathbb{R}$ and $x = (x_1, \dots, x_p) \in \mathbb{C}^p$ be such that $\|x\|_2 = 1$ and

$$\forall \ell \le p: \quad \sum_{k=1}^{p} x_k Y^{(k)}_{u_\ell} = \xi.$$

The vector x and ξ can be taken as follows: if the matrix $M := (Y^{(k)}_{u_\ell})_{1\le\ell,k\le p}$ is of full rank, then take $x = y/\|y\|_2$ and $\xi = 1/\|y\|_2$, where $y = M^{-1}\mathbf{1}$, otherwise take any unit vector in the kernel of M and set $\xi = 0$. Denote $Z := Y(x) = \sum_{k=1}^{p} x_k Y^{(k)}$. Observe that deterministically

$$Z_{u_1} = \dots = Z_{u_p} = \xi.$$

We then have

$$\begin{aligned}
&\mathbb{P}\Big\{|Z_j - Z_{u_1}| \ge \alpha p \text{ for at least half of indices } j \in \bigcup_{\ell=1}^{p} U_\ell\Big\}\\
&\quad\le \sum_{\ell=1}^{p} \mathbb{P}\{|Z_j - Z_{u_\ell}| \ge \alpha p \text{ for at least half of indices } j \in U_\ell\}\\
&\quad\le \sum_{\ell=1}^{p}\sum_{k=1}^{p} \mathbb{P}\{|Y^{(k)}_j - Y^{(k)}_{u_\ell}| \ge \alpha \text{ for at least } |U_\ell|/(2p) \text{ indices } j \in U_\ell\}\\
&\quad\le \sum_{\ell=1}^{p}\sum_{k=1}^{p} \frac{2p}{|U_\ell|} \sum_{j\in U_\ell} \mathbb{P}\{|Y^{(k)}_j - Y^{(k)}_{u_\ell}| \ge \alpha\} \le 2p^3\beta,
\end{aligned}$$

where the first inequality follows by the union bound; the second one by a combination of the triangle inequality, the fact that $\|x\|_2 = 1$, the pigeonhole principle, and the union bound; the third one from Markov's inequality; and the last one from the definition of (α, β)-strongly correlated pairs. This together with the assumptions on p and β implies

$$\mathbb{P}\{\exists \lambda \in \mathbb{C} : |Z_j - \lambda| \le \alpha p \text{ for more than } V/2 \text{ indices } j \in [n]\} \ge 1 - 2p^3\beta \ge 1/2.$$

On the other hand, applying Propositions 4.4 and 4.7 we find that with probability at least $1 - 9/\sqrt{n}$ one has

$$Z^*_{c_{4.4} m} \ge c_{4.4} m/n \quad \text{and} \quad \forall i \le m/4: \quad Z^*_i \le C_{4.7}\, p\sqrt{\ln(np/i)}.$$

Intersecting the previous events we deduce that there exists a realization of Z (which will give the required vector w) satisfying

$$\left|\left\{i \le n : |Z_i - \lambda| \le \frac{\alpha p n}{c_{4.4} m} Z^*_{c_{4.4}m}\right\}\right| > V/2 = \left(\frac{m}{n}\right)^{\gamma/2} n$$

for some $\lambda \in \mathbb{C}$ and

$$\forall i \le cm : \quad Z_i^* \le \frac{C_{4.7} n\, p\sqrt{\ln(np/i)}}{c_{4.4} m} Z^*_{c_{4.4}m} \le 0.9 \left(\frac{n}{i}\right)^3 Z^*_{c_{4.4}m},$$

where $c > 0$ is a small enough absolute constant and where we have used $p \le 2n/m$. To complete the proof, we choose $c_{4.10} = \min(c, c_{4.4})$ and note that

$$\frac{\alpha p n}{c_{4.4} m} \le \exp(-2(n/m)^\gamma)$$

for an appropriate choice of the constant C (we may take $C = -\ln c_{4.4}$). □

4.3. *Proof of Theorem 4.2*

Let $d, n, z, \gamma, I, E, G, A_n$ be as in the statement of Theorem 4.2, and Y as above. We may assume without loss of generality that $\dim E = |I|$ a.s.; otherwise, we complement E to form a subspace E_0 of dimension $|I|$. In this case orthogonality to E_0 will imply orthogonality to E, therefore the proof below will not be affected. Let $m = |I^c|$. Denote

$$s = \frac{1}{2}\left(\frac{m}{n}\right)^{3/2}, \quad \alpha = \exp\left(-C_\gamma \left(\frac{n}{m}\right)^\gamma\right), \quad \beta = \frac{1}{4}\left(\frac{m}{4n}\right)^3, \quad V = 2n\left(\frac{m}{n}\right)^{\gamma/2},$$

where C_γ is the constant from Lemma 4.10. Let the sequence $(U_\ell)_{\ell\ge1}$ be constructed as above. If $|U_1| < m/2$ set $p = 0$, otherwise let p be the largest integer such that $|U_p| \ge m/2$. Since $(|U_\ell|)_{\ell\ge1}$ is non-increasing, we have $p \le 2n/m$. Notice that $(U_\ell)_{\ell\ge1}$ and p inherit randomness only from E. Let

$$J := \bigcup_{\ell=1}^{p} U_\ell \subset [n]$$

(if $p = 0$ then $J = \emptyset$). Consider the event $\mathcal{E} := \{|J| > V\}$ (depending only on E). Lemma 4.10 implies that $\mathcal{E} \subset \mathcal{E}_1^c$, where $\mathcal{E}_1$ denotes the event appearing in Theorem 4.1. Denoting by $\mathcal{E}_2$ the event of Theorem 4.2 and applying Theorem 4.1, we get

$$\mathbb{P}(\mathcal{E}_2) \le \mathbb{P}(\mathcal{E}_2 \cap \mathcal{E}^c) + \mathbb{P}(\mathcal{E}) \le \mathbb{P}(\mathcal{E}_2 \cap \mathcal{E}^c) + 1/n.$$

Now note that once in $\mathcal{E}^c$, we have $|J| \le V$. Therefore, since $J^c = \bigcup_{\ell=p+1}^n U_\ell$,

$$\mathbb{P}(\mathcal{E}_2 \cap \mathcal{E}^c) \le \mathbb{P}\left\{\exists\lambda \in \mathbb{C} : \left|\{j \in J^c : |Y_j - \lambda| \le \exp(-C(n/m)^\gamma)\}\right| \ge m\right\}.$$

Since $n/m \geq d^3$ and d is large enough, there exists a sufficiently large absolute constant C satisfying

$$\exp(-C(n/m)^{\gamma}) \leq \alpha s/\sqrt{4\ln(1/\beta)}.$$

Applying Lemma 4.9 with $k = p+1$ and $b = m/2$ (then $U_k \leq b$), we obtain

$$\mathbb{P}(\mathcal{E}_2 \cap \mathcal{E}^c) \leq \mathbb{P}\big\{\exists\lambda \in \mathbb{C} : \big|\{j \in J^c : |Y_j - \lambda| \leq \alpha s/\sqrt{4\ln(1/\beta)}\}\big| \geq m\big\}$$
$$\leq \frac{(sn)^2}{2(m/2)^2} = \frac{m}{2n}.$$

Since $1/n \leq m/(2n)$, this completes the proof. □

5. Intermediate singular values

The goal of this section is to establish the bounds on the intermediate singular values stated in the introduction (see Theorem 1.6). We first briefly show how to derive the estimates on the singular values far from the lower edge of the spectrum. As mentioned in the introduction, these follow from the work of Cook [12]. The majority of the section is devoted to the complementary regime, that is, to bounding the singular values closer to the edge.

5.1. Higher end of the spectrum

Following the comparison strategy described in the introduction, the following proposition was proved by Cook [12, Proposition 7.3].

