The Annals of Probability
2019, Vol. 47, No. 4, 2359–2416
https://doi.org/10.1214/18-AOP1310

THE CIRCULAR LAW FOR SPARSE NON-HERMITIAN MATRICES

BY ANIRBAN BASAK[*,†,1] AND MARK RUDELSON[‡,2]

ICTS–TIFR[*], Weizmann Institute of Science[†] and University of Michigan[‡]

For a class of sparse random matrices of the form $A_n = (\xi_{i,j}\delta_{i,j})_{i,j=1}^n$, where $\{\xi_{i,j}\}$ are i.i.d. centered sub-Gaussian random variables of unit variance, and $\{\delta_{i,j}\}$ are i.i.d. Bernoulli random variables taking value 1 with probability p_n, we prove that the empirical spectral distribution of $A_n/\sqrt{np_n}$ converges weakly to the circular law, in probability, for all p_n such that $p_n = \omega(\log^2 n/n)$. Additionally if p_n satisfies the inequality $np_n > \exp(c\sqrt{\log n})$ for some constant c, then the above convergence is shown to hold almost surely. The key to this is a new bound on the smallest singular value of complex shifts of real valued sparse random matrices. The circular law limit also extends to the adjacency matrix of a directed Erdős–Rényi graph with edge connectivity probability p_n.

CONTENTS

1. Introduction 2359
Outline of the paper 2365
2. Preliminaries and proof outline 2366
2.1. Smallest singular value 2367
2.2. Intermediate singular values 2372
2.3. Weak convergence 2372
3. The structure of the kernel: Vectors with nondominated real part 2373
4. Net construction: Genuinely complex case 2378
5. The structure of the kernel in the genuinely complex case 2387
6. Construction of the net and the structure of the kernel in the essentially real case 2394
7. Proof of Theorem 2.2 2400
8. Intermediate singular values 2404
9. Weak convergence 2406
10. Proof of Theorem 1.3 2408
Acknowledgments 2414
References 2414

1. Introduction. For a $n \times n$ matrix B_n let $\lambda_1(B_n), \lambda_2(B_n), \ldots, \lambda_n(B_n)$ be its eigenvalues. The empirical spectral distribution (ESD) of B_n, denoted hereafter by

Received July 2017; revised June 2018.
[1]Supported in part by Israel Science Foundation Grant 147/15 and ICTS–Infosys Excellence Grant.
[2]Supported in part by NSF Grant DMS 1464514.
MSC2010 subject classifications. 15B52, 60B10, 60B20.
Key words and phrases. Random matrix, sparse matrix, smallest singular value, circular law.

L_{B_n}, is given by

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

where δ_x is the dirac measure at x. If B_n is a nonnormal matrix (i.e., $B_nB_n^* \neq B_n^*B_n$), then its eigenvalues are complex valued, resulting in L_{B_n} being supported on the complex plane. Furthermore, when B_n is random its ESD is a random probability measure. Thus, to study the asymptotics of ESDs of random matrices one needs to define appropriate notions of convergence. This can be done in one of the two following ways: If $\{\mu_n\}$ is a sequence of random probability measures such that for every $f \in C_b(\mathbb{C})$, that is, $f : \mathbb{C} \mapsto \mathbb{R}$ is bounded, $\int_{\mathbb{C}} f\,d\mu_n \to \int_{\mathbb{C}} f\,d\mu$ in probability, for some probability measure μ (possibly random), then μ_n is said to converge weakly to μ, in probability. If $\int_{\mathbb{C}} f\,d\mu_n \to \int_{\mathbb{C}} f\,d\mu$ almost surely, then μ_n is said to converge to μ weakly, almost surely.

The study of the ESD for random matrices can be traced back to Wigner [42, 43] who showed that the ESD's of $n \times n$ Hermitian matrices with i.i.d. centered entries of variance $1/n$ (modulo symmetry) satisfying appropriate moment bounds (e.g., Gaussian) converge to the *semicircle distribution*. The conditions on the finiteness of the moments were relaxed in subsequent works; see [4, 32] and the references therein.

The rigorous study of non-Hermitian matrices, in particular nonnormal matrices, emerged much later. The main difficulties were the sensitivity of the eigenvalues of nonnormal matrices under small perturbations and the lack of appropriate tools. For example, Wigner's proof employed the method of moments. Noting that the moments of the semicircle law determine it, one computes by combinatorial means the expectation and the variance of the normalized trace of powers of the matrix to find the asymptotics of the moments of the ESDs. The analogue of this for nonnormal matrices is to compute the mixed moments, that is, compute

$$\int_{\mathbb{C}} w^k \bar{w}^\ell\, dL_{B_n}(w) = \frac{1}{n}\sum_{i=1}^{n}\lambda_i(B_n)^k\bar{\lambda}_i(B_n)^\ell. \tag{1.1}$$

For B_n nonnormal, the RHS of (1.1) cannot be expressed as powers of traces of B_n and B_n^*. So the method of moment approach does not work. Another technique that works well for Hermitian matrices is the method of evaluating limiting *Stieltjes transform* (see [6]). Since the Stieltjes transform of a probability measure is well defined outside its support, and the ESDs of nonnormal matrices are supported on $\mathbb{C}$, their Stieltjes transforms fail to capture the spectral measure.

In the 1950s, based on numerical evidences, it was conjectured that the ESD of $B_n/\sqrt{n}$, where B_n is an $n \times n$ matrix with i.i.d. entries of zero mean and unit variance, converges to the circular law, the uniform distribution on the unit disk in the complex plane. In random matrix literature, this conjecture is commonly known as the *circular law conjecture*.

Using the formula for the joint density function of the eigenvalues, Ginibre [24] solved the case when the entries have a complex Gaussian distribution. The case of real Gaussian entries, where a similar formula is available, was settled by Edelman [21]. For the general case when there is no such formula, the problem remained open for a very long time. An approach to the problem, which eventually played an important role in the resolution of the conjecture, was suggested by Girko in the 1980s [25]; however, mathematically it contained significant gaps. The first non-Gaussian case (assuming the existence of density for the entries) was rigorously treated by Bai [5], and the first result without the density assumption was obtained by Götze and Tikhomirov [26]. After a series of partial results (see [14] and the references therein), the circular law conjecture was established in its full generality in the seminal work of Tao and Vu [40].

THEOREM 1.1 (Circular law for i.i.d. entries [40], Theorem 1.10). *Let M_n be an $n \times n$ random matrix whose entries are i.i.d. complex random variables with zero mean and unit variance. Then the* ESD *of $\frac{1}{\sqrt{n}}M_n$ converges weakly to the uniform distribution on the unit disk on the complex plane, both in probability and in the almost sure sense.*

A remarkable feature of Theorem 1.1 is its *universality*. The asymptotic behavior of the ESD of $M_n/\sqrt{n}$ is insensitive to the specific details of the entry distributions as long as they are i.i.d. and have zero mean and unit variance. Since the work of Tao and Vu, there have been numerous attempts to extend the universality of Theorem 1.1 for a wider class of random matrix ensembles. A natural extension would be to prove Theorem 1.1 for matrix ensembles with dependent entries. This has been shown in [1, 2, 12, 30, 31].

Another direction to pursue is to study the asymptotics of the ESD of sparse matrices. Sparse matrices are abundant in statistics, neural network, financial modeling, electrical engineering, wireless communications, neuroscience and in many other fields. We refer the reader to [4], Chapter 7, for other examples, and their relevant references. One model for sparse random matrices is the adjacency matrices of random d-regular directed graphs with $d = o(n)$ (for $\{a_n\}$ and $\{b_n\}$, sequences of positive reals, the notation $a_n = o(b_n)$ means $\lim_{n\to\infty} a_n/b_n = 0$). Recently, in [7, 20, 29], the circular law conjecture was established for two different models of random d-regular directed graphs.

One of the most natural models for sparse random matrices is the Hadamard product of a matrix of i.i.d. entries with zero mean and unit variance, and a matrix of i.i.d. Ber(p_n) entries, with $p_n = o(1)$. In this paper, we focus on the limiting spectral distribution of this class of sparse matrices. When $p_n = n^{\alpha-1}$ for some $\alpha \in (0, 1)$, it has been shown that, under the assumption of the existence of $(2 + \delta)$th moment of the entries, the ESD of these sparse matrices (properly scaled) converges weakly to the circular law, in probability and almost surely (see [26, 39]). Later in

[44] the assumption on the existence of $(2+\delta)$th moment was removed but the convergence was shown to hold in probability.

In this paper, we prove that the circular law limit continues hold when np_n grows at a subpolynomial rate (in n). Namely, we show that under certain moment assumptions of the entries the circular law limit holds for sparse non-Hermitian matrices whenever np_n grows at a polylogarithmic rate. Under an additional assumption on p_n (see (1.2) below), the convergence is shown to hold almost surely.

Before stating our result, let us recall the well-known definition of sub-Gaussian random variables.

DEFINITION 1.2. For a random variable ξ, the sub-Gaussian norm of ξ, denoted by $\|\xi\|_{\psi_2}$, is defined as

$$\|\xi\|_{\psi_2} := \sup_{k\geq 1} k^{-1/2}\|\xi\|_k \qquad \text{where } \|\xi\|_k := \left(\mathbb{E}|\xi|^k\right)^{1/k}.$$

If the sub-Gaussian norm is finite, the random variable ξ is called sub-Gaussian.

We use the following standard notation: for two sequences positive reals $\{a_n\}$ and $\{b_n\}$ we write $a_n = \omega(b_n)$ if $b_n = o(a_n)$ and $a_n = O(b_n)$ if $\limsup_{n\to\infty} a_n/b_n < \infty$.

The following is the main result of this article.

THEOREM 1.3. *Let A_n be an $n\times n$ matrix with i.i.d. entries $a_{i,j} = \delta_{i,j}\xi_{i,j}$, where $\delta_{i,j}$, $i,j\in[n]$ are independent Bernoulli random variables taking value* 1 *with probability $p_n \in (0,1]$, and $\xi_{i,j}$, $i,j\in[n]$ are real-valued i.i.d. centered sub-Gaussian with unit variance*:

(i) *If p_n is such that $np_n = \omega(\log^2 n)$, the* ESD *of $A_n/\sqrt{np_n}$ converges weakly to the circular law, as $n\to\infty$, in probability.*

(ii) *There exists a constant $c_{1.3}$, depending only on the sub-Gaussian norm of $\{\xi_{i,j}\}$, such that if p_n satisfies the inequality*

$$np_n > \exp(c_{1.3}\sqrt{\log n}) \tag{1.2}$$

then the conclusion of part (i) *holds almost surely.*

It will be seen in Section 2 that a key to the proof of Theorem 1.3 is a uniform bound on $s_{\min}(A_n - w\sqrt{np_n}I_n)$ for Lebesgue a.e. $w\in\mathbb{C}$, where $s_{\min}(\cdot)$ denotes the smallest singular value. We initiated this work in [9] and showed that the desired bound holds when $w\in\mathbb{R}$. The result of [9] relied on identifying the obstacles of arithmetic nature by methods of Fourier analysis, and using geometry to show that with high probability none of these obstacles realizes. However, even if the matrix A_n is real valued, the extension to $w\in\mathbb{C}\backslash\mathbb{R}$ makes the set of potential arithmetic obstacles so rich that it cannot be handled within the framework of the previous

argument. This required developing new methods providing both a more precise identification of the arithmetic obstacles arising from the complex structure and entropy bounds showing that with high probability these obstacles are avoided.

The main part of this paper is devoted to find the desired bound on $s_{\min}$ with a probability bound strong enough to apply the Borel–Cantelli lemma in order to deduce the almost sure convergence of Theorem 1.3(ii). To remove the condition (1.2), one needs an improvement of [9], Proposition 3.1. See Remark 7.4 for more details.

It is easy to see that if $p_n = \frac{\log n}{n}$, then the number of zero columns of A_n is positive (and hence $s_{\min}(A_n) = 0$) with probability bounded away from zero. So $\frac{\log n}{n}$ is a natural barrier in this set-up. To extend the bound on $s_{\min}$ beyond this barrier, one needs to analyze the smallest singular value of the adjacency matrix of "core of the graph," when A_n is viewed as the adjacency matrix of directed random weighted graph. We leave this effort to future ventures.

Another key ingredient for the proof of Theorem 1.3 is the bound on the smallish singular values of $(A_n - w\sqrt{np_n}I_n)$ (see Theorem 2.10). This is derived by relating the inverse second moment of the singular values to that of the distance of a random vector from a random subspace. This idea originated in [40] and was crucial in relaxing the assumption on the existence of the $(2+\delta)$th moment and prove Theorem 1.1 only under the second moment assumption. To carry out a similar scheme in the sparse set-up, one needs to consider random subspaces of dimension $n - m$ with $m = o(n/\log n)$. Concentration inequalities yield a probability bound $\exp(-cmp_n)$, for some $c > 0$. To accommodate a union bound, we then need $np_n = \omega(\log^2 n)$ which translates to the required lower bound on p_n in Theorem 1.3(i).

REMARK 1.4 (Sub-Gaussianity assumption). The sub-Gaussianity assumption of Theorem 1.3 is used in Theorem 2.2 to show that $\|A_n\| = O(\sqrt{np_n})$, where $\|\cdot\|$ denotes the operator norm. From [9], Remark 1.9, we note that if $\{\xi_{i,j}\}$ are such that

$$\mathbb{E}|\xi_{i,j}|^h \le C^h h^{\beta h} \qquad \text{for all } h \ge 1, \text{ and for some constants } C \text{ and } \beta, \tag{1.3}$$

then $\|A_n\| = O(\sqrt{np_n})$, for all p_n satisfying $np_n = \Omega((\log n)^{2\beta})$ (for two sequences of positive reals $\{a_n\}$ and $\{b_n\}$ we write $a_n = \Omega(b_n)$ if $b_n = O(a_n)$). The case $\beta = 1/2$ corresponds to the sub-Gaussian random variables. So we conclude that if $\{\xi_{i,j}\}$ satisfies the moment assumption (1.3), for some $\beta \ge 1/2$, then the conclusion of Theorem 1.3(i) holds for all p_n satisfying $np_n \ge \omega(\log^2 n), \Omega((\log n)^{2\beta})$. It is easy to check that p_n satisfies (1.2) whenever $np_n = \omega((\log n)^{2\beta})$. Hence, the conclusion of Theorem 1.3(ii) also holds under the moment assumption (1.3). To retain the clarity of presentation we prove Theorem 1.3 for sub-Gaussian random variables.

REMARK 1.5 (Circular law limit for shifted sparse matrices). It is well known that the spectrum of normal matrices is stable under small perturbations (see [3], Lemma 2.1.19, and [6], Lemma 2.2). However, for a general nonnormal matrix its spectrum is highly sensitive to small perturbations; for example, see [38], Section 2.8.1. So there are no analogues of [3], Lemma 2.1.19 and [6], Lemma 2.2, for an arbitrary nonnormal matrix. Nevertheless, in [44] it was shown that if D_n is any $n \times n$ matrix with $\operatorname{rank}(D_n) = o(n)$ and $\operatorname{Tr}(D_n D_n^*) = O(n^2 p_n)$ then the ESD of $(A_n + D_n)/\sqrt{np_n}$ admit a circular law limit. Investigating our proof, one can deduce that the ESD of $(A_n + D_n)/\sqrt{np_n}$ have a circular law limit for any sequence real diagonal matrices $\{D_n\}$ such that $\|D_n\| = O(\sqrt{np_n})$ and $\operatorname{Tr}(D_n^2) = o(n^2 p_n)$. It is possible to modify the proof of Theorem 1.3 to establish the circular law limit for general shifts. We do not pursue this direction here.

We next show that the circular law limit holds for the adjacency matrix of a directed Erdős–Rényi random graph which may be of interest in computer science and graph theory. Let us begin with the relevant definitions.

DEFINITION 1.6. Let G_n be a random directed graph on n vertices, with vertex set $[n]$, such that for every $i \neq j$, a directed edge from i to j is present with probability p, independently of everything else. Assume that the graph G_n is simple, that is, no self-loops or multiple edges are present. We call this graph G_n a directed Erdős–Rényi graph with edge connectivity probability p. For any such graph G_n, we denote $\mathrm{Adj}_n := \mathrm{Adj}(\mathsf{G}_n)$ to be its adjacency matrix. That is, for any $i, j \in [n]$,

$$\mathrm{Adj}_n(i,j) = \begin{cases} 1 & \text{if a directed edge from } i \text{ to } j \text{ is present in } \mathsf{G}_n, \\ 0 & \text{otherwise.} \end{cases}$$

THEOREM 1.7. *Let* Adj_n *be the adjacency matrix of a directed Erdős–Rényi graph, with edge connectivity probability* $p_n \in (0,1)$. *Denote* $\bar{p}_n := \min\{p_n, 1 - p_n\}$:

(i) *If* p_n *is such that* $n\bar{p}_n = \omega(\log^2 n)$ *the* ESD *of* $\mathrm{Adj}_n/\sqrt{np_n(1-p_n)}$ *converges weakly to the circular law, as* $n \to \infty$, *in probability.*

(ii) *There exists an absolute constant* $c_{1.7}$ *such that if* p_n *satisfies the inequality*

$$n\bar{p}_n > \exp(c_{1.7}\sqrt{\log n}) \tag{1.4}$$

then the conclusion of part (i) *holds almost surely.*

The proof of Theorem 1.7 follows from a relatively standard modification of that of Theorem 1.3. We refer the reader to the arXiv version of this paper [8].

REMARK 1.8. Theorem 1.3 and Theorem 1.7 find the asymptotics of the eigenvalues of a large class of sparse non-Hermitian random matrices at a macroscopic level. An interesting question would be to prove the universality of the eigenvalue distribution at the microscopic level. This has been shown for a wide class of Hermitian random matrices (see [11] and references therein). For dense non-Hermitian random matrices, it was shown in [15] that the local circular law holds. In a forthcoming article [10], we establish the same for sparse non-Hermitian random matrices.

Outline of the paper. Section 2 provides a brief outline of the proof techniques of Theorem 1.3. We begin Section 2 with a replacement principle (see Lemma 2.1) which is a consequence of Girko's method. The replacement principle allows us to focus only on the integrability of $\log(\cdot)$ with respect to the ESD of $\widetilde{\boldsymbol{B}}_n^w := [(B_n - wI_n)^*(B_n - wI_n)]^{1/2}$ for $w \in \mathbb{C}$, where B_n is any random matrix. To implement this scheme, one requires a good control on $s_{\min}(\widetilde{\boldsymbol{B}}_n^w)$ as well as on its *smallish* singular values. One also needs to establish weak convergence of the ESDs of $\widetilde{\boldsymbol{B}}_n^w$.

The required control on $s_{\min}$ and smallish singular values are derived in Theorem 2.2 and Theorem 2.10, and we outline of their proofs in Section 2.1 and Section 2.2, respectively. The limit of the ESDs of $\widetilde{\boldsymbol{B}}_n^w$ is derived in Theorem 2.11 with the outline of the proof appearing in Section 2.3.

Section 3 through Section 7 are devoted to the proof of Theorem 2.2. Since $s_{\min}(M_n)$ equals the infimum of $\|M_n u\|_2$ ($\|\cdot\|_2$ denotes the Euclidean norm) over all vectors u of unit ℓ_2 norm, we split the unit sphere into three parts: *compressible vectors*, *dominated vectors* and the complement of their union. The compressible vectors and dominated vectors are treated with results from [9]. The majority of the work is to control infimum over the vectors that are neither compressible nor dominated. Using a result of [34] (see Lemma 3.5 there), this boils down to controlling the inner product of the first column of $(A_n - w\sqrt{np_n}I_n)$ and the vector normal to H_n^w, the subspace spanned by the last $(n-1)$ columns of $(A_n - w\sqrt{np_n}I_n)$. In Section 7, it is shown that the last assertion can be proved using Berry–Esséen theorem. However, the probability bounds obtained from Berry–Esséen theorem is too weak to prove the almost sure convergence of Theorem 1.3(ii).

In Section 3 through Section 6, we derive a better probability bound that is suitable for the proof of Theorem 1.3(ii). We split the set of vectors into two categories: *genuinely complex* and *essentially real*. Roughly speaking, the set of essentially real vectors are those for which the real and the imaginary parts are almost linearly dependent, and its complement is the set of genuinely complex vectors.

In Section 3, we show that the vector normal to H_n^w has a nondominated real part, with high probability. We construct a net of small cardinality for the genuinely complex vectors in Section 4. We then use this net in Section 5 and results of Section 3 to show that with high probability, the normal vector cannot be a genuinely complex vector with a subexponential (in $\sqrt{np_n}$) LCD. A similar result for

essentially real vectors is obtained in Section 6. Then we complete the proof of Theorem 2.2 in Section 7.

In Section 8, we prove Theorem 2.10. The key idea is to show that the distance of any row of A_n from any given subspace of relatively small dimension cannot be too small with large probability. This observation together with [40], Lemma A.4, completes the proof.

Section 9 is devoted to the proof of Theorem 2.11, which establishes the weak convergence of the empirical measure of the singular values of $(A_n/\sqrt{np_n} - wI_n)$. The weak convergence is established by appealing to the Lindeberg replacement lemma, which was introduced by Chatterjee in [17], in conjunction with the standard concentration inequalities for the spectral measure of a Hermitian matrix and a truncation argument.

Finally in Section 10, combining the results of Section 7 through Section 9, we complete the proof of Theorem 1.3.

2. Preliminaries and proof outline. In this section we provide an outline of the proof of Theorem 1.3 and introduce necessary definitions and notation. As mentioned in Section 1, the standard technique to find the limiting spectral distribution of a nonnormal matrix is the Girko's method. We refer the reader to [7] for a detailed description of it. The utility of Girko's method, in the context of our set-up, can be captured by the following *replacement principle*. The replacement principle, which has its origin in [40], gives a sufficient criterion for the ESD of two random matrix ensembles to have the same limit. To state the relevant result, first we introduce a few definitions. A sequence of random variables $\{X_n\}$ is said to be bounded in probability if

$$\lim_{K\to\infty} \limsup_{n\to\infty} \mathbb{P}\big(|X_n| \le K\big) = 1$$

and $\{X_n\}$ is said to be almost surely bounded if

$$\mathbb{P}\Big(\limsup_n |X_n| < \infty\Big) = 1.$$

Next, for a matrix B_n, we denote $\|B_n\|_2$ its Frobenius norm, that is, $\|B_n\|_2 := \sqrt{\mathrm{Tr}(B_n^* B_n)}$.

LEMMA 2.1 (Replacement lemma). (a) *Let $B_n^{(1)}$ and $B_n^{(2)}$ be two sequences of $n \times n$ random matrices, such that*:

(i) *The expression*

$$\frac{1}{n}\big\|B_n^{(1)}\big\|_2^2 + \frac{1}{n}\big\|B_n^{(2)}\big\|_2^2 \qquad \textit{is bounded in probability}, \tag{2.1}$$

and

(ii) *For Lebesgue almost all* $w \in \mathbb{D} \subset B_{\mathbb{C}}(0, R) \subset \mathbb{C}$, *for some domain* $\mathbb{D}$ *and some finite* R,

$$\frac{1}{n}\log|\det(B_n^{(1)} - wI_n)| - \frac{1}{n}\log|\det(B_n^{(2)} - wI_n)| \to 0 \qquad \text{in probability.} \tag{2.2}$$

Then for every $f \in C_c^2(\mathbb{C})$ *supported on* $\mathbb{D}$,

$$\int f(z)\,dL_{B_n^{(1)}}(w) - \int f(z)\,dL_{B_n^{(2)}}(w) \to 0 \qquad \text{in probability.} \tag{2.3}$$

(b) *If* (2.1) *is almost surely bounded and* (2.2) *holds almost surely then* (2.3) *holds almost surely as well.*

The replacement principle of [40] requires a uniform control on $s_{\min}(A_n - w\sqrt{np}I_n)$ for Lebesgue almost every $w \in \mathbb{C}$. Theorem 2.2 (see below) provides such a control only when w is away from the real line. Therefore, we need to use Lemma 2.1, borrowed from [7], instead of [40], Theorem 2.1.

Lemma 2.1(a) was proved in [7]. Repeating the proof of [7], Lemma 10.1, one can derive Lemma 2.1(b). Details are left to the reader.

We apply Lemma 2.1 with $B_n^{(1)} = \frac{1}{\sqrt{np_n}}A_n$ and $B_n^{(2)}$ which is the matrix of i.i.d. centered complex Gaussian entries with variance $1/n$. The assumption (i) is straightforward to verify: it follows from laws of large numbers. It is well known that $\frac{1}{n}\log|\det(B_n^{(2)} - wI_n)|$ admits a limit. Hence, establishing assumption (ii) of Lemma 2.1 boils down to showing that $\log(\cdot)$ is integrable with respect to the empirical measure of the singular values of $B_n^{(1)} - wI_n$. As $\log(\cdot)$ is unbounded near zero, one needs to establish the weak convergence of the empirical measure of the singular values, find bounds on $s_{\min}$, and show that there are not many singular values in an interval near zero (the unboundedness of $\log(\cdot)$ near infinity is not a problem since the maximal singular value of $B_n^{(1)} - wI_n$ is $O(1)$ with probability $1 - o(1)$). These are the three ingredients of the proof of Theorem 1.3.

2.1. *Smallest singular value.* The desired bound on $s_{\min}(A_n - \sqrt{np_n}wI_n)$ is derived in the theorem below.

