THE CIRCULAR LAW FOR SPARSE NON-HERMITIAN MATRICES

ANIRBAN BASAK* AND MARK RUDELSON†

Abstract. 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 $\log(1/p_n) < c(\log np_n)^2$ 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.

1. Introduction

For a $n \times n$ matrix B_n denote by $\lambda_1(B_n), \lambda_2(B_n), \ldots, \lambda_n(B_n)$ its eigenvalues. The empirical spectral distribution (ESD) of B_n, denoted hereafter by 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 non-normal matrix (i.e. $B_n B_n^* \neq B_n^* B_n$) then its eigenvalues are complex valued, resulting in L_{B_n} being supported in 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 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 $f \in C_b(\mathbb{C})$, i.e. $f : \mathbb{C} \mapsto \mathbb{R}$ 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 [39, 40] 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 (for example, Gaussian) converge to the *semicircle distribution*. The conditions on finiteness of moments were relaxed in subsequent works, see [6, 27] and the references therein.

The rigorous study of non-Hermitian matrices, in particular non-normal matrices, emerged much later. The main difficulties were the sensitivity of the eigenvalues of non-normal 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

---

Date: July 19, 2017.

2010 *Mathematics Subject Classification.* 15B52, 60B10, 60B20.

Key words and phrases. Random matrix, sparse matrix, smallest singular value, circular law.

*Research partially supported by grant 147/15 from the Israel Science Foundation.

†Research partially supported by NSF grant DMS 1464514.

matrix to find the asymptotics of the moments of the ESDs. The analogue of this for non-normal matrices is to compute the mixed moments, i.e. 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 non-normal, 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 [5]). Since Stieltjes transform of a probability measure is well defined outside its support, and the ESDs of non-normal matrices are supported on $\mathbb{C}$, their Stieltjes transform fail to capture the spectral measure.

In the 1950's, 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 Ginibre's formula for the joint density function of the eigenvalues Mehta [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 [18]. 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 1980's [22]. However mathematically it contained significant gaps. The first non-Gaussian case (assuming existence of density for the entries) was rigorously treated by Bai [4], and the first result without the density assumption was obtained by Götze and Tikhomirov [23]. After a series of partial results (see [13] and the references therein), the circular law conjecture was established in its full generality in the seminal work of Tao and Vu [37]:

Theorem 1.1 (Circular law for i.i.d. entries [37, 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, 11, 25, 26].

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 [6, 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, 16] 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. $\text{Ber}(p_n)$ 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 [23, 36]). Later in [41] 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 p_n decays faster than a polynomial (in n) rate. 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 poly-logarithmic 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

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

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

We will use the 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 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*

$$\log(1/p_n) < c_{1.3}(\log np_n)^2 \tag{1.2}$$

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 [8] and showed that the desired bound holds when $w\in\mathbb{R}$. The result of [8] 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 that 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 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 [8, 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 the smallest singular 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 2.12). 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 where the dimension of the random subspace is $n-m$ with $m = o(n/\log n)$. Concentration inequalities yield a probability bound $\exp(-cmp)$, for some $c > 0$. To accommodate an union bound we then need $np = \omega(\log^2 n)$. Alternatively, one can try to prove an analogue of [34, Theorem 2.4] for the sparse set-up. But this would again require an improvement of [8, Proposition 3.1].

Remark 1.4 (Sub-Gaussianity assumption I)**.** 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 [8, Remark 1.9] we note that if $\{\xi_{i,j}\}$ are such that

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

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 if p_n satisfies (1.2) then $np_n = \omega((\log n)^{2\beta})$. 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 (Sub-Gaussianity assumption II)**.** Similar to here, in [31] bounds on $s_{\min}$ for dense matrices were derived using the sub-Gaussianity assumption where it was used to find a bound on $\|A_n\|$. In a recent work of Rebrova and Tikhomirov [29] the sub-Gaussianity assumption was removed and it was shown that the bounds of [31] continue to hold only under the finiteness of the second moment assumption (see [29, Theorem B]). One may adapt the techniques of [29] to the sparse set-up to remove the sub-Gaussianity assumption from Theorem 2.2 and hence from Theorem 1.3. We refrain from doing it here in order to keep this already long paper to a manageable length.

Remark 1.6 (Circular law limit for shifted sparse matrices)**.** It is well known that the spectrum of normal matrices is stable under small perturbation (see [3, Lemma 2.1.19] and [5, Lemma 2.2]). However, for a general non-normal matrix its spectrum highly sensitive to small perturbations, for example see [35, Section 2.8.1]. So there are no analogues of [3, Lemma 2.1.19] and [5, Lemma 2.2] for an arbitrary non-normal matrix. Nevertheless, in [41] it was shown that if D_n any $n \times n$ matrix with $\text{rank}(D_n) = o(n)$ and $\text{Tr}(D_n D_n^*) = O(n^2 p_n)$ 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 $\text{Tr}(D_n^2) = o(n^2 p_n)$. 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 1.7. 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, i.e. 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 $\text{Adj}_n := \text{Adj}(\mathsf{G}_n)$ to be its adjacency matrix. That is, for any $i, j \in [n]$,

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

Theorem 1.8. *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 $\text{Adj}_n/\sqrt{np_n(1-p_n)}$ converges weakly to the circular law, as $n \to \infty$, in probability.

(ii) There exists an absolute constant $c_{1.8}$ such that if p_n satisfies the inequality

$$\log(1/\bar{p}_n) < c_{1.8}(\log n\bar{p}_n)^2 \tag{1.4}$$

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

Remark 1.9. Theorem 1.3 and Theorem 1.8 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 [10] and references therein). For dense non-Hermitian random matrices, it was shown in [14] that the local circular law holds. In a forthcoming article [9] 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 2.12, 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 2.13 with the outline of the proof appearing in Section 2.3.

Section 3 - 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 [8]. So the majority of the work is to control infimum over the vectors that are neither compressible nor dominated. Using a result of [31] (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 - 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 non-dominated 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 sub-exponential (in $\sqrt{np_n}$) LCD. A similar result for essentially real vectors is obtained in Section 6. Then we finish the proof of Theorem 2.2 in Section 7.

In Section 8 we prove Theorem 2.12. 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 [37, Lemma A.4] finishes the proof.

Section 9 is devoted to the proof of Theorem 2.13, 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 showing the Stieltjes transform converges to the appropriate limit. To prove the latter convergence, we use the formula for the inverse of a block matrix and Talagrand's concentration inequality to show that the Stieltjes transform satisfies an approximate fixed point equation. From this the result follows using a uniqueness property of the limiting fixed point equation.

In Section 10, combining the results of Section 7 - Section 9, we finish the proof of Theorem 1.3. Then extending Theorem 2.2, Theorem 2.12, and Theorem 2.13, applicable to set-up of Theorem 1.8, we complete the proof of Theorem 1.8 in Section 11.