Proposition 5.1 (Anti-concentration of the spectrum). *Assume $d = o(\sqrt{n})$ and $d \to \infty$ as $n \to \infty$. Then with probability at least $1 - C_0\exp(-n)$, for all $\eta \in (0, 1]$ one has*

$$\nu_{B_z}([0,\eta]) < C_0(\eta + d^{-1/48}),$$

where C_0 is an absolute positive constant.

Based on this, it is easy to derive Proposition 1.5.

Proof of Proposition 1.5. For $k \leq n - 2C' nd^{-1/48}$, set $\eta_k := (n-k)/(2C'n) \geq d^{-1/48}$. Proposition 5.1 applied with $\eta = \eta_k$ implies that with probability $1 - \exp(-n)$, for any $k \leq n - 2C_0 nd^{-1/48}$ the number of singular values smaller than η_k is less than $2C_0\eta_k n$. This yields that $s_k = s_{n-2C_0\eta_i n} \geq \eta_k$. Setting $C = 2C'$ and $c = 1/(2C_0)$ we complete the proof. □

Remark 5.2. Proposition 5.1 is Proposition 7.3 from [12] stated there for $d \geq \ln^4 n$. Let us indicate the changes needed to cover our range of interest, that is, $d = o(\sqrt{n})$ and $d \to \infty$ as $n \to \infty$ (without the restriction $d \geq \ln^4 n$).

The proof of Proposition 7.3 in [12] combines three lemmas: Lemmas 8.1, 8.2, and 8.4 there. Lemma 8.4 establishes bounds on the intermediate singular values for shifts of

Gaussian matrices and does not demand d to be polylogarithmic in n. Lemma 8.2 compares the expectation of the Stieltjes transforms of the Bernoulli model (with parameter d/n) with its Gaussian counterpart. Here as well, no restriction on d is required and one only needs that $d \to \infty$ as $n \to \infty$ for the approximation to be effective.

The last piece of the procedure, Lemma 8.1, compares the uniform d-regular model with the Bernoulli matrix. Its proof uses a general concentration inequality for linear eigenvalue statistics of Hermitian random matrices [12, Lemma 9.1] and an estimate of the probability that a Bernoulli matrix with parameter d/n is d-regular [12, Lemma 9.2]. The latter indeed requires $d \geq \ln^4 n$ as stated, since it covers also large values of d. Since in our regime we suppose that $d = o(\sqrt{n})$, we could replace the estimate of Lemma 9.2 by a bound proved by McKay and Wang [30], which is also mentioned in [12, Remark 9.3]. This implies the validity of Lemma 8.1 for any $d = o(\sqrt{n})$ with the term $\exp(-O(d^{2/3} n \ln n))$ in the probability bound replaced with $\exp(-O(n \ln d))$. This affects the proof of Proposition 7.3 in a trivial way, as one would change the choice of ε there to be $(\ln d/d)^{1/4}$ and carry out the remaining part of the proof in exactly the same way as before.

Note that the same change in Lemma 8.1 is sufficient to extend the proof of Proposition 7.2 in [12] to our range of d, which gives our Proposition 1.3.

5.2. Lower end of the spectrum. Proof of Theorem 1.6

We first relate the intermediate singular values to separation estimates between the rows of the matrix. As an important technical ingredient, we use the so-called negative second moment identity, which was employed earlier in papers on the circular law (see [39, 12]). We note that one could also use the restricted invertibility principle instead (see [31]).

Lemma 5.3. *Let B be an $n \times n$ complex random matrix with distribution invariant under permutation of rows. Let $m \leq n$ be positive integers and $\rho, \delta > 0$ be such that*

$$\mathbb{P}\Big\{\mathrm{dist}\big(R_m(B), \mathrm{span}\,\{R_j(B)\}_{j \leq m-1}\big) < \rho\Big\} \leq \delta.$$

Then for every $1 \leq L \leq \frac{1}{2\delta}$,

$$\mathbb{P}\{s_{(1-2L\,\delta)m}(B) \geq \rho\sqrt{L\,\delta}\} \geq 1 - 1/L.$$

Proof. For each $i \leq m$, let χ_i be the characteristic function of the event

$$\Big\{\mathrm{dist}\big(R_i(B), \mathrm{span}\,\{R_j(B)\}_{j \in [m]\setminus\{i\}}\big) < \rho\Big\}.$$

By the conditions of the lemma (including the permutation invariance), we have $\mathbb{E}\,\chi_i \leq \delta$, hence, by Markov's inequality, the event

$$\mathcal{E} := \Big\{\sum_{i=1}^{m} \chi_i > L\delta m\Big\}$$

has probability at most $1/L$. Conditioning on the complement $\mathcal{E}^c$, we can find a set of indices $I \subset [m]$ of cardinality at least $m - L\delta m$ such that for every $i \in I$ one has

$$\operatorname{dist}\big(R_i(B), \operatorname{span}\{R_j(B)\}_{j\in[m]\setminus\{i\}}\big) \geq \rho.$$

Passing to the $|I| \times n$ submatrix B' with rows $R_j(B)$, $j \in I$, we obviously have, for $i \leq |I|$,

$$\operatorname{dist}\big(R_i(B'), \operatorname{span}\{R_j(B')\}_{j\neq i}\big) \geq \rho.$$

Applying the negative second moment identity (see, e.g., [39, Lemma A.4]), we obtain

$$\sum_{i=1}^{|I|} s_i(B')^{-2} = \sum_{i=1}^{|I|} \operatorname{dist}\big(R_i(B'), \operatorname{span}\{R_j(B')\}_{j\neq i}\big)^{-2} \leq |I|\rho^{-2}.$$

Therefore,

$$L\delta m s_{m-2L\,\delta m}(B')^{-2} \leq \sum_{j=m-2L\delta m}^{m-L\delta m} s_j(B')^{-2} \leq \sum_{j=1}^{|I|} s_j(B')^{-2} \leq m\rho^{-2},$$

which implies

$$s_{m-2L\delta m}(B') \geq \rho\sqrt{L\delta}.$$

Clearly, we deterministically have

$$s_{m-2L\delta m}(B) \geq s_{m-2L\delta m}(B').$$

Thus, $s_{m-2L\,\delta m}(B) \geq \rho\sqrt{L\,\delta}$ everywhere on $\mathcal{E}^c$, which yields the desired result. □

We now provide bounds on the distances under consideration.

Lemma 5.4. *Let d, n be large enough integers such that $d \leq \ln^{96} n$, let $z \in \mathbb{C}$ be such that $|z| \leq \sqrt{d}\ln d$, and set $\gamma = 1/288$. Let σ_n denote the uniform random permutation of $[n]$ independent of A_n and, as before, $B_z = d^{-1/2}A_n - z\,\mathrm{Id}$. Then for every i satisfying*

$$2n/\ln^{1/\gamma} n \leq n - i \leq d^{-3}n$$

one has

$$\mathbb{P}\left\{\operatorname{dist}\big(R_{\sigma_n(i)}(B_z), \operatorname{span}\{R_{\sigma_n(j)}(B_z)\}_{j\leq i-1}\big) < \exp\left(-C\left(\frac{n}{n-i}\right)^{\gamma}\right)\right\} \leq C\frac{n-i}{n},$$

where C is a positive universal constant.

Since the proof of Lemma 5.4 requires developing certain anti-concentration tools, we postpone it and turn to the proof of Theorem 1.6.