THEOREM 2.2. *Let* A_n *be an* $n \times n$ *matrix with i.i.d. entries* $a_{i,j} = \delta_{i,j}\xi_{i,j}$, *where* $\{\delta_{i,j}\}$ *are independent Bernoulli random variables taking value* 1 *with probability* $p_n \in (0, 1]$, *and* $\{\xi_{i,j}\}$ *are i.i.d. real-valued centered sub-Gaussian with unit variance. Fix* $R \geq 1$, $r \in (0, 1]$ *and let* D_n *be a diagonal matrix such that* $\|D_n\| \leq R\sqrt{np_n}$ *and* $\mathrm{Im}(D_n) = r'\sqrt{np_n}I_n$ *for some* r' *with* $|r'| \in [r, 1]$. *Then there exist constants* $0 < c_{2.2}, \bar{c}_{2.2}, c'_{2.2}, C_{2.2}, C'_{2.2}, \overline{C}_{2.2} < \infty$, *depending only on* R, r *and the sub-Gaussian norm of* $\{\xi_{i,j}\}$, *such that for any* $\varepsilon > 0$ *we have the following*:

(i) *If*

$$p_n \geq \frac{\overline{C}_{2.2} \log n}{n},$$

then

$$\mathbb{P}\left(s_{\min}(A_n + D_n) \leq c_{2.2}\varepsilon \exp\left(-C_{2.2} \frac{\log(1/p_n)}{\log(np_n)} \right) \sqrt{\frac{p_n}{n}} \right) \leq \varepsilon + \frac{C'_{2.2}}{\sqrt{np_n}}.$$

(ii) *Additionally, if*

$$\log(1/p_n) < \bar{c}_{2.2} (\log np_n)^2, \tag{2.4}$$

then

$$\begin{aligned} \mathbb{P}\left(s_{\min}(A_n + D_n) \leq c_{2.2}\varepsilon \exp\left(-C_{2.2} \frac{\log(1/p_n)}{\log(np_n)} \right) \sqrt{\frac{p_n}{n}} \right) \\ \leq \varepsilon + \exp(-c'_{2.2}\sqrt{np_n}). \end{aligned} \tag{2.5}$$

REMARK 2.3. It is easy to check that if $np_n > \exp(c_\star \sqrt{\log n})$ for some constant $c_\star$ then p_n satisfies (2.4). Therefore, it is enough to prove Theorem 2.2(ii) under the assumption (2.4) in order to apply it to the proof of Theorem 1.3(ii).

REMARK 2.4 (Optimality of Theorem 2.2). It is believed that for A_n and D_n as in Theorem 2.2 one has that $s_{\min}(A_n + D_n) \sim p_n^{1/2} n^{-1/2}$. Hence, for $p_n \sim n^{\alpha - 1}$ for some $\alpha > 0$, Theorem 2.2 gives an optimal lower bound on $s_{\min}(A_n + D_n)$. However, when np_n grows at a rate subpolynomial in n, we get an additional factor $n^{-o(1)}$. This is due to the fact that one needs $\rho = o(1)$ in [9], Proposition 3.1. To obtain the optimal lower bound on $s_{\min}(A_n + D_n)$, one needs $\rho = \Omega(1)$ there.

We also add that the optimal probability bound for the event on the LHS of (2.5) is $\varepsilon + \exp(-c'_{2.2} np_n)$. The suboptimality of the probability bound in (2.5) is again due to the fact that we can only allow $\rho = o(1)$ in [9], Proposition 3.1.

Similar to [9], without loss of generality, we can and will assume that $p \leq cR^{-2}$, for some small positive constant c. For larger values of p, the entries $a_{i,j}$ have variance bounded below by an absolute constant. In such case, we can ignore sparsity and regard entries $a_{i,j}$ as i.i.d. centered sub-Gaussian random variables whose variance is bounded below.

To prove Theorem 2.2, we follow the same scheme as in [9] and borrow some of its results. Recalling the definition of the smallest singular value we have

$$s_{\min}(A_n + D_n) = \inf_{z \in S_{\mathbb{C}}^{n-1}} \|(A_n + D_n)z\|_2,$$

where $S_{\mathbb{C}}^{n-1} := \{z \in \mathbb{C}^n : \|z\|_2 = 1\}$. Thus, to bound $s_{\min}$ we need a lower bound on this infimum. To obtain such a bound, we decompose the unit sphere into *compressible*, *dominated* and *incompressible* vectors, and obtain necessary bounds on

the infimum on each of these parts separately. The definitions of compressible, dominated and incompressible vectors are borrowed from [9]. However, we now need to treat complex shifts of the matrix A_n which necessitates a straightforward modification of those definitions to accommodate vectors with complex valued entries. We start with the definition of compressible and incompressible vectors.

DEFINITION 2.5. Fix $m < n$. The set of m-sparse vectors is given by

$$\mathrm{Sparse}(m) := \{z \in \mathbb{C}^n \mid |\mathrm{supp}(z)| \leq m\},$$

where $|S|$ denotes the cardinality of a set S and $\mathrm{supp}(\cdot)$ denotes the support. Furthermore, for any $\delta > 0$, the vectors which are δ-close to m-sparse vectors in Euclidean norm, are called (m, δ)-compressible vectors. The set of all such vectors, hereafter will be denoted by $\mathrm{Comp}(m, \delta)$. Thus,

$$\mathrm{Comp}(m, \delta) := \{z \in S_{\mathbb{C}}^{n-1} \mid \exists y \in \mathrm{Sparse}(m) \text{ such that } \|z - y\|_2 \leq \delta\}.$$

The vectors in $S_{\mathbb{C}}^{n-1}$ which are not compressible are defined to be incompressible, and the set of all incompressible vectors is denoted as $\mathrm{Incomp}(m, \delta)$.

Next, we define the dominated vectors. These are close to sparse vectors but in a different sense.

DEFINITION 2.6. For any $z \in S_{\mathbb{C}}^{n-1}$, let $\pi_z : [n] \to [n]$ be a permutation which arranges the absolute values of the coordinates of z in a nonincreasing order. For $1 \leq m \leq m' \leq n$, denote by $z_{[m:m']} \in \mathbb{C}^n$ the vector with coordinates

$$z_{[m:m']}(j) := z(j) \cdot \mathbf{1}_{[m:m']}(\pi_z(j)).$$

In other words, we include in $z_{[m:m']}$ the coordinates of z which take places from m to m' in the nonincreasing rearrangement of its absolute values. For $\alpha < 1$ and $m \leq n$, define the set of vectors with dominated tail as follows:

$$\mathrm{Dom}(m, \alpha) := \{z \in S_{\mathbb{C}}^{n-1} \mid \|z_{[m+1:n]}\|_2 \leq \alpha\sqrt{m}\|z_{[m+1:n]}\|_\infty\}.$$

Note that by definition, $\mathrm{Sparse}(m) \cap S_{\mathbb{C}}^{n-1} \subset \mathrm{Dom}(m, \alpha)$, since for m-sparse vectors, $z_{[m+1:n]} = 0$.

While studying the behavior of $s_{\min}$ of real shifts of A_n in [9], we noted that the control of the infimum over compressible and dominated vectors can be extended when they are viewed as subsets of $S_{\mathbb{C}}^{n-1}$ (cf. [9], Remark 3.10). So we only need to control the infimum over vectors that are neither compressible nor dominated. The infimum over the incompressible vectors is tackled by associating it with the average distance of a column of the matrix A_n from the subspace spanned by the rest of the columns. We use the following result.

LEMMA 2.7 (Invertibility via distance [34], Lemma 3.5). *Let $\tilde{A}_n$ be any $n \times n$ random matrix. For $j \in [n]$, let $\tilde{A}_{n,j} \in \mathbb{C}^n$ be the jth column of $\tilde{A}_n$, and let $H_{n,j}$ be the subspace of $\mathbb{C}^n$ spanned by $\{\tilde{A}_{n,i}, i \in [n]\backslash\{j\}\}$. Then for any $\varepsilon, \rho > 0$, and $M < n$,*

$$\mathbb{P}\Bigg(\inf_{z \in \mathrm{Incomp}(M,\rho)} \|\tilde{A}_n z\|_2 \le \varepsilon \rho^2 \sqrt{\frac{p}{n}}\Bigg) \le \frac{1}{M} \sum_{j=1}^{n} \mathbb{P}\big(\mathrm{dist}(\tilde{A}_{n,j}, H_{n,j}) \le \rho\sqrt{p}\varepsilon\big). \tag{2.6}$$

We should mention here that Lemma 2.7 can be extended to the case when the event on the LHS of (2.6) is intersected with an event Ω, and in that case Lemma 2.7 continues to hold if the RHS of (2.6) is replaced by intersecting each of the event under the summation sign with the same event Ω (see also [9], Remark 2.5). We will actually use this generalized version of Lemma 2.7.

In order to apply Lemma 2.7 in our set-up, denote by $B^{D,n-1}$ the $(n-1) \times n$ matrix obtained by collecting the last $(n-1)$ rows of $(A_n + D_n)^{\mathsf{T}}$. Hereafter, for brevity, we will often write B^D instead of $B^{D,n-1}$. We note that any unit vector z such that $B^D z = 0$ is a vector normal to the subspace spanned by the last $(n-1)$ columns of $(A_n + D_n)$. Thus, applying Lemma 2.7 and the fact that the columns of $\bar{A}_n$ are i.i.d., we see that it is enough to find bounds on $\langle A^D_{n,1}, z\rangle$, such that $B^D z = 0$, where $A^D_{n,1}$ is the first column of $(A_n + D_n)$.

The small ball probability bounds on $\langle A^D_{n,1}, z\rangle$ depend on the additive structure of the vector z. Following [9], we see that with high probability, we can assume that any $z \in \mathrm{Ker}(B^D)$ is neither compressible nor dominated, where $\mathrm{Ker}(B^D) := \{u \in \mathbb{C}^n : B^D u = 0\}$. Therefore, it is enough to obtain estimates on the small ball probability for incompressible and nondominated vectors. To this end, we define the following notion of *Lévy concentration function*.

DEFINITION 2.8. Let Z be a random variable in $\mathbb{C}^n$. For every $\varepsilon > 0$, the Lévy concentration function of Z is defined as

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

The Berry–Esséen bound of [37], Theorem 2.2.17, yields a weak control on Lévy concentration function which is enough to prove Theorem 2.2(i). To prove Theorem 2.2(ii), a significant amount of additional work is needed which is the key contribution of this paper.

To obtain a strong probability bound on the Lévy concentration function, the standard approach is to first quantify the additive structure present in an incompressible vector via the definition of *least common denominator* (LCD). When the LCD is large, one can derive a good bound on the Lévy concentration function

using Esséen's inequality [22] (see also [41], Theorem 6.3). However, Esséen's inequality does not yield a strong small ball probability estimate for vectors with small values of LCD. Nevertheless, these vectors are shown to admit a special net of small cardinality and, therefore, one can still apply the union bound to complete the proof; for example, see [9, 34, 35]. One would hope to carry out the same program here. However, when we view the incompressible and nondominated vectors of small LCD as a subset of $S_{\mathbb{C}}^{n-1}$, its real dimension is twice as large as in the proof [9], Proposition 4.1. On the other hand, for the real-valued random variables in A_n, one does not expect to obtain better control on the Lévy concentration function. Thus the proof of [9], Proposition 4.1, breaks down as the bounds on the Lévy concentration function and the size of the net do not match (see also [9], Remark 4.5).

To tackle this obstacle, we decompose the vectors according to the angle between their real and imaginary parts. More precisely, we define the *real-imaginary decorrelation* as follows.

DEFINITION 2.9. Let $z \in \mathbb{C}^m$ for some positive integer m. Denote $V := V(z) := \binom{x^{\mathsf{T}}}{y^{\mathsf{T}}}$, where $z = x + \mathrm{i}y$. Then we denote the real-imaginary de-correlation of z by

$$d(z) := \bigl(\det\bigl(VV^{\mathsf{T}}\bigr)\bigr)^{1/2}.$$

This notion of real-imaginary decorrelation was introduced in [36] to study the no-gap delocalization property of the eigenvectors of a wide class of random matrices. In [36], it is termed as "real-complex correlation." Here we deviate from that terminology upon noting that small values of $d(z)$ indicate that the real and the imaginary part of z are close to being linearly dependent.

If a vector $z \in S_{\mathbb{C}}^{n-1}$ has a large value of $d(z)$, then we call this vector *genuinely complex*, whereas vectors with small real-imaginary decorrelations are termed *essentially real* vectors (see (5.1) and (6.1) for a precise formulation). The real and imaginary parts of essentially real vectors being almost linearly dependent it allows us to construct a net whose cardinality is a polynomial of degree n in terms of the mesh. Therefore, one can use the small ball probability estimates from [9] to show that with high probability, there does not exist any essentially real vector in the kernel of B^D with a small LCD.

The analysis of genuinely complex vectors is more delicate. Following the recent work of [36], we define a notion of a *two-dimensional* LCD. Roughly speaking the two-dimensional LCD $D_2(\cdot)$ of $z = x + \mathrm{i}y$ identifies a nontrivial $\theta^\star(z) := (\theta_1^\star(z), \theta_2^\star(z)) \in \mathbb{R}^2$ such that $\theta_1^\star(z)x + \theta_2^\star(z)y$ is close to an integer point and $\theta^\star(z)$ has the least possible Euclidean norm among all such choices. See Definition 4.3 for a precise formulation. Using a result of [36] (Theorem 7.5 there), we show that the small ball probability bound of genuinely complex vectors decays roughly as the inverse of the $(2n)$th power of $D_2(\cdot)$ (see the bound in (5.7)). This

probability bound balances the cardinality of the net of genuinely complex vectors for which $\Delta(z_{\text{small}}) := \|\theta_1^\star(z_{\text{small}})x_{\text{small}} + \theta_2^\star(z_{\text{small}})y_{\text{small}}\|_2$ (precise definition of $\Delta(\cdot)$ can be found in Definition 4.4) is large, where $z_{\text{small}} = x_{\text{small}} + \mathrm{i}y_{\text{small}}$ is the part of z containing the coordinates of small modulus. It allows us to take the union bound over the net of such vectors. To treat the remaining set of genuinely complex vectors, using results from [9], we show that, with high probability, there cannot exist a vector $z \in \mathrm{Ker}(B^D)$ with a dominated real part. This additional observation then shows that for any $z \in \mathrm{Ker}(B^D)$ the quantity $\Delta(z_{\text{small}})$ must also be large. This completes the outline of the proof of Theorem 2.2.

2.2. *Intermediate singular values.* We also need to show that there are not too many singular values of $(A_n - w\sqrt{np}I_n)$ in a small interval around zero. The following theorem does that job. Before stating the theorem, for $i \in [n]$, let us denote $s_i(\cdot)$ to be the ith largest singular value.

THEOREM 2.10. *Let A_n be an $n \times n$ matrix whose entries are $\{\xi_{i,j}\delta_{i,j}\}_{i,j=1}^n$ where $\{\xi_{i,j}\}_{i,j=1}^n$ are i.i.d. real-valued random variables with zero mean and unit variance, and $\{\delta_{i,j}\}_{i,j=1}^n$ are i.i.d.* $\mathrm{Ber}(p_n)$ *random variables. There exist constants $c_{2.10}$ and $C_{2.10}$[3] such that the following holds: Let $\psi : \mathbb{N} \mapsto \mathbb{N}$ be such that $\psi(n) < n$ and $\min\{p_n\psi(n), \psi^2(n)/n\} \geq C_{2.10}\log n$. Then for any $w \in B_{\mathbb{C}}(0,1)$ we have*

$$\mathbb{P}\left(\bigcup_{i=3\psi(n)}^{n-1}\left\{s_{n-i}\left(\frac{A_n}{\sqrt{np_n}} - wI_n\right) \leq c_{2.10}\frac{i}{n}\right\}\right) \leq \frac{2}{n^2}.$$

To prove Theorem 2.10, we follow the approach of [40], which was adapted to the sparse case in [13, 44]. We first show that the distance of any row of A_n from any given subspace of not very large dimension cannot be too small with large probability. This observation together with a result from [40] completes the proof.

2.3. *Weak convergence.* Recall that to show the integrability of $\log(\cdot)$ we further need to establish the weak convergence of the empirical measure of the singular values of $\frac{1}{\sqrt{np}}A_n - wI_n$. Define

$$\boldsymbol{A}_n^w := \begin{bmatrix} 0 & \frac{1}{\sqrt{np}}A_n - wI_n \\ \frac{1}{\sqrt{np}}A_n^* - \bar{w}I_n & 0 \end{bmatrix} \tag{2.7}$$

and denote by ν_n^w the ESD of $\boldsymbol{A}_n^w$. It can be easily checked that ν_n^w is the symmetrized version of the empirical measure of the singular values $\frac{1}{\sqrt{np}}A_n - wI_n$. Thus, it is enough to prove the weak convergence of ν_n^w.

[3]The constants $c_{2.10}$ and $C_{2.10}$ can potentially depend on the tail of the distribution of $\{\xi_i\}_{i=1}^n$.

THEOREM 2.11. (i) *Let A_n be an $n \times n$ matrix with entries $a_{i,j} = \delta_{i,j} \cdot \xi_{i,j}$, where $\delta_{i,j}$ are i.i.d. Bernoulli random variables with $\mathbb{P}(\delta_{i,j} = 1) = p_n$, and $\xi_{i,j}$ are centered i.i.d. real-valued random variables with unit variance. Assume $p_n = \omega(\frac{\log n}{n})$. Fix any $w \in B_{\mathbb{C}}(0,1)$. Then there exists a probability measure ν_∞^w such that ν_n^w converges weakly to ν_∞^w, in probability.*

(ii) *If additionally $\{\xi_{i,j}\}_{i,j=1}^n$ have finite fourth moment and $\sum_{n=1}^\infty (n^2 p_n)^{-1} < \infty$ then the above convergence holds almost surely.*

To prove Theorem 2.11, we first apply a standard truncation technique which shows that it is enough to prove the weak convergence of ν_n^w to ν_∞^w for bounded $\{\xi_{i,j}\}_{i,j=1}^n$ (see Lemma 9.1). This truncation argument requires the additional assumptions of part (ii) of Theorem 2.11 to establish the almost sure convergence.

It is well known that $\nu_{G_n}^w$, the symmetrized version of the empirical law of the singular values of $\frac{1}{\sqrt{n}} G_n - w I_n$, where G_n is a complex Ginibre matrix, converges weakly, almost surely to ν_∞^w. Therefore, to obtain Theorem 2.11 it is enough to show that for bounded $\{\xi_{i,j}\}_{i,j=1}^n$, the signed measure $\nu_n^w - \nu_{G_n}^w$ converges weakly to the point mass to zero, as $n \to \infty$, in probability or almost surely, depending on the sparsity parameter p.

This is done in Section 9 using the following two-fold argument. First, we establish that both the random probability measures ν_n^w and $\nu_{G_n}^w$ are close to their expectations, denoted hereafter by $\mathbb{E}\nu_n^w$ and $\mathbb{E}\nu_{G_n}^w$, respectively. This step uses standard concentration inequalities for the spectral measure of Hermitian random matrices.

To complete the proof of Theorem 2.11, we then need to show that $\mathbb{E}\nu_n^w$ and $\mathbb{E}\nu_{G_n}^w$ are themselves close to each other. Here we appeal to the Lindeberg replacement principle which was introduced to the random matrix theory in [17, 18] to prove the semicircle law for random symmetric matrices with exchangeable entries on and above the diagonal. Subsequently, this technique has been used on numerous occasions in the random matrix theory literature.

3. The structure of the kernel: Vectors with nondominated real part. Recall from Section 2.1 that the main challenge in proving Theorem 2.2 is to show that there does not exist a genuinely complex vector $z \in \mathrm{Ker}(B^D)$ with a small two-dimensional LCD. As a first step, we show that for any $z \in \mathrm{Ker}(B^D)$, its real part must have a nondominated component with high probability. This is shown in the following result, which is the main result of this section. Before stating the result, let us introduce some notation, which is borrowed from [36]: For a number $M < n/2$, to be determined during the course of the proof, we denote by $\mathrm{small}(z)$ the set of the $(n-M)$ coordinates of z having the smallest absolute values. The ties are broken arbitrarily. We also write $z_{\mathrm{small}} = x_{\mathrm{small}} + \mathrm{i} y_{\mathrm{small}} := z_{\mathrm{small}(z)}$. Hereafter, we drop the subscript in p_n and for ease we write p instead.

PROPOSITION 3.1. *Let A_n be a matrix with i.i.d. entries $a_{i,j} = \xi_{ij}\delta_{ij}$, where $\{\xi_{ij}\}$ are i.i.d. centered real-valued random variables with unit variance and finite fourth moment, and $\{\delta_{ij}\}$ are i.i.d.* Ber(p) *random variables. Set*

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

Fix $r \in (0,1]$ and $R \geq 1$ such that $\mathrm{Im}(D_n) = r'\sqrt{np}I_n$ with $|r'| \in [r,1]$ and $\|D_n\| \leq R\sqrt{np}$. Fix another positive real $K \geq 1$. Then there exist constants $C_{3.1}, \widetilde{C}_{3.1}, c_{3.1}$ and $\bar{c}_{3.1}$, depending only on r, R, K and the fourth moment of $\{\xi_{ij}\}$, such that the following holds. Denote $\rho := (\widetilde{C}_{3.1}(K+R))^{-\ell_0-6}$ and assume that

$$c_{3.1}\rho^5 pn > 1. \tag{3.1}$$

Set $M = C_{3.1}\rho^{-4}p^{-1}$. Then

$$\mathbb{P}\left(\exists z \in \mathrm{Ker}(B^D) \cap S_{\mathbb{C}}^{n-1} : \left\| \frac{x_{\mathrm{small}}}{\|x_{\mathrm{small}}\|_2} \right\|_\infty \geq \rho p^{1/2} \text{ and } \|A_n\| \leq K\sqrt{np}\right) \leq \exp(-\bar{c}_{3.1}np).$$

REMARK 3.2. For clarity, we only prove Proposition 3.1 for $r' \in [r,1]$. It will be evident that the proof of the case $r' \in [-1,-r]$ is exactly the same. We spare the details.

The key to the proof of Proposition 3.1 is in showing that if the real part of a vector z is compressible then $\|B^D z\|_2$ cannot be too small. This is derived in the following lemma.

LEMMA 3.3. *Let B^D, A_n, ρ, K, R, r and r' be as in Proposition 3.1. Then there exist constants $0 < c_{3.3}, c'_{3.3}, c''_{3.3}, \overline{c}_{3.3} < \infty$, depending only on K, R, r and the fourth moment of $\{\xi_{i,j}\}$, such that for any $p^{-1} \leq M \leq c'_{3.3}n/\log(1/\rho)$,*

$$\mathbb{P}\big(\exists z = x + \mathrm{i}y \in S_{\mathbb{C}}^{n-1} : \|B^D z\|_2 \leq c_{3.3}\rho\sqrt{np}, \|x_{\mathrm{small}}\|_2 \leq c''_{3.3}\rho, \text{ and } \|A_n\| \leq K\sqrt{np}\big) \leq \exp(-\overline{c}_{3.3}np).$$

To prove Lemma 3.3, we borrow results from [9]. In [9], Proposition 3.1, we showed that, with high probability, there does not exist any real-valued compressible or dominated vector z such that $\|\mathrm{Re}(B^D)z\|_2$ is small, where $\mathrm{Re}(B^D)$ denotes the real part of the matrix B^D. In [9], Remark 3.10, it was also argued that the same conclusion holds for $\|B^D z\|_2$ when z is now allowed to be complex valued. We will need this result to prove Lemma 3.3. For completeness, we state it below.

PROPOSITION 3.4 ([9], Proposition 3.1, Remark 3.10). *Let A_n be as in Proposition 3.1. Fix $K, R \geq 1$, and assume that D_n is a nonrandom diagonal matrices with complex entries such that $\|D_n\| \leq R\sqrt{pn}$. Then there exist constants*

$0 < c_{3.4}, \overline{c}_{3.4}, c'_{3.4}, C_{3.4}, \widetilde{C}_{3.4}, \overline{C}_{3.4} < \infty$, *depending only on* K, R *and the fourth moment of* $\{\xi_{ij}\}$, *such that for*

$$\frac{\overline{C}_{3.4}\log n}{n} \le p \le \frac{1}{10}, \tag{3.2}$$

and any $p^{-1} \le M \le c_{3.4}n$, *we have*

$$\begin{aligned}\mathbb{P}\big(&\exists z \in \mathrm{Dom}\big(M, \big(C_{3.4}(K+R)\big)^{-4}\big) \cup \mathrm{Comp}(M,\rho)\\ &\big\|(A_n + D_n)z\big\|_2 \le c'_{3.4}(K+R)\rho\sqrt{np} \text{ and } \|A_n\| \le K\sqrt{pn}\big)\\ &\le \exp(-\overline{c}_{3.4}pn),\end{aligned}$$

where $\rho = (\widetilde{C}_{3.4}(K+R))^{-\ell_0-6}$ *and* ℓ_0 *are as in Proposition* 3.1.