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 non-normal 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*. First we introduce few definitions. A sequence of random variables $\{X_n\}$ is said to be bounded in probability if

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

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

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

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

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

$$\frac{1}{n}\left\|B_n^{(1)}\right\|_2^2 + \frac{1}{n}\left\|B_n^{(2)}\right\|_2^2 \quad \text{is bounded in probability,} \tag{2.1}$$

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,

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

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

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

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

Lemma 2.1(a) was proved in [7]. Repeating the proof of [7, Lemma 10.1] one can derive Lemma 2.1(b). The replacement principle had already appeared in random matrix literature (see [37, Theorem 2.1]). However, [37, Theorem 2.1] 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 instead of [37, Theorem 2.1].

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 handled by assumption (i) of Lemma 2.1). 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 $\bar{A}_n$ be an $n \times n$ matrix with zero diagonal and i.i.d. off-diagonal 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 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 $\operatorname{Im}(D_n) = r'\sqrt{np_n}I_n$ 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
$$\mathbb{P}\left(s_{\min}(\bar{A}_n + D_n) \leq c_{2.2}\varepsilon \exp\left(-C_{2.2}\frac{\log(1/p_n)}{\log(np_n)}\right)\sqrt{\frac{p_n}{n}}\right) \leq \varepsilon + \frac{C'_{2.2}}{\sqrt{np_n}}.$$

(ii) Additionally, if
$$\log(1/p_n) < \bar{c}_{2.2}(\log np_n)^2,$$
then
$$\mathbb{P}\left(s_{\min}(\bar{A}_n + D_n) \leq c_{2.2}\varepsilon \exp\left(-C_{2.2}\frac{\log(1/p_n)}{\log(np_n)}\right)\sqrt{\frac{p_n}{n}}\right) \leq \varepsilon + \exp(-c'_{2.2}\sqrt{np_n}).$$

Since the diagonal of the matrix Adj_n is zero (recall Definition 1.7), in Theorem 2.2 we have taken the diagonal of $\bar{A}_n + D_n$ to be non-random, so that the set-up of both Theorem 1.3 and Theorem 1.8 fits this general framework. To apply Theorem 2.2 in the proof of Theorem 1.3, we simply condition on the diagonal and later take an average over the diagonal entries. Note that from Theorem 2.2 we obtain a uniform control on $s_{\min}(A_n - \sqrt{np}wI_n)$ for w's satisfying $|\operatorname{Im} w| \in [r,1]$ (hereafter, for brevity, we will often write p instead of p_n).

Similar to [8], without loss of generality, we can assume that $p \le c(K+R)^{-2}$, 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 sparcity and regard entries $A_{i,j}$ as i.i.d. centered subgaussian random variables whose variance is bounded below.

To prove Theorem 2.2 we follow the same scheme as in [8] and borrow some of its results. Recalling the definition of the smallest singular value we have

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

where $S_{\mathbb{C}}^{n-1} := \{z \in \mathbb{C}^n : \|z\|_2 = 1\}$. Thus, to bound $s_{\min}$ we need 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 [8]. However, we now need to treat complex shifts of the matrix $\bar{A}_n$ 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 2.3. Fix $m < n$. The set of m-sparse vectors is given by

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

where $|S|$ denotes the cardinality of a set S and $\text{supp}(\cdot)$ 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 $\text{Comp}(m, \delta)$. Thus,

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

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 $\text{Incomp}(m, \delta)$.

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

Definition 2.4. For any $z \in S_{\mathbb{C}}^{n-1}$, let $\pi_z : [n] \to [n]$ be a permutation which arranges the moduli of the coordinates of z in an non-increasing order. For $1 \le m \le m' \le n$, 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 non-increasing rearrangement of its moduli. For $\alpha < 1$ and $m \le n$ define the set of vectors with dominated tail as follows:

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

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

While studying the behavior of $s_{\min}$ of real shifts of $\bar{A}_n$ in [8], 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. [8, 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 $\bar{A}_n$ from the subspace spanned by the rest of the columns. We use the following result:

Lemma 2.5 (Invertibility via distance [31, Lemma 3.5])**.** *For $j \in [n]$, let $\tilde{A}_{n,j} \in \mathbb{C}^n$ be the j-th 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 $M < n$,*

$$\mathbb{P}\left(\inf_{z \in \text{Incomp}(M,\rho)} \left\| \tilde{A}_n z \right\|_2 \leq \varepsilon \rho^2 \sqrt{\frac{p}{n}}\right) \leq \frac{1}{M} \sum_{j=1}^{n} \mathbb{P}\left(\text{dist}(\tilde{A}_{n,j}, H_{n,j}) \leq \rho\sqrt{p}\varepsilon\right). \tag{2.4}$$

We should mention here that Lemma 2.5 can be extended to the case when the event on the LHS of (2.4) is intersected with an event Ω, and in that case Lemma 2.5 continues to hold if the RHS of (2.4) is replaced by intersecting each of the event under the summation sign with the same event Ω (see also [8, Remark 2.5]). We will actually use this generalized version of Lemma 2.5.

In order to apply Lemma 2.5 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 $(\bar{A}_n + D_n)^{\mathsf{T}}$. 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 the vector normal to the subspace spanned by the last $(n-1)$ columns of $(\bar{A}_n + D_n)$. Thus, applying Lemma 2.5 and the fact that the columns of $\bar{A}_n$ are i.i.d (ignoring the zero diagonal), 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 $(\bar{A}_n + D_n)$.

The small ball probability bounds on $\langle A^D_{n,1}, z\rangle$ depend on the additive structure of the vector z. Following [8], we see that with high probability, we can assume that any $z \in \text{Ker}(B^D)$ is neither compressible nor dominated, where $\text{Ker}(B^D) := \{u \in \mathbb{C}^n : B^D u = 0\}$. Therefore, it is enough to obtain estimates on the small ball probability for incompressible and non-dominated vectors. To this end, we define the following notion of *Lévy concentration function*:

Definition 2.6. Let Z be random variable in $\mathbb{C}^n$. For every $\varepsilon > 0$, the Lévy concentration function of Z is defined as

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

The Berry-Esséen bound of [28, 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 [20] (see also [38, 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 [8, 31, 32]. One would hope to carry out the same program here. However, when we view the incompressible and non-dominated vectors of small LCD as a subset of $S^{n-1}_{\mathbb{C}}$, its real dimension is twice as large as in the proof [8, Proposition 4.1]. On the other hand, for the real-valued random variables in $\bar{A}_n$, one does not expect to obtain better control on the Lévy concentration function. Thus the proof of [8, Proposition 4.1] breaks down as the bounds on the Lévy concentration function and the size of the net do not match (see also [8, 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 correlation* as follows:

Definition 2.7. Let $z \in \mathbb{C}^m$ for some positive integer m. Denote $V := V(z) := \begin{pmatrix} x^\mathsf{T} \\ y^\mathsf{T} \end{pmatrix}$, where $z = x + \mathrm{i}y$. Then we denote the real-imaginary correlation of z by

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

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 correlations are termed *essentially real* vectors (See (5.1) and (6.1) for a precise formulation). The real and imaginary parts of essentially real vectors are almost linearly dependent. This fact 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 [8] 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 [33] we define a notion of a *two-dimensional* LCD and using [33, Theorem 7.5] obtain better bounds on the Lévy concentration function.