Proof of Theorem 1.6. Let i satisfy $2n/\ln^{1/\gamma} n \le n - i \le d^{-3}n$ and let σ_n, B_z, C be as in Lemma 5.4. Denote $\varepsilon = (n-i)/n$. Then

$$\mathbb{P}\big\{\mathrm{dist}\big(R_{\sigma_n(i)}(B_z), \mathrm{span}\,\{R_{\sigma_n(j)}(B_z)\}_{j\le i-1}\big) < \exp(-C\varepsilon^{-\gamma})\big\} \le C\varepsilon.$$

Let B be the matrix obtained from B_z by permuting its rows according to σ_n. Then B has the same singular values as B_z and the distribution of B is invariant under permutation of rows. Therefore applying Lemma 5.3 with

$$\rho = \rho(\varepsilon) = \exp(-C\varepsilon^{-\gamma}), \qquad \delta = \delta(\varepsilon) = C\varepsilon \quad \text{and} \quad L = \frac{1}{2\sqrt{C\delta}}$$

(then $(1-\sqrt{\varepsilon})i \le (1-2L\delta)i$), we obtain

$$\mathbb{P}\{s_{(1-\sqrt{\varepsilon})i}(B) \ge (\varepsilon/4)^{1/4}\exp(-C\varepsilon^{-\gamma})\} \ge 1 - 2C\sqrt{\varepsilon}.$$

Using $(1-\sqrt{\varepsilon})i \ge (1-2\sqrt{\varepsilon})n$ and $(\varepsilon/4)^{1/4} \ge \exp(\varepsilon^{-\gamma})$ when d is large enough (recall $i \ge n - n/d^3$), we deduce that for an appropriate absolute constant $C_1 > 0$,

$$\mathbb{P}\{s_{(1-2\sqrt{\varepsilon})n}(B) \ge \exp(-C_1\varepsilon^{-\gamma})\} \ge 1 - 2C\sqrt{\varepsilon}.$$

Writing $k = (1-2\sqrt{\varepsilon})n$ (with a slight adjustment to make it an integer), so that $\varepsilon = \frac{(n-k)^2}{(2n)^2}$, we clearly have

$$n - 2d^{-3/2}n \le k \le n - 2\sqrt{2}n/\ln^{144} n.$$

Using the facts that σ_n is independent of A_n, that B and B_z have the same singular values, and that $(s_i)_i$ is increasing, we obtain the desired result. □

5.3. *Anti-concentration*

To state the main result of the subsection, we need to define a special distribution on the set of n-dimensional $0/1$ vectors. For any matrix $M \in \mathcal{M}_{n,d}$ and for any non-empty subset $K \subset [n]$ denote

$$\mathbb{M}_{M,K} := \{M' \in \mathcal{M}_{n,d} : R_i(M') = R_i(M) \text{ for all } i \notin K\}.$$

Now, fix $J \subset [n]$ of cardinality at least $n/2$. In this section, we denote by $\mathcal{I} = \mathcal{I}(J)$ a uniform random subset of J of cardinality $\lfloor n^{1/4} \rfloor$. Next, fix an index $u \in [n] \setminus J$ and a matrix $M \in \mathcal{M}_{n,d}$ and define a random vector $X_{M,J,u}$ via its conditional distribution with respect to $\mathcal{I}$; namely, we postulate that, conditioned on a realization I_0 of the set $\mathcal{I}$, the vector $X_{M,J,u}$ takes values in the set

$$Q_{M,J,u} := \{R_u(M') : M' \in \mathbb{M}_{M,I_0\cup\{u\}}\}$$

and

$$\forall x \in Q_{M,J,u} : \quad \mathbb{P}\{X_{M,J,u} = x \mid \mathcal{I} = I_0\} = \frac{|\{M' \in \mathbb{M}_{M,I_0\cup\{u\}} : R_u(M') = x\}|}{|\mathbb{M}_{M,I_0\cup\{u\}}|}.$$

Proposition 5.5. *Let d, n be large enough positive integers such that $d \le n^{1/8}$. Let J be a subset of $[n]$ of cardinality at least $n/2$, $u \in [n] \setminus J$, and let M be a fixed matrix in $\mathcal{M}_{n,d}$. Further, let $\delta, \rho > 0$, and let y be a fixed vector in $\mathbb{C}^n$ such that for some subset $\widetilde{J} \subset [n]$ we have*

$$\forall \lambda \in \mathbb{C}: \quad \left|\{j \in [n] \setminus \widetilde{J}: |y_j - \lambda| \le \rho\}\right| \le \delta n.$$

Then

$$\forall \lambda \in \mathbb{C}: \quad \mathbb{P}\{|\langle y, X_{M,J,u}\rangle - \lambda| \le \rho/4\} \le (8|\widetilde{J}|/n)^d + 144\delta + n^{-1/10}.$$

To prove this proposition we need several lemmas.

Lemma 5.6. *Let d, n be large enough positive integers such that $d \le n^{1/8}$ and let M be a fixed matrix in $\mathcal{M}_{n,d}$. Further, let $J \subset [n]$ be a fixed subset of cardinality at least $n/2$, $u \in [n] \setminus J$ and $\mathcal{I} = \mathcal{I}(J)$. Then with probability at least $1 - 2n^{-1/4}$ the supports of the rows $R_i(M)$, $i \in \mathcal{I} \cup \{u\}$, are pairwise disjoint.*

Proof. Denote by $Q \subset (J \cup \{u\}) \times (J \cup \{u\})$ the subset of all pairs (i, j) such that

$$\operatorname{supp} R_i(M) \cap \operatorname{supp} R_j(M) \neq \emptyset.$$

By d-regularity we observe that for any $i \in J \cup \{u\}$ there are less than d^2 indices j with $(i, j) \in Q$. Thus, $|Q| \le d^2(|J| + 1)$. On the other hand, an easy computation shows that for any $(i_1, i_2) \in Q$ with $i_1 \neq i_2$, the probability that both i_1 and i_2 belong to $\mathcal{I}$ is

$$\binom{|J| + 1 - 2}{\lfloor n^{1/4} \rfloor - 2} \binom{|J| + 1}{\lfloor n^{1/4} \rfloor}^{-1} = \frac{\lfloor n^{1/4} \rfloor (\lfloor n^{1/4} \rfloor - 1)}{|J|(|J| + 1)}.$$

Hence,

$$\mathbb{P}\{\mathcal{I} \text{ contains a disjoint pair in } Q\} \le |Q|\sqrt{n}/(|J|(|J| + 1)) \le d^2\sqrt{n}/|J|.$$

The assumptions on $|J|$ and d imply the result. □

Lemma 5.7. *Let $d \le n$ be large enough positive integers and M be a fixed matrix in $\mathcal{M}_{n,d}$. Further, let $J \subset [n]$ be a subset of cardinality at least $n/2$, and let $\mathcal{I} = \mathcal{I}(J)$. Then for every subset $L \subset [n]$ with probability at least $1 - 1/n^2$ we have*

$$\left| \Big(\bigcup_{i \in \mathcal{I}} \operatorname{supp} R_i(M) \Big) \cap L \right| \le 14 d^2 \ln n + 4 d n^{-3/4} |L|.$$

Proof. Without loss of generality we assume that $|L| \ge d^2$. Fix a partition $(L_k)_{k=1}^{d^2}$ of L such that for every $k \le d^2$ and for every $i \neq j \in L_k$ there is no row of M such that i, j are simultaneously contained in its support. Such a partition can be constructed as follows: take a graph Γ on L without loops such that $i \neq j \in L$ are connected by an edge whenever there is a row of M whose support contains both i and j. The d-regularity immediately implies that the maximum vertex degree of this graph is strictly less than d^2

(in fact, not greater than $d(d-1)$). Therefore, by Brook's theorem, the chromatic number of Γ does not exceed d^2, which justifies the number of sets in the required partition of L.