Observe that Proposition 3.4 is stated for the square matrix A_n. To prove Lemma 3.3, we need a version of Proposition 3.4 for $(n-1)\times n$ matrices. As noted in [9], Remark 3.9, this follows from an easy adaptation. So, without loss of generality we will use Proposition 3.4 also for $(n-1)\times n$ matrices. The final ingredient for the proof of Lemma 3.3 is an estimate on the Lévy concentration function for incompressible and nondominated vectors. Such an estimate was derived in [9], Corollary 3.7, for real valued vectors and matrices with zero diagonal and i.i.d. off-diagonal entries. One can investigate its proof to convince oneself that the same proof works for complex valued vectors and matrices with i.i.d. entries. We state this modified version below.

LEMMA 3.5 ([9], Corollary 3.7). *Let* A_n *be as in Proposition* 3.1. *Then for any* $\alpha > 1$, *there exist* $\beta, \gamma > 0$, *depending on* α *and the fourth moment of* $\{\xi_{ij}\}$, *such that for* $z \in \mathbb{C}^n$, *satisfying* $\|z\|_\infty/\|z\|_2 \le \alpha\sqrt{p}$, *we have*

$$\mathcal{L}\big(A_n z, \beta\cdot\sqrt{pn}\|z\|_2\big) \le \exp(-\gamma n).$$

We now proceed to the proof of Lemma 3.3.

PROOF OF LEMMA 3.3. The proof is based on ideas from [23]. For ease of writing, let us write $c_0 := (C_{3.4}(K+R))^{-4}$. We also denote

$$\begin{aligned}\Omega_{D,C} := \big\{&\mathrm{Re}\big(B^D\big) : \exists z \in \mathrm{Dom}(M, c_0) \cup \mathrm{Comp}(M,\rho)\\ &\big\|B^D z\big\|_2 \le c'_{3.4}(K+R)\rho\sqrt{np} \text{ and } \|A_n\| \le K\sqrt{np}\big\}.\end{aligned}$$

Using Proposition 3.4, we see that $\mathbb{P}(\Omega_{D,C}) \le \exp(-\overline{c}_{3.4}np)$. We now make the following claim.

CLAIM. *Fix any* $J \subset [n]$ *of cardinality* M *and let*

$$\mathcal{Z}'_J := \big\{z = x + \mathrm{i}y : \|x_{\mathrm{small}}\|_2 \le c''\rho \text{ and } \mathrm{supp}(x_{[1:M]}) \subset J\big\},$$

for some small constant c'' to be determined later. Then

$$\mathbb{P}\big(\{\exists z \in \mathcal{Z}'_J \text{ such that } \|B^D z\|_2 \le c\rho\sqrt{np}\} \cap \Omega^c_{D,C}\big) \le \exp(-\bar{c}n),$$

for some small constants c and $\bar{c}$.

The conclusion of the lemma immediately follows from the claim by taking a union bound over $J \subset [n]$, such that $|J| = M$. Thus we now only need to prove the claim.

To prove this claim, we will first show that if $z \in \mathcal{Z}'_J$ such that $\|B^D z\|_2$ is small, then y, the imaginary part of z, belongs to a small neighborhood of a linear image of the subspace spanned by the largest M coordinates of x, the real part of z. This together with the fact that $\|x_{\text{small}}\|_2$ is small enables us to obtain a net of $\mathcal{Z}'_J$ with small cardinality. Finally, using the estimate on Lévy concentration function of Lemma 3.5 and the union bound, we complete the proof of the claim. Below we carry out the details.

Fix any $J \subset [n]$ and let $\mathrm{Re}(B^D)|_J$ denote the submatrix induced by the columns of $\mathrm{Re}(B^D)$ indexed by J. We first condition on a realization of $\mathrm{Re}(B^D)|_J$ and show that for every such realization the conditional probability of the event in the claim is less than $e^{-\bar{c}n}$. Then taking an average over the realizations of $\mathrm{Re}(B^D)|_J$, the proof will be completed.

So let us assume that $z \in \mathcal{Z}'_J$ be such that $\|B^D z\|_2 \le c\rho\sqrt{np}$. Then we see that

$$\big\|\mathrm{Re}(B^D)x - \mathrm{Im}(B^D)y\big\|_2 \le \|B^D z\|_2 \le c\rho\sqrt{np}. \tag{3.3}$$

Notice that $\|x_{[M+1:n]}\|_2 \le \|x_{\text{small}}\|_2$ as $x_{[M+1:n]}$ consists of the smallest in the absolute value coordinates of x. Since

$$\big\|\mathrm{Re}(B^D)|_{J^c}\big\| \le \big\|\mathrm{Re}(B^D)\big\| \le \|B^D\| \le \|A_n\| + \|D_n\| \le (K+R)\sqrt{np},$$

applying the triangle inequality we further deduce that

$$\begin{aligned}\big\|\mathrm{Im}(B^D)y - \mathrm{Re}(B^D)|_J x_{[1:M]}\big\|_2 &\le c\rho\sqrt{np} + \big\|\mathrm{Re}(B^D)|_{J^c} x_{[M+1:n]}\big\|_2 \\ &\le c\rho\sqrt{np} + \big\|\mathrm{Re}(B^D)|_{J^c}\big\| \cdot \|x_{\text{small}}\|_2 \\ &\le 2c\rho\sqrt{np},\end{aligned}$$

where in the last step we chose c'' so that $c''(K+R) \le c$.

Hereafter, we write $\mathrm{Im}(B^D)$ to denote the imaginary part of the matrix B^D. Hence $\mathrm{Im}(B^D)$ is a $(n-1) \times n$ matrix whose first column is zero and the last $(n-1)$ columns form a diagonal matrix whose entries are all equal to $r'\sqrt{np}$. Therefore, denoting $y|_{[2:n]}$ to be the $(n-1)$ dimensional vector consisting of the last $(n-1)$ coordinates of y we further have that

$$\left\| y|_{[2:n]} - \frac{1}{r'\sqrt{np}}\,\mathrm{Re}(B^D)|_J x_{[1:M]} \right\|_2 \le 2r'^{-1}c\rho \le 2r^{-1}c\rho. \tag{3.4}$$

Thus (3.4) implies that the vector $y|_{[2:n]}$ belongs to a $(2r^{-1}c\rho)$-neighborhood of the linear subspace $\mathscr{E}'_J := \operatorname{span}(\operatorname{Re}(B^D)|_J\mathbb{R}^J) \subset \mathbb{R}^{n-1}$. Since $\|x_{[M+1:n]}\|_2 \le c''\rho \le r^{-1}c\rho$, we have that for any $z \in \mathcal{Z}'_J$, such that $\|B^D z\|_2 \le c\rho\sqrt{np}$, belongs to a $(3r^{-1}c\rho)$-neighborhood of the set

$$\mathscr{E}_J := \{x + \mathrm{i}y : \operatorname{supp}(x) \subset J,\ y|_{[2:n]} \in \mathscr{E}'_J,\ y_1 \in [-1,1]\},$$

with $\dim(\mathscr{E}_J) \le 2M+1$. Since $\mathcal{Z}'_J \subset S^{n-1}_{\mathbb{C}}$, applying the triangle inequality and choosing $c \le r/3$ we further see that every vector in $z \in \mathcal{Z}'_J$, such that $\|B^D z\|_2 \le c\rho\sqrt{np}$, belongs to a $(3r^{-1}c\rho)$-neighborhood of $(2B^n_{\mathbb{C}}) \cap \mathscr{E}_J$. Therefore, we can choose a $(r^{-1}c\rho)$-net $\mathcal{N} \subset (2B^n_{\mathbb{C}}) \cap \mathscr{E}_J$ of cardinality

$$|\mathcal{N}| \le \left(\frac{12}{c\rho}\right)^{2M+1} \le \exp\bigl(3M\log(12/(c\rho))\bigr). \tag{3.5}$$

Note that, using the triangle inequality we see that $\mathcal{N}$ is $(4r^{-1}c\rho)$-net of the set of all vectors $z \in \mathcal{Z}'_J$ such that $\|B^D z\|_2 \le c\rho\sqrt{np}$. Thus, for a $z \in \mathcal{Z}'_J$ with $\|B^D z\|_2 \le c\rho\sqrt{np}$, there must exist at least one $w \in \mathcal{N}$ such that $\|B^D w\|_2 \le 5r^{-1}(K+R)c\rho\sqrt{np}$. Now shrink c such that $10r^{-1}c \le c'_{3.4}$. With this choice of the constant c, we see that $w \notin \operatorname{Dom}(M, c_0) \cup \operatorname{Comp}(M, \rho)$ on the event $\Omega^c_{D,C}$.

However, for any $w \notin \operatorname{Dom}(M, c_0)$ we have

$$\frac{\|w_{[M+1:n]}\|_\infty}{\|w_{[M+1:n]}\|_2} \le (c_0\sqrt{M})^{-1} \le c_0^{-1}\sqrt{p},$$

where in the last step we used the fact that $M \ge p^{-1}$. Thus applying Lemma 3.5, there exists constants $c_\star$ and $\overline{c}$ such that

$$\begin{aligned}\mathbb{P}\bigl(\|B^D w\|_2 \le c_\star\rho\sqrt{np}\,\big|\,\operatorname{Re}(B^D)|_J\bigr) &\\ \le \mathcal{L}\bigl(\operatorname{Re}(B^D)|_{J^c} w_{[M+1:n]}, c_\star\|w_{[M+1:n]}\|_2\sqrt{np}\bigr) &\le \exp(-2\overline{c}n).\end{aligned}$$

Hence, by the union bound,

$$\mathbb{P}\bigl(\exists w \in \mathcal{N} : \|B^D w\|_2 \le c_\star\rho\sqrt{np}\,\big|\,\operatorname{Re}(B^D)|_J\bigr) \le |\mathcal{N}| \cdot \exp(-2\overline{c}n) \le \exp(-\overline{c}n),$$

where the last step follows from the bound (3.5) and the fact that $M\log(1/\rho) < c'n$ for a sufficiently chosen small constant c'. Thus shrinking c again such that $5r^{-1}(K+R)c \le c_\star$ we obtain that

$$\mathbb{P}\bigl(\{\exists z \in \mathcal{Z}'_J \text{ such that } \|B^D z\|_2 \le c\rho\sqrt{np}\} \cap \Omega^c_{D,C}\,\big|\,\operatorname{Re}(B^D)|_J\bigr) \le \exp(-\overline{c}n).$$

Finally, taking an average over all the realizations of $\operatorname{Re}(B^D)|_J$ completes the proof. □

We are now ready to prove Proposition 3.1.

PROOF OF PROPOSITION 3.1. Let $z \in \mathrm{Ker}(B^D) \cap S_{\mathbb{C}}^{n-1}$. Assume that the event $\Omega_{D,C}^c$ defined above occurs, so $z \notin \mathrm{Dom}(M, c_0) \cup \mathrm{Comp}(M, \rho)$. We note that if $c_{3.1}$ is chosen sufficiently small then the assumption (3.1) implies that

$$\frac{M \log(1/\rho)}{c'_{3.3} n} < 1,$$

whenever n is large enough. So Lemma 3.3 can be applied, which implies that, with high probability, $\|x_{\mathrm{small}}\|_2 > c''_{3.3}\rho$. On the other hand,

$$\|x_{\mathrm{small}}\|_\infty \le \|z_{\mathrm{small}}\|_\infty \le \frac{1}{c_0\sqrt{M}}.$$

Combining the last two inequalities, we show that on the event $\Omega_{D,C}^c$,

$$\frac{\|x_{\mathrm{small}}\|_\infty}{\|x_{\mathrm{small}}\|_2} \le \frac{1}{c''_{3.3}\rho c_0\sqrt{M}}, \tag{3.6}$$

and the result follows upon choosing $C_{3.1}$ sufficiently large. □

REMARK 3.6. Note that the inequality (3.6) continues to hold even if the constant $C_{3.1}$ is increased without changing other constants $\widetilde{C}_{3.1}$, $c_{3.1}$ and $\bar{c}_{3.1}$, appearing in Proposition 3.1. This implies that, if needed, we can arbitrarily increase the constant $C_{3.1}$. This observation will be used later in the paper.

4. Net construction: Genuinely complex case. In this section we show that the set of genuinely complex vectors admits a net of small cardinality. We begin with the relevant definitions.

DEFINITION 4.1. For $y > 0$, denote $\log_1(y) := \log y \cdot \mathbb{I}(y \ge e)$. Fixing $L \ge 1$, for a nonzero vector $x \in \mathbb{R}^m$, we set

$$D_1(x) := \inf\left\{\theta > 0 : \mathrm{dist}(\theta x, \mathbb{Z}^m) < 2^5 L\sqrt{\log_1 \frac{\|\theta x\|_2}{2^6 L}}\right\}. \tag{4.1}$$

If V is a $2 \times m$ matrix, define

$$D_2(V) := \inf\left\{\|\theta\|_2 : \theta \in \mathbb{R}^2, \mathrm{dist}(V^{\mathsf{T}}\theta, \mathbb{Z}^m) < L\sqrt{\log_1 \frac{\|V^{\mathsf{T}}\theta\|_2}{2^8 L}}\right\}. \tag{4.2}$$

We will call the first version of the LCD one-dimensional, and the second one two-dimensional. Note that $D_1(\cdot)$ matches with the definition of the LCD used in [9] up to constants.

REMARK 4.2. The different powers of 2 appearing in the definitions (4.1) and (4.2) play only a technical role. They do not affect most of the proof, and they will be needed in Section 6 to compare the one and the two-dimensional LCD for almost real vectors (see Lemma 6.1).

Observe that $D_1(\cdot)$ and $D_2(\cdot)$ are defined for real-valued vectors and matrices, respectively. However, both these notions can be extended for complex valued vectors by the following simple adaptation.

DEFINITION 4.3. Consider a complex vector $z = x + \mathrm{i}y \in \mathbb{C}^m$. Denote $\tilde{z} := \tilde{z}(z) := \binom{x}{y} \in \mathbb{R}^{2m}$, and define a $2 \times m$ matrix $V := V(z) := \binom{x^{\mathsf{T}}}{y^{\mathsf{T}}}$. Using these two different representations of $z \in \mathbb{C}^m$, we now define

$$D_2(z) := D_2(V) \quad \text{and} \quad D_1(z) := D_1(\tilde{z}).$$

Let us assume that the infimum in (4.2) is attained at $\theta^\star$. Then from Definition 4.1, we have that $D_2(V)$ equals $\|\theta^\star\|_2$. We will see below that the cardinality of the desired net for the genuinely complex vectors also depends on $\|V^{\mathsf{T}}\theta^\star\|_2$. However, the infimum in (4.2) is not always achieved. Hence, we have the following definition.

DEFINITION 4.4. For a real-valued $2 \times m$ matrix V, define

$$\Delta(V) := \liminf_{\tau \to 1+} \Big\{ \|V^{\mathsf{T}}\theta\|_2 : \operatorname{dist}(V^{\mathsf{T}}\theta, \mathbb{Z}^m) < L\sqrt{\log_1 \frac{\|V^{\mathsf{T}}\theta\|_2}{2^8 L}}, \\ \|\theta\|_2 \le \tau D_2(V) \Big\}.$$

As before, for a $z \in \mathbb{C}^m$, we define $\Delta(z) := \Delta(V)$ where $V = V(z)$. For later use, let us note that for any $z \in S_{\mathbb{C}}^{m-1}$,

$$d(z)D_2(z) \le \Delta(z) \le D_2(z), \tag{4.3}$$

where $d(z)$ denotes the real-imaginary decorrelation of z appearing in Definition 2.9. Indeed, the inequalities (4.3) are immediate from the fact that the singular values of V^{T} are bounded by one and $d(z)$ is the product of the singular values of V^{T}.

REMARK 4.5. We take $L = (\delta_0 p)^{-1/2}$, where $\delta_0 \in (0, 1)$ is a universal constant as in [9], Remark 2.7.

Equipped with the above definitions, we consider the following simple reduction. Fix $M < n/2$, $z \in S_{\mathbb{C}}^{n-1}$, and let $J = \operatorname{small}(z)$. It can be easily verified that for any $z \in \mathbb{C}^n$ there exists a $\tau \in [0, 2\pi)$ such that $z_J = e^{\mathrm{i}\tau}(w_1 + \mathrm{i}w_2)$, where $w_1 \perp w_2$ and $\|w_2\|_2 \le \|w_1\|_2$. As $z_J \in \operatorname{Ker}(B)$ if and only if $e^{-\mathrm{i}\tau} z_J \in \operatorname{Ker}(B)$, without loss of generality, we can only consider the following set:

$$\mathcal{Z} := \big\{ z \in S_{\mathbb{C}}^{n-1} \setminus \big(\operatorname{Dom}\big(M, \big(C_{3.4}(K+R)\big)^{-4}\big) \cup \operatorname{Comp}(M, \rho) \big) : \\ z_{\text{small}} = w_1 + \mathrm{i}w_2,\ w_1 \perp w_2,\ \|w_1\|_2 \ge \|w_2\|_2 \big\} \tag{4.4}$$

instead of $S_{\mathbb{C}}^{n-1}\setminus(\mathrm{Dom}(M,(C_{3.4}(K+R))^{-4})\cup\mathrm{Comp}(M,\rho))$. Therefore, our revised goal is to show that the set of genuinely complex vectors, when specialized to $\mathcal{Z}$, admits a net of small cardinality.

To this end, fixing a set $J\subset[n]$, we start with constructing a small net for the set of pairs (ϕ,ψ) with $\phi\perp\psi$ in the unit sphere of $\mathbb{R}^J\times\mathbb{R}^J$ for which the value of the two-dimensional LCD, the auxiliary parameter $\Delta(\cdot)$, and the decorrelation $d(\phi,\psi)=\|\phi\|_2\|\psi\|_2$ are approximately constant. The condition on the two-dimensional LCD means that there exists a linear combination of the vectors ϕ and ψ which is close to an integer point. Our aim is to use this linear combination to construct separate approximations of ϕ and ψ.

For any $\gamma>0$, let us denote $\mathbb{Z}_\gamma^J:=\mathbb{Z}^J\cap\gamma B_2^{|J|}$. Using a simple volumetric comparison argument, we have following estimate on $|\mathbb{Z}_\gamma^J|$:

$$|\mathbb{Z}_\gamma^J|\le\left(C_0\left(\frac{\gamma}{\sqrt{|J|}}+1\right)\right)^{|J|}, \tag{4.5}$$

for some absolute constant C_0. The main technical result of this section is the following lemma.

LEMMA 4.6. *Let $d\in(0,1)$, and $0<\alpha\le d\mathfrak{D}\le\Delta\le\mathfrak{D}$. Define the set*

$$\begin{aligned}S_J(\mathfrak{D},\Delta,d):=\big\{&(\phi,\psi)\in\mathbb{R}^J\times\mathbb{R}^J:\phi\perp\psi,\ \|\phi\|_2\in[1/2,1],\ \|\psi\|_2\in[d,3d]\\&\exists\zeta\in\mathbb{R}^2\|\zeta\|_2\in[\mathfrak{D},2\mathfrak{D}],\ \|\zeta_1\phi+\zeta_2\psi\|_2\in[\Delta,2\Delta],\\&\text{and }\operatorname{dist}(\zeta_1\phi+\zeta_2\psi,\mathbb{Z}^J)<\alpha\big\}.\end{aligned}$$

Then there exists a $(\frac{C_{4.5}\alpha}{\mathfrak{D}})$-net $\mathcal{M}_J(\mathfrak{D},\Delta,d)\subset S_J(\mathfrak{D},\Delta,d)$ such that

$$|\mathcal{M}_J(\mathfrak{D},\Delta,d)|\le\left(\bar{C}_{4.5}\frac{d\mathfrak{D}^2}{\alpha}\cdot\left(\frac{1}{\sqrt{|J|}}+\frac{1}{\Delta}\right)\right)^{|J|}\cdot\left(\frac{\mathfrak{D}}{\alpha}\right)^2,$$

for some absolute constants $C_{4.5}$, and $\bar{C}_{4.5}$.

This lemma provides a significant improvement over the standard volumetric estimate yielding $(c\mathfrak{D}^2/\alpha^2)^{|J|}$. This improved bound precisely balances the term appearing in the small ball probability estimate. Note that the bounds on $\|\phi\|_2$ and $\|\psi\|_2$ imply that the decorrelation $d(\phi,\psi)$ is approximately constant in the set $S_J(\mathfrak{D},\Delta,d)$, whereas the bounds on $\|\zeta\|_2$ and $\operatorname{dist}(\zeta_1\phi+\zeta_2\psi,\mathbb{Z}^J)$ ensure that the two-dimensional LCD and the auxiliary parameter $\Delta(\cdot)$ are approximately constant. Lemma 4.6 deals with the case when the decorrelation between ϕ and ψ is relatively large, represented by the assumption $d\ge\alpha/\mathfrak{D}$, which in turn implies that the angle between the real and the imaginary part of the vectors is nonnegligible. In Section 5, we use this criterion to formally define the notion of genuinely complex vectors.

We also point out to the reader that a net for the genuinely complex vectors was constructed in [36] (see Lemma 11.2 there). Using that net and the bound on the Lévy concentration function (see [36], Theorem 10.3), it was then showed that there are no vectors with LCD less than $O(n)$ in the kernel of the matrix in context, with high probability. Repeating the same argument here, one can at best hope to show that there does not exist any vector in $\mathrm{Ker}(B^D)$ with LCD $O(np)$, with high probability. To treat the remaining vectors again, one needs to apply bounds on Lévy concentration function (e.g., see the bound derived in Proposition 5.3). However, for such vectors the bound is too weak to deduce almost sure convergence of the ESD of $\frac{1}{\sqrt{np}}A_n$.

Hence, we need to proceed differently. In particular, we use all the parameters D, Δ, and d to find a net of appropriate size such that its cardinality balances with the small ball probability derived in Proposition 5.5 so that we are able to show that there are no vectors in $\mathrm{Ker}(B^D)$ with LCD less than $\exp(O(np\rho^4))$ with high probability.

PROOF OF LEMMA 4.6. Assume that there exists $\zeta := (\zeta_1, \zeta_2) \in \mathbb{R}^2$ and $q \in \mathbb{Z}^J$ satisfying

$$\|\zeta_1\phi + \zeta_2\psi\|_2 \in [\Delta, 2\Delta] \quad \text{and} \quad \|\zeta_1\phi + \zeta_2\psi - q\|_2 < \alpha. \tag{4.6}$$

We consider two cases depending on the size of ζ_1. Let us start with the case when this value is small. Consider the set

$$\begin{aligned} S_J^0(\mathfrak{D}, \Delta, d) := \Big\{ &(\phi, \psi) \in S_J(\mathfrak{D}, \Delta, d) : \exists (\zeta_1, \zeta_2) \in \mathbb{R}^2, \|(\zeta_1, \zeta_2)\|_2 \in [\mathfrak{D}, 2\mathfrak{D}], \\ &|\zeta_1| \le \frac{1}{2} d\mathfrak{D}, \|\zeta_1\phi + \zeta_2\psi\|_2 \in [\Delta, 2\Delta], \text{ and } \exists q \in \mathbb{Z}^J \text{ such that} \\ &\|\zeta_1\phi + \zeta_2\psi - q\|_2 < \alpha \Big\}. \end{aligned}$$

Since $d < 1$, note that the condition on ζ_1 implies that $\mathfrak{D}/2 \le |\zeta_2| \le 2\mathfrak{D}$. Hence

$$\Delta \le \|\zeta_1\phi + \zeta_2\psi\|_2 \le \frac{1}{2} d\mathfrak{D}\|\phi\|_2 + 2\mathfrak{D}\|\psi\|_2 \le 7d\mathfrak{D}. \tag{4.7}$$

We will approximate ϕ using the standard volumetric net and use (4.6) to construct a small net for ψ. To this end, consider $(\phi, \psi) \in S_J^0(\mathfrak{D}, \Delta, d)$ and let $(\zeta_1, \zeta_2) \in \mathbb{R}^2$ be the corresponding vector (i.e., for which (4.6) holds). Then, by the triangle inequality,

$$\|q\|_2 < \alpha + 2\Delta \le 3\Delta,$$

that is, $q \in \mathbb{Z}_{3\Delta}^J$. Denote by $\mathcal{N}_\phi$ an $(\alpha/\mathfrak{D})$-net in B_2^J with

$$|\mathcal{N}_\phi| \le \left(\frac{3\mathfrak{D}}{\alpha}\right)^{|J|}.$$

Choose $\phi' \in \mathcal{N}_\phi$ such that $\|\phi - \phi'\|_2 < \alpha/\mathfrak{D}$. Then

$$\|\zeta_1\phi' + \zeta_2\psi - q\|_2 < \alpha + |\zeta_1| \cdot \|\phi - \phi'\|_2 < 2\alpha,$$

as $|\zeta_1| \le \frac{1}{2}d\mathfrak{D} \le \mathfrak{D}$. Therefore,

$$\left\|\psi + \frac{\zeta_1}{\zeta_2}\phi' - \frac{\mathfrak{D}/2}{\zeta_2} \cdot \frac{q}{\mathfrak{D}/2}\right\|_2 < \frac{2\alpha}{|\zeta_2|} < \frac{4\alpha}{\mathfrak{D}}.$$

We observe that

$$\left|\frac{\zeta_1}{\zeta_2}\right| \vee \left|\frac{\mathfrak{D}/2}{\zeta_2}\right| \le 1, \qquad \|\phi'\|_2 \le 1 \quad \text{and} \quad \left\|\frac{q}{\mathfrak{D}/2}\right\|_2 \le 6\frac{\Delta}{\mathfrak{D}} \le 6, \tag{4.8}$$

where the last inequality follows from our assumption $\Delta \le \mathfrak{D}$. Next, let $\mathcal{N}_\square$ be an $(\alpha/\mathfrak{D})$-net in the unit square in $\mathbb{R}^2$ with $|\mathcal{N}_\square| \le (6\mathfrak{D}/\alpha)^2$. Using (4.8), and applying the triangle inequality, we now see that there exists $(x_1, x_2) \in \mathcal{N}_\square$ such that

$$\left\|\psi - x_1\phi' - x_2\frac{q}{\mathfrak{D}/2}\right\|_2 < \frac{11\alpha}{\mathfrak{D}}.$$

Hence,

$$\mathcal{M}_J^0(\mathfrak{D}, \Delta, d) := \left\{\left(\phi', x_1\phi' + x_2\frac{q}{\mathfrak{D}/2}\right) : \phi' \in \mathcal{N}_\phi, q \in \mathbb{Z}_{3\Delta}^J, (x_1, x_2) \in \mathcal{N}_\square\right\},$$

is a $\frac{12\alpha}{\mathfrak{D}}$-net of $S_J^0(\mathfrak{D}, \Delta, d)$, with

$$\begin{aligned}|\mathcal{M}_J^0(\mathfrak{D}, \Delta, d)| &\le |\mathcal{N}_\phi| \cdot |\mathbb{Z}_{3\Delta}^J| \cdot |\mathcal{N}_\square| \le \left(\frac{3C_0\mathfrak{D}}{\alpha} \cdot \left(\frac{3\Delta}{\sqrt{|J|}} + 1\right)\right)^{|J|} \cdot \left(\frac{6\mathfrak{D}}{\alpha}\right)^2 \\ &\le \left(63C_0\frac{d\mathfrak{D}^2}{\alpha} \cdot \left(\frac{1}{\sqrt{|J|}} + \frac{1}{\Delta}\right)\right)^{|J|} \cdot \left(\frac{6\mathfrak{D}}{\alpha}\right)^2,\end{aligned}$$

where (4.5) has been used to bound $|\mathbb{Z}_{3\Delta}^J|$ and (4.7) has been used to replace Δ by $d\mathfrak{D}$ in the last inequality.