Definition 2.8. For $y > 0$, denote $\log_1(y) := \log y \cdot \mathbb{I}(y \geq e)$. Fixing $L \geq 1$, for a non-zero vector $x \in \mathbb{R}^m$, we set

$$D_1(x) := \inf \left\{ \theta > 0 : \ \operatorname{dist}(\theta x, \mathbb{Z}^m) < 2^5 L \sqrt{\log_1 \frac{\|\theta x\|_2}{2^6 L}} \right\}. \tag{2.5}$$

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

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

We will call the first version of the LCD one-dimensional, and the second one two-dimensional. This explains the subscripts 1 and 2 in the notation above. Also note that $D_1(\cdot)$ matches with the definition of the LCD used in [8] up to constants.

Note 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 2.9. Consider a complex vector $z = x + \mathrm{i}y \in \mathbb{C}^m$. Denote $\tilde{z} := \tilde{z}(z) := \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^{2m}$, and define a $2 \times m$ matrix $V := V(z) := \begin{pmatrix} x^\mathsf{T} \\ y^\mathsf{T} \end{pmatrix}$. Using these two different representations of $z \in \mathbb{C}^m$, we now define:

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

From Definition 2.8 we see that $D_2(V)$ is defined in terms of two different norms: $\|\theta\|_2$ and $\left\|V^\mathsf{T}\theta\right\|_2$. To take advantage of both norms, we introduce the following auxiliary quantity.

Definition 2.10. For a real-valued $2 \times m$ matrix V, define

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

As before, for a $z \in \mathbb{C}^m$, we define $\Delta(z) := \Delta(V)$ where $V = V(z)$. It is easy to see that

$$d(z) D_2(z) \leq \Delta(z) \leq D_2(z). \tag{2.6}$$

Remark 2.11. Similar to [8] we will take $L = (\delta_0 p)^{-1/2}$, where $\delta_0 \in (0,1)$ is a constant such that $\mathcal{L}(\xi\delta, \varepsilon) \le 1 - \delta_0 p$, for all $\varepsilon < \varepsilon_0'$, ξ is a random variable with unit variance and finite fourth moment, $\delta \sim \mathrm{Ber}(p)$ independent of ξ, and $\varepsilon_0' \in (0,1)$. See [8, Remark 2.7] for more details.

Equipped with the definitions of LCDs we can now finish the outline of the proof of Theorem 2.2. Using a recent result of [33] (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 the two-dimensional LCD (see the bound in (5.8)). This probability bound nicely balances with the cardinality of the net of genuinely complex vectors for which $\Delta(z_{\text{small}})/L$ is large, where z_{small} is the part of z containing the "smallest" coordinates in the absolute value. 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 [8], we show that there cannot exist a vector $z \in \mathrm{Ker}(B^D)$ with a dominated real part, with high probability. This additional observation then shows that for any $z \in \mathrm{Ker}(B^D)$ the quantity $\Delta(z_{\text{small}})/L$ must also be large. This finishes 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 i-th largest singular value.

Theorem 2.12. *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. with zero mean and unit variance, and $\{\delta_{i,j}\}_{i,j=1}^n$ are i.i.d.* $\mathrm{Ber}(p)$ *random variables. There exist constants $c_{2.12}$ and $C_{2.12}$*[1] *such that the following holds: Let $\psi : \mathbb{N} \mapsto \mathbb{N}$ be such that $\psi(n) < n$ and $\min\{p\psi(n), \psi^2(n)/n\} \ge C_{2.12} \log n$. Then for any $w \in B_{\mathbb{C}}(0,1)$ we have*

$$\mathbb{P}\left(\bigcup_{i=3\psi(n)}^{n-1} \left\{ s_{n-i}\left(\frac{A_n}{\sqrt{np_n}} - wI_n \right) \le c_{2.12}\frac{i}{n} \right\} \right) \le \frac{2}{n^2}.$$

To prove Theorem 2.12 we follow the approach of [12, 37, 41]. 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 [37] finishes 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.

Theorem 2.13. *(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 $\mathbb{P}(\delta_{i,j} = 1) = p$, and $\xi_{i,j}$ are centered i.i.d. random variables with unit variance. Assume $p = \omega(\frac{\log n}{n})$. 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 $\sum_{n=1}^\infty (n^2p)^{-1} < \infty$ then the above convergence holds almost surely.

[1] the constants $c_{2.12}$ and $C_{2.12}$ can potentially depend on the tail of the distribution of $\{\xi_i\}_{i=1}^n$.

To prove Theorem 2.13 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 2.13 to establish the almost sure convergence. Now turning to the case of bounded $\{\xi_{i,j}\}_{i,j=1}^n$, we see that $\boldsymbol{A}_n^w$ is a Hermitian matrix. Therefore, it suffices to show the convergence of Stieltjes transform of the ESD of $\boldsymbol{A}_n^w$ (for example, see [3, Theorem 2.4.4]). To this end, recall the definition of the Stieltjes transform of a probability measure.

Definition 2.14. 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 will write m_n to denote the Stieltjes transform of ν_n^w. Therefore we see that $m_n(\zeta) := m_n(\zeta, w) := \frac{1}{2n} \operatorname{Tr}\left(\boldsymbol{A}_n^w - \zeta I_{2n}\right)^{-1}$. We then need to identify the limit of $m_n(\cdot)$. When B_n is a matrix consisting of i.i.d. entries of zero mean and unit variance, with certain moment assumptions, the Stieltjes transform of the ESD of

$$\begin{bmatrix} 0 & \frac{1}{\sqrt{n}} B_n - wI_n \\ \frac{1}{\sqrt{n}} B_n^* - \bar{w} I_n & 0 \end{bmatrix}$$

is known to converge to $m_\infty(\cdot)$, where $m_\infty(\cdot)$ is the unique root of the equation

$$P(m) := m(m+\zeta)^2 + m(1-|w|^2) + \zeta = 0, \tag{2.8}$$

with positive imaginary part on $\mathbb{C}^+ := \{\xi \in \mathbb{C} : \operatorname{Im} \xi > 0\}$ (see [4, 23]). It is also known that the ESD of $B_n/\sqrt{n}$ converges to the circular law. Since the ESD of $A_n/\sqrt{np}$ also admits the same limit, the Stieltjes transforms $\{m_n(\zeta)\}_{n\in\mathbb{N}}$ should have the same limit $m_\infty(\zeta)$. This is shown in the theorem below.