Further, let $\widetilde{\mathcal{I}}$ be a random subset of J such that each index $i \in J$ is included into $\widetilde{\mathcal{I}}$ with probability $\lfloor n^{1/4} \rfloor / |J|$ independently of the others. Fix $k \le d^2$ for a moment. For any $i \in L_k$, let η_i^k be the indicator function of the event that

$$i \in \bigcup_{j \in \widetilde{\mathcal{I}}} \operatorname{supp} R_j(M).$$

Note that by our construction $(\eta_i^k)_{i \in L_k}$ are jointly independent, and for all $i \in L_k$,

$$\mathbb{E}\, \eta_i^k = \mathbb{E}\, (\eta_i^k)^2 = \mathbb{P}\{\eta_i^k = 1\} \le d n^{1/4}/|J| := \delta.$$

Applying Bernstein's inequality with $t = \delta |L_k| + 14 \ln n$, we obtain

$$\mathbb{P}\Big\{\Big|L_k \cap \bigcup_{j \in \widetilde{\mathcal{I}}} \operatorname{supp} R_j(M)\Big| \ge 2\delta|L_k| + 14 \ln n\Big\} \le \mathbb{P}\Big\{\sum_{i \in L_k} (\eta_i^k - \mathbb{E}\, \eta_i^k) \ge t\Big\}$$
$$\le \exp\Big(-\frac{3t^2}{2(t + 3\delta|L_k|)}\Big) \le \exp\Big(-\frac{3t}{8}\Big) \le n^{-5}.$$

Then the union bound implies that with probability at least $1 - d^2 n^{-5}$ one has

$$\Big|L \cap \bigcup_{j \in \widetilde{\mathcal{I}}} \operatorname{supp} R_j(M)\Big| \le \sum_{k=1}^{d^2} \Big(14 \ln n + \frac{2dn^{1/4}|L_k|}{|J|}\Big) = 14 d^2 \ln n + \frac{2dn^{1/4}|L|}{|J|}.$$

Finally, note that the cardinality of $\widetilde{\mathcal{I}}$ equals exactly $m := \lfloor n^{1/4} \rfloor$ with probability

$$\binom{|J|}{m} \Big(\frac{m}{|J|}\Big)^m \Big(1 - \frac{m}{|J|}\Big)^{|J|-m} \ge \Big(1 - \frac{m}{|J|}\Big)^{|J|} \ge \exp(-2m) \ge n^{-1/4}.$$

Therefore

$$\mathbb{P}\Big\{\Big|L \cap \bigcup_{j \in \widetilde{\mathcal{I}}} \operatorname{supp} R_j(M)\Big| \le 14 d^2 \ln n + \frac{2dn^{1/4}|L|}{|J|} \;\Big|\; |\widetilde{\mathcal{I}}| = \lfloor n^{1/4} \rfloor\Big\}$$
$$\ge 1 - d^2 n^{-4} \ge 1 - \frac{1}{n^2},$$

which implies the desired result, since $|J| \ge n/2$. □

Lemma 5.8. *Let $d < n$ be positive integers. Let M be a fixed matrix in $\mathcal{M}_{n,d}$, $J \subset [n]$ be a subset of cardinality at least $n/2$, and $\mathcal{I} = \mathcal{I}(J)$. Let $u \in [n] \setminus J$ and let $I_0 \subset J$ of size $\lfloor n^{1/4} \rfloor$ be such that the supports of the rows $R_i(M)$, $i \in I_0 \cup \{u\}$, are pairwise disjoint. Then, conditioned on $\mathcal{I} = I_0$, the support of the random vector $X_{M,J,u}$ is a uniformly distributed d-subset of*

$$S := \bigcup_{i \in I_0 \cup \{u\}} \operatorname{supp} R_i(M).$$

Proof. We first show that for any two 0/1 vectors x, y satisfying

$$\operatorname{supp} x, \operatorname{supp} y \subset S, \quad |\operatorname{supp} x| = |\operatorname{supp} y| = d, \quad \text{and} \quad |\operatorname{supp} x \setminus \operatorname{supp} y| = 1,$$

the sets

$$S_x := \{M' \in \mathbb{M}_{M,I_0\cup\{u\}} : R_u(M') = x\} \quad \text{and} \quad S_y := \{M'' \in \mathbb{M}_{M,I_0\cup\{u\}} : R_u(M'') = y\}$$

have the same cardinality. Without loss of generality, assume that $x_1 = y_2 = 1$ and $x_2 = y_1 = 0$. Then $\{1, 2\} \subset S$. For every matrix $M' \in S_x$ we construct a matrix $M'' \in S_y$ as follows. Since $\{1, 2\} \subset S$ and the rows indexed by $I_0 \cup \{u\}$ are pairwise disjoint, there exists a unique index $i = i(M') \in I_0 \cup \{u\}$ such that $M'_{i,1} = 0$ and $M'_{i,2} = 1$. Let M'' be obtained by performing the simple switching operation on M' which interchanges the entries $M'_{u,1}$ and $M'_{u,2}$ with $M'_{i,1}$ and $M'_{i,2}$ respectively. Clearly, $M'' \in S_y$, and it is not difficult to see that the constructed mapping is injective. Therefore, $|S_x| \le |S_y|$. Reversing the argument, we find that $|S_x| = |S_y|$. Since for every 0/1 vector z satisfying supp $z \subset S$ and $|\text{supp } z| = d$ one can construct a sequence of vectors $x_0 = x, x_1, \ldots, x_k = z$ with $\operatorname{supp} x_i \subset S$, $|\operatorname{supp} x_i| = d$, and such that two vectors x_{i-1}, x_i differ in exactly two coordinates for every $1 < i \le k$, we obtain $|S_x| = |S_z|$. Thus

$$\mathbb{P}\{X_{M,J,u} = x \mid \mathcal{I} = I_0\} = \mathbb{P}\{X_{M,J,u} = z \mid \mathcal{I} = I_0\},$$

which means that, conditioned on $\mathcal{I} = I_0$, the support of the random vector $X_{M,J,u}$ is uniformly distributed on the set of d-subsets of S. □

Lemma 5.9 (Coupling). *Let d, n be large enough positive integers such that $d \le n^{1/8}$ and let M be a fixed matrix in $\mathcal{M}_{n,d}$. Let $J \subset [n]$ be a subset of cardinality at least $n/2$ and $\mathcal{I} = \mathcal{I}(J)$. Assume that $u \in [n] \setminus J$ and let $I_0 \subset J$ be of size $\lfloor n^{1/4} \rfloor$ and such that the supports of rows $R_i(M)$, $i \in I_0 \cup \{u\}$, are pairwise disjoint. Let $\xi_1, \ldots, \xi_d$ be i.i.d. random variables uniformly distributed on*

$$S := \bigcup_{i\in I_0\cup\{u\}} \operatorname{supp} R_i(M), \quad \textit{and set} \quad Y_\xi := \sum_{i=1}^{d} e_{\xi_i}.$$

Then there is a coupling (X, Y_ξ), with X distributed as $X_{M,J,u}$, such that

$$\mathbb{P}\{X = Y_\xi \mid \mathcal{I} = I_0\} \ge 1 - n^{-1/8}.$$

Proof. Note that, conditioned on the event

$$\mathcal{E} := \{\xi_i \ne \xi_j \text{ for all } i \ne j\},$$

the random set $X := \{\xi_1, \ldots, \xi_d\}$ is a uniformly distributed d-subset of S. Therefore, by Lemma 5.8, the distribution of $X_{M,J,u}$ conditioned on $I = I_0$ agrees with the distribution of Y_ξ conditioned on $\mathcal{E}$. Since $R_i(M)$, $i \in I_0 \cup \{u\}$, are pairwise disjoint, we have $|S| \ge dn^{1/4}$, hence

$$\mathbb{P}\{\xi_i = \xi_j \text{ for some } i \ne j\} \le d^2\, \mathbb{P}\{\xi_1 = \xi_2\} \le d^2/|S| \le n^{-1/8}.$$

This implies the desired result. □

Lemma 5.10. *Let $\delta, \rho > 0$, $\widetilde{J} \subset [n]$, and y be a fixed vector in $\mathbb{C}^n$ such that*

$$\forall \lambda \in \mathbb{C}: \quad \left|\{j \in [n] \setminus \widetilde{J} : |y_j - \lambda| \le \rho\}\right| \le \delta n.$$