Turning to prove the case of $|\zeta_1| > \frac{1}{2}d\mathfrak{D}$, we denote

$$S_J^1(\mathfrak{D}, \Delta, d) := S_J(\mathfrak{D}, \Delta, d) \backslash S_J^0(\mathfrak{D}, \Delta, d).$$

That is,

$$\begin{aligned}S_J^1(\mathfrak{D}, \Delta, d) := \Big\{&(\phi, \psi) \in S_J(\mathfrak{D}, \Delta, d) : \exists(\zeta_1, \zeta_2) \in \mathbb{R}^2, \|(\zeta_1, \zeta_2)\|_2 \in [\mathfrak{D}, 2\mathfrak{D}], \\ &|\zeta_1| \in \left[\frac{1}{2}d\mathfrak{D}, 2\mathfrak{D}\right], \|\zeta_1\phi + \zeta_2\psi\|_2 \in [\Delta, 2\Delta] \text{ and } \exists q \in \mathbb{Z}^J \\ &\text{such that } \|\zeta_1\phi + \zeta_2\psi - q\|_2 < \alpha\Big\}.\end{aligned}$$

Now let us construct a net in $S_J^1(\mathfrak{D}, \Delta, d)$. Our strategy here is opposite to what we used in the previous case. Namely, we use the volumetric approximation for ψ and then use (4.6) to approximate ϕ. To this end, consider any $(\phi, \psi) \in S_J^1(\mathfrak{D}, \Delta, d)$ and let $(\zeta_1, \zeta_2) \in \mathbb{R}^2$ be the corresponding vector. As in the previous case, we see $\|q\|_2 < 3\Delta$, i.e. $q \in \mathbb{Z}_{3\Delta}^J$. Since $|\zeta_1| \geq \frac{1}{2}d\mathfrak{D}$ and $|\zeta_2| \leq 2\mathfrak{D}$, we also see that $24\|\zeta_1\phi\|_2 \geq 6d\mathfrak{D} \geq \|\zeta_2\psi\|_2$. Therefore,

$$\Delta \leq \|\zeta_1\phi\|_2 + \|\zeta_2\psi\|_2 \leq 25\|\zeta_1\phi\|_2 \leq 25|\zeta_1|. \tag{4.9}$$

Recall that by assumption, $\alpha/\mathfrak{D} \leq d$. Hence, we see that

$$|\mathcal{N}_\psi| \leq \left(\frac{9d\mathfrak{D}}{\alpha}\right)^{|J|},$$

where $\mathcal{N}_\psi$ is an $(\alpha/\mathfrak{D})$-net in $3dB_2^J$. Since $\|\psi\|_2 \leq 3d$, there exists $\psi' \in \mathcal{N}_\psi$ such that $\|\psi - \psi'\|_2 < \alpha/\mathfrak{D}$. As in the previous case, this yields

$$\|\zeta_1\phi + \zeta_2\psi' - q\|_2 < \alpha + |\zeta_2| \cdot \|\psi - \psi'\|_2 \leq 3\alpha,$$

and so

$$\left\|\phi + \frac{\Delta\zeta_2}{50\mathfrak{D}\zeta_1} \cdot \frac{50\mathfrak{D}\psi'}{\Delta} - \frac{\Delta}{25\zeta_1} \cdot \frac{25q}{\Delta}\right\|_2 < \frac{3\alpha}{|\zeta_1|} \leq \frac{75\alpha}{\Delta},$$

where we have used (4.9) in the last step. Note that

$$\left|\frac{\Delta\zeta_2}{50\mathfrak{D}\zeta_1}\right| \vee \left|\frac{\Delta}{25\zeta_1}\right| \leq 1, \qquad \left\|\frac{50\mathfrak{D}\psi'}{\Delta}\right\|_2 \leq \frac{150d\mathfrak{D}}{\Delta} \leq 150 \quad \text{and}$$
$$\left\|\frac{25q}{\Delta}\right\|_2 \leq 75.$$

Let $\mathcal{N}_\square$ be the same $(\alpha/\mathfrak{D})$-net in the unit square as in the previous case. Since $\Delta \leq \mathfrak{D}$, combining the previous estimates with the triangle inequality, we have that there exists a $(x_1, x_2) \in \mathcal{N}_\square$ such that

$$\left\|\phi - x_1 \cdot \frac{50\mathfrak{D}\psi'}{\Delta} - x_2 \cdot \frac{25q}{\Delta}\right\|_2 < \frac{300\alpha}{\Delta}.$$

Using the fact $\Delta \leq \mathfrak{D}$ again, we now obtain an $(\alpha/\mathfrak{D})$-net $\mathcal{M}_\phi$ in $(\frac{300\alpha}{\Delta}) \cdot B_2^J$ with

$$|\mathcal{M}_\phi| \leq \left(\frac{900\mathfrak{D}}{\Delta}\right)^{|J|}.$$

Thus there exists $\nu \in \mathcal{M}_\phi$ such that

$$\left\|\phi - x_1 \cdot \frac{50\mathfrak{D}\psi'}{\Delta} - x_2 \cdot \frac{25q}{\Delta} - \nu\right\|_2 < \frac{\alpha}{\mathfrak{D}}.$$

This implies that the set

$$\mathcal{M}_J^1(\mathfrak{D},\Delta,d) := \Big\{\Big(x_1\cdot\frac{50\mathfrak{D}\psi'}{\Delta}+x_2\cdot\frac{25q}{\Delta}+\nu,\psi'\Big): \psi'\in\mathcal{N}_\psi, q\in\mathbb{Z}_{3\Delta}^{\gamma}, \nu\in\mathcal{M}_\phi, (x_1,x_2)\in\mathcal{N}_\square\Big\}$$

is a $(2\alpha/\mathfrak{D})$-net for $S_J^1(\mathfrak{D},\Delta,d)$. We observe that

$$\begin{aligned}|\mathcal{M}_J^1(\mathfrak{D},\Delta,d)| &\le |\mathcal{N}_\psi|\cdot|\mathbb{Z}_{3\Delta}|\cdot|\mathcal{M}_\phi|\cdot|\mathcal{N}_\square|\\ &\le \Big(\bar{C}\frac{d\mathfrak{D}}{\alpha}\cdot\Big(\frac{\Delta}{\sqrt{|J|}}+1\Big)\cdot\frac{\mathfrak{D}}{\Delta}\Big)^{|J|}\cdot\Big(\frac{\mathfrak{D}}{\alpha}\Big)^2\\ &\le \Big(\bar{C}\frac{d\mathfrak{D}^2}{\alpha}\cdot\Big(\frac{1}{\sqrt{|J|}}+\frac{1}{\Delta}\Big)\Big)^{|J|}\cdot\Big(\frac{\mathfrak{D}}{\alpha}\Big)^2,\end{aligned}$$

where $\bar{C}$ is some absolute constant.

Since $S_J(\mathfrak{D},\Delta,d)=S_J^0(\mathfrak{D},\Delta,d)\cup S_J^1(\mathfrak{D},\Delta,d)$, it therefore means that

$$\mathcal{M}_J(\mathfrak{D},\Delta,d) := \mathcal{M}_J^0(\mathfrak{D},\Delta,d)\cup\mathcal{M}_J^1(\mathfrak{D},\Delta,d)$$

is a $(C\alpha/\mathfrak{D})$-net for the set $S_J(\mathfrak{D},\Delta,d)$, where C is an absolute constant.

The net $\mathcal{M}_J(\mathfrak{D},\Delta,d)$ constructed above is not necessarily contained in $S_J(\mathfrak{D},\Delta,d)$. However, we can construct a new net by replacing each point of this net by a point of the set $S_J(\mathfrak{D},\Delta,d)$ which is within distance $C\alpha/\mathfrak{D}$ from this point. If a $(C\alpha/\mathfrak{D})$-close point does not exist, we skip the original point. Such process creates a $(2C\alpha/\mathfrak{D})$-net contained in $S_J(\mathfrak{D},\Delta,d)$ without increasing the cardinality. Thus the lemma is proved. □

Building on Lemma 4.6, we now obtain a net with small cardinality for the collection of vectors z for which $D_2(z_{\text{small}}/\|z_{\text{small}}\|_2)\approx\mathfrak{D}$, $\Delta(z_{\text{small}}/\|z_{\text{small}}\|_2)\approx\Delta$, and $d(z_{\text{small}}/\|z_{\text{small}}\|_2)\approx d$, where we recall that the vector z_{small} contains $n-M>n/2$ coordinates of z having the smallest magnitude. To this end, let us define the following set:

$$\begin{aligned}\mathcal{Z}(\mathfrak{D},\Delta,d) := \big\{&z\in\mathcal{Z}: D_2\big(z_{\text{small}}/\|z_{\text{small}}\|_2\big)\in\big[\mathfrak{D},(3/2)\mathfrak{D}\big],\\ &\Delta\big(z_{\text{small}}/\|z_{\text{small}}\|_2\big)\in\big[\Delta,(3/2)\Delta\big],\\ &d\big(z_{\text{small}}/\|z_{\text{small}}\|_2\big)\in\big[d,(3/2)d\big]\big\}.\end{aligned}\tag{4.10}$$

As will be seen in Section 5, the small ball probability for the images of such vectors is controlled by the values of the two-dimensional LCD and the real-imaginary decorrelation. So we partition this set according to $D_2(\cdot)$, $\Delta(\cdot)$, and $d(\cdot)$. The net $\mathcal{M}_J(\mathfrak{D},\Delta,d)$ provides a net for the vectors which have $D_2(\cdot)\approx\mathfrak{D}$, $\Delta(\cdot)\approx\Delta$, and $d(\cdot)\approx d$. This is shown in the proposition below.

PROPOSITION 4.7. *Let $d \in (0,1)$, $\mathfrak{D}, \Delta > 1$, and denote*

$$\alpha := L\sqrt{\log_1 \frac{\Delta}{2^7 L}}. \tag{4.11}$$

Assume that $\alpha \le d\mathfrak{D} \le \Delta \le \mathfrak{D}$. Then there exist absolute constants $C_{4.7}$, $\bar{C}_{4.7}$ and a set $\mathcal{N}(\mathfrak{D}, \Delta, d) \subset \mathcal{Z}(\mathfrak{D}, \Delta, d)$ with

$$|\mathcal{N}(\mathfrak{D}, \Delta, d)| \le \bar{C}_{4.7}^n \left(\frac{n}{\rho M} \cdot \frac{\mathfrak{D}}{\alpha}\right)^{5M} \cdot \left(\frac{d\mathfrak{D}^2}{\alpha} \cdot \left(\frac{1}{\sqrt{n}} + \frac{1}{\Delta}\right)\right)^{n-M}$$

having the following approximation properties: Let $z \in \mathcal{Z}(\mathfrak{D}, \Delta, d)$ be any vector and denote $J = \operatorname{small}(z)$. Then there exists $w \in \mathcal{N}(\mathfrak{D}, \Delta, d)$ such that

$$\left\| \frac{z_J}{\|z_J\|_2} - \frac{w_J}{\|w_J\|_2} \right\|_2 < C_{4.7} \frac{\alpha}{\mathfrak{D}}, \qquad \|z_{J^c} - w_{J^c}\|_2 \le C_{4.7} \frac{\rho\alpha}{\mathfrak{D}},$$
$$\big| \|z_J\|_2 - \|w_J\|_2 \big| \le C_{4.7} \frac{\rho\alpha}{\mathfrak{D}}.$$

REMARK 4.8. Note that Lemma 4.6 holds also for any subset of $S_J(\mathfrak{D}, \Delta, d)$. That is, given any $\mathcal{S} \subset S_J(\mathfrak{D}, \Delta, d)$ there exists a net $\mathcal{M}_J^{\mathcal{S}}(\mathfrak{D}, \Delta, d) \subset \mathcal{S}$ with the same properties as in Lemma 4.6. We use this version of Lemma 4.6 to prove Proposition 4.7. Similarly, we will see that given any $\mathcal{S} \subset \mathcal{Z}(\mathfrak{D}, \Delta, d)$ there exists a set $\mathcal{N}^{\mathcal{S}}(\mathfrak{D}, \Delta, d) \subset \mathcal{S}$ with the same approximation properties and the cardinality bound as in Proposition 4.7. This version of Proposition 4.7 will be used in Section 5.

In proving Proposition 4.7, our strategy will be to use the net $\mathcal{M}_J^{\mathcal{S}}(\mathfrak{D}, \Delta, d)$, for some suitable choice of $\mathcal{S}$, obtained from Lemma 4.6, to approximate the small coordinates. The cardinality of the net to approximate the large ones will be obtained by a simple volumetric estimate.

PROOF OF PROPOSITION 4.7. Fix a set $J \subset [n]$, $|J| = n - M$, and denote

$$\mathcal{Z}_J(\mathfrak{D}, \Delta, d) := \{ z \in \mathcal{Z}(\mathfrak{D}, \Delta, d) : \operatorname{small}(z) = J \}.$$

Let us now construct an approximating set for this subset. Denote $\phi + \mathrm{i}v = \phi(z) + \mathrm{i}\psi(z) := z_J / \|z_J\|_2 \in \mathbb{C}^J$. Recalling the definition of $\Delta(\phi + \mathrm{i}\psi)$, we see that there exists $\zeta \in \mathbb{R}^2$ such that

$$\|\zeta\|_2 \le (4/3) D_2(\phi + \mathrm{i}\psi) \le 2\mathfrak{D},$$
$$\|\zeta_1\phi + \zeta_2\psi\|_2 \le (4/3)\Delta(\phi + \mathrm{i}\psi) \le 2\Delta \quad \text{and}$$
$$\operatorname{dist}(\zeta_1\phi + \zeta_2\psi, \mathbb{Z}^n) < L\sqrt{\log_1 \frac{\|\zeta_1\phi + \zeta_2\psi\|_2}{2^8 L}} \le L\sqrt{\log_1 \frac{\Delta}{2^7 L}} = \alpha.$$

Further recall that $d(z_{\text{small}}/\|z_{\text{small}}\|_2) = \|\phi\|_2\|\psi\|_2$ and note that by our convention we have $\|\phi\|_2 \in [1/2, 1]$. Thus we deduce that $\|\psi\|_2 \in [d, 3d]$. Hence $(\phi, \psi) \in S_J(\mathfrak{D}, \Delta, d)$, and in particular $(\phi, \psi) \in \mathcal{S}$ where $\mathcal{S} := \{(\phi(z), \psi(z)) : z \in \mathcal{Z}(\mathfrak{D}, \Delta, d)\}$. So it can be approximated by an element of $\mathcal{M}_J^{\mathcal{S}}(\mathfrak{D}, \Delta, d)$. Set

$$\mathcal{M}_J := \{\phi + \mathrm{i}\psi : (\phi, \psi) \in \mathcal{M}_J^{\mathcal{S}}(\mathfrak{D}, \Delta, d)\}.$$

Then for any $z \in \mathcal{Z}_J(\mathfrak{D}, \Delta, d)$, there exists $w' \in \mathcal{M}_J$ such that

$$\left\| \frac{z_J}{\|z_J\|_2} - w' \right\|_2 < C_{4.5}\frac{\alpha}{\mathfrak{D}}.$$

For the set J^c, we will use a net satisfying the volumetric estimate. Since $z \in S_{\mathbb{C}}^{n-1}$, there exists a set $\mathcal{N}_{J^c}$ with

$$|\mathcal{N}_{J^c}| \le \left(\frac{3}{C_{4.5}\rho} \cdot \frac{\mathfrak{D}}{\alpha}\right)^{2M},$$

such that for every $z \in \mathcal{Z}_J(D, \Delta, d)$ there exists a $w_{J^c} \in \mathcal{N}_{J^c}$ for which

$$\|z_{J^c} - w_{J^c}\|_2 \le C_{4.5}\frac{\rho\alpha}{\mathfrak{D}}.$$

Finally, consider a net $\mathcal{N}_{[0,1]}$ with $|\mathcal{N}_{[0,1]}| \le 3\mathfrak{D}/(C_{4.5}\rho\alpha)$ such that for every $z \in \mathcal{Z}_J(\mathfrak{D}, \Delta, d)$ there exists a $\rho' \in [0, 1]$ for which

$$\big|\|z_J\|_2 - \rho'\big| \le C_{4.5}\frac{\rho\alpha}{\mathfrak{D}}.$$

Now let us define

$$\mathcal{N}(\mathfrak{D}, \Delta, d) := \bigcup_{|J|=n-M} \{\rho' w' + w'' : \rho' \in \mathcal{N}_{[0,1]}, w' \in \mathcal{M}_J, w'' \in \mathcal{N}_{J^c}\}.$$

Then for any $z \in \mathcal{Z}(\mathfrak{D}, \Delta, d)$ there exists a $w = w_J + w_{J^c} \in \mathcal{N}(\mathfrak{D}, \Delta, d)$ such that

$$\left\| \frac{z_J}{\|z_J\|_2} - \frac{w_J}{\|w_J\|_2} \right\|_2 < C_{4.5}\frac{\alpha}{\mathfrak{D}}, \qquad \|z_{J^c} - w_{J^c}\|_2 \le C_{4.5}\frac{\rho\alpha}{\mathfrak{D}},$$
$$\big|\|z_J\|_2 - \|w_J\|_2\big| \le C_{4.5}\frac{\rho\alpha}{\mathfrak{D}}.$$

The set $\mathcal{N}(\mathfrak{D}, \Delta, d)$ thus constructed may not be contained in $\mathcal{Z}(\mathfrak{D}, \Delta, d)$. However, as in the proof of Lemma 4.5 this can be rectified easily. It thus remains to bound the cardinality of $\mathcal{N}(\mathfrak{D}, \Delta, d)$. By Lemma 4.6, we have

$$\begin{aligned}
|\mathcal{N}(\mathfrak{D}, \Delta, d)| &\le \sum_{|J|=n-M} |\mathcal{N}_{J^c}| \cdot |\mathcal{N}_{[0,1]}| \cdot |\mathcal{M}_J| \\
&\le \binom{n}{M} \cdot \left(\frac{3}{C_{4.6}\rho} \cdot \frac{\mathfrak{D}}{\alpha}\right)^{2M+1} \\
&\quad \cdot \left(\bar{C}_{4.6}\frac{d\mathfrak{D}^2}{\alpha} \cdot \left(\frac{1}{\sqrt{n-M}} + \frac{1}{\Delta}\right)\right)^{n-M} \cdot \left(\frac{\mathfrak{D}}{\alpha}\right)^2.
\end{aligned}$$

Since $1 \le M < n/2$, the required estimate follows from a straightforward calculation. This completes the proof. $\square$

5. The structure of the kernel in the genuinely complex case. In this section our goal is to show that with high probability, any genuinely complex vector in $\operatorname{Ker}(B^D)$ has a large two-dimensional LCD. Before proceeding any further let us formally define the notion of genuinely complex vectors:

$$\begin{aligned}\operatorname{Compl}(\mathcal{Z}) &:= \Bigg\{ z \in \mathcal{Z} : d\big(z_{\mathrm{small}}/\|z_{\mathrm{small}}\|_2\big) \\ &\ge \frac{4L}{D_2(z_{\mathrm{small}}/\|z_{\mathrm{small}}\|_2)} \sqrt{\log_1 \frac{\Delta(z_{\mathrm{small}}/\|z_{\mathrm{small}}\|_2)}{2^7 L}} \Bigg\},\end{aligned} \tag{5.1}$$

where we recall the definition of $\mathcal{Z}$ from (4.4). Equipped with the notion of genuinely complex vectors we state the main result of this section.

THEOREM 5.1. *Let B^D, A_n, ρ, K, R, r and r' be as in Proposition 3.1. Then there exist constants $c_{5.1}, c'_{5.1}$, depending only on K, R, r and the fourth moment of $\{\xi_{ij}\}$, such that if p satisfies the inequality*

$$c_{5.1}\rho^5 pn > 1, \tag{5.2}$$

then we have

$$\mathbb{P}\Bigg(\exists z \in \operatorname{Compl}(\mathcal{Z}) \cap \operatorname{Ker}\big(B^D\big) : D_2\big(z_{\mathrm{small}}/\|z_{\mathrm{small}}\|_2\big) \le \exp\Big(c'_{5.1} \frac{n}{M} \Big), \\ \|A_n\| \le K\sqrt{pn} \Bigg) \le \exp(-\bar{c}_{5.1} np),$$

where $M = C_{3.1}\rho^{-4}p^{-1}$ and z_{small} is the smallest $(n - M)$ coordinates of z in modulus.

The proof of Theorem 5.1 is carried out by the following two-fold argument. Using Proposition 4.7, we show that the subset of vectors in $\operatorname{Compl}(\mathcal{Z})$ that have a large value of $\Delta(z_{\mathrm{small}}/\|z_{\mathrm{small}}\|_2)$ admits a net of small cardinality. This observation together with an estimate on the small ball probability, obtained from [36], Theorem 7.5, yields the desired conclusion for vectors $z \in \operatorname{Compl}(\mathcal{Z})$ which possess a large value of $\Delta(z_{\mathrm{small}}/\|z_{\mathrm{small}}\|_2)$ (see Proposition 5.2). For the other case, we first show that such vectors, upon rotation, must have a dominated real part. Applying Proposition 3.1, we show that this is impossible with high probability, which completes the proof of Theorem 5.1. The rest of this section is devoted to implementing this idea.

First, let us consider the case of large $\Delta(z_{\mathrm{small}}/\|z_{\mathrm{small}}\|_2)$. For such vectors, we prove that the following holds.

PROPOSITION 5.2. *Let $s \geq 1$. Define the set $\mathcal{Z}(s)$ by*

$$\mathcal{Z}(s) := \{z \in \mathrm{Compl}(\mathcal{Z}) : \Delta(z_{\mathrm{small}}/\|z_{\mathrm{small}}\|_2) \geq sL\}.$$

Let A_n, B^D, K, R and ρ be as in Theorem 5.1. Then there exist $s_{5.2}, r_\star > 1$, and $c'_{5.2} > 0$, depending only on K, R and the fourth moment of $\{\xi_{ij}\}$ such that for any $r_\star^2 p^{-1} \leq M \leq \rho n$ we have

$$\mathbb{P}\big(\exists z \in \mathcal{Z}(s_{5.2}) \cap \mathrm{Ker}(B^D) : D_2(z_{\mathrm{small}}/\|z_{\mathrm{small}}\|_2) \leq \exp(c'_{5.2} n/M) \text{ and } \|A_n\| \leq K\sqrt{pn}\big) \leq e^{-n}.$$

To prove Proposition 5.2, we use bounds on Lévy concentration function. Using such bounds, we show that for any vector $z \in \mathcal{Z}(s)$ the ℓ_2 norm of $B^D z$ cannot be too small with large probability. From the net constructed in Section 4, it follows that $\mathcal{Z}(s)$ admits a net of small cardinality which enables us to take the union bound and complete the proof of Proposition 5.2.

As mentioned above to prove Proposition 5.2, we need to derive bounds on the small ball probability, in particular on Lévy concentration function. Before deriving the bounds such bounds we need to fix some notation. Let $z \in \mathbb{C}^m$ and $J \subset [m]$. Denote $z_J := (z_i)_{i \in J} \in \mathbb{C}^J$ and $V_J := V(z_J)$, where we recall that for any $z' = x + \mathrm{i}y \in \mathbb{C}^{m'}$ we define $V(z') := \binom{x^{\mathsf{T}}}{y^{\mathsf{T}}}$. Further denote the real-imaginary decorrelation of V_J by

$$d(V_J) := d(z_J) = \big(\det(V_J V_J^{\mathsf{T}})\big)^{1/2}.$$

This parameter, along with the LCD of $z_J/\|z_J\|_2$ controls the Lévy concentration function of $\sum_{j=1}^m \Xi_j z_j$, for a sequence of independent random variables $\{\Xi_j\}_{j \in [m]}$. Below is the desired estimate on the Lévy concentration function, which is a direct corollary of [36], Theorem 7.5.