Theorem 2.15. *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 bounded i.i.d. random variables with zero mean and unit variance. Fix any $C \geq 1$ and denote*

$$\mathscr{T}_C := \{\zeta \in \mathbb{C}^+ \cap B_{\mathbb{C}}(0, 4C) : \operatorname{Im} \zeta \geq C\}.$$

If $np = \omega(\log n)$, then there exists an absolute constant $C_{2.15}$ such that for any $\zeta \in \mathscr{T}_{C_{2.15}}$ we have $m_n(\zeta) \to m_\infty(\zeta)$ almost surely.

To prove Theorem 2.15 we use the formula for the inverse of a block matrix to first derive equations involving entries of the inverse of $(\boldsymbol{A}_n^w - \zeta I_{2n})$. Then using Talagrand's concentration inequality we identify the dominant and negligible terms from those equations, which allows us to deduce that $P(m_n(\zeta)) = o(1)$ for all large n. Finally using a uniqueness property of the roots of the equation $P(m) = 0$ (see Lemma 9.10) we deduce that $|m_n(\zeta) - m_\infty(\zeta)| = o(1)$ completing the proof of Theorem 2.15.

Finally to deduce the weak convergence of ν_n^w to ν_∞^w from Theorem 2.15 we need the following proposition.

Proposition 2.16. *Let $\{\mu_n\}$ be a collection of probability measures on $\mathbb{R}$ (possibly random) and $\widehat{m}_n(\cdot)$ be the Stieltjes transform of μ_n. Suppose $\widehat{m}_n(\zeta) \to \widehat{m}(\zeta)$ in probability (almost surely), as $n \to \infty$, for all $\zeta \in \mathscr{T}_C$ and some $C > 0$, where $\widehat{m}(\cdot)$ is the Stieltjes transform of a (non-random) probability measure μ on $\mathbb{R}$. Then μ_n converges weakly to μ in probability (almost surely).*

The proof of Proposition 2.16 is standard. Indeed, from the proof of [3, Theorem 2.4.4(c)] it follows that it is enough to show the convergence of $m_n(\zeta)$ to $m_\infty(\zeta)$ for all $\zeta \in \mathscr{C}$, where $\mathscr{C} \subset \mathbb{C}\backslash\mathbb{R}$ is such that it has an accumulation point $c \in \mathbb{C}\backslash\mathbb{R}$.

3. The structure of the kernel: Vectors with non-dominated real part

Recall from Section 2.1 that to prove Theorem 2.2 the main challenge is to show that there does not exist a genuinely complex vector $z \in \text{Ker}(B^D)$ with a small two-dimensional LCD. As a first step we show that for any $z \in \text{Ker}(B^D)$, its real part must have a non-dominated 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 the following notation: For some $M < n/2$, to be determined during the course of the proof, we denote by small(z) the set of the $(n-M)$ coordinates of z having the smallest absolute values. The ties are broken arbitrarily. We also write $z_{\text{small}} = x_{\text{small}} + \mathrm{i}y_{\text{small}} := z_{\text{small}(z)}$.

Proposition 3.1. *Let $\bar{A}_n$ be a matrix with zero diagonal and i.i.d. off-diagonals $a_{i,j} = \xi_{ij}\delta_{ij}$, where $\{\xi_{ij}\}$ are i.i.d. centered 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 $\text{Im}\,(D_n) = r'\sqrt{np}I_n$ 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: Let $1 \leq \mu \leq \rho^{-1}$, and set $M = C_{3.1}\mu^2\rho^{-2}p^{-1}$. If

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

then, for $\rho := (\widetilde{C}_{3.1}(K+R))^{-\ell_0-6}$, we have

$$\mathbb{P}\left(\exists z \in \text{Ker}(B^D) \cap S_{\mathbb{C}}^{n-1} : \left\| \frac{\text{Re}(z_{\text{small}})}{\|\text{Re}(z_{\text{small}})\|_2} \right\|_\infty \geq \mu^{-1}p^{1/2} \text{ and } \|\bar{A}_n\| \leq K\sqrt{np}\right) \leq \exp(-\bar{c}_{3.1}np). \tag{3.2}$$

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

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

Before going to the proof of Lemma 3.3 let us introduce a notation. Write Re$\,(B^D)$ to denote the real part of the matrix B^D. Recalling the definition of B^D from Section 2.1 we see that its rows are the last $(n-1)$ columns of $\bar{A}_n + D_n$. Since $\bar{A}_n$ is a real valued matrix we see that the rows of $\mathrm{Re}\,(B^D)$ are the last $(n-1)$ columns of $\bar{A}_n + \mathrm{Re}\,(D_n)$. This, in particular, implies that $\mathrm{Im}\,(B^D)$, the imaginary part B^D is non-random.

To prove Lemma 3.3 we borrow results from [8]. In [8, Proposition 3.1] we showed that, with high probability, there does not exist any real-valued compressible or dominated vector z such that $\left\|\mathrm{Re}\,(B^D)z\right\|_2$ is small. In [8, Remark 3.10] it was also argued that the same conclusion holds for $\left\|B^D z\right\|_2$ when z is now allowed to be complex valued. We will need both these results to prove Lemma 3.3. For completeness we state it below.

Proposition 3.4 ([8, Proposition 3.1, Remark 3.10]). *Let $\bar{A}_n$ be as in Proposition 3.1. Fix $K, R \ge 1$, and assume that $D_n^{(1)}$ and $D_n^{(2)}$ are two non-random diagonal matrices with real and complex entries, respectively, such that $\|D_n^{(1)}\|, \|D_n^{(2)}\| \le R\sqrt{pn}$. Then there exist constants $0 < c_{3.4}, \bar{c}_{3.4}, c'_{3.4}, C_{3.4}, \widetilde{C}_{3.4}, \overline{C}_{3.4} < \infty$, depending only on K, R, and the fourth moment of $\{\xi_{ij}\}$, such that for*

$$\frac{\overline{C}_{3.4}\log n}{n} \le p \le \frac{1}{10}, \tag{3.3}$$

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

$$\mathbb{P}\Big(\exists x \in (\mathrm{Dom}(M, (C_{3.4}(K+R))^{-4}) \cup \mathrm{Comp}(M,\rho)) \cap S_{\mathbb{R}}^{n-1} \\ \Big\|(\bar{A}_n + D_n^{(1)})x\Big\|_2 \le c'_{3.4}(K+R)\rho\sqrt{np} \text{ and } \left\|\bar{A}_n\right\| \le K\sqrt{pn}\Big) \le \exp(-\bar{c}_{3.4}pn)$$

and

$$\mathbb{P}\Big(\exists z \in \mathrm{Dom}(M, (C_{3.4}(K+R))^{-4}) \cup \mathrm{Comp}(M,\rho) \\ \Big\|(\bar{A}_n + D_n^{(2)})z\Big\|_2 \le c'_{3.4}(K+R)\rho\sqrt{np} \text{ and } \left\|\bar{A}_n\right\| \le K\sqrt{pn}\Big) \le \exp(-\bar{c}_{3.4}pn),$$

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

Observe that Proposition 3.4 is stated for the square matrix $\bar{A}_n$. To prove Lemma 3.3 we need a version of Proposition 3.4 for $(n-1)\times n$ matrices. As noted in [8, 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 non-dominated vectors. Such an estimate was derived in [8, Corollary 3.7] for real valued vectors. One can investigate its proof to convince oneself that the same proof works for complex valued vectors. We state this modified version below.