Then there exists a partition $(U_{ij})_{i\le 9,\, j\le n}$ of $[n] \setminus \widetilde{J}$ such that $|U_{ij}| \le \delta n$ for all $i \le 9$, $j \le n$, and

$$\forall i \le 9\, \forall j \ne j' \in [n]\, \forall s \in U_{ij}\, \forall s' \in U_{ij'}: \quad |y_s - y_{s'}| \ge \rho.$$

Proof. We identify $\mathbb{C}$ with $\mathbb{R}^2$. Consider the following nine points:

$$a_1 = (0,0), \quad a_2 = (1,0), \quad a_3 = (2,0), \quad a_4 = (0,1), \quad a_5 = (0,2),$$
$$a_6 = (1,1), \quad a_7 = (2,1), \quad a_8 = (1,2), \quad a_9 = (2,2).$$

For $i \le 9$, set

$$\mathcal{V}_i := \rho(a_i + 3\mathbb{Z} \times 3\mathbb{Z}).$$

Note that any two points in $\mathcal{V}_i$ are at distance at least 3ρ from each other and the union of the $\mathcal{V}_i$'s is $\mathbb{C}$. We first construct a partition $(\mathcal{V}_{ij})_{i\le 9,\, j\in\mathbb{Z}^2}$ of the complex plane as follows. First, set $\mathcal{V}_{1j}$ to be the Euclidean balls of radius ρ centered at $\rho(a_1 + 3j) \in \mathcal{V}_1$. Observe that the balls are necessarily pairwise disjoint. Further, assuming that $\mathcal{V}_{\ell j}$, $\ell < i$, $j \in \mathbb{Z}^2$, are constructed (for some $1 < i \le 9$), define $\mathcal{V}_{ij}$ as the set difference of the Euclidean ball of radius ρ centered at $\rho(a_i + 3j) \in \mathcal{V}_i$ and the union of $\mathcal{V}_{\ell j'}$, $\ell < i$, $j' \in \mathbb{Z}^2$. Then $(\mathcal{V}_{ij})_{i\le 9,\, j\in\mathbb{Z}^2}$ is a partition and moreover, for any $i \le 9$ and any $j \ne j' \in \mathbb{Z}^2$, one has $|x - x'| \ge \rho$ for any $x \in \mathcal{V}_{ij}$, $x' \in \mathcal{V}_{ij'}$. Indeed, this follows by an application of the triangle inequality together with the fact that the centers of these two balls are at distance at least 3ρ. Therefore, one can partition the coordinates of y by intersecting the above partition of $\mathbb{C}$ with $\{y_i\}_{i\le n}$. This naturally defines a partition of $[n] \setminus \widetilde{J}$ by setting the sets of the partition to be the indices of the corresponding coordinates of y. The assumption on y implies that each set in the partition contains at most δn elements. □

Proof of Proposition 5.5. Fix $\lambda \in \mathbb{C}$. Then

$$\mathbb{P}\{|\langle y, X_{M,J,u}\rangle - \lambda| \le \rho/4\} \le \sum_{\substack{I_0\subset J\\ |I_0|=\lfloor n^{1/4}\rfloor}} \mathbb{P}\{|\langle y, X_{M,J,u}\rangle - \lambda| \le \rho/4 \mid \mathcal{I} = I_0\}\, \mathbb{P}\{\mathcal{I} = I_0\}.$$

Let $(U_{ij})_{i\le 9,\, j\le n}$ be the partition of $[n] \setminus \widetilde{J}$ given by Lemma 5.10, in particular $|U_{ij}| \le \delta n$ for all i, j. Let T be the collection of all subsets I_0 of J of cardinality $\lfloor n^{1/4}\rfloor$ satisfying the following three conditions:

$$\text{the rows } R_i(M) \text{ for } i \in I_0 \cup \{u\} \text{ are pairwise disjoint;} \tag{7}$$

$$\left|\widetilde{J} \cap \bigcup_{i\in I_0} \operatorname{supp} R_i(M)\right| \le 14d^2 \ln n + 4dn^{-3/4}|\widetilde{J}|; \tag{8}$$

$$\left|U_{ij} \cap \bigcup_{i\in I_0} \operatorname{supp} R_i(M)\right| \le 14d^2 \ln n + 4\delta dn^{1/4}. \tag{9}$$

By Lemmas 5.6 and 5.7 and the union bound, the event $\{\mathcal{I} \in T\}$ has probability at least $1 - 3n^{-1/4}$. Thus, we have

$$\mathbb{P}\left\{|\langle y, X_{M,J,u}\rangle - \lambda| \le \frac{\rho}{4}\right\} \le \sum_{I_0 \in T} \mathbb{P}\left\{|\langle y, X_{M,J,u}\rangle - \lambda| \le \frac{\rho}{4} \,\middle|\, \mathcal{I} = I_0\right\} \mathbb{P}\{\mathcal{I} = I_0\} + \frac{3}{n^{1/4}}.$$

Further, fix any I_0 in T. Let S, $\xi_1, \dots, \xi_d$, and Y_ξ be defined in Lemma 5.9. Note that by (7), $|S| \ge dn^{1/4}$. Lemma 5.9 implies

$$\mathbb{P}\{|\langle y, X_{M,J,u}\rangle - \lambda| \le \rho/4 \mid \mathcal{I} = I_0\} \le \mathbb{P}\{|\langle y, Y_\xi\rangle - \lambda| \le \rho/4\} + n^{-1/8}.$$

Denote

$$S_0 := \bigcup_{i \in I_0} \operatorname{supp} R_i(M) \setminus \widetilde{J}, \qquad S_1 := \Big(\widetilde{J} \cap \bigcup_{i \in I_0} \operatorname{supp} R_i(M)\Big) \cup \operatorname{supp} R_u(M),$$

and $\xi = \{\xi_1, \dots, \xi_d\}$. Note that by properties (7) and (8) and assuming that $|\widetilde{J}| \le n/8$ (otherwise the bound for the probability in Proposition 5.5 is trivial), one has

$$|S| \ge dn^{1/4}, \qquad \frac{|S_1|}{|S|} \le \frac{15d \ln n}{n^{1/4}} + \frac{4|\widetilde{J}|}{n} \le \frac{3}{4}, \qquad \frac{|S_0|}{|S|} = 1 - \frac{|S_1|}{|S|} \ge \frac{1}{4}. \tag{10}$$

Consider the events

$$\mathcal{E}_1 := \{\xi \cap S_0 = \emptyset\} = \{\xi \subset S_1\} \quad \text{and} \quad \mathcal{E}_2 := \{\xi \cap S_0 \ne \emptyset\}.$$

Using property (10) and independence of ξ_i's, we clearly have

$$\mathbb{P}(\mathcal{E}_1) = (|S_1|/|S|)^d \le \left(\frac{30d \ln n}{n^{1/4}}\right)^d + \left(\frac{8|\widetilde{J}|}{n}\right)^d.$$

To estimate the remaining probability we split $\mathcal{E}_2$ into the disjoint union of the events

$$\mathcal{E}_W := \{\xi_i \in S_0 \text{ for all } i \in W \text{ and } \xi_i \notin S_0 \text{ for all } i \notin W\},$$

where W runs over all non-empty subsets of $[d]$. Then

$$\mathbb{P}\{|\langle y, Y_\xi\rangle - \lambda| \le \rho/4 \mid E_2\} \le \sup_W \mathbb{P}\{|\langle y, Y_\xi\rangle - \lambda| \le \rho/4 \mid \mathcal{E}_W\}.$$

Fix a non-empty $W \subset [d]$ and $m \in W$. Since ξ_i's are i.i.d., we observe that

$$\begin{aligned}\mathbb{P}\{|\langle y, Y_\xi\rangle - \lambda| \le \rho/4 \mid \mathcal{E}_W\} &\le \sup_{\widetilde{\lambda} \in \mathbb{C}} \mathbb{P}\{|\langle y, e_{\xi_1}\rangle - \widetilde{\lambda}| \le \rho/4 \mid \xi_m \in S_0\} \\ &= \sup_{\widetilde{\lambda} \in \mathbb{C}} \mathbb{P}\{|\langle y, e_{\xi_1}\rangle - \widetilde{\lambda}| \le \rho/4 \mid \xi_1 \in S_0\}.\end{aligned}$$