PROPOSITION 5.3. *Fix any positive integer m and let $\Xi := (\Xi_1, \dots, \Xi_m) \in \mathbb{R}^m$, $\Xi_j := d_j \xi_j$, $j = 1, \dots, m$, where $d_1, \dots, d_m$ are i.i.d.* $\mathrm{Ber}(p)$, *and ξ_j are i.i.d. random variables satisfying*

$$\mathcal{L}(\xi_j, 1) \leq 1 - c_1 \quad \text{and} \quad \mathbb{P}(|\xi_j| > C_1) \leq c_1/2 \tag{5.3}$$

for some absolute constants C_1 and $c_1 \in (0,1)$. Then for any $z \in \mathbb{C}^m$, $J \in [m]$ such that $z_J \neq 0$, and $\varepsilon > 0$, we have

$$\mathcal{L}(V\Xi, \varepsilon\sqrt{p}\|V_J\|) \leq \frac{C_{5.3}}{d(V_J)/\|V_J\|}\left(\varepsilon + \frac{1}{\sqrt{p} D_2(V_J/\|V_J\|)}\right)^2 \tag{5.4}$$

and if $\|\mathrm{Re}(z_J)\|_2 \geq \|\mathrm{Im}(z_J)\|_2$, then

$$\mathcal{L}(V\Xi, \varepsilon\sqrt{p}\|V_J\|) \leq \overline{C}_{5.3}\left(\varepsilon + \frac{1}{\sqrt{p} D_1(\mathrm{Re}(z_J))/\|V_J\|}\right), \tag{5.5}$$

where $V := V(z)$, $C_{5.3}$ and $\overline{C}_{5.3}$ are some constants, depending only on c_1 and C_1.

REMARK 5.4. We point out to the reader that the definition of LCD in [36] and that of ours are slightly different from each other. For example, to define LCD in [36] the function $\log_+(x) := \max\{\log x, 0\}$ was used instead of $\log_1(\cdot)$. Moreover, the constants appearing in front of L are different (compare Definition 4.1 with [36], Definition 7.1). However, upon investigating the proof of [36], Theorem 7.5, it becomes evident that the same proof can be carried out for the LCDs $D_2(\cdot)$ and $D_1(\cdot)$ to obtain the same estimates on the Lévy concentration function. It only changes the constant that appears in [36], equation (7.3). Below we apply this version of [36], Theorem 7.5.

PROOF OF PROPOSITION 5.3. As mentioned above, the proof is a straightforward application of [36], Theorem 7.5. Indeed, we note that $\mathcal{L}(V\Xi, t) \le \mathcal{L}(V_J \Xi_J, t)$, for any $t > 0$, where $\Xi_J := (\Xi_j)_{j\in J}$. The assertion (5.3) implies that

$$\mathcal{L}(\Xi_j, 1) \le 1 - pc_1 \quad \text{and} \quad \mathbb{P}\big(|\Xi_j| > C_1\big) \le pc_1/2. \tag{5.6}$$

Since $L = (\delta_0 p)^{-1/2}$ (see Remark 4.5), shrinking δ_0 if necessary, the inequality (5.4) follows directly from [36], Theorem 7.5, applied with $m = 2$. To prove (5.5), using the triangle inequality we further note that $\mathcal{L}(V_J \Xi_J, t) \le \mathcal{L}(x_J^{\mathsf{T}} \Xi_J, t)$. Thus applying [36], Theorem 7.5, with $m = 1$ we obtain (5.5). □

Applying Proposition 5.3 and standard tensorization techniques, we obtain the following result, the proof of which is omitted.

PROPOSITION 5.5. *Let B^D be as in Proposition 3.1. Fix any $z \in \mathbb{C}^n$ and $J :=$ small(z) such that $z_J \neq 0$. Then for any $\varepsilon > 0$, we have*

$$\begin{aligned}\mathcal{L}\big(B^D z, \varepsilon\sqrt{p(n-1)}\|z_J\|_2\big) \\ \le \left[\frac{C_{5.5}}{d(z_J/\|z_J\|_2)}\left(\varepsilon + \frac{1}{\sqrt{p}D_2(z_J/\|z_J\|_2)}\right)^2\right]^{n-1},\end{aligned} \tag{5.7}$$

and if $\|\operatorname{Re}(z_J)\|_2 \ge \|\operatorname{Im}(z_J)\|_2$, then

$$\begin{aligned}\mathcal{L}\big(B^D z, \varepsilon\sqrt{p(n-1)}\|z_J\|_2\big) \\ \le \left[\overline{C}_{5.5}\left(\varepsilon + \frac{1}{\sqrt{p}D_1(\operatorname{Re}(z_J)/\|z_J\|_2)}\right)\right]^{n-1},\end{aligned} \tag{5.8}$$

for some constants $C_{5.5}$ and $\overline{C}_{5.5}$, depending only on $\mathbb{E}|\xi_{ij}|$ and $\mathbb{E}(\xi_{ij}^4)$.

REMARK 5.6. The inequality (5.8) provides bounds on Lévy concentration function based on one-dimensional LCD. It will be used later in Section 6 to treat essentially real vectors.

To prove Proposition 5.2 we also need the following elementary lower bound on the LCD of nondominated vectors. Its proof follows from [36], Proposition 7.4, and the definition of dominated vectors.

LEMMA 5.7. *If $z \notin \mathrm{Dom}(M, \alpha_\star)$ for some $M < n$ and $\alpha_\star > 0$, then we have*

$$D_2\big(z_{\mathrm{small}}/\|z_{\mathrm{small}}\|_2\big) \ge \frac{\alpha_\star\sqrt{M}}{2}.$$

We are now ready to prove Proposition 5.2.

PROOF OF PROPOSITION 5.2. The set in question can be partitioned into the subsets of $\mathcal{Z}(\mathfrak{D}, \Delta, d)$ appearing in Section 4. Indeed, using Lemma 5.7 we note that for any $z \in \mathcal{Z}(s)$ we have $D_2(z_{\mathrm{small}}/\|z_{\mathrm{small}}\|_2) \ge \frac{1}{2}\alpha_\star\sqrt{M} \ge \frac{\alpha_\star r_\star}{2}p^{-1/2}$ for some $\alpha_\star > 0$. Since $L = (\delta_0 p)^{-1/2}$, choosing $r_\star$ sufficiently large, we therefore obtain that

$$\big\{z \in \mathcal{Z}(s) : D_2\big(z_{\mathrm{small}}/\|z_{\mathrm{small}}\|_2\big) \le \exp\big(c'n/M\big)\big\} \subset \bigcup_{\mathfrak{D},\Delta,d} \mathcal{Z}(\mathfrak{D}, \Delta, d) \cap \mathcal{Z}(s),$$

where the union is taken over all $\mathfrak{D} = 2^k$, $C_\star L \le \mathfrak{D} \le \exp(c'n/M)$, for some large constant $C_\star$, and over all $\Delta = 2^m$, $d = 2^{-\ell}$ satisfying $d\mathfrak{D} \le \Delta \le \mathfrak{D}$. Also note that for any $z \in \mathrm{Compl}(\mathcal{Z})$ we have

$$d\big(z_{\mathrm{small}}/\|z_{\mathrm{small}}\|_2\big) \ge \frac{4L}{D_2(z_{\mathrm{small}}/\|z_{\mathrm{small}}\|_2)}\sqrt{\log_1 \frac{\Delta(z_{\mathrm{small}}/\|z_{\mathrm{small}}\|_2)}{2^7 L}}. \tag{5.9}$$

If $z \in \mathcal{Z}(\mathfrak{D}, \Delta, d)$, we further have that $D_2(z_{\mathrm{small}}/\|z_{\mathrm{small}}\|_2) \le 2\mathfrak{D}$, $d(z_{\mathrm{small}}/\|z_{\mathrm{small}}\|_2) \le 2d$, and $\Delta(z_{\mathrm{small}}/\|z_{\mathrm{small}}\|_2) \ge \Delta$. Therefore, it follows from (5.9) that

$$\alpha := L\sqrt{\log_1 \frac{\Delta}{2^7 L}} \le d\mathfrak{D}.$$

So it allows us to use Proposition 4.7. Recalling that $M \ge r_\star^2 p^{-1} \ge p^{-1}$, we see that the number of different values of $\mathfrak{D}$ appearing in the partitions is bounded by $c'pn$. Using the fact that $\alpha \ge L$, we see that the number of different values of d is bounded by the same number, and so is the number of different values of Δ. Therefore, using the union bound, we deduce that it is enough to show that

$$\begin{aligned}&\mathbb{P}\big(\exists z \in \mathcal{Z}(D, \Delta, d) \cap \mathcal{Z}(s) : B^D z = 0, \text{ and}\\ &\qquad \|B^D\| \le (K+R)\sqrt{pn}\big) \le e^{-2n},\end{aligned} \tag{5.10}$$

for each such triple $(\mathfrak{D}, \Delta, d)$.

To this end, we note that $\mathcal{Z}(\mathfrak{D}, \Delta, d) \cap \mathcal{Z}(s)$ admits a net $\mathcal{N}(\mathfrak{D}, \Delta, d) \subset \mathcal{Z}(\mathfrak{D}, \Delta, d) \cap \mathcal{Z}(s)$. Therefore, from Proposition 5.5 it follows that

$$\mathcal{L}\big(B^D w, \varepsilon\sqrt{p(n-1)}\|w_{\mathrm{small}}\|_2\big) \le \left[\frac{C_{5.5}}{d}\left(\varepsilon + \frac{1}{\sqrt{p}\mathfrak{D}}\right)^2\right]^{n-1},$$

for any $w \in \mathcal{N}(\mathfrak{D}, \Delta, d)$ and $\varepsilon > 0$. Set

$$\varepsilon_0 := 40C_{4.7}(K+R)\frac{\alpha}{\mathfrak{D}}. \tag{5.11}$$

Since $\alpha \geq L = (\delta_0 p)^{-1/2} \geq p^{-1/2}$ and $K, R \geq 1$ we note that $\varepsilon_0 \geq \frac{1}{\sqrt{p}\mathfrak{D}}$. Therefore,

$$\mathbb{P}\bigg(\|B^D w\|_2 \leq \frac{\varepsilon_0}{2}\|w_{\text{small}}\|_2 \cdot \sqrt{pn}\bigg) \leq \bigg[\frac{C}{d}\bigg((K+R)\frac{\alpha}{\mathfrak{D}}\bigg)^2\bigg]^{n-1},$$

for some positive constant C. Hence, by the union bound and applying Proposition 4.7,

$$\begin{aligned}
&\mathbb{P}\bigg(\exists w \in \mathcal{N}(\mathfrak{D}, \Delta, d) : \|B^D w\|_2 \leq \frac{\varepsilon_0}{2}\|w_{\text{small}}\|_2 \cdot \sqrt{pn}\bigg) \\
&\quad \leq |\mathcal{N}(\mathfrak{D}, \Delta, d)| \cdot \bigg[\frac{C}{d}\bigg((K+R)\frac{\alpha}{\mathfrak{D}}\bigg)^2\bigg]^{n-1} \\
&\quad \leq \bigg[\frac{C}{d}\bigg((K+R)\frac{\alpha}{\mathfrak{D}}\bigg)^2\bigg]^{n-1} \cdot \bar{C}_{4.7}^n \bigg(\frac{n}{\rho M} \cdot \frac{\mathfrak{D}}{\alpha}\bigg)^{5M} \\
&\qquad \cdot \bigg(\frac{d\mathfrak{D}^2}{\alpha} \cdot \bigg(\frac{1}{\sqrt{n}} + \frac{1}{\Delta}\bigg)\bigg)^{n-M}.
\end{aligned} \tag{5.12}$$

Recalling the definition of α and using the inequalities $L \leq \alpha \leq d\mathfrak{D}$ and $\Delta \leq \mathfrak{D}$ we note that

$$\frac{1}{d} \cdot \bigg(\frac{\alpha}{\mathfrak{D}}\bigg)^2 \leq \frac{\alpha}{\mathfrak{D}} = \frac{L}{\mathfrak{D}}\sqrt{\log_1 \frac{\Delta}{2^7 L}} \leq \frac{L}{\mathfrak{D}}\sqrt{\log_1 \frac{\mathfrak{D}}{2^7 L}} \leq \frac{1}{2^7}\sqrt{\frac{2^7 L}{\mathfrak{D}}},$$

where the last step follows from the fact that $\log x \leq x$ for any $x \geq e$. As we have already noted that $\mathfrak{D} \geq C_\star L$, for some large $C_\star$, we can enlarge $C_\star$ further (i.e., we increase $r_\star$) so that $\frac{C}{d}((K+R)\frac{\alpha}{\mathfrak{D}})^2 \bar{C}_{4.7} < 1$. This means that we can drop the term

$$\bigg[\frac{C}{d}\bigg((K+R)\frac{\alpha}{\mathfrak{D}}\bigg)^2\bigg]^{M-1} \cdot \bar{C}_{4.7}^M < 1$$

in (5.12). Therefore, from (5.12) we obtain

$$\mathbb{P}\bigg(\exists w \in \mathcal{N}(\mathfrak{D}, \Delta, d) : \|B^D w\|_2 \leq \frac{\varepsilon_0}{2}\|w_{\text{small}}\|_2 \cdot \sqrt{pn}\bigg) \leq \exp(-\Gamma n), \tag{5.13}$$

where

$$\begin{aligned}
\Gamma := &-\bigg(1 - \frac{M}{n}\bigg) \cdot \log\bigg(\frac{C'(K+R)^2\alpha}{\sqrt{n}} + \frac{C'(K+R)^2\alpha}{\Delta}\bigg) \\
&- \frac{5M}{n}\log\bigg(\frac{n}{\rho M} \cdot \frac{\mathfrak{D}}{\alpha}\bigg),
\end{aligned}$$

and $C' := C \cdot \bar{C}_{4.7}$. To complete the proof, we need to show that $\Gamma \geq 2$.

Turning to proof of this, we recall that $\frac{n}{M} \geq \frac{1}{\rho}$, $\alpha \geq L \geq 1$ and $\mathfrak{D} \leq \exp(c'\frac{n}{M})$. So, choosing c' sufficiently small we obtain

$$\frac{5M}{n}\log\Big(\frac{n}{\rho M}\cdot\frac{\mathfrak{D}}{\alpha}\Big) \leq \frac{10M}{n}\log\Big(\frac{n}{M}\Big) + \frac{5M}{n}\log\mathfrak{D} \leq 20.$$

Next, recall that $L = (d_0 p)^{-1/2}$ and $\Delta \leq \mathfrak{D} \leq \exp(c'n/M)$. So

$$\alpha = L\sqrt{\log_1 \frac{\Delta}{2^7 L}} \leq \sqrt{\frac{c'n}{d_0 Mp}}.$$

Therefore, using the fact that $M < n/2$ we obtain

$$\Gamma \geq -\frac{1}{2}\cdot\log\Big(\frac{C'(K+R)^2\sqrt{c'}}{\sqrt{d_0 Mp}} + \frac{C'(K+R)^2\alpha}{\Delta}\Big) - 20.$$

We claim that by choosing s to be a sufficiently large constant, we can guarantee that

$$\frac{C'(K+R)^2\alpha}{\Delta} < e^{-50}. \tag{5.14}$$

Thus choosing c' small enough and recalling that $Mp \geq r_\star^2$, from claim (5.14) we see that $\Gamma \geq 2$, providing the required bound for the probability.

Now let us check our claim (5.14). Using the definition of α, choosing $s_{5.2}$ sufficiently large, and using the fact that the function $f(x) := x^{-1}\sqrt{\log_1 x}$ tends to 0 as $x \to \infty$, we note that

$$e^{50}\frac{C'(K+R)^2\alpha}{\Delta} = e^{50}C'(K+R)^2\frac{L}{\Delta}\sqrt{\log_1\frac{\Delta}{2^7 L}} \leq 1,$$

for any Δ such that $\Delta \geq s_{5.2}L$. This proves the claim (5.14).

Thus we have shown that for a sufficiently large value of $s_{5.2}$,

$$\mathbb{P}\Big(\exists w \in \mathcal{N}(\mathfrak{D},\Delta,d) : \|B^D w\|_2 \leq \frac{\varepsilon_0}{2}\|w_{\mathrm{small}}\|_2\cdot\sqrt{pn}\Big) \leq \exp(-2n). \tag{5.15}$$

To deduce (5.10) from (5.15), we simply use the property of the net $\mathcal{N}(\mathfrak{D},\Delta,d)$. Indeed, let us assume that there exists a $z \in \mathcal{Z}(s_{5.2}) \cap \mathcal{Z}(\mathfrak{D},\Delta,d)$ so that $B^D z = 0$. Denoting $J = \mathrm{small}(z)$, using Proposition 4.7, and the triangle inequality we see that there exists a $w \in \mathcal{N}(\mathfrak{D},\Delta,d)$ such that

$$\begin{aligned}
\|B^D w\|_2 &= \|B^D(w-z)\|_2 \\
&\leq \|B^D\|_2 \\
&\quad\cdot\Big(\|z_{J^c} - w_{J^c}\|_2 + \|w_J\|_2\Big\|\frac{z_J}{\|z_J\|_2} - \frac{w_J}{\|w_J\|_2}\Big\|_2 + \big|\|w_J\| - \|z_J\|_2\big|\Big) \\
&\leq (K+R)\sqrt{np}\cdot\Big(2C_{4.7}\frac{\rho\alpha}{\mathfrak{D}} + C_{4.7}\frac{\alpha}{\mathfrak{D}}\|w_J\|_2\Big) \\
&\leq 3C_{4.7}(K+R)\frac{\alpha}{\mathfrak{D}}\|w_J\|_2\sqrt{np},
\end{aligned}$$

where the last inequality follows from the fact that $w \in \mathcal{N}(\mathfrak{D}, \Delta, d) \subset \mathcal{Z}(s_{5.2}) \subset \mathrm{Comp}(M, \rho)^c$. To complete the proof, let us show that $\|w_J\|_2 \le 4\|w_{\mathrm{small}}\|_2$. Assume for a moment that the opposite inequality holds. Since $\mathfrak{D} \ge \frac{\alpha_\star r_\star}{2} p^{-1/2}$, and $L = (\delta_0 p)^{-1/2}$, choosing a sufficiently large $r_\star$, we may assume that

$$C_{4.7}\frac{\alpha}{\mathfrak{D}} \le C_{4.7}\frac{L}{\mathfrak{D}}\sqrt{\log_1 \frac{\mathfrak{D}}{2^7 L}} \le \frac{1}{8}.$$

Denote $\mathrm{small}(w) = I$. Combining the estimate above and Proposition 4.7, we see that

$$\begin{aligned}\|z_I\|_2 &\le \|w_I\|_2 + \|z_{I\cap J} - w_{I\cap J}\|_2 + \|z_{I\setminus J} - w_{I\setminus J}\|_2\\ &\le \frac{1}{4}\|w_J\|_2 + \|z_J - w_J\|_2 + \|z_{J^c} - w_{J^c}\|_2\\ &\le \frac{1}{2}\|z_J\|_2 + \frac{1}{4}\|z_J\|_2 + \frac{1}{8}\rho < \|z_J\|_2,\end{aligned}$$

where we used $\|z_J\|_2 \ge \rho$ in the last inequality. This contradicts the definition of J as $J = \mathrm{small}(z)$ and proves the desired inequality $\|w_J\|_2 \le 4\|w_{\mathrm{small}}\|_2$.

Recalling the definition of ε_0 in (5.11), we now deduce (5.10) from (5.15). This completes the proof. □

Using Proposition 5.2, we now complete the proof of Theorem 5.1. The final ingredient for the proof of Theorem 5.1 is a lower bound on the one-dimensional LCD, the proof of which follows from [41], Lemma 6.2.

LEMMA 5.8. *For any $x \in \mathbb{R}^m$, $D_1(x) \ge \frac{1}{2\|x\|_\infty}$.*

PROOF OF THEOREM 5.1. We first claim that

$$\begin{aligned}&\mathbb{P}\Big(\exists z \in \mathrm{Ker}\big(B^D\big) \cap S_{\mathbb{C}}^{n-1} : \Delta\big(z_{\mathrm{small}}/\|z_{\mathrm{small}}\|_2\big) \le \frac{\rho^{-1}}{4\sqrt{p}}\Big)\\ &\qquad \le \exp(-\bar{c}_{3.1} np).\end{aligned} \tag{5.16}$$

The probability bound above would follow from Proposition 3.1 if for any such z, we find a number $\nu \in \mathbb{C}$, $|\nu| = 1$ such that the vector νz_{small} has a dominated real part. To implement this idea and show (5.16), we fix $z \in \mathrm{Ker}(B^D) \cap S_{\mathbb{C}}^{n-1}$ and denote $z_{\mathrm{small}}/\|z_{\mathrm{small}}\|_2 = \phi + \mathrm{i}\psi$, where $\phi, \psi \in \mathbb{R}^J$, $J = \mathrm{small}(z)$. Let $\theta = (\theta_1, \theta_2) \in \mathbb{R}^2$ be such that

$$\mathrm{dist}\big(\theta_1\phi + \theta_2\psi, \mathbb{Z}^J\big) < L\sqrt{\log_1 \frac{\|\theta_1\phi + \theta_2\psi\|_2}{2^8 L}} \tag{5.17}$$

and $\|\theta_1\phi + \theta_2\psi\|_2 \le 2\Delta(z_{\mathrm{small}}/\|z_{\mathrm{small}}\|_2)$. Denote

$$w := \frac{\theta_1 - \mathrm{i}\theta_2}{|\theta_1 - \mathrm{i}\theta_2|} z.$$

Then $w \in \mathrm{Ker}(B^D) \cap S_{\mathbb{C}}^{n-1}$ and $w_{\mathrm{small}} = z_{\mathrm{small}}$. Therefore, (5.17) implies that

$$D_1\big(\mathrm{Re}(w_{\mathrm{small}})/\big\|\mathrm{Re}(w_{\mathrm{small}})\big\|_2\big) \le 2\Delta\big(z_{\mathrm{small}}/\|z_{\mathrm{small}}\|_2\big).$$

Upon applying Lemma 5.8, we find that

$$\left\| \frac{\mathrm{Re}(w_{\mathrm{small}})}{\|\mathrm{Re}(w_{\mathrm{small}})\|_2} \right\|_\infty \ge \left[4\Delta\left(\frac{z_{\mathrm{small}}}{\|z_{\mathrm{small}}\|_2}\right)\right]^{-1} \ge \rho p^{1/2}.$$

Since $w \in \mathrm{Ker}(B^D) \cap S_{\mathbb{C}}^{n-1}$, the claim (5.16) follows from Proposition 3.1.

Next, recalling the fact that $L = (\delta_0 p)^{-1/2}$, and shrinking ρ, if necessary, we obtain that

$$s_{5.2} L < \frac{\rho^{-1}}{4\sqrt{p}}.$$

The desired result then follows upon combining Proposition 5.2 and (5.16). □

6. Construction of the net and the structure of the kernel in the essentially real case. In this section we consider the class of vectors whose real and imaginary parts are almost linearly dependent. Namely, we introduce the set of *essentially real* vectors $\mathrm{Real}(\mathcal{Z})$ defined by

$$\mathrm{Real}(\mathcal{Z}) := \mathcal{Z} \setminus \mathrm{Compl}(\mathcal{Z}), \tag{6.1}$$

where we recall the definitions of $\mathcal{Z}$ and $\mathrm{Compl}(\mathcal{Z})$ from (4.4) and (5.1) respectively. Having shown that there does not exist any vector in $\mathrm{Compl}(\mathcal{Z}) \cap \mathrm{Ker}(B^D)$ such that its two-dimensional LCD is small, it remains to show the same for $\mathrm{Real}(\mathcal{Z}) \cap \mathrm{Ker}(B^D)$. For essentially real vectors, the real-imaginary decorrelation $d(\cdot)$ is very small which precludes using (5.7). Instead we have to rely on the probability bound obtained in (5.8), which depends on the one-dimensional LCD. As the bound on $D_1(u)$ implies a much more rigid arithmetic structure than a bound on $D_2(u)$, construction of a net of real($\mathcal{Z}$) would be easier. To construct such a net, we will follow the method of [36]. Before finding a net, let us remind the reader that the definition of $\mathrm{Compl}(\mathcal{Z})$, and hence that of $\mathrm{Real}(\mathcal{Z})$, depends on the two-dimensional LCD (see (5.1)). Since the bound on Lévy concentration function, for vectors in $\mathrm{Real}(\mathcal{Z})$, depends on the one-dimensional LCD, we need a result that connects $D_1(\cdot)$ with $D_2(\cdot)$. The lemma below does that job and this is the sole reason of introducing different powers of 2 in the definitions (4.1) and (4.2) (recall Remark 4.2).