Lemma 3.5 ([8, Corollary 3.7]). *Let $\bar{A}_n$ be as in Proposition 3.1. For every $z \in \mathbb{C}^n$ and $i \in [n]$, define $z_{(i)}$ to be the vector obtained from z by setting its i-th coordinate to be zero. 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 $\sup_{i\in[n]}\left(\left\|z_{(i)}\right\|_\infty / \left\|z_{(i)}\right\|_2\right) \le \alpha\sqrt{p}$, we have*

$$\mathcal{L}\left(\bar{A}_n z, \beta\cdot\sqrt{pn}\inf_{i\in[n]}\left\|z_{(i)}\right\|_2\right) \le \exp(-\gamma n).$$

We now proceed to the proof of Lemma 3.3.

Proof of Lemma 3.3. The proof is based on ideas from [21]. For ease of writing let us write $c_0 := (C_{3.4}(K+R))^{-4}$. We also denote

$$\Omega_{D,C} := \{\mathrm{Re}\,(B^D) : \exists z \in \mathrm{Dom}(M, c_0) \cup \mathrm{Comp}(M,\rho) \quad \left\|B^D z\right\|_2 \le c'_{3.4}(K+R)\rho\sqrt{np} \text{ and } \left\|\bar{A}_n\right\| \le K\sqrt{np}\},$$

$$\Omega_C := \{\mathrm{Re}\,(B^D) : \exists u \in \mathrm{Comp}(M,\rho) \cap S_{\mathbb{R}}^{n-1} \quad \left\|\mathrm{Re}\,(B^D) u\right\|_2 \le c'_{3.4}(K+R)\rho\sqrt{np} \text{ and } \left\|\bar{A}_n\right\| \le K\sqrt{np}\}.$$

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

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

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

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

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

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

The conclusion of the lemma immediately follows from the claim by taking an 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 finish the proof of the claim. Below we carry out the details.

To this end, fix any $J \subset [n]$ and let $\mathrm{Re}\,(B^D)|_J$ denote the sub-matrix 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 $\left\|B^D z\right\|_2 \le c\rho\sqrt{np}$. Then we see that

$$\left\|\mathrm{Re}\,(B^D)x - \mathrm{Im}\,(B^D)y\right\|_2 \le \left\|B^D z\right\|_2 \le c\rho\sqrt{np}. \tag{3.4}$$

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

$$\left\|\mathrm{Re}\,(B^D)|_{J^c}\right\| \le \left\|\mathrm{Re}\,(B^D)\right\| \le \left\|B^D\right\| \le \left\|\bar{A}_n\right\| + \|D_n\| \le (K+R)\sqrt{np},$$

applying the triangle inequality we further deduce that

$$\begin{aligned}\left\|\mathrm{Im}\,(B^D)y - \mathrm{Re}\,(B^D)|_J x_{[1:M]}\right\|_2 &\le c\rho\sqrt{np} + \left\|\mathrm{Re}\,(B^D)|_{J^c} x_{[M+1:n]}\right\|_2 \\ &\le c\rho\sqrt{np} + \left\|\mathrm{Re}\,(B^D)|_{J^c}\right\| \cdot \|x_{\text{small}}\|_2 \le 2c\rho\sqrt{np},\end{aligned} \tag{3.5}$$

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

Recall that $\mathrm{Im}\,(B^D)$ 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

$$\left\| y|_{[2:n]} - \frac{1}{r'\sqrt{np}} \mathrm{Re}\,(B^D)|_J x_{[1:M]} \right\|_2 \le 2r'^{-1} c\rho \le 2r^{-1} c\rho. \tag{3.6}$$

Thus (3.6) implies that the vector $y|_{[2:n]}$ belongs to a $(2r^{-1}c\rho)$-neighborhood of the linear subspace $\mathscr{E}'_J := \mathrm{span}(\mathrm{Re}\,(B^D)|_J \mathbb{R}^J) \subset \mathbb{R}^{n-1}$. 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 space

$$\mathscr{E}_J := \{x + \mathrm{i}y : \ \mathrm{supp}(x) \subset J, \ y|_{[2:n]} \in \mathscr{E}'_J, y_1 \in [-1,1]\}, \quad \text{with} \quad \dim(\mathscr{E}_J) \le 2M+1. \tag{3.7}$$

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 $(2B^n_{\mathbb{C}}) \cap \mathcal{E}_J$. Therefore we can choose a $(r^{-1}c\rho)$-net $\mathcal{N} \subset (2B^n_{\mathbb{C}}) \cap \mathscr{E}_J$ of cardinality

$$|\mathcal{N}| \le \left(\frac{12}{c\rho}\right)^{2M+1} \le \exp(3M\log(12/(c\rho))). \tag{3.8}$$

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 $\left\|B^D w\right\|_2 \le 5r^{-1}(K+R)c\rho\sqrt{np}$. Now shrink c such that $10r^{-1}c \le c'_{3.4}$. With this choice of the constant c we see that we $w \notin \mathrm{Dom}(M, c_0) \cup \mathrm{Comp}(M,\rho)$ on the event $\Omega^c_{D,C}$.

However, for any w we have $\|w_{[M+1:n]}\|_\infty \le 1/\sqrt{M}$, and moreover for any $w \notin \mathrm{Comp}(M,\rho)$, one has $\|w_{[M+1:n]}\|_2 \ge \rho$. Therefore,

$$\left\|w_{[M+1:n]\setminus\{i\}}\right\|_2 \ge \left\|w_{[M+1:n]}\right\|_2 - 1/\sqrt{M} \ge \frac{1}{2}\left\|w_{[M+1:n]}\right\|_2. \tag{3.9}$$

Here the last inequality follows from the definition of ρ and the assumption $p \le \widetilde{c}(K+R)^{-2}$, for sufficiently small $\widetilde{c}$, which we made at the beginning of Section 2. Therefore using the fact that $w \notin \mathrm{Dom}(M, c_0)$, we have

$$\sup_{i\in[n]} \frac{\left\|w_{[M+1:n]\setminus\{i\}}\right\|_\infty}{\left\|w_{[M+1:n]\setminus\{i\}}\right\|_2} \le 2(c_0\sqrt{M})^{-1} \le 2c_0^{-1}\sqrt{p},$$