This implies

$$p_0 := \mathbb{P}\{\mathcal{E}_2 \text{ and } |\langle y, Y_\xi\rangle - \lambda| \le \rho/4\} \le \sup_{\widetilde{\lambda} \in \mathbb{C}} \mathbb{P}\{|\langle y, e_{\xi_1}\rangle - \widetilde{\lambda}| \le \rho/4 \mid \xi_1 \in S_0\}.$$

Fix $\widetilde{\lambda} \in \mathbb{C}$. By Lemma 5.10 for every $i \le 9$ there exists at most one $j(i) \le n$ such that

$$\xi_1 \in S_0 \text{ and } |\langle y, e_{\xi_1}\rangle - \widetilde{\lambda}| \le \rho/4 \quad \text{implies} \quad \xi_1 \in S_0 \cap \bigcup_{i=1}^{9} U_{ij(i)}.$$

Using this, (9) and (10), we observe

$$p_0 \le \frac{1}{\mathbb{P}\{\xi_1 \in S_0\}} \sum_{i=1}^{9} \mathbb{P}\{\xi_1 \in S_0 \cap U_{ij(i)}\} \le \frac{|S|}{|S_0|} \sum_{i=1}^{9} \frac{|S_0 \cap U_{ij(i)}|}{|S|} \le \frac{540\, d \ln n}{n^{1/4}} + 144\delta.$$

Since $\mathbb{P}\{|\langle y, Y_\xi\rangle - \lambda| \le \rho/4\} \le \mathbb{P}(\mathcal{E}_1) + p_0$, $d \le n^{1/8}$, and n is large enough, this completes the proof. □

5.4. *Distance estimates. Proof of Lemma 5.4*

The goal of this subsection is to prove Lemma 5.4.

Fix $z \in \mathbb{C}$, $\gamma = 1/288$, and $i \in [n]$ satisfying $n/\ln^{1/\gamma} n \le n - i \le d^{-3}n$. Recall that σ_n denotes the uniform random permutation of $[n]$ independent of A_n, and $B_z = d^{-1/2}A_n - z\,\mathrm{Id}$. Denote $E_i := E(B_z, \sigma([i-1])$, i.e., the random subspace spanned by the rows $R_{\sigma_n(j)}(B_z)$, $j \le i-1$.

We now define a random triple (A_n, A_n', σ_n) in the following way (the choice of notation will be justified after the construction). For each matrix $M \in \mathcal{M}_{n,d}$ and a permutation $\sigma \in \Pi_n$ let

$$\mathbb{M}_{M,\sigma} := \{M' \in \mathcal{M}_{n,d} : R_{\sigma(j)}(M') = R_{\sigma(j)}(M) \text{ for all } j \notin [i - \lfloor n^{1/4} \rfloor, i]\}.$$

Define

$$U := \bigcup_{\sigma \in \Pi_n} \bigcup_{M \in \mathcal{M}_{n,d}} \{(M, M', \sigma) : M' \in \mathbb{M}_{M,\sigma}\}.$$

Further, define a probability measure η on U by

$$\forall (M, M', \sigma) \in U : \quad \eta(\{(M, M', \sigma)\}) = \frac{1}{n!|\mathcal{M}_{n,d}|} \frac{1}{|\mathbb{M}_{M,\sigma}|}.$$

We postulate that the triple (A_n, A_n', σ_n) takes values in U and is distributed according to the measure η. It is not difficult to see that (individual) marginal distributions of A_n and A_n' are uniform on $\mathcal{M}_{n,d}$, and that σ_n is uniformly distributed on Π_n. Moreover, A_n and σ_n are independent, as also are A_n' and σ_n. This justifies our choice of notation for A_n and σ_n (which otherwise would come into conflict with our "old" notions of A_n and σ_n). As usual, below we assume that G is independent of the triple (A_n, A_n', σ_n) and that all random variables are defined on the same probability space.

Fix a matrix $M \in \mathcal{M}_{n,d}$, a subset $J \subset [n]$ of cardinality $i-1$ and an index $u \in [n] \setminus J$. Define the event

$$\mathcal{E}_{M,J,u} := \big\{A_n = M,\ \{\sigma_n(r) : r \le i-1\} = J,\ \sigma_n(i) = u\big\}.$$

Observe that, conditioned on $\mathcal{E}_{M,J,u}$, the set

$$W := \{\sigma_n(j) : j = i - \lfloor n^{1/4} \rfloor, \dots, i-1\}$$

is a uniform random $\lfloor n^{1/4} \rfloor$-subset of J. Let $W_0 \subset J$ be any realization of W and set

$$\mathcal{E}_{M,J,u,W_0} := \mathcal{E}_{M,J,u} \cap \{W = W_0\}.$$

Conditioned on $\mathcal{E}_{M,J,u,W_0}$, A'_n takes values in the set of matrices $\mathbb{M}_{M,W_0\cup\{u\}}$ defined as in Section 5.3, and the u-th row of A'_n has conditional distribution defined by

$$\mathbb{P}\{R_u(A'_n) = x \mid \mathcal{E}_{M,J,u,W_0}\} = \frac{|\{M' \in \mathbb{M}_{M,W_0\cup\{u\}} : R_u(M') = x\}|}{|\mathbb{M}_{M,W_0\cup\{u\}}|}.$$

In other words, conditioned on $\mathcal{E}_{M,J,u}$, the u-th row of A'_n is distributed exactly as the random vector $X_{M,J,u}$ defined in Section 5.3. Now, let $\mathcal{E}'_{M,J,u} \subset \mathcal{E}_{M,J,u}$ be the event that the uniform random normal $P_{E_i^\perp} G$ satisfies the following condition:

$$\exists \widetilde{J} \subset [n],\ |\widetilde{J}| \le 2\Big(\frac{n-i}{n}\Big)^{\gamma/2} n, \text{ such that}$$
$$\forall \lambda \in \mathbb{C} : \Big|\Big\{ j \in [n] \setminus \widetilde{J} : |\langle P_{E_i^\perp} G, e_j\rangle - \lambda| \le \exp\Big(-C_0\Big(\frac{n}{n-i}\Big)^{\gamma}\Big)\Big\}\Big| \le n-i,$$

where C_0 is the constant from Theorem 4.2. Note that conditioned on the event $\mathcal{E}_{M,J,u}$, the subspace E_i is completely determined by M and J, in particular it is fixed within the event $\mathcal{E}_{M,J,u}$. Therefore, by the independence of G from the triple (A_n, A'_n, σ_n), we see that $P_{E_i^\perp} G$ and the u-th row of A'_n are independent conditioned on $\mathcal{E}_{M,J,u}$. Then, conditioning on the event $\mathcal{E}'_{M,J,u}$ and denoting

$$B'_z := d^{-1/2} A'_n - z\,\mathrm{Id},$$

we apply Proposition 5.5 with $y = P_{E_i^\perp} G$ and $\lambda = d^{1/2}\langle y, R_u(z\,\mathrm{Id})\rangle$ to get

$$\mathbb{P}\Big\{ |\langle P_{E_i^\perp} G, R_u(B'_z)\rangle| \le (16d)^{-1/2} \exp\Big(-C_0\Big(\frac{n}{n-i}\Big)^{\gamma}\Big) \ \Big|\ \mathcal{E}'_{M,J,u}\Big\}$$
$$\le 144\,\frac{n-i}{n} + \Big(16\Big(\frac{n-i}{n}\Big)^{\gamma/2}\Big)^{d} + n^{-1/10} \le 145\frac{n-i}{n} \qquad (11)$$

provided that d is large enough. For convenience, we denote $q := i - \lfloor n^{1/4} \rfloor$. Define another (the last) auxiliary event

$$\widetilde{\mathcal{E}}_{M,J,u} := \mathcal{E}_{M,J,u} \cap \{\ln(n/(n-i))\|P_{E_q^\perp} R_u(B'_z)\|_2 \ge |\langle R_u(B'_z), P_{E_i^\perp} G\rangle|\}.$$