LEMMA 6.1. *Fix* $z \in \mathrm{Real}(\mathcal{Z})$ *and let* $z_{\mathrm{small}}/\|z_{\mathrm{small}}\|_2 =: \phi + \mathrm{i}\psi$. *Then* $D_1(\phi) \le 2D_2(z_{\mathrm{small}}/\|z_{\mathrm{small}}\|_2)$. *In particular, if* $D_2(z_{\mathrm{small}}/\|z_{\mathrm{small}}\|_2) \le \mathfrak{D}$ *then* $D_1(\phi) \le 2\mathfrak{D}$.

PROOF. Let us denote $J = \text{small}(z)$. Denoting $\mathfrak{D} = D_2(\phi + \mathrm{i}\psi)$, we see that there exists $\theta = (\theta_1, \theta_2) \in \mathbb{R}^2$ with $\|\theta\|_2 \le 2\mathfrak{D}$ and $\|\phi\theta_1 + \psi\theta_2\|_2 \ge \Delta(z_J/\|z_J\|_2)/\sqrt{2}$, and $q \in \mathbb{Z}^J$ such that

$$\|\theta_1\phi + \theta_2\psi - q\|_2 < L\sqrt{\log_1 \frac{\|\theta_1\phi + \theta_2\psi\|_2}{2^8 L}}. \tag{6.2}$$

Using the triangle inequality, and the facts that $|\theta_2| \le \|\theta\|_2$, $\|\phi\|_2 \cdot \|\psi\|_2 = d(z_J/\|z_J\|_2)$, and $\|\phi\|_2 \ge 1/2$, we also obtain

$$\|\theta_1\phi + \theta_2\psi\|_2 \le \|\theta_1\phi\|_2 + 4d\big(z_J/\|z_J\|_2\big)D_2\big(z_J/\|z_J\|_2\big). \tag{6.3}$$

Since $\phi + \mathrm{i}\psi \in \text{Real}(\mathcal{Z})$, we further note that

$$d\big(z_J/\|z_J\|_2\big)D_2\big(z_J/\|z_J\|_2\big) \le 4L\sqrt{\log_1 \frac{\Delta(z_J/\|z_J\|_2)}{2^7 L}}$$

(see (5.1) and (6.1)). Therefore, denoting

$$\alpha_0 := L\sqrt{\log_1 \frac{\|\theta_1\phi + \theta_2\psi\|_2}{2^6\sqrt{2}L}},$$

from (6.2)–(6.3) we note that

$$\|\theta_1\phi + \theta_2\psi - q\|_2 < \alpha_0 \le L\sqrt{\log_1 \frac{\|\theta_1\phi\|_2 + 16\alpha_0}{2^6\sqrt{2}L}}. \tag{6.4}$$

It is easy to check that

$$s \le \sqrt{\log_1(t + s/4\sqrt{2})},\ s > 0 \text{ and } t \ge 0 \quad \Rightarrow \quad s \le \sqrt{\log_1(\sqrt{2}t)}.$$

Hence we deduce that

$$\|\theta_1\phi + \theta_2\psi - q\|_2 < L\sqrt{\log_1 \frac{\|\theta_1\phi\|_2}{2^6 L}}.$$

As we have already noted $\|\theta_2\psi\|_2 \le 4d(z_J/\|z_J\|_2)D_2(z_J/\|z_J\|_2)$, using the fact $z \in \text{Real}(\mathcal{Z})$, the triangle inequality, and (6.4), we conclude

$$\|\theta_1\phi - q\|_2 \le \|\theta_1\phi + \theta_2\psi - q\|_2 + \|\theta_2\psi\|_2 \le 17\alpha_0 < 2^5 L\sqrt{\log_1 \frac{\|\theta_1\phi\|_2}{2^6 L}}.$$

Since $|\theta_1| \le \|\theta\|_2 \le 2\mathfrak{D}$, the proof of the lemma is now complete. □

Next, we find a small net for $\text{Real}(\mathcal{Z})$. As in the genuinely complex case, we start with constructing a small net for the set of the small coordinates.

LEMMA 6.2. *Fix $J \subset [n]$ and $0 < \widetilde{\alpha} \le \mathfrak{D}$. Define*

$$S_J(\mathfrak{D}) := \big\{(u,v) \in \mathbb{R}^J \times \mathbb{R}^J : \|u\|_2^2 + \|v\|_2^2 = 1,\ \|u\|_2 \ge \|v\|_2,\\ d(u,v) \le \widetilde{\alpha}/\mathfrak{D}, \text{ and } \exists \theta \in [\mathfrak{D}, 3\mathfrak{D}], \text{ such that } \operatorname{dist}\big(\theta u, \mathbb{Z}^J\big) < \widetilde{\alpha}\big\}.$$

Then there exists a $(\frac{C_{6.2}\widetilde{\alpha}}{\mathfrak{D}})$-net $\mathcal{M}_J(\mathfrak{D}) \subset S_J(\mathfrak{D})$ with

$$|\mathcal{M}_J(\mathfrak{D})| \le \frac{\mathfrak{D}}{\widetilde{\alpha}} \cdot \left(\bar{C}_{6.2}\Big(\frac{\mathfrak{D}}{\sqrt{|J|}} + 1\Big)\right)^{|J|},$$

where $C_{6.2}$ and $\bar{C}_{6.2}$ are some absolute constants.

PROOF. Let $(u,v) \in S_J(\mathfrak{D})$, and let $\theta \in [\mathfrak{D}, 3\mathfrak{D}]$, $q \in \mathbb{Z}^J$ be such that

$$\|\theta u - q\|_2 < \widetilde{\alpha}.$$

Then, using the triangle inequality, $\|q\|_2 < \widetilde{\alpha} + |\theta| \le 4\mathfrak{D}$, and so $q \in \mathbb{Z}_{4\mathfrak{D}}$. This implies that

$$\left\|u - \frac{\mathfrak{D}}{\theta} \cdot \frac{q}{\mathfrak{D}}\right\|_2 < \frac{\alpha}{\mathfrak{D}} \qquad \text{where } \left|\frac{\mathfrak{D}}{\theta}\right| \le 1,\ \left\|\frac{q}{\mathfrak{D}}\right\|_2 \le 4. \tag{6.5}$$

From the definition of real-imaginary decorrelation, it also follows that

$$\|v\|_2 \le 2d(u,v) \le \frac{2\widetilde{\alpha}}{\mathfrak{D}}. \tag{6.6}$$

Let $\mathcal{N}_1$ be an $(\widetilde{\alpha}/\mathfrak{D})$-net in $[-1,1]$ with $|\mathcal{N}_1| \le 2\mathfrak{D}/\widetilde{\alpha}$. Define $\mathcal{M}_J^1(\mathfrak{D})$ by

$$\mathcal{M}_J^1(\mathfrak{D}) := \left\{\left(x\frac{q}{\mathfrak{D}}, 0\right) : q \in \mathbb{Z}_{4\mathfrak{D}}, x \in \mathcal{N}_1\right\}.$$

Then from (6.5)–(6.6) we deduce that $\mathcal{M}_J^1(\mathfrak{D})$ is a $(7\widetilde{\alpha}/\mathfrak{D})$-net for $S_J(\mathfrak{D})$ and $|\mathcal{M}_J^1(\mathfrak{D})| = |\mathbb{Z}_{4\mathfrak{D}}| \cdot |\mathcal{N}_1|$. This in combination with the bound in (4.5) yields the required estimate for the cardinality of the net. To complete the proof, we have to replace the constructed set of vectors by a subset of $S_J(\mathfrak{D})$. This is done in the same way as in Lemma 4.6. We skip the details. □

Now we use Lemma 6.2 to construct a small net in the set of essentially real vectors with an approximately constant value of the one-dimensional LCD. Define the set $\widetilde{\mathcal{Z}}(\mathfrak{D})$ by

$$\widetilde{\mathcal{Z}}(\mathfrak{D}) := \left\{z \in \operatorname{Real}(\mathcal{Z}) : \frac{z_{\text{small}}}{\|z_{\text{small}}\|_2} = \phi + \mathrm{i}\psi,\ \|\phi\|_2 \ge \|\psi\|_2,\ D_1(\phi) \in [\mathfrak{D}, 2\mathfrak{D}],\\ d(\phi,\psi) \le \widetilde{\alpha}/\mathfrak{D}\right\},$$

where

$$\widetilde{\alpha} := 2^5 L \sqrt{\log_1 \frac{\mathfrak{D}}{2^5 L}}. \tag{6.7}$$

The set $\widetilde{\mathcal{Z}}(\mathfrak{D})$ is the collection of vectors in $\mathrm{Real}(\mathcal{Z})$ for which $D_1(z_{\mathrm{small}}/\|z_{\mathrm{small}}\|_2) \approx \mathfrak{D}$. The condition $d(\phi,\psi) \le \widetilde{\alpha}/\mathfrak{D}$ ensures that the real-imaginary decorrelation is small.

PROPOSITION 6.3. *Fix $\mathfrak{D} > 1$. Let $\widetilde{\alpha}$ be as in (6.7) and assume $0 < \widetilde{\alpha} \le \mathfrak{D}$. Then there exist absolute constants $C_{6.3}$, $\bar{C}_{6.3}$ and a set $\widetilde{\mathcal{N}}(\mathfrak{D}) \subset \widetilde{\mathcal{Z}}(\mathfrak{D})$ with*

$$|\widetilde{\mathcal{N}}(\mathfrak{D})| \le \bar{C}_{6.3}^n \left(\frac{n}{\rho M} \cdot \frac{\mathfrak{D}}{\widetilde{\alpha}}\right)^{4M} \cdot \left(\frac{\mathfrak{D}}{\sqrt{n}} + 1\right)^{n-M}$$

having the following approximation property: Let $z \in \widetilde{\mathcal{Z}}(D)$ be any vector and denote $J = \mathrm{small}(z)$. Then there exists $w \in \widetilde{\mathcal{N}}(D)$ such that

$$\left\| \frac{z_J}{\|z_J\|_2} - \frac{w_J}{\|w_J\|_2} \right\|_2 < C_{6.3} \frac{\widetilde{\alpha}}{\mathfrak{D}}, \qquad \|z_{J^c} - w_{J^c}\|_2 \le C_{6.3} \frac{\rho\widetilde{\alpha}}{\mathfrak{D}},$$

$$\big|\|z_J\|_2 - \|w_J\|_2\big| \le C_{6.3} \frac{\rho\widetilde{\alpha}}{\mathfrak{D}}.$$

Proposition 6.3 is derived from Lemma 6.2 in the same way as Proposition 4.7 was derived from Lemma 4.6. We omit the details.

Now, we are ready to prove the main result of this section which shows that with high probability, there are no essentially real vectors with a subexponential LCD in the kernel of B^D.

PROPOSITION 6.4. *Let B^D, A_n, ρ, K, R, r and r' be as in Proposition 3.1. Then there exists a positive constant $c'_{6.4}$, depending only on K, R and the fourth moment of $\{\xi_{ij}\}$, such that*

$$\mathbb{P}\big(\exists z \in \mathrm{Real}(\mathcal{Z}) \cap \mathrm{Ker}\big(B^D\big) : D_2\big(z_{\mathrm{small}}/\|z_{\mathrm{small}}\|_2\big) \le \exp\big(c'_{6.4} n/M\big) \text{ and } \|A_n\| \le K\sqrt{pn}\big) \le e^{-n},$$

where $M = C_{3.1}\rho^{-4}p^{-1}$.

PROOF. The proof of this proposition is very similar to that of Proposition 5.2. First, we note that using Lemma 6.1 it follows that it is enough to show that, with high probability, there does not exist $z \in \mathrm{Ker}(B^D) \cap \mathrm{Real}(\mathcal{Z})$ such that $D_1(\phi(z)) \le \exp(c'n/M)$ for some small constant c', where $z_{\mathrm{small}}/\|z_{\mathrm{small}}\|_2 =: \phi(z) + \mathrm{i}\psi(z)$ with $\|\phi(z)\|_2 \ge \|\psi(z)\|_2$. We then claim that the subset of $\mathrm{Real}(\mathcal{Z})$ in context can be partitioned into the sets $\widetilde{\mathcal{Z}}(\mathfrak{D})$ as follows:

$$\big\{z \in \mathrm{Real}(\mathcal{Z}) : D_1\big(\phi(z)\big) \le \exp\big(c'n/M\big)\big\} \subset \bigcup_{\mathfrak{D}} \widetilde{\mathcal{Z}}(\mathfrak{D}), \tag{6.8}$$

where the union is taken over all $\mathfrak{D} = 2^k$, $\mathfrak{D} \le \exp(c'n/M)$. Note that the claim in (6.8) is obvious if we drop the requirement $d(z_{\text{small}}/\|z_{\text{small}}\|_2) = d(\phi(z), \psi(z)) \le \widetilde{\alpha}/\mathfrak{D}$ from the definition of $\widetilde{\mathcal{Z}}(\mathfrak{D})$. We show that the required condition on the real-imaginary decorrelation is automatically satisfied for all $z \in \text{Real}(\mathcal{Z})$. Indeed, recalling the definition of $\text{Real}(\mathcal{Z})$, and the fact that

$$\Delta\big(z_{\text{small}}/\|z_{\text{small}}\|_2\big) \le D_2\big(z_{\text{small}}/\|z_{\text{small}}\|_2\big)$$

we see that for any $z \in \text{Real}(\mathcal{Z})$,

$$\begin{aligned} d\big(z_{\text{small}}/\|z_{\text{small}}\|_2\big) &\le \frac{4L}{D_2(z_{\text{small}}/\|z_{\text{small}}\|_2)}\sqrt{\log_1 \frac{D_2(z_{\text{small}}/\|z_{\text{small}}\|_2)}{2^7 L}} \\ &\le \frac{8L}{D_1(\phi(z))}\sqrt{\log_1 \frac{D_1(\phi(z))}{2^8 L}}, \end{aligned} \tag{6.9}$$

where the last inequality is obtained upon noting that $x\sqrt{\log_1(1/x)}$ is an increasing function for $x \in (0, e^{-1}]$ together with an application of Lemma 6.1. If $z \in \text{Real}(\mathcal{Z})$ such that $D_1(\phi(z)) \in [\mathfrak{D}, 2\mathfrak{D}]$ then recalling the definition of $\widetilde{\alpha}$, from (6.9) we see that

$$d\big(\phi(z), \psi(z)\big) = d\big(z_{\text{small}}/\|z_{\text{small}}\|_2\big) \le \widetilde{\alpha}/\mathfrak{D},$$

which in turn proves the claim (6.8). We further claim that the lower bound on $\mathfrak{D}$ in (6.8) can be improved to

$$\mathfrak{D}_0 := C_0 \rho^{-1} p^{-1/2},$$

where $C_0 := \sqrt{C_{3.1}}/2$. To see this we note that $\text{Real}(\mathcal{Z}) \subset \text{Incomp}(M, \rho)$. Therefore, for any $z \in \text{Real}(\mathcal{Z})$ we have

$$\|\phi(z)\|_\infty \le 2\frac{\|z_{\text{small}}\|_\infty}{\|z_{\text{small}}\|_2} \le \frac{2}{\rho\sqrt{M}} = \frac{2\sqrt{p}}{\rho^{-1}\sqrt{C_{3.1}}},$$

where the last step follows from our choice of M. Hence, using Lemma 5.8 we see that for any $z \in \text{Real}(\mathcal{Z})$ we must have $D_1(\phi(z)) \ge \mathfrak{D}_0$. This establishes that the union in the RHS of (6.8) can be taken over all $\mathfrak{D} = 2^k$, $\mathfrak{D}_0 \le \mathfrak{D} \le \exp(c'n/M)$. So using the union bound, we deduce that it is enough to show that

$$\mathbb{P}\big(\exists z \in \widetilde{\mathcal{Z}}(\mathfrak{D}) : B^D z = 0, \text{ and } \|B^D\| \le (K+R)\sqrt{pn}\big) \le e^{-2n},$$

for each such $\mathfrak{D}$.

To this end, using Proposition 5.5 we see that for any $w \in \widetilde{\mathcal{N}}(\mathfrak{D})$ we have

$$\mathcal{L}\big(B^D w, \varepsilon\sqrt{p(n-1)}\|w_{\text{small}}\|_2\big) \le \left[\overline{C}_{5.5}\left(\varepsilon + \frac{1}{\sqrt{p}\mathfrak{D}}\right)\right]^{n-1}.$$

Now set

$$\widetilde{\varepsilon}_0 := 40 C_{6.3}(K+R)\frac{\widetilde{\alpha}}{\mathfrak{D}}.$$

Since the fact $\widetilde{\alpha} \geq L = (\delta_0 p)^{-1/2}$ implies that $\widetilde{\varepsilon}_0 \geq \frac{1}{\sqrt{p}\mathfrak{D}}$, we obtain that for any $w \in \widetilde{\mathcal{N}}(\mathfrak{D})$,

$$\mathbb{P}\Big(\|B^D w\|_2 \leq \frac{\varepsilon_0}{2} \|w_{\text{small}}\|_2 \cdot \sqrt{pn} \Big) \leq \Big[\overline{C}_{5.5} \Big(\widetilde{\varepsilon}_0 + \frac{1}{\sqrt{p}\mathfrak{D}} \Big) \Big]^{n-1} \leq \Big(\widetilde{C}(K+R) \frac{\widetilde{\alpha}}{\mathfrak{D}} \Big)^{n-1},$$

for some constant $\widetilde{C}$. Hence, by the union bound and applying Proposition 6.3 we obtain

$$\begin{aligned} \mathbb{P}\Big(\exists w \in \widetilde{\mathcal{N}}(\mathfrak{D}) : \|B^D w\|_2 \leq \frac{\varepsilon_0}{2} \|w_{\text{small}}\|_2 \cdot \sqrt{pn} \Big) &\leq |\widetilde{\mathcal{N}}(\mathfrak{D})| \cdot \Big(\widetilde{C}(K+R) \frac{\widetilde{\alpha}}{\mathfrak{D}} \Big)^{n-1} \\ &\leq \Big(C'(K+R) \frac{\widetilde{\alpha}}{\mathfrak{D}} \Big)^{n-1} \Big(\frac{n}{\rho M} \cdot \frac{\mathfrak{D}}{\widetilde{\alpha}} \Big)^{4M} \Big(\frac{\mathfrak{D}}{\sqrt{n}} + 1 \Big)^{n-M}, \end{aligned}$$

where C' is some large constant. Next recalling the definitions of $\widetilde{\alpha}$ and $\mathfrak{D}_0$, using the facts that $\mathfrak{D} \geq \mathfrak{D}_0$, $L = (\delta_0 p)^{-1/2}$ and the function $f(x) := x\sqrt{\log_1(1/x)}$ is increasing for $x \in (0, e^{-1})$ we find that

$$\frac{\widetilde{\alpha}}{\mathfrak{D}} = \frac{2^5 L}{\mathfrak{D}} \sqrt{\log_1 \frac{\mathfrak{D}}{2^5 L}} \leq \frac{2^5 L}{\mathfrak{D}_0} \sqrt{\log_1 \frac{\mathfrak{D}_0}{2^5 L}} = f\Big(\frac{2^5}{C_0 \rho^{-1} \delta_0^{1/2}} \Big). \tag{6.10}$$

Recalling the definition of C_0 and enlarging $C_{3.1}$ we therefore note from above that we can assume $\widetilde{C}(K+R)\widetilde{\alpha}/\mathfrak{D} < 1$. This yields

$$\mathbb{P}\Big(\exists w \in \widetilde{\mathcal{N}}(\mathfrak{D}) : \|B^D w\|_2 \leq \frac{\varepsilon_0}{2} \|w_{\text{small}}\|_2 \cdot \sqrt{pn} \Big) \leq \exp(-\widetilde{\Gamma} n),$$

where

$$\widetilde{\Gamma} := -\Big(1 - \frac{M}{n} \Big) \cdot \log\Big(\frac{C'(K+R)\widetilde{\alpha}}{\sqrt{n}} + \frac{C'(K+R)\widetilde{\alpha}}{\mathfrak{D}} \Big) - \frac{4M}{n} \log\Big(\frac{n}{\rho M} \cdot \frac{\mathfrak{D}}{\widetilde{\alpha}} \Big).$$

We next show that $\widetilde{\Gamma} \geq 2$ which allows us to deduce that

$$\mathbb{P}\Big(\exists w \in \widetilde{\mathcal{N}}(\mathfrak{D}) : \|B^D w\|_2 \leq \frac{\varepsilon_0}{2} \|w_{\text{small}}\|_2 \cdot \sqrt{pn} \Big) \leq \exp(-2n). \tag{6.11}$$

To prove that $\widetilde{\Gamma} \geq 2$, we recall that $\frac{n}{M} \geq \frac{1}{\rho}$ and $L \leq \widetilde{\alpha} \leq \mathfrak{D} \leq \exp(c' \frac{n}{M})$. Therefore,

$$\frac{4M}{n} \log\Big(\frac{n}{\rho M} \cdot \frac{\mathfrak{D}}{\widetilde{\alpha}} \Big) \leq 10,$$

upon choosing c' sufficiently small. Using the fact $M \le n/2$, this yields

$$\widetilde{\Gamma} \ge -\frac{1}{2} \cdot \log\left(\frac{C'(K+R)\widetilde{\alpha}}{\sqrt{n}} + \frac{C'(K+R)\widetilde{\alpha}}{\mathfrak{D}} \right) - 10.$$

Recalling (6.10), we see that we may enlarge $C_{3.1}$ (and thus, the minimal value of $\mathfrak{D}$) further so that $\widetilde{C}(K+R)\widetilde{\alpha}/\mathfrak{D} < e^{-30}$. Using the upper bound for $\mathfrak{D}$, we also note that

$$\frac{\widetilde{\alpha}}{\sqrt{n}} \le \frac{2^5 L\sqrt{\log_1 \frac{\mathfrak{D}}{2^5 L}}}{\sqrt{n}} \le \frac{2^5 L\sqrt{c' \frac{n}{M}}}{\sqrt{n}} = \frac{2^5\sqrt{c'}}{\sqrt{d_0 M p}} \le \frac{2^5 \rho^2 \sqrt{c'}}{\sqrt{C_{3.1} d_0}} \le e^{-30},$$

where the second last inequality follows from our choice of M, and the last inequality results from enlarging $C_{3.1}$ once more. This completes the proof of the claim that $\widetilde{\Gamma} \ge 2$. Thus we have shown that (6.11) holds. The rest of the proof relies on the approximation of a general point of $\widetilde{\mathcal{Z}}(\mathfrak{D})$ by a point of the set $\widetilde{\mathcal{N}}(\mathfrak{D})$, and is exactly the same as that of Proposition 5.2. We leave the details to the reader. This completes the proof. □

7. Proof of Theorem 2.2. In this section our goal is to combine the results of previous sections and complete the proof of Theorem 2.2. First, let us state the following general result from which Theorem 2.2 follows.

THEOREM 7.1. *Let A_n be an $n \times n$ matrix with i.i.d. entries $a_{i,j} = \delta_{i,j}\xi_{i,j}$, where $\{\delta_{i,j}\}$ are independent Bernoulli random variables taking value 1 with probability $p_n \in (0,1]$, and $\{\xi_{i,j}\}$ are i.i.d. centered real-valued random variables with unit variance and finite fourth moment. Fix $K, R \ge 1$, and $r \in (0,1]$ and let $\Omega_K := \{\|A_n\| \le K\sqrt{np_n}\}$. Assume that D_n is a diagonal matrix such that $\|D_n\| \le R\sqrt{np_n}$ and $\mathrm{Im}(D_n) = r'\sqrt{np_n} I_n$ with $|r'| \in [r,1]$. Then there exists constants $0 < c_{7.1}, \bar{c}_{7.1}, c'_{7.1}, C_{7.1}, C'_{7.1}, \overline{C}_{7.1} < \infty$, depending only on K, R, r and the fourth moment of $\{\xi_{i,j}\}$, such that for any $\varepsilon > 0$ we have the following*:

(i) *If*

$$p_n \ge \frac{\overline{C}_{7.1} \log n}{n},$$

then

$$\mathbb{P}\left(\left\{ s_{\min}(A_n + D_n) \le c_{7.1}\varepsilon \exp\left(-C_{7.1} \frac{\log(1/p_n)}{\log(np_n)} \right) \sqrt{\frac{p_n}{n}} \right\} \cap \Omega_K \right)$$
$$\le \varepsilon + \frac{C'_{7.1}}{\sqrt{np_n}}.$$

(ii) *Additionally, if*

$$\log(1/p_n) < \bar{c}_{7.1} (\log np_n)^2, \tag{7.1}$$

then

$$\mathbb{P}\Bigg(\Bigg\{s_{\min}(A_n + D_n) \le c_{7.1}\varepsilon \exp\Bigg(-C_{7.1}\frac{\log(1/p_n)}{\log(np_n)}\Bigg)\sqrt{\frac{p_n}{n}}\Bigg\} \cap \Omega_K\Bigg) \le \varepsilon + \exp\bigl(-c'_{7.1}\sqrt{np_n}\bigr).$$

The proof of part (i) of Theorem 7.1 follows from Berry–Esséen theorem and Proposition 3.4. The proof of part (ii) uses results from Section 5 and Section 6. Recall that in Section 5 and Section 6 we have shown that there does not exist any vector in $\mathrm{Ker}(B^D)$ with a subexponential two-dimensional LCD, with high probability. To prove the second part of Theorem 7.1, we use LCD based bounds on Lévy concentration function. At this moment, we know that with high probability, any vector in in $\mathrm{Ker}(B^D)$ has an exponential two-dimensional LCD. However, we do not have any control the real-imaginary decorrelation of this vector. This means that we cannot use the bound (5.7), and have to rely on (5.8).