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 $\bar{c}$ such that

$$\begin{aligned} &\mathbb{P}\left(\left\|B^D w\right\|_2 \le c_\star\rho\sqrt{np} \,\Big|\, \mathrm{Re}\,(B^D)|_J\right) \\ \le\, &\mathcal{L}\left(\mathrm{Re}\,(B^D)|_{J^c} w_{[M+1:n]}, 2c_\star \inf_{i\in[n]} \left\|w_{[M+1:n]\setminus\{i\}}\right\|_2 \sqrt{np}\right) \le \exp(-2\bar{c}n), \end{aligned}$$

Hence, by the union bound,

$$\mathbb{P}\left(\exists w \in \mathcal{N} : \ \left\|B^D w\right\|_2 \le c_\star\rho\sqrt{np} \,\Big|\, \mathrm{Re}\,(B^D)|_J\right) \le |\mathcal{N}| \cdot \exp(-2\bar{c}n) \le \exp(-\bar{c}n),$$

where the last step follows from the bound (3.8) 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

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

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

We are now ready to prove Proposition 3.1.

Proof of Proposition 3.1. 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, $\|x_{\text{small}}\|_2 > c''_{3.3}\rho$. Denote $z_{\text{small}} = \text{Re}(z_{\text{small}}) + \mathrm{i}\,\text{Im}(z_{\text{small}}) =: u + \mathrm{i}v$. Recall that $x = \text{Re}\,(z)$. Therefore

$$\|u\|_2 \geq \left\|x_{[M+1:n]}\right\|_2 \geq c''_{3.3}\rho,$$

with high probability. On the other hand,

$$\|u\|_\infty \leq \|z_{\text{small}}\|_\infty \leq \frac{1}{\sqrt{M}}.$$

Combining last two inequalities, we obtain

$$\frac{\|u\|_\infty}{\|u\|_2} \leq \frac{1}{c''_{3.3}\rho\sqrt{M}}, \tag{3.10}$$

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

Remark 3.6. Note that the inequality (3.10) 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 admit a net of small cardinality. Let us start with a simple reduction. Fix $M < n/2$, $z \in S_{\mathbb{C}}^{n-1}$, and let $J = \text{small}(z)$. Denote $V_J := V(z_J)$, where $z_J \in \mathbb{C}^J$ (for ease of writing we write $\mathbb{C}^J$ instead of $\mathbb{C}^{|J|}$) is the vector obtained from z by keeping the indices corresponding to J and recall the definition of $V(\cdot)$ from Definition 2.9. Now consider the singular value decomposition of the matrix V_J:

$$V_J = U_1 S U_2 = U_1 \begin{pmatrix} w_1^{\mathsf{T}} \\ w_2^{\mathsf{T}} \end{pmatrix},$$

where U_1 is a 2×2 orthogonal matrix, S is a 2×2 diagonal matrix of singular values, and U_2 is a $2\times|J|$ isometry matrix. The vectors $w_1,\ w_2 \in \mathbb{R}^J$ are scaled copies of the right singular vectors of the matrix V_J. This means that $w_1 \perp w_2$, and without loss of generality, we can assume that $\|w_2\|_2 \leq \|w_1\|_2$. Translating this back to the complex notation, we find a $\tau \in [0, 2\pi]$ such that $z_J = e^{\mathrm{i}\tau}(w_1 + \mathrm{i}w_2)$. As $z_J \in \text{Ker}(B)$ if and only if $e^{-\mathrm{i}\tau}z_J \in \text{Ker}(B)$, without loss of generality, we can only consider the following set

$$\begin{aligned} \mathcal{Z} :=\{ z \in S_{\mathbb{C}}^{n-1}\backslash(\text{Dom}(M, (C_{3.4}(K+R))^{-4}) \cup \text{Comp}(M,\rho)) : z_{\text{small}} = w_1 + \mathrm{i}w_2, \\ w_1 \perp w_2,\ \|w_1\|_2 \geq \|w_2\|_2\} \end{aligned} \tag{4.1}$$

instead of $S_{\mathbb{C}}^{n-1}\backslash(\text{Dom}(M, (C_{3.4}(K+R))^{-4}) \cup \text{Comp}(M,\rho))$. Therefore our revised goal is to show that the set of genuinely complex vectors, when viewed as a subset of $\mathcal{Z}$, 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 correlation $d(\phi,\psi) = \|\phi\|_2 \|\psi\|_2$ 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 $\mathbb{Z}^J_\gamma := \mathbb{Z}^J \cap \gamma B_2^{|J|}$. Using a simple volumetric comparison argument we have following estimate on $|\mathbb{Z}^J_\gamma|$:

$$|\mathbb{Z}^J_\gamma| \le \left(C_0 \left(\frac{\gamma}{\sqrt{|J|}} + 1 \right) \right)^{|J|}, \tag{4.2}$$

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

Lemma 4.1. *Let $d \in (0,1)$, and $0 < \alpha \le dD \le \Delta \le D$. Define the set*

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

Then, there exists a $\left(\frac{C_{4.2}\alpha}{D}\right)$-net $\mathcal{M}_J(D,\Delta,d) \subset S_J(D,\Delta,d)$ such that

$$|\mathcal{M}_J(D,\Delta,d)| \le \left(\bar{C}_{4.2} \frac{dD^2}{\alpha} \cdot \left(\frac{1}{\sqrt{|J|}} + \frac{1}{\Delta} \right) \right)^{|J|} \cdot \left(\frac{D}{\alpha} \right)^2,$$

for some absolute constants $C_{4.2}$, and $\bar{C}_{4.2}$.

This lemma provides a significant improvement over the standard volumetric estimate yielding $\left(cD^2/\alpha^2\right)^{|J|}$. As we will see below, 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 correlation $d(\phi,\psi)$ is approximately constant in the set $S_J(D,\Delta,d)$, 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. We also note that Lemma 4.1 deals with the case when the correlation between ϕ and ψ is relatively large, which is represented by the assumption $d \ge \alpha/D$. Under this assumption, the angle between the real and the imaginary part of the vectors are non-negligible. In Section 5 we will use this criteria to formally define the notion of genuinely complex vectors.