Using the deterministic relation

$$\|P_{E_i^\perp} R_u(B'_z)\|_2 \le \|P_{E_q^\perp} R_u(B'_z)\|_2,$$

the independence of $R_u(A'_n)$ and $P_{E_i^\perp} G$ conditioned on $\mathcal{E}_{M,J,u}$, and (5) applied with $t = \ln(n/(n-i))$, we obtain

$$\mathbb{P}(\widetilde{\mathcal{E}}_{M,J,u} \mid \mathcal{E}_{M,J,u}) \geq 1 - \frac{n-i}{n},$$

and thus

$$\mathbb{P}(\widetilde{\mathcal{E}}^c_{M,J,u} \mid \mathcal{E}'_{M,J,u}) \leq \frac{\mathbb{P}(\mathcal{E}^c_{M,J,u} \cap \mathcal{E}_{M,J,u})}{\mathbb{P}(\mathcal{E}'_{M,J,u})} \leq \frac{n-i}{n} \cdot \frac{\mathbb{P}(\mathcal{E}_{M,J,u})}{\mathbb{P}(\mathcal{E}'_{M,J,u})}.$$

Together with (11) and using

$$4\sqrt{d}\, \ln(n/(n-i)) \leq \exp\left(\left(\frac{n}{n-i}\right)^\gamma\right)$$

for sufficiently large d, we get, for an appropriate choice of the constant $\widetilde{C}$,

$$\begin{aligned}
&\mathbb{P}\left\{\|P_{E_q^\perp} R_{\sigma_n(i)}(B'_z)\|_2 \leq \exp\left(-\widetilde{C}\left(\frac{n}{n-i}\right)^\gamma\right) \,\middle|\, \mathcal{E}'_{M,J,u}\right\} \\
&\quad \leq \mathbb{P}\left\{|\langle P_{E_i^\perp} G, R_u(B'_z)\rangle| \leq c\, d^{-1/2} \exp\left(-C_0\left(\frac{n}{n-i}\right)^\gamma\right) \,\middle|\, \mathcal{E}'_{M,J,u}\right\} \\
&\quad\quad + \mathbb{P}(\widetilde{\mathcal{E}}^c_{M,J,u} \mid \mathcal{E}'_{M,J,u}) \\
&\quad \leq \frac{n-i}{n}\left(145 + \frac{\mathbb{P}(\mathcal{E}_{M,J,u})}{\mathbb{P}(\mathcal{E}'_{M,J,u})}\right) \leq 146\frac{n-i}{n}\,\frac{\mathbb{P}(\mathcal{E}_{M,J,u})}{\mathbb{P}(\mathcal{E}'_{M,J,u})}.
\end{aligned}$$

Using the independence of G and (A_n, A'_n, σ_n) and applying Theorem 4.2 with $I = E_i$, which is fixed within the event $\mathcal{E}_{M,J,u}$, we observe

$$\mathbb{P}\Big(\bigcup_{M,J,u} \mathcal{E}'_{M,J,u}\Big) \geq 1 - \frac{n-i}{n}.$$

Note also that the events $\mathcal{E}_{M,J,u}$ are pairwise disjoint, so that $\sum_{M,J,u} \mathbb{P}(\mathcal{E}_{M,J,u}) \leq 1$. Therefore, using $\mathcal{E}'_{M,J,u} \subset \mathcal{E}_{M,J,u}$ we obtain

$$\begin{aligned}
&\mathbb{P}\left\{\|P_{E_q^\perp} R_{\sigma_n(i)}(B'_z)\|_2 \leq \exp\left(-\widetilde{C}\left(\frac{n}{n-i}\right)^\gamma\right)\right\} \\
&\quad \leq \sum_{M,J,u} \mathbb{P}\left\{\|P_{E_q^\perp} R_{\sigma_n(i)}(B'_z)\|_2 \leq \exp\left(-\widetilde{C}\left(\frac{n}{n-i}\right)^\gamma\right) \,\middle|\, \mathcal{E}'_{M,J,u}\right\} \mathbb{P}(\mathcal{E}'_{M,J,u}) \\
&\quad + \mathbb{P}\left(\Big[\bigcup_{M,J,u} \mathcal{E}'_{M,J,u}\Big]^c\right) \leq 146\frac{n-i}{n} \sum_{M,J,u} \mathbb{P}(\mathcal{E}_{M,J,u}) + \frac{n-i}{n} \leq 147\frac{n-i}{n}.
\end{aligned}$$

Note that for any realization (M, M', σ) of (A_n, A'_n, σ_n) we have $R_{\sigma(j)}(M') = R_{\sigma(j)}(M)$ for all $j < q$, therefore

$$E_q = \operatorname{span}\{R_{\sigma_n(j)}(B_z)\}_{j<q} = \operatorname{span}\{R_{\sigma_n(j)}(B'_z)\}_{j<q}.$$

Thus

$$\mathbb{P}\left\{\mathrm{dist}(R_{\sigma_n(i)}(B'_z), \mathrm{span}\,\{R_{\sigma_n(j)}(B'_z)\}_{j<q}) \le \exp\left(-\widetilde{C}\left(\frac{n}{n-i}\right)^{\gamma}\right)\right\} \le 147\frac{n-i}{n}.$$

In view of the independence of σ_n and A'_n, we can replace the row $R_{\sigma_n(i)}(B'_z)$ in the above formula with $R_{\sigma_n(q)}(B'_z)$ with no change to the probability estimates. Since A'_n and A_n are equidistributed, we can also replace $R_{\sigma_n(q)}(B'_z)$ and $R_{\sigma_n(j)}(B'_z)$ with $R_{\sigma_n(q)}(B_z)$ and $R_{\sigma_n(j)}(B_z)$. Finally, note that in our range of i, $\frac{n-i}{n}$ is equivalent to $\frac{n-q}{n}$ up to a constant 2 and that $n - n/d^3 \le q \le n - n/\ln^{1/\gamma} n - n^{1/4}$. This completes the proof of Lemma 5.4. □

Acknowledgments. P.Y. was supported by grant ANR-16-CE40-0024-01. A.L. was supported by grant no. 2018/31/B/ST1/03937, National Science Centre, Poland. A significant part of this work was completed while the last three named authors were in residence at the Mathematical Sciences Research Institute in Berkeley, California, supported by NSF grant DMS-1440140, and the first two named authors visited the Institute. The hospitality of MSRI and of the organizers of the program on Geometric Functional Analysis and Applications is gratefully acknowledged.