To apply (5.8), we therefore need to show that any vector with a large two-dimensional LCD must also admit a large value of one-dimensional LCD. This calls for another modification to the definition of the one-dimensional LCD.

DEFINITION 7.2. For a nonzero vector $x \in \mathbb{R}^m$, we set

$$\widehat{D}_1(x) := \inf\Bigg\{\theta > 0 : \mathrm{dist}\bigl(\theta x, \mathbb{Z}^m\bigr) < L\sqrt{\log_1 \frac{\|\theta x\|_2}{2^8 L}}\Bigg\}.$$

The advantage of working with this one-dimensional LCD $\widehat{D}_1(\cdot)$ can be seen from the following result.

LEMMA 7.3. *For $z := x + \mathrm{i}y \in \mathbb{C}^m$, we have $\widehat{D}_1(x) \ge D_2(z)$.*

PROOF. The proof follows by simply noting that if there exists a $\theta' > 0$ such that

$$\mathrm{dist}\bigl(\theta' x, \mathbb{Z}^m\bigr) < L\sqrt{\log_1 \frac{\|\theta' x\|_2}{2^8 L}},$$

then for $\theta = (\theta', 0)$ we also have that

$$\mathrm{dist}\bigl(V^{\mathsf{T}}\theta, \mathbb{Z}^m\bigr) < L\sqrt{\log_1 \frac{\|V^{\mathsf{T}}\theta\|_2}{2^8 L}}. \qquad \square$$

Now we are ready to prove Theorem 7.1.

PROOF OF THEOREM 7.1. The proof is similar to that of [9], Theorem 1.1. We include it for completeness. Note that for any $\vartheta > 0$,

$$\begin{aligned}\mathbb{P}\big(\{s_{\min}(A_n + D_n) \le \vartheta\} \cap \Omega_K\big) &\le \mathbb{P}\Big(\Big\{\inf_{x\in V^c} \|(A_n + D_n)x\|_2 \le \vartheta\Big\} \cap \Omega_K\Big) \\ &\quad + \mathbb{P}\Big(\Big\{\inf_{x\in V} \|(A_n + D_n)x\|_2 \le \vartheta\Big\} \cap \Omega_K\Big),\end{aligned} \tag{7.2}$$

where

$$V := S_{\mathbb{C}}^{n-1} \setminus \big(\mathrm{Comp}(c_{3.4}n, \rho) \cup \mathrm{Dom}\big(c_{3.4}n, (C_{3.4}(K + R))^{-4}\big)\big),$$

and ρ as in Proposition 3.4. Using Proposition 3.4 with $M = c_{3.4}n$, we obtain that

$$\mathbb{P}\Big(\inf_{x\in V^c} \|(A_n + D_n)x\|_2 \le c'_{3.4}(K + R)\rho\sqrt{np}, \|A_n\| \le K\sqrt{pn}\Big) \le \exp(-\overline{c}_{3.4}np).$$

Therefore, it only remains to find an upper bound on the second term in the RHS of (7.2). Applying Lemma 2.7, we see that to find an upper bound of

$$\mathbb{P}\left(\left\{\inf_{x\in V} \|(A_n + D_n)x\|_2 \le \varepsilon\rho^2\sqrt{\frac{p}{n}}\right\} \cap \Omega_K\right)$$

is enough to find the same for

$$\mathbb{P}\big(\{\mathrm{dist}(A_{n,j}, H_{n,j}) \le \rho\sqrt{p}\varepsilon\} \cap \Omega_K\big) \qquad \text{for a fixed } j,$$

where $A_{n,j}$ are columns of $(A_n + D_n)$. As these estimates are the same for different j's, we only need to consider the case $j = 1$. Recall that B^D is the matrix whose rows are the columns $A_{n,2}, \dots, A_{n,n}$. Therefore,

$$\mathrm{dist}(A_{n,1}, H_{n,1}) \ge |\langle v, A_{n,1}\rangle|,$$

for any $v \in S_{\mathbb{C}}^{n-1} \cap \mathrm{Ker}(B^D)$. Thus it is enough to find an upper bound on

$$\mathbb{P}\big(\{\exists v \in \mathcal{Z} \cap \mathrm{Ker}(B^D) : |\langle A_{n,1}, v\rangle| \le \rho\varepsilon\sqrt{p}\} \cap \Omega_K\big). \tag{7.3}$$

First, we obtain a bound on (7.3) under the assumption of part (i). This follows from a simple Berry–Esséen bound.

Since $v \in S_{\mathbb{C}}^{n-1} \cap \mathrm{Ker}(B^D)$ using Proposition 3.4 again, we may assume that $v \notin \mathrm{Comp}(c_{3.4}n, \rho) \cup \mathrm{Dom}(c_{3.4}n, (C_{3.4}(K + R))^{-4})$. Let $J = \mathrm{supp}(v_{[c_{3.4}n+1,n]})$. Then

$$\mathbb{P}\big(|\langle A_{n,1}, v\rangle| \le \rho\varepsilon\sqrt{p}\big) \le \mathcal{L}\Big(\sum_{i\in J} v_i d_i \xi_i, \rho\sqrt{p}\varepsilon\Big).$$

Since $v \notin \mathrm{Comp}(c_{3.4}n, \rho) \cup \mathrm{Dom}(c_{3.4}n, (C_{3.4}(K + R))^{-4})$, we have

$$\|v_J\|_\infty \le \frac{C_{3.4}(K + R))^4}{\sqrt{c_{3.4}n}}\|v_J\|_2 \quad \text{and} \quad \|v_J\|_2 \ge \rho.$$

The Berry–Esséen theorem (see [37], Theorem 2.2.17) then yields that

$$\begin{aligned}\mathcal{L}\Big(\sum_{i\in J} v_i d_i \xi_i, \rho\sqrt{p}\varepsilon\Big) &\le C\varepsilon + C'\frac{p\|v_J\|_3^3}{p^{3/2}\|v_J\|_2^3} \\ &\le C\varepsilon + C'\frac{\|v_J\|_\infty}{p^{1/2}\|v_J\|_2} \le C\varepsilon + \frac{C''}{\sqrt{pn}},\end{aligned} \tag{7.4}$$

where C is an absolute constant, the constant C' depends only on the fourth moment of $\{\xi_{i,j}\}$, and

$$C'' = \frac{C_{3.4}(K+R))^4}{\sqrt{c_{3.4}}}\cdot C'.$$

Replacing ε by ε/C completes the proof of part (i) of the theorem.

It remains to prove part (ii). As seen above, we only need to obtain a bound on (7.3) under the stronger assumption of p_n of part (ii). To this end, we apply Proposition 3.4 again. Setting $M_0 = C_{3.1}\rho^{-4}p^{-1}$ from Proposition 3.4, we find that it is enough to bound

$$\mathbb{P}\big(\{\exists v \in V_0 \cap \operatorname{Ker}(B^D) : |\langle A_{n,1}, v\rangle| \le \rho\varepsilon\sqrt{p}\} \cap \Omega_K\big), \tag{7.5}$$

where

$$V_0 := S_{\mathbb{C}}^{n-1}\backslash\big(\operatorname{Comp}(M_0,\rho)\cup \operatorname{Dom}(M_0,(C_{3.4}(K+R))^{-4})\big).$$

Further denote

$$V_1 := \{w \in V_0 : D_2(w_{\text{small}}/\|w_{\text{small}}\|_2) \le \exp(c'n/M_0)\} \quad \text{and} \quad V_2 := V_0\backslash V_1,$$

where $c' := \min\{c'_{5.1}, c'_{6.4}\}$. We will show that

$$\mathbb{P}\big(\{\exists v \in V_1 \cap \operatorname{Ker}(B^D)\} \cap \Omega_K\big) \le \exp(-\bar{c}np), \tag{7.6}$$

for some $\bar{c} > 0$. Since $\operatorname{Ker}(B^D)$ is invariant under rotation, recalling the definition of the set $\mathcal{Z}$ (see (4.4)), we see that it is enough to show that

$$\begin{aligned}&\mathbb{P}\big(\{\exists v \in \mathcal{Z} \cap \operatorname{Ker}(B^D) : D_2(v_{\text{small}}/\|v_{\text{small}}\|_2) \le \exp(c'n/M_0)\} \cap \Omega_K\big) \\ &\quad\le \exp(-\bar{c}np).\end{aligned}$$

Note that, if p satisfies (7.1) with a sufficiently small $\bar{c}_{7.1}$, then it also satisfies the assumption (5.2). So we can apply Theorem 5.1. Applying Theorem 5.1 and Proposition 6.4, we then immediately obtain our claim (7.6). Therefore, it only remains to find an upper bound on

$$\begin{aligned}&\mathbb{P}\big(\{\exists v \in \mathcal{Z} \cap \operatorname{Ker}(B^D) : D_2(v_{\text{small}}/\|v_{\text{small}}\|_2) > \exp(c'n/M_0) \text{ and} \\ &\qquad |\langle A_{n,1}, v\rangle| \le \rho\varepsilon\sqrt{p}\} \cap \Omega_K\big).\end{aligned} \tag{7.7}$$

To obtain the desired bound we condition on B^D, which fixes the vector v for which

$$D_2\big(v_{\mathrm{small}}/\|v_{\mathrm{small}}\|_2\big) > \exp\big(c'n/M_0\big).$$

Lemma 7.3 implies that

$$\widehat{D}_1\big(\phi(v)\big) > \exp\big(c'n/M_0\big),$$

where we recall that $v_{\mathrm{small}}/\|v_{\mathrm{small}}\|_2 = \phi(v) + \mathrm{i}\psi(v)$. Inequality (5.8) holds with $\widehat{D}_1(\cdot)$ instead of $D_1(\cdot)$ if the constant $\overline{C}_{5.5}$ is appropriately adjusted. Recalling the definition of M_0 we deduce that

$$\begin{aligned}\mathbb{P}\big(\big|\langle A_{n,1}, v\rangle\big| \le \varepsilon\rho\sqrt{p}\big) &\le \bar{C}\bigg(\varepsilon + \frac{1}{\sqrt{p}\,\widehat{D}_1(\phi(v))}\bigg)\\ &\le \bar{C}\bigg(\varepsilon + \frac{1}{\sqrt{p}}\exp\big(-c''np\rho^4\big)\bigg),\end{aligned}$$

for some constants $\bar{C}$ and c''. Choosing $\bar{c}_{7.1}$ sufficiently small and recalling the definition of ρ, we further deduce that

$$\frac{1}{\sqrt{p}}\exp\big(-c''np\rho^4\big) \le \exp\big(-c''\sqrt{np}\big).$$

Therefore, replacing ε by $\varepsilon/\bar{C}$ we conclude that (7.7) is bounded by

$$\varepsilon + \bar{C}\exp\big(-c''\sqrt{np}\big).$$

This completes the proof of the theorem. □

PROOF OF THEOREM 2.2. The proof immediately follows from Theorem 7.1, [9], Theorem 1.7, and the triangle inequality. □

REMARK 7.4. From the proof of Theorem 7.1, we note that the assumption (1.2) (equivalently (7.1)) was needed to show that the assumption (5.2) holds. From [9], Proposition 3.1, we have $\rho = \exp(-C\log(1/p)/\log(np))$, for some large C. If one can improve the conclusion of [9], Proposition 3.1, to accommodate $\rho = \Omega(1)$ then it is obvious that (5.2) holds without the assumption (1.2) and, therefore, Theorem 1.3(ii) can be extended without any extra assumption.

8. Intermediate singular values. The goal of this short section is to prove Theorem 2.10 which shows that there are not too many singular values of the matrix $\frac{1}{\sqrt{np}}A_n - wI_n$ near zero. In proving Theorem 2.10, we employ the same strategy as in [13, 40, 44]. Namely, we first show that the distance of any row of A_n from any given subspace of not very large dimension cannot be too small with large probability.

LEMMA 8.1. *Let* $\boldsymbol{a} := (\xi_i \delta_i)_{i=1}^n$ *be an* n*-dimensional vector where* $\{\xi_i\}_{i=1}^n$ *are i.i.d. with zero mean and unit variance and* $\{\delta\}_{i=1}^n$ *are i.i.d.* Ber(p). *Let* $\psi : \mathbb{N} \mapsto \mathbb{N}$ *be such that* $\psi(n) \to \infty$ *and* $\psi(n) < n$. *Then there exists a positive finite constant* $c_{8.1}$[4] *such that for every subspace* H *of* $\mathbb{C}^n$ *with* $1 \le \dim(H) \le n - \psi(n)$, *we have*

$$\mathbb{P}\big(\operatorname{dist}(\boldsymbol{a}, H) \le c_{8.1}\sqrt{p\big(n - \dim(H)\big)}\big) \le \exp\big(-c_{8.1} p\psi(n)\big) + \exp\big(-c_{8.1}\psi^2(n)/n\big).$$

A result similar to Lemma 8.1 was obtained in [40] (see Proposition 5.1 there) for the dense case. Later in [13] (and [44]) it was improved for the sparse case. Our Lemma 8.1 follows from [13], Lemma 3.5, when applied to the set-up of this paper. So we omit the proof and refer the reader to the proof of [13], Lemma 3.5.

We now complete the proof of Theorem 2.10 using Lemma 8.1. We use same approach as in [40], pages 2055–2056 (see also the proof of [13], Lemma 3.14).

PROOF OF THEOREM 2.10. To lighten the notation, let us denote by $s_1 \ge s_2 \ge \cdots \ge s_n$ the singular values of $(A_n - \sqrt{np}wI_n)$. Fix i such that $3\psi(n) \le i \le n - 1$ and denote by $A_n^{m,w}$ the submatrix formed by first m rows of the matrix $(A_n - \sqrt{np}wI_n)$, where $m = n - \lceil i/2 \rceil$. Further denote by $s_1' \ge s_2' \ge \cdots \ge s_m'$ the singular values of $A_n^{m,w}$. Using Cauchy's interlacing inequality, we see that

$$s_{n-i}' \le s_{n-i}. \tag{8.1}$$

Next, from [40], Lemma A.4, it follows that

$$s_1'^{-2} + s_2'^{-2} + \cdots + s_m'^{-2} = \operatorname{dist}_1'^{-2} + \operatorname{dist}_2'^{-2} + \cdots + \operatorname{dist}_m'^{-2}, \tag{8.2}$$

where $\operatorname{dist}_j' := \operatorname{dist}(\boldsymbol{a}_j - w\sqrt{np}e_j, H_{j,n}^{m,w})$, $\boldsymbol{a}_j^{\mathsf{T}}$ is the jth row of the matrix A_n, $H_{j,n}^{m,w}$ is the subspace spanned by all the rows of $A_n^{m,w}$ except the jth row, and e_j is the jth canonical basis. We also note that $\operatorname{dist}_j \le \operatorname{dist}_j'$, where $\operatorname{dist}_j := \operatorname{dist}(\boldsymbol{a}_j, \operatorname{span}(H_{j,n}^{m,w}, e_j))$. Thus from (8.1)–(8.2), we deduce

$$\frac{i}{2n} s_{n-i}^{-2} \le \frac{1}{n} \sum_{j=n-i}^{m} s_j'^{-2} \le \frac{1}{n} \sum_{j=1}^{m} \operatorname{dist}_j^{-2}. \tag{8.3}$$

It is easy to note that $\dim(\operatorname{span}(H_{j,n}^{m,w}, e_j)) \le m + 1 \le n - \psi(n)$ for all $j = 1, 2, \ldots, m$. Therefore, from Lemma 8.1 we further obtain

$$\mathbb{P}(\operatorname{dist}_j \le c_{8.1}\sqrt{p \cdot i/3}) \le 2n^{-4}, \qquad j = 1, 2, \ldots, m,$$

[4]The constant $c_{8.1}$ may depend on the tail of the distribution of $\{\xi_i\}_{i=1}^n$.

where we used the fact that $n - \dim(\mathrm{span}(H_{j,n}^{m,w}, e_j)) \geq n - (m+1) \geq i/3$ and chose $C_{2.10} \geq 4c_{8.1}^{-1}$. Hence, from (8.3) we see that

$$\mathbb{P}\left(s_{n-i} \leq \frac{c_{8.1}}{\sqrt{3}} \cdot \sqrt{np} \cdot \frac{i}{n} \right) \leq 2n^{-3},$$

for all i such that $3\psi(n) \leq i \leq n-1$. After taking the union over i, the proof of the theorem completes. □

9. Weak convergence. Here, our goal is to prove Theorem 2.11. As mentioned in Section 2.3, using a truncation argument, we first show that it is enough to restrict to the case of bounded $\{\xi_{i,j}\}_{i,j=1}^n$. To this end, we have the following lemma.

LEMMA 9.1. *If the conclusion of Theorem 2.11 holds for $\{\xi_{i,j}\}_{i,j=1}^n$ bounded, then it continues to hold without the boundedness assumption.*

The proof of the truncation argument has now become standard in the random matrix literature which follows from an application of Hoffman–Wielandt inequality (see [3], Lemma 2.1.19) upon using the fact that the bounded Lipschitz metric on the space of all probability measures on $\mathbb{R}$ metrizes the weak convergence of probability measures (see [3], Theorem C.8). We refer the reader to [3], Appendix C.2, for a definition of the bounded Lipschitz metric. Using the above two ingredients and proceeding similarly as in the the proof of [16], Proposition 4.1, one can complete the proof of Lemma 9.1. Further details are omitted.

Equipped with Lemma 9.1, hereafter we may and will assume that $\{\xi_{i,j}\}_{i,j=1}^n$ are bounded by some constant $\mathfrak{K}_0$. It is well known that the conclusion of Theorem 2.11 holds for $\nu_{G_n}^w$, the symmetrized version of the empirical law of the singular values of $(n^{-1/2}G_n - wI_n)$, where G_n is a real Ginibre matrix. Thus to prove Theorem 2.11 it is enough to show that

$$\int f(x)\, d\nu_{G_n}^w(x) - \int f(x)\, d\nu_n^w(x) \to 0 \qquad \text{as } n \to \infty, \text{ almost surely,} \tag{9.1}$$

for every $f \in C_c(\mathbb{R})$, the set of all continuous and compactly supported functions on $\mathbb{R}$, and $w \in \mathbb{C}$. To prove (9.1), we first show that both the random probability measures ν_n^w and $\nu_{G_n}^w$ are close to their expectations, $\mathbb{E}\nu_n^w$ and $\mathbb{E}\nu_{G_n}^w$, respectively, and then we establish that $\mathbb{E}\nu_n^w$ and $\mathbb{E}\nu_{G_n}^w$ are close to each other as well. To carry out the first step, we use the following concentration inequality.

LEMMA 9.2. *Let $H_n := (h_{i,j})_{i,j=1}^n$ be a Hermitian random matrix with entries on and above the diagonal are jointly independent. Let $f : \mathbb{R} \mapsto \mathbb{R}$ be an L-Lipschitz function supported on a compact interval $I \subset \mathbb{R}$, with $\|f\|_\infty := \sup_{x \in I} |f(x)| \leq 1$. Further, let H_0 be an arbitrary $n \times n$ deterministic matrix and denote $\mathsf{H}_n := H_n + H_0$. Fix an $\varepsilon > 0$.*

(i) *If the entries of H_n are uniformly bounded by $\mathfrak{K}/\sqrt{n}$ for some $\mathfrak{K} < \infty$ then*

$$\mathbb{P}\left(\left|\int f(x)\,dL_{\mathsf{H}_n}(x) - \mathbb{E}\int f(x)\,dL_{\mathsf{H}_n}(x)\right| \geq \varepsilon\right) \leq \frac{C_{9.2}L|I|}{\varepsilon}\exp\left(-\frac{c_{9.2}n^2\varepsilon^4}{\mathfrak{K}^2L^4|I|^2}\right),$$

for some absolute constants $c_{9.2}, C_{9.2} > 0$.

(ii) *If the entries of H_n satisfy the logarithmic Sobolev inequality with uniform constant $\mathfrak{L}$ then*

$$\mathbb{P}\left(\left|\int f(x)\,dL_{\mathsf{H}_n}(x) - \mathbb{E}\int f(x)\,dL_{\mathsf{H}_n}(x)\right| \geq \varepsilon\right) \leq 2\exp\left(-\frac{c_{9.2}n^2\varepsilon^2}{8\mathfrak{L}L^2}\right).$$

PROOF. Part (i) is a consequence of [19], Lemma 3.2. For $H_0 = 0$, the proof part (ii) is immediate from [28], Theorem 1.1(b). A key step in the proof [28], Theorem 1.1(b), is to show that for any Lipschitz function $f : \mathbb{R} \mapsto \mathbb{R}$ the map $H \mapsto \int f(x)d_{L_H}(x)$ is also a Lipschitz function (see [28], Lemma 1.2(b)). The same proof shows that for any deterministic H_0 the map $H \mapsto \int f(x)d_{L_{H+H_0}}(x)$ is again Lipschitz. Therefore, the general case follows. □

Next, we need to show that $\mathbb{E}\nu_n^w$ and $\mathbb{E}\nu_{G_n}^w$ are close to each other. This will be obtained upon showing that the corresponding Stieltjes transforms asymptotically are the same. Before stating the relevant result, let us define the Stieltjes transform of a probability measure on $\mathbb{R}$.

DEFINITION 9.3. Let μ be a probability measure on $\mathbb{R}$. Its Stieltjes transform is given by

$$G_\mu(\zeta) := \int_{\mathbb{R}} \frac{1}{x-\zeta}\,d\mu(x), \qquad \zeta \in \mathbb{C}\backslash\mathbb{R}.$$

We write $m_n(\zeta) := m_n(\zeta, w)$ and $m_{G_n}(\zeta) := m_{G_n}(\zeta, w)$ to denote the Stieltjes transform of ν_n^w and $\nu_{G_n}^w$, respectively. We now have the following lemma.

LEMMA 9.4. *Let A_n be an $n \times n$ matrix with entries $a_{i,j} = \delta_{i,j} \cdot \xi_{i,j}$, where $\delta_{i,j}$ are i.i.d. Bernoulli random variables with $\mathbb{P}(\delta_{i,j} = 1) = p$, and $\xi_{i,j}$ are centered i.i.d. random variables with unit variance bounded by $\mathfrak{K}_0$ for some $\mathfrak{K}_0$. Fix any $w \in \mathbb{C}$ and $\zeta \in \mathbb{C}^+$. Then*

$$|\mathbb{E}m_n(\zeta) - \mathbb{E}m_{G_n}(\zeta)| \leq \frac{C_{9.4}}{\sqrt{np}(\operatorname{Im}\zeta)^4}\left(1 + \frac{1}{(n\operatorname{Im}\zeta)^2}\right),$$

for some constant $C_{9.4}$ depending only on $\mathfrak{K}_0$.

A version of Lemma 9.4 was proved in [20] when A_n is replaced by B_n, a matrix of i.i.d. centered Ber(p) random variables. A crucial step in the proof of [20], Lemma 8.2, is the Lindeberg replacement strategy of [17]. The latter depends only the bounds on the first three moments of the entries of B_n. Since $\{\xi_{i,j}\}_{i,j=1}^n$ are uniformly bounded by $\mathfrak{K}_0$, one can obtain the bounds necessary to apply [17], Theorem 1.1, Corollary 1.2. Hence, following the same lines of arguments as in [20], Lemma 8.2, one completes the proof of Lemma 9.4.

Equipped with all the required ingredients we now complete the proof of Theorem 2.11.

PROOF OF THEOREM 2.11. It is known that for any $w \in \mathbb{C}$ the random probability measure $\nu_{G_n}^w$ converges weakly to ν_∞^w, for some probability measure, almost surely; for example, see [5]. Therefore, it is enough show that

$$\int f(x)\,d\nu_n^w(x) - \int f(x)\,d\nu_{G_n}^w(x) \to 0 \qquad \text{almost surely, as } n \to \infty, \tag{9.2}$$

for every bounded continuous function $f : \mathbb{R} \mapsto \mathbb{R}$. Since $\{\xi_{i,j}\}_{i,j=1}^n$ are uniformly bounded, by Chernoff's bound and the Borel–Cantelli lemma it follows that the sequence of probability measures $\{\nu_n^w\}_{n\in\mathbb{N}}$ are almost surely compactly supported, whenever $np \geq C_0 \log n$ for some large constant C_0. The strong law of large number further shows that $\{\nu_{G_n}^w\}_{n\in\mathbb{N}}$ are almost surely compactly supported. Hence, we only need to show that (9.2) holds for continuous and compactly supported functions f.

Now, given any $\delta > 0$, and a continuous and compactly supported function f, one can construct a Lipschitz function f_δ, such that $\|f - f_\delta\|_\infty \leq \delta$ with the Lipschitz constant of f_δ depending only on f and δ. Hence, it suffices to establish (9.2) only for compactly supported Lipschitz functions.

To this end, recalling that Gaussian random variables satisfy the logarithmic Sobolev inequality (see [27]), from Lemma 9.2 we deduce that

$$\int f(x)\,d\mathfrak{u}_n(x) - \mathbb{E}\int f(x)\,d\mathfrak{u}_n(x) \to 0 \qquad \text{as } n \to \infty, \tag{9.3}$$

almost surely, for $\mathfrak{u}_n = \nu_n^w$ and $\nu_{G_n}^w$. Since $np \to \infty$ as $n \to \infty$, from Lemma 9.4 it also follows that (one can argue similarly as in the proof of [3], Theorem 2.4.3)

$$\mathbb{E}\int f(x)\,d\nu_n^w(x) - \mathbb{E}\int f(x)\,d\nu_{G_n}^w(x) \to 0 \qquad \text{as } n \to \infty. \tag{9.4}$$

Combining (9.3)–(9.4), we establish that (9.2) holds for any Lipschitz and compactly supported function f. This completes the proof. □

10. Proof of Theorem 1.3. In this section we combine Theorem 2.2, Theorem 2.10 and Theorem 2.11 to prove Theorem 1.3. As already mentioned in Section 2, to prove Theorem 1.3 we need to invoke the replacement principle.