Proof of Lemma 4.1. Assume that there exists $\zeta := (\zeta_1, \zeta_2) \in \mathbb{R}^2$ and $q \in \mathbb{Z}^J$ satisfying

$$\|\zeta_1\phi + \zeta_2\psi\|_2 \in [\Delta, 2\Delta] \quad \text{and} \quad \|\zeta_1\phi + \zeta_2\psi - q\|_2 < \alpha. \tag{4.3}$$

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

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

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

$$\Delta \le \|\zeta_1\phi + \zeta_2\psi\|_2 \le \frac{1}{2}dD\,\|\phi\|_2 + 2D\,\|\psi\|_2 \le 7dD. \tag{4.4}$$

We will approximate ϕ using the standard volumetric net and use (4.3) to construct a small net for ψ. To this end, consider $(\phi, \psi) \in S_J^0(D, \Delta, d)$ and let $(\zeta_1, \zeta_2) \in \mathbb{R}^2$ be the corresponding vector (i.e. for which (4.3) holds). Then, by the triangle inequality,

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

i.e. $q \in \mathbb{Z}^J_{3\Delta}$. Denote by $\mathcal{N}_\phi$ an (α/D)-net in B_2^J with

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

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

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

as $|\zeta_1| \le \frac{1}{2}dD \le D$. Therefore

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

We observe that

$$\left|\frac{\zeta_1}{\zeta_2}\right| \bigvee \left|\frac{D/2}{\zeta_2}\right| \le 1, \quad \left\|\phi'\right\|_2 \le 1, \quad \text{and} \quad \left\|\frac{q}{D/2}\right\|_2 \le 6\frac{\Delta}{D} \le 6, \tag{4.5}$$

where the last inequality follows from our assumption $\Delta \le D$. Next let $\mathcal{N}_\square$ be an (α/D)-net in the unit square in $\mathbb{R}^2$ with $|\mathcal{N}_\square| \le (6D/\alpha)^2$. Using (4.5), and applying the triangle inequality, we now see that there exists $(x_1, x_2) \in \mathcal{N}_\square$ such that

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

Hence,

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

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

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

where (4.2) has been used to bound $|\mathbb{Z}^J_{3\Delta}|$ and (4.4) has been used to replace Δ by dD in the last inequality.

Turning to prove the case of $|\zeta_1| > \frac{1}{2}dD$ we denote $S_J^1(D, \Delta, d) := S_J(D, \Delta, d) \backslash S_J^0(D, \Delta, d)$. That is,

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

Now let us construct a net in $S_J^1(D, \Delta, d)$. Our strategy here is opposite to what we used in the previous case. Namely, we use the volumetric approximation for ψ and then use (4.3) to approximate

ϕ. To this end, consider any $(\phi, \psi) \in S_J^1(D, \Delta, d)$ 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 $|\zeta_1| \geq \frac{1}{2}dD$ and $|\zeta_2| \leq 2D$ we also see that $24\|\zeta_1\phi\|_2 \geq 6dD \geq \|\zeta_2\psi\|_2$. Therefore

$$\Delta \leq \|\zeta_1\phi\|_2 + \|\zeta_1\psi\|_2 \leq 25\|\zeta_1\phi\|_2 \leq 25|\zeta_1|. \tag{4.6}$$

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

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

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

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

and so

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

where we have used (4.6) in the last step. Note that

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

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

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

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

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

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

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

This implies that the set

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

is a $(2\alpha/D)$-net in $S_J(D, \Delta, d)$. We observe that

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

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

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

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

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

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

Remark 4.2. Note that Lemma 4.1 holds also for any subset of $S_J(D,\Delta,d)$. That is, given any $\mathcal{S} \subset S_J(D,\Delta,d)$ there exists a net $\mathcal{M}_J^{\mathcal{S}}(D,\Delta,d) \subset \mathcal{S}$ with the same properties as in Lemma 4.1. We use this version of Lemma 4.1 to prove Proposition 4.3.

Building on Lemma 4.1 we now obtain a net with small cardinality for the collection of vectors z for which $D_2(z_{\text{small}}/\|z_{\text{small}}\|_2) \approx D, \Delta(z_{\text{small}}/\|z_{\text{small}}\|_2) \approx \Delta$, 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:

$$\begin{aligned}(4.7)\quad \mathcal{Z}(D,\Delta,d) := \{z \in \mathcal{Z} :\ & D_2(z_{\text{small}}/\left\|z_{\text{small}}\right\|_2) \in [D,(3/2)D], \\ & \Delta(z_{\text{small}}/\left\|z_{\text{small}}\right\|_2) \in [\Delta,(3/2)\Delta], d(z_{\text{small}}/\left\|z_{\text{small}}\right\|_2) \in [d,(3/2)d]\}.\end{aligned}$$

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 correlation. So we partition this set according to $D_2(\cdot)$, $\Delta(\cdot)$, and $d(\cdot)$. The net $\mathcal{M}_J(D,\Delta,d)$ provides a net for the vectors which have $D_2(\cdot) \approx D, \Delta(\cdot) \approx \Delta$, and $d(\cdot) \approx d$. This is shown in the proposition below.

Proposition 4.3. *Let $d \in (0,1)$, $D, \Delta > 1$, and denote*

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

Assume that $\alpha \le dD \le \Delta \le D$. Then there exist absolute constants $C_{4.3}$, $\bar{C}_{4.3}$, and a set $\mathcal{N}(D,\Delta,d) \subset \mathcal{Z}(D,\Delta,d)$ with

$$|\mathcal{N}(D,\Delta,d)| \le \bar{C}_{4.3}^n \left(\frac{n}{\rho M}\cdot\frac{D}{\alpha}\right)^{5M} \cdot \left(\frac{dD^2}{\alpha}\cdot\left(\frac{1}{\sqrt{n}}+\frac{1}{\Delta}\right)\right)^{n-M}$$

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

$$\left\|\frac{z_J}{\|z_J\|_2} - \frac{w_J}{\|w_J\|_2}\right\|_2 < C_{4.3}\frac{\alpha}{D},\ \|z_{J^c} - w_{J^c}\|_2 \le C_{4.3}\frac{\rho\alpha}{D}, |\|z_J\|_2 - \|w_J\|_2| \le C_{4.3}\frac{\rho\alpha}{D}.$$

To prove Proposition 4.3 our strategy will be to use the net $\mathcal{M}_J^{\mathcal{S}}(D,\Delta,d)$, for some suitable choice of $\mathcal{S}$, obtained from Lemma 4.1, to approximate the small coordinates. The cardinality of the net to approximate the large ones will be obtained by the simple volumetric estimate.