References

[1] Adamczak, R., Guédon, O., Latała, R., Litvak, A. E., Oleszkiewicz, K., Pajor, A., Tomczak-Jaegermann, N.: Moment estimates for convex measures. Electron. J. Probab. **17**, no. 101, 19 pp. (2012) Zbl 1286.46014 MR 3005719

[2] Adamczak, R., Latała, R., Litvak, A. E., Oleszkiewicz, K., Pajor, A., Tomczak-Jaegermann, N.: A short proof of Paouris' inequality. Canad. Math. Bull. **57**, 3–8 (2014) Zbl 1329.46014 MR 3150710

[3] Alesker, S.: ψ_2-estimate for the Euclidean norm on a convex body in isotropic position. In: Geometric Aspects of Functional Analysis (Israel, 1992–1994), Oper. Theory Adv. Appl. 77, Birkhäuser, Basel, 1–4 (1995) Zbl 0834.52004 MR 1353444

[4] Anderson, G. W., Guionnet, A., Zeitouni, O.: An Introduction to Random Matrices. Cambridge Stud. Adv. Math. 118, Cambridge Univ. Press, Cambridge (2010) Zbl 1184.15023 MR 2760897

[5] Bai, Z. D.: Circular law. Ann. Probab. **25**, 494–529 (1997) Zbl 0871.62018 MR 1428519

[6] Bai, Z., W. Silverstein, J.: Spectral Analysis of Large Dimensional Random Matrices. 2nd ed., Springer Ser. Statist., Springer, New York (2010) Zbl 1301.60002 MR 2567175

[7] Basak, A., Cook, N., Zeitouni, O.: Circular law for the sum of random permutation matrices. Electron. J. Probab. **23**, art. 33, 51 pp. (2018) Zbl 1386.15066 MR 3798243

[8] Basak, A., Rudelson, M.: The circular law for sparse non-Hermitian matrices. Ann. Probab. **47**, 2359–2416 (2019) Zbl 07114719 MR 3980923

[9] Bauerschmidt, R., Huang, J., Yau, H.-T.: Local Kesten–McKay law for random regular graphs. Comm. Math. Phys. **369**, 523–636 (2019) Zbl 1414.05260 MR 3962004

[10] Bordenave, C., Chafaï, D.: Around the circular law. Probab. Surv. **9**, 1–89 (2012) Zbl 1243.15022 MR 2908617

[11] Cook, N. A.: On the singularity of adjacency matrices for random regular digraphs. Probab. Theory Related Fields **167**, 143–200 (2017) Zbl 1365.05260 MR 3602844

[12] Cook, N. A.: The circular law for random regular digraphs. Ann. Inst. H. Poincaré Probab. Statist. **55**, 2111–2167 (2019) Zbl 1432.15035 MR 4029149

[13] Edelman, A.: The probability that a random real Gaussian matrix has k real eigenvalues, related distributions, and the circular law. J. Multivariate Anal. **60**, 203–232 (1997) Zbl 0886.15024 MR 1437734

[14] Erdős, L., Yau, H.-T.: A Dynamical Approach to Random Matrix Theory. Courant Lecture Notes in Math. 28, Courant Inst. Math. Sci., New York (2017) Zbl 1379.15003 MR 3699468

[15] Ginibre, J.: Statistical ensembles of complex, quaternion, and real matrices. J. Math. Phys. **6**, 440–449 (1965) Zbl 0127.39304 MR 0173726

[16] Girko, V. L.: The circular law. Teor. Veroyatnost. i Primenen. **29**, 669–679 (1984) (in Russian) Zbl 0565.60034 MR 0773436

[17] Girko, V. L.: The strong circular law. Twenty years later. II. Random Oper. Stochastic Equations **12**, 255–312 (2004) Zbl 1065.60026 MR 2085255

[18] Girko, V. L.: The circular law. Thirty year later. Random Oper. Stochastic Equations **20**, 143–187 (2012) Zbl 1278.15041 MR 2980360

[19] Götze, F., Tikhomirov, A.: The circular law for random matrices. Ann. Probab. **38**, 1444–1491 (2010) Zbl 1203.60010 MR 2663633

[20] Kesten, H.: Symmetric random walks on groups. Trans. Amer. Math. Soc. **92**, 336–354 (1959) Zbl 0092.33503 MR 0109367

[21] Latała, R., Mankiewicz, P., Oleszkiewicz, K., Tomczak-Jaegermann, N.: Banach–Mazur distances and projections on random subgaussian polytopes. Discrete Comput. Geom. **38**, 29–50 (2007) Zbl 1134.52016 MR 2322114

[22] Litvak, A. E., Lytova, A., Tikhomirov, K., Tomczak-Jaegermann, N., Youssef, P.: Anti-concentration property for random digraphs and invertibility of their adjacency matrices. C. R. Math. Acad. Sci. Paris **354**, 121–124 (2016) Zbl 1388.60039 MR 3456885

[23] Litvak, A. E., Lytova, A., Tikhomirov, K., Tomczak-Jaegermann, N., Youssef, P.: Adjacency matrices of random digraphs: singularity and anti-concentration. J. Math. Anal. Appl. **445**, 1447–1491 (2017) Zbl 1344.05126 MR 3545253

[24] E. Litvak, A., Lytova, A., Tikhomirov, K., Tomczak-Jaegermann, N., Youssef, P.: The smallest singular value of a shifted d-regular random square matrix. Probab. Theory Related Fields **173**, 1301–1347 (2019) Zbl 1411.60015 MR 3936157

[25] Litvak, A. E., Lytova, A., Tikhomirov, K., Tomczak-Jaegermann, N., Youssef, P.: Structure of eigenvectors of random regular digraphs. Trans. Amer. Math. Soc. **371**, 8097–8172 (2019) Zbl 07062707 MR 3955544

[26] Litvak, A. E., Rivasplata, O.: Smallest singular value of sparse random matrices. Studia Math. **212**, 195–218 (2012) Zbl 1277.60016 MR 3009072

[27] Marchenko, V. A., Pastur, L. A.: Distribution of eigenvalues in certain sets of random matrices. Mat. Sb. (N.S.) **72 (114)**, 507–536 (1967) (in Russian) Zbl 0152.16101 MR 0208649

[28] McKay, B. D.: The expected eigenvalue distribution of a large regular graph. Linear Algebra Appl. **40**, 203–216 (1981) Zbl 0468.05039 MR 0629617

[29] Mehta, M. L.: Random Matrices. 3rd ed., Pure Appl. Math. (Amsterdam) 142, Elsevier/Academic Press, Amsterdam (2004) Zbl 1107.15019 MR 2129906

[30] McKay, B. D., Wang, X.: Asymptotic enumeration of 0-1 matrices with equal row sums and equal column sums. Linear Algebra Appl. **373**, 273–287 (2003) Zbl 1026.05015 MR 2022291

[31] Naor, A., Youssef, P.: Restricted invertibility revisited. In: A Journey through Discrete Mathematics, Springer, Cham, 657–691 (2017) Zbl 1403.46013 MR 3726618

[32] Pan, G., Zhou, W.: Circular law, extreme singular values and potential theory. J. Multivariate Anal. **101**, 645–656 (2010) Zbl 1203.60011 MR 2575411

[33] Paouris, G.: Concentration of mass on convex bodies. Geom. Funct. Anal. **16**, 1021–1049 (2006) Zbl 1114.52004 MR 2276533

[34] Pastur, L., Shcherbina, M.: Eigenvalue Distribution of Large Random Matrices. Math. Surveys Monogr. 171, Amer. Math. Soc., Providence, RI (2011) Zbl 1244.15002 MR 2808038

[35] Rudelson, M., Vershynin, R.: Hanson–Wright inequality and sub-gaussian concentration. Electron. Comm. Probab. **18**, art. 82, 9 pp. (2013) Zbl 1329.60056 MR 3125258

[36] Tao, T.: Topics in Random Matrix Theory. Grad. Stud. Math. 132, Amer. Math. Soc., Providence, RI (2012) Zbl 1256.15020 MR 2906465

[37] Tao, T., Vu, V.: Random matrices: the circular law. Comm. Contemp. Math. **10**, 261–307 (2008) Zbl 1156.15010 MR 2409368

[38] Tao, T., H. Vu, V.: Inverse Littlewood–Offord theorems and the condition number of random discrete matrices. Ann. of Math. **169**, 595–632 (2009) Zbl 1250.60023 MR 2480613

[39] Tao, T., Vu, V.: Random matrices: universality of ESDs and the circular law. Ann. Probab. **38**, 2023–2065 (2010) Zbl 1203.15025 MR 2722794

[40] Tikhomirov, K., Youssef, P.: The spectral gap of dense random regular graphs. Ann. Probab. **47**, 362–419 (2019) Zbl 1407.05212 MR 3909972

[41] Yin, Y. Q.: Limiting spectral distribution for a class of random matrices. J. Multivariate Anal. **20**, 50–68 (1986) Zbl 0614.62060 MR 0862241