We fix $r \in (0,1)$ and define $\mathbb{D}_r := \{w \in B_{\mathbb{C}}(0,1-r) : |\operatorname{Im} w| \geq r\}$. Then applying Lemma 2.1, we show that for every $f \in C_c^2(\mathbb{C})$ supported on $\mathbb{D}_r$, we have $\int f(w)\,dL_n(w) \to \int f(w)\,d\mathfrak{m}(w)$ in probability or almost surely, depending on the choice of the sparsity parameter p, where L_n is the ESD of $\frac{1}{\sqrt{np}}A_n$. Afterward, letting $r \to 0$, we establish the circular law limit. Below we make this idea precise.

Before we prove Theorem 1.3, we need some properties of the probability measure ν_∞^w. Recall that ν_∞^w is the limit of the ESD of $\boldsymbol{A}_n^w$ where $\boldsymbol{A}_n^w$ was defined in (2.7).

LEMMA 10.1. (i) *For any $w \in B_{\mathbb{C}}(0,1)$, the probability measure ν_∞^w is supported on $[-\sqrt{\lambda_+}, \sqrt{\lambda_+}]$, where*

$$\lambda_+ := \lambda_+(w) := \frac{(\sqrt{1+8|w|^2}+3)^3}{8(\sqrt{1+8|w|^2}+1)}.$$

(ii) *There exists some absolute constant $r_0 \in (0,1)$ such that for all $r \in (0, r_0)$, $\tau \in (0,1)$, and $w \in B_{\mathbb{C}}(0,1-r)$ we have*

$$\int_{-\tau}^{\tau} \big|\log|x|\big|\,d\nu_\infty^w(x) \leq C_{10.1}\,\tau\,|\log\tau|,$$

for some positive constant $C_{10.1}$ which depends only on r.

PROOF. In [5], Lemma 4.2, it was shown that for any $w \in B_{\mathbb{C}}(0,1)$ the probability measure $\widetilde{\nu}_\infty^w$ is supported on $[0, \lambda_+(w)]$ where for any $t > 0$, $\widetilde{\nu}_\infty^w((0,t^2)) = \nu_\infty^w((-t,t))$. From this part, (i) of the lemma follows.

Turning to prove (ii), using integration by parts we note that for any probability measure μ on $\mathbb{R}$ and $0 < \tau < 1$,

$$\int_0^\tau \big|\log(x)\big|\,d\mu(x) \leq \big|\log(\tau)\big|\mu\big((0,\tau)\big) + \int_0^\tau \frac{\mu((0,t))}{t}\,dt. \tag{10.1}$$

Using [7], Lemma 7.8 and [7], Lemma 8.3, we see that for any $t \in (0,1)$,

$$\nu_\infty^w\big((0,t)\big) \leq \nu_\infty^w\big((-t,t)\big) \leq 2t \cdot \big(\operatorname{Im} m_\infty(\mathrm{i}t)\big) \leq 2Ct,$$

for some large constant C depending on r. The rest follows from (10.1). □

We are now ready to prove Theorem 1.3.

PROOF OF THEOREM 1.3. We first prove part (i). That is we show that the ESD of $A_n/\sqrt{np}$ converges weakly to the circular law, in probability. Fix $r \in (0,1/2)$ and denote $\mathbb{D}_r := \{w \in B_{\mathbb{C}}(0,1-r) : |\operatorname{Im} w| \geq r\}$. Let us also fix $w \in \mathbb{D}_r$. Define

$$\Omega_n' := \left\{ s_{\min}\left(\frac{A_n}{\sqrt{np}} - wI_n\right) \geq c_n \right\},$$

where

$$c_n := c_{2.2} \exp\left(-C_{2.2} \frac{\log(1/p)}{\log(np)}\right) \frac{1}{n\sqrt{np}}.$$

Setting $D_n := -w\sqrt{np}\, I_n$ and applying Theorem 2.2 we deduce that

$$\mathbb{P}(\Omega_n') \geq 1 - \frac{1 + C_{2.2}'}{\sqrt{np}}. \tag{10.2}$$

Fix any $\delta \in (0,1)$ and let $\tau := \tau(\delta) := c_{2.10}\delta$. Further denote $\psi(n) := \max\{\sqrt{n/p}, n/(\log n)^2\}$. Since $np = \omega(\log^2 n)$, we note that $\psi(n) = o(n/\log n)$. Equipped with these notation, we recall the definition of ν_n^w to see that

$$\begin{aligned} &\int_{-\tau}^{\tau} \big|\log(|x|)\big|\, d\nu_n^w(x) \\ &\quad = \frac{1}{n} \sum_{i=1}^{n-3\psi(n)} \big|\log(s_i)\big| \mathbb{I}(s_i \leq \tau) + \frac{1}{n} \sum_{i=n-3\psi(n)+1}^{n} \big|\log(s_i)\big| \mathbb{I}(s_i \leq \tau). \end{aligned} \tag{10.3}$$

We evaluate each term of the RHS of (10.3) separately. Focusing on the second term, we see that on the event Ω_n'

$$\begin{aligned} \frac{1}{n} \sum_{i=n-3\psi(n)+1}^{n} \big|\log(s_i)\big| \mathbb{I}(s_i \leq \tau) &\leq \big|\log(s_n)\big| \cdot \frac{3\psi(n)}{n} \\ &\leq \log c_n^{-1} \cdot \frac{3\psi(n)}{n} = o(1). \end{aligned} \tag{10.4}$$

We next consider the first term of (10.3). Denote the event described in Theorem 2.10 by Ω_n''. Since $\min\{p\psi(n), \psi^2(n)/n\} \geq C_{2.10} \log n$, on the event Ω_n'' we have

$$\begin{aligned} &\frac{1}{n} \sum_{i=1}^{n-3\psi(n)} \big|\log(s_i)\big| \mathbb{I}(s_i \leq \tau) \\ &\quad = \frac{1}{n} \sum_{i=3\psi(n)}^{n-1} \big|\log(s_{n-i})\big| \mathbb{I}(s_{n-i} \leq \tau) \\ &\quad \leq \frac{\log(1/c_{2.10})}{n} \sum_{i=3\psi(n)}^{n-1} \mathbb{I}(s_{n-i} \leq \tau) \\ &\qquad + \frac{1}{n} \sum_{i=3\psi(n)}^{n-1} \log\left(\frac{n}{i}\right) \mathbb{I}(s_{n-i} \leq \tau), \end{aligned} \tag{10.5}$$

and by Theorem 2.10, $\mathbb{P}((\Omega_n'')^c) \le 2/n^2$. Recalling the definition of τ, from Theorem 2.10 it also follows that

$$s_{n-i} \le \tau \Rightarrow i \le \delta n$$

on the event Ω_n''. So from (10.5), we deduce that

$$\begin{aligned}\frac{1}{n}\sum_{i=1}^{n-3\psi(n)} |\log(s_i)|\mathbb{I}(s_i \le \tau) &\le \delta \cdot \log(1/c_{2.10}) + \frac{1}{n}\sum_{i=3\psi(n)}^{\delta n} \log\left(\frac{n}{i}\right) \\ &\le \delta \cdot \log(1/c_{2.10}) - 2\int_0^{\delta} \log x \, dx,\end{aligned} \tag{10.6}$$

for all large n. Hence, denoting $\Omega_n := \Omega_n' \cup \Omega_n''$, from (10.3)–(10.4) and (10.6) we obtain that

$$\int_{-\tau(\delta)}^{\tau(\delta)} \bigl|\log(|x|)\bigr| \, d\nu_n^w(x) \le \kappa(\delta),$$

for all large n, on the event Ω_n, where $\kappa(\delta) := 2\delta \cdot \log(1/c_{2.10}) - 2\int_0^{\delta} \log x \, dx$. Note that $\lim_{\delta \to 0} \kappa(\delta) = 0$. Therefore, given any $\kappa > 0$ there exists $\tau_\kappa := \tau(\kappa)$, with the property $\lim_{\kappa \to 0} \tau_k = 0$, such that

$$\begin{aligned}&\limsup_{n\to\infty} \mathbb{P}\left(\int_{-\tau_\kappa}^{\tau_\kappa} \bigl|\log|x|\bigr| \, d\nu_n^w(x) \ge \kappa\right) \\ &\quad \le \limsup_{n\to\infty} \mathbb{P}\left(\left\{\int_{-\tau_\kappa}^{\tau_\kappa} \bigl|\log|x|\bigr| \, d\nu_n^w(x) \ge \kappa\right\} \cap \Omega_n\right) = 0.\end{aligned} \tag{10.7}$$

Next, noting that $\log(\cdot)$ is a bounded function on a compact interval that is bounded away from zero, we apply Theorem 2.11 to deduce that

$$\int_{(-R,-\tau_\kappa)\cup(\tau_\kappa,R)} \bigl|\log|x|\bigr| \, d\nu_n^w(x) \to \int_{(-R,-\tau_\kappa)\cup(\tau_\kappa,R)} \bigl|\log|x|\bigr| \, d\nu_\infty^w(x) \tag{10.8}$$

in probability, for any $R \ge 1$. Recall that for $w \in \mathbb{D}_r$ the support of ν_∞^w is contained in $[-6, 6]$ (see Lemma 10.1(i)). On the other hand, from [9], Theorem 1.7, it follows that there exists a $K_0 < \infty$ such that

$$\mathbb{P}\bigl(\|A_n\| \ge K_0\sqrt{np}\bigr) \le \exp(-cnp),$$

for some constant $c > 0$. Therefore, using the Borel–Cantelli lemma we deduce that

$$\begin{aligned}&\int_{(-R_0,R_0)^c} \bigl|\log|x|\bigr| \, d\nu_n^w(x) \\ &\quad \to \int_{(-R_0,R_0)^c} \bigl|\log|x|\bigr| \, d\nu_\infty^w(x) \qquad \text{almost surely,}\end{aligned} \tag{10.9}$$

for some $R_0 > 0$. From Lemma 10.1(ii), we also have that

$$\int_{-\tau_\kappa}^{\tau_\kappa} |\log|x|| \, d\nu_\infty^w(x) \le C\tau_\kappa |\log \tau_\kappa|, \tag{10.10}$$

for some constant C. As $\kappa > 0$ is arbitrary and $\tau_\kappa \to 0$ as $\kappa \to 0$, combining (10.7)–(10.10) we deduce that

$$\begin{aligned} \frac{1}{n}\log|\det(A_n/\sqrt{np} - wI_n)| &= \int_{-\infty}^{\infty} \log|x| \, d\nu_n^w(x) \\ &\to \int_{-\infty}^{\infty} \log|x| \, d\nu_\infty^w(x), \end{aligned} \tag{10.11}$$

in probability. Now the rest of the proof is completed using Lemma 2.1. Indeed, consider $\mathfrak{G}_n$ the $n \times n$ matrix with i.i.d. centered Gaussian entries with variance one. It is well known that, for Lebesgue almost all w,

$$\frac{1}{n}\log|\det(\mathfrak{G}_n/\sqrt{n} - wI_n)| \to \int_{-\infty}^{\infty} \log|x| \, d\nu_\infty^w(x) \qquad \text{almost surely.} \tag{10.12}$$

For example, one can obtain a proof of (10.12) using [14], Lemma 4.11, Lemma 4.12, [15], Theorem 3.4, and [33], Lemma 3.3.

Thus setting $\mathbb{D} = \mathbb{D}_r$, $B_n^{(1)} = A_n/\sqrt{np}$, and $B_n^{(2)} = \mathfrak{G}_n/\sqrt{n}$ in Lemma 2.1(a), we see that assumption (ii) there is satisfied. The assumption (i) of Lemma 2.1(a) follows from weak laws of large numbers for triangular arrays. Hence, using Lemma 2.1(i) and the circular law for i.i.d. Gaussian matrix of unit variance (e.g., [40], Theorem 1.13), we obtain that for every $r > 0$ and every $f_r \in C_c^2(\mathbb{C})$, supported on $\mathbb{D}_r$,

$$\int f_r(w) \, dL_n(w) \to \frac{1}{\pi}\int f_r(w) \, d\mathfrak{m}(w) \qquad \text{in probability.} \tag{10.13}$$

To complete the proof, it now remains to show that one can extend the convergence of (10.13) to all $f \in C_c^2(\mathbb{C})$. It follows from a standard argument.

Indeed, for any $r > 0$ define $i_r \in C_c^2(\mathbb{C})$ such that i_r is supported on $\mathbb{D}_r$ and $i_r \equiv 1$ on $\mathbb{D}_{2r}$, and $i_r(\mathbb{D}_r \setminus \mathbb{D}_{2r}) \subset [0,1]$. Fixing any $\delta > 0$ choose an $r > 0$ such that $r \le \delta/64\mathfrak{K}$ where $\mathfrak{K} := \|f\|_\infty := \sup_{w\in\mathbb{C}} |f(w)|$. Denote $f_r := f i_r$ and $\bar{f}_r := f - f_r$. Applying (10.13) for the function i_r and the triangle inequality we find that

$$\begin{aligned} &\mathbb{P}\left(\left|\int \bar{f}_r(w) \, dL_n(w)\right| \ge \delta/4\right) \\ &\le \mathbb{P}\left(\left|\int (1 - i_r(w)) \, dL_n(w)\right| \ge \frac{\delta}{4\mathfrak{K}}\right) \\ &\le \mathbb{P}\left(\left|\int i_r(w) \, dL_n(w) - \frac{1}{\pi}\int i_r(w) \, d\mathfrak{m}(w)\right| \ge \frac{\delta}{8\mathfrak{K}}\right) \to 0, \end{aligned} \tag{10.14}$$

as $n \to \infty$, where the last step follows from the fact that

$$\left|\frac{1}{\pi}\int_{B_{\mathbb{C}}(0,1)} \big(1 - i_r(w)\big)\, d\mathfrak{m}(w)\right| \le \frac{1}{\pi}\int_{B_{\mathbb{C}}(0,1)\setminus\mathbb{D}_{2r}} d\mathfrak{m}(w) \le \frac{\delta}{8\mathfrak{K}}, \tag{10.15}$$

by our choice of r. Thus combining (10.13)–(10.15) and the triangle inequality we find that for any $f \in C_c^2(\mathbb{C})$

$$\begin{aligned}
&\mathbb{P}\left(\left|f(w)\,dL_n(w) - \frac{1}{\pi}\int f(w)\,d\mathfrak{m}(w)\right| \ge \delta\right)\\
&\quad\le \mathbb{P}\left(\left|f_r(w)\,dL_n(w) - \frac{1}{\pi}\int f_r(w)\,d\mathfrak{m}(w)\right| \ge \delta/2\right)\\
&\qquad + \mathbb{P}\left(\left|\int \bar{f}_r(w)\,dL_n(w)\right| \ge \delta/4\right) \to 0,
\end{aligned}$$

as $n \to \infty$. This completes the proof of the first part of the theorem. To prove the second part of the theorem, we note that under the assumption (1.2), using Theorem 2.2 it follows

$$\mathbb{P}(\tilde{\Omega}_n') \ge 1 - O\left(\frac{1}{n^2}\right),$$

where

$$\tilde{\Omega}_n' := \left\{s_{\min}\left(\frac{A_n}{\sqrt{np}} - wI_n\right) \ge \tilde{c}_n\right\} \quad \text{and}$$

$$\tilde{c}_n := c_{2.2}\exp\left(-C_{2.2}\frac{\log(1/p_n)}{\log(np_n)}\right)\cdot n^{-3}.$$

Therefore, proceeding similarly as above, applying the Borel–Cantelli lemma, and using Theorem 2.11(ii) we see that the conclusions of (10.7)–(10.8) hold almost surely.

Thus under the assumption (1.2) we have shown that (10.11) holds almost surely. Therefore, proceeding same as above and using Lemma 2.1(ii), we obtain that for every $r > 0$ and every $f \in C_c^2(\mathbb{C})$,

$$\int f_r(w)\,dL_n(w) \to \frac{1}{\pi}\int f_r(w)\,d\mathfrak{m}(w) \qquad \text{almost surely,} \tag{10.16}$$

where we recall that $f_r = f \cdot i_r$. The same proof shows that

$$\limsup_{n\to\infty} \int \big(1 - i_r(w)\big)\,dL_n(w) \le \frac{1}{\pi}\int_{B_{\mathbb{C}}(0,1)} \big(1 - i_r(w)\big)\,d\mathfrak{m}(w) \le 8r,$$

almost surely. Therefore,

$$\limsup_{n\to\infty}\left|\int f(w)\,dL_n(w) - \frac{1}{\pi}\int_{B_{\mathbb{C}}(0,1)} f(w)\,d\mathfrak{m}(w)\right| \le 16\mathfrak{K}r \qquad \text{almost surely,}$$

where we recall $\mathfrak{K} = \|f\|_\infty$. Since r is arbitrary, the result follows. □

Acknowledgments. We thank the anonymous referee for her/his comments that lead to an improvement of the presentation and shortening of the paper. A.B. thanks Amir Dembo and Ofer Zeitouni for suggesting the problem and helpful conversations.

REFERENCES

[1] ADAMCZAK, R. (2011). On the Marchenko–Pastur and circular laws for some classes of random matrices with dependent entries. *Electron. J. Probab.* **16** 1068–1095. MR2820070

[2] ADAMCZAK, R., CHAFAÏ, D. and WOLFF, P. (2016). Circular law for random matrices with exchangeable entries. *Random Structures Algorithms* **48** 454–479. MR3481269

[3] ANDERSON, G. W., GUIONNET, A. and ZEITOUNI, O. (2010). *An Introduction to Random Matrices*. *Cambridge Studies in Advanced Mathematics* **118**. Cambridge Univ. Press, Cambridge. MR2760897

[4] BAI, Z. and SILVERSTEIN, J. W. (2010). *Spectral Analysis of Large Dimensional Random Matrices*, 2nd ed. *Springer Series in Statistics*. Springer, New York. MR2567175

[5] BAI, Z. D. (1997). Circular law. *Ann. Probab.* **25** 494–529. MR1428519

[6] BAI, Z. D. (1999). Methodologies in spectral analysis of large-dimensional random matrices, a review. *Statist. Sinica* **9** 611–677. MR1711663

[7] BASAK, A., COOK, N. and ZEITOUNI, O. (2018). Circular law for the sum of random permutation matrices. *Electron. J. Probab.* **23** Paper No. 33, 51 pp. MR3798243

[8] BASAK, A. and RUDELSON, M. (2017). The circular law for sparse non-Hermitian matrices. Preprint. Available at arXiv:1707.03675v1.

[9] BASAK, A. and RUDELSON, M. (2017). Invertibility of sparse non-Hermitian matrices. *Adv. Math.* **310** 426–483. MR3620692

[10] BASAK, A. and RUDELSON, M. (2019). The local circular law for sparse non-Hermitian matrices. In preparation.

[11] BENAYCH-GEORGES, F. and KNOWLES, A. (2018). Local semicircle law for Wigner matrices. In *Advanced Topics in Random Matrices*. *Panoramas et Synthéses* **53**.

[12] BORDENAVE, C., CAPUTO, P. and CHAFAÏ, D. (2012). Circular law theorem for random Markov matrices. *Probab. Theory Related Fields* **152** 751–779. MR2892961

[13] BORDENAVE, C., CAPUTO, P. and CHAFAÏ, D. (2014). Spectrum of Markov generators on sparse random graphs. *Comm. Pure Appl. Math.* **67** 621–669. MR3168123

[14] BORDENAVE, C. and CHAFAÏ, D. (2012). Around the circular law. *Probab. Surv.* **9** 1–89. MR2908617

[15] BOURGADE, P., YAU, H.-T. and YIN, J. (2014). Local circular law for random matrices. *Probab. Theory Related Fields* **159** 545–595. MR3230002

[16] BRYC, W., DEMBO, A. and JIANG, T. (2006). Spectral measure of large random Hankel, Markov and Toeplitz matrices. *Ann. Probab.* **34** 1–38. MR2206341

[17] CHATTERJEE, S. (2005). A simple invariance theorem. Preprint. Available at arXiv:0508.213.

[18] CHATTERJEE, S. (2006). A generalization of the Lindeberg principle. *Ann. Probab.* **34** 2061–2076. MR2294976

[19] COOK, N. (2017). The circular law for random regular digraphs with random edge weights. *Random Matrices Theory Appl.* **6** 1750012, 23. MR3686490

[20] COOK, N. (2017). The circular law for random regular digraphs. Preprint. Available at arXiv:1703.05839.

[21] EDELMAN, A. (1988). Eigenvalues and condition numbers of random matrices. *SIAM J. Matrix Anal. Appl.* **9** 543–560. MR0964668

[22] ESSEEN, C. G. (1966). On the Kolmogorov–Rogozin inequality for the concentration function. *Z. Wahrsch. Verw. Gebiete* **5** 210–216. MR0205297

[23] GE, S. (2017). Eigenvalue spacing of i.i.d. random matrices. Preprint.

[24] GINIBRE, J. (1965). Statistical ensembles of complex, quaternion, and real matrices. *J. Math. Phys.* **6** 440–449. MR0173726

[25] GIRKO, V. L. (1984). The circular law. *Teor. Veroyatn. Primen.* **29** 669–679. MR0773436

[26] GÖTZE, F. and TIKHOMIROV, A. (2010). The circular law for random matrices. *Ann. Probab.* **38** 1444–1491. MR2663633

[27] GROSS, L. (1975). Logarithmic Sobolev inequalities. *Amer. J. Math.* **97** 1061–1083. MR0420249

[28] GUIONNET, A. and ZEITOUNI, O. (2000). Concentration of the spectral measure for large matrices. *Electron. Commun. Probab.* **5** 119–136. MR1781846

[29] LITVAK, A., LYTOVA, A., TIKHOMIROV, K., TOMCZAK-JAEGERMANN, N. and YOUSSEF, P. (2018). Circular law for sparse random regular digraphs. Preprint. Available at arXiv:1801.05576.

[30] NGUYEN, H. H. (2014). Random doubly stochastic matrices: The circular law. *Ann. Probab.* **42** 1161–1196. MR3189068

[31] NGUYEN, H. H. and VU, V. H. (2013). Circular law for random discrete matrices of given row sum. *J. Comb.* **4** 1–30. MR3064040

[32] PASTUR, L. A. (1972). On the spectrum of random matrices. *Theoret. Math. Phys.* **10** 67–74.

[33] RUDELSON, M. (2008). Invertibility of random matrices: Norm of the inverse. *Ann. of Math.* (2) **168** 575–600. MR2434885

[34] RUDELSON, M. and VERSHYNIN, R. (2008). The Littlewood–Offord problem and invertibility of random matrices. *Adv. Math.* **218** 600–633. MR2407948

[35] RUDELSON, M. and VERSHYNIN, R. (2009). Smallest singular value of a random rectangular matrix. *Comm. Pure Appl. Math.* **62** 1707–1739. MR2569075

[36] RUDELSON, M. and VERSHYNIN, R. (2016). No-gaps delocalization for general random matrices. *Geom. Funct. Anal.* **26** 1716–1776. MR3579707

[37] STROOCK, D. W. (2011). *Probability Theory*: *An Analytic View*, 2nd ed. Cambridge Univ. Press, Cambridge. MR2760872

[38] TAO, T. (2012). *Topics in Random Matrix Theory*. *Graduate Studies in Mathematics* **132**. Amer. Math. Soc., Providence, RI. MR2906465

[39] TAO, T. and VU, V. (2008). Random matrices: The circular law. *Commun. Contemp. Math.* **10** 261–307. MR2409368

[40] TAO, T. and VU, V. (2010). Random matrices: Universality of ESDs and the circular law. *Ann. Probab.* **38** 2023–2065. MR2722794

[41] VERSHYNIN, R. (2014). Invertibility of symmetric random matrices. *Random Structures Algorithms* **44** 135–182. MR3158627

[42] WIGNER, E. P. (1955). Characteristic vectors of bordered matrices with infinite dimensions. *Ann. of Math.* (2) **62** 548–564. MR0077805

[43] WIGNER, E. P. (1958). On the distribution of the roots of certain symmetric matrices. *Ann. of Math.* (2) **67** 325–327. MR0095527

[44] WOOD, P. M. (2012). Universality and the circular law for sparse random matrices. *Ann. Appl. Probab.* **22** 1266–1300. MR2977992

INTERNATIONAL CENTRE FOR
THEORETICAL SCIENCES
TATA INSTITUTE OF FUNDAMENTAL
RESEARCH
SURVEY NO. 151, SHIVAKOTE
HESARAGHATTA HOBLI
BANGALORE 560089
INDIA
AND
DEPARTMENT OF MATHEMATICS
WEIZMANN INSTITUTE OF SCIENCE
234 HERZL STREET, POB 26
REHOVOT 76100
ISRAEL
E-MAIL: anirban.basak@icts.res.in

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF MICHIGAN
EAST HALL, 530 CHURCH STREET
ANN ARBOR, MICHIGAN 48109
USA
E-MAIL: rudelson@umich.edu