Proof of Proposition 4.3. Fix a set $J \subset [n]$, $|J| = n - M$, and denote

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

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

$$\|\zeta\|_2 \le (4/3)D_2(\phi + i\psi) \le 2D, \qquad \|\zeta_1\phi + \zeta_2\psi\|_2 \le (4/3)\Delta(\phi + i\psi) \le 2\Delta,$$

$$\text{and } \operatorname{dist}(\zeta_1\phi + \zeta_2\psi, \mathbb{Z}^n) < L\sqrt{\log_1 \frac{\|\zeta_1\phi + \zeta_2\psi\|_2}{2^8 L}} \le L\sqrt{\log_1 \frac{\Delta}{2^7 L}} = \alpha.$$

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 $(\phi, \psi) \in S_J(D, \Delta, d)$, and in particular $(\phi, \psi) \in \mathcal{S}$ where $\mathcal{S} := \{(\phi(z), \psi(z)) : z \in \mathcal{Z}(D, \Delta, d)\}$. So it can be approximated by an element of $\mathcal{M}_J^{\mathcal{S}}(D, \Delta, d)$. To this end, set

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

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

$$\left\| \frac{z_J}{\|z_J\|_2} - w' \right\|_2 < C_{4.2} \frac{\alpha}{D}.$$

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

$$|\mathcal{N}_{J^c}| \le \left(\frac{3}{C_{4.2}\rho} \cdot \frac{D}{\alpha} \right)^{2M}.$$

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

$$\|z_{J^c} - w_{J^c}\|_2 \le C_{4.2} \frac{\rho\alpha}{D}.$$

Finally we obtain a net $\mathcal{N}_{[0,1]}$ with $|\mathcal{N}_{[0,1]}| \le 3D/(C_{4.2}\rho\alpha)$ such that for every $z \in \mathcal{Z}_J(D, \Delta, d)$ there exists a $\rho' \in [0, 1]$ for which

$$\big| \|z_J\|_2 - \rho' \big| \le C_{4.2} \frac{\rho\alpha}{D}.$$

Now let us define

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

Then setting $\|w_J\|_2 = \rho'$ and $w_J = \rho' \cdot w'$ we see that for any $z \in \mathcal{Z}(D, \Delta, d)$ there exists a $w \in \mathcal{N}(D, \Delta, d)$ such that

$$\left\| \frac{z_J}{\|z_J\|_2} - \frac{w_J}{\|w_J\|_2} \right\|_2 < C_{4.2} \frac{\alpha}{D}, \ \|z_{J^c} - w_{J^c}\|_2 \le C_{4.2} \frac{\rho\alpha}{D}, |\|z_J\|_2 - \|w_J\|_2| \le C_{4.2} \frac{\rho\alpha}{D}.$$

The set $\mathcal{N}(D, \Delta, d)$ thus constructed may not be contained in $\mathcal{Z}(D, \Delta, d)$. However, as in the proof of Lemma 4.2 this can be rectified easily. It thus remains to bound the cardinality of $\mathcal{N}(D, \Delta, d)$. By Lemma 4.1, we have

$$|\mathcal{N}(D, \Delta, d)| \le \sum_{|J| = n - M} |\mathcal{N}_{J^c}| \cdot |\mathcal{N}_{[0,1]}| \cdot |\mathcal{M}_J|$$

$$\le \binom{n}{M} \cdot \left(\frac{3}{C_{4.1}\rho} \cdot \frac{D}{\alpha} \right)^{2M+1} \cdot \left(\bar{C}_{4.1} \frac{dD^2}{\alpha} \cdot \left(\frac{1}{\sqrt{n - M}} + \frac{1}{\Delta} \right) \right)^{n-M} \cdot \left(\frac{D}{\alpha} \right)^2.$$

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

Remark 4.4. Similar to Remark 4.2 we note again that given any $\mathcal{S} \subset \mathcal{Z}(D,\Delta,d)$ there exists a net $\mathcal{N}^{\mathcal{S}}(D,\Delta,d) \subset \mathcal{S}$ with the same approximation properties and the cardinality bound as in Proposition 4.3. In Section 5 this version of Proposition 4.3 will be used.

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 $\text{Ker}(B^D)$ has a large two-dimensional LCD. Before proceeding any further let us formally define the notion of genuinely complex vectors:

$$\text{Compl}(\mathcal{Z}) := \left\{ z \in \mathcal{Z} : d(z_{\text{small}}/\left\|z_{\text{small}}\right\|_2) \geq \frac{4L}{D_2(z_{\text{small}}/\left\|z_{\text{small}}\right\|_2)} \sqrt{\log_1 \frac{\Delta(z_{\text{small}}/\left\|z_{\text{small}}\right\|_2)}{2^7 L}} \right\}. \tag{5.1}$$

Equipped with the notion of genuinely complex vectors we state the main result of this section.

Theorem 5.1. *Let $B^D, \bar{A}_n, \rho, K, R, r,$ and r' be as in Proposition 3.1. Then there exist constants $c_{5.1}, c'_{5.1}$, and $1 \leq \mu_{5.1} \leq \rho^{-1}$, 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

$$\mathbb{P}\Big(\exists z \in \text{Compl}(\mathcal{Z}) \cap \text{Ker}(B^D) : D_2(z_{\text{small}}/\left\|z_{\text{small}}\right\|_2) \leq \exp(c'_{5.1}\frac{n}{M}), \left\|\bar{A}_n\right\| \leq K\sqrt{pn}\Big) \leq \exp(-\bar{c}_{5.1}np),$$

where $M = C_{3.1}\mu_{5.1}^2\rho^{-2}p^{-1}$.

The proof of Theorem 5.1 is carried out by the following two-fold argument. Using Proposition 4.3 we show that the subset of vectors in $\text{Compl}(\mathcal{Z})$ that have a large value of $\Delta(z_{\text{small}}/\left\|z_{\text{small}}\right\|_2)$ admits a net of small cardinality. This observation together with an estimate on the small ball probability, derived from [33, Theorem 7.5], yields the desired conclusion for vectors $z \in \text{Complex}(\mathcal{Z})$ which possess a large value of $\Delta(z_{\text{small}}/\left\|z_{\text{small}}\right\|_2)$ (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 finishes the proof of Theorem 5.1. The rest of this section is devoted to make the above idea precise.

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

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

$$\mathcal{Z}(s) := \left\{ z \in \text{Compl}(\mathcal{Z}) : \ \Delta(z_{\text{small}}/\left\|z_{\text{small}}\right\|_2) \geq sL \text{ and } d(z_{\text{small}}/\left\|z_{\text{small}}\right\|_2) \leq 4 \inf_{j\in[n]} d\left(z_{\text{small}\setminus\{j\}}/\|z_{\text{small}\setminus\{j\}}\|_2\right) \right\}.$$

Let $\bar{A}_n, B^D, K, R,$ 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,

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

To prove Proposition 5.2 we use estimates on Lévy concentration function. Since the diagonal entries of $\bar{A}_n + D_n$ are non-random the bound on the Lévy concentration function depends on $\inf_{j\in[n]} d\left(z_{\text{small}\setminus\{j\}}/\|z_{\text{small}\setminus\{j\}}\|_2\right)$ rather than the real-imaginary correlation of $z_{\text{small}}/\|z_{\text{small}}\|_2$ (see Proposition 5.6). This forces the additional constraint on the real-imaginary correlation in the definition of $\mathcal{Z}(s)$. During the proof of Theorem 5.1 we will remove this constraint by another application Proposition 3.1.

As mentioned above, the proof of Proposition 5.2 requires bounds on the 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. Before deriving the bounds on the Lévy concentration function 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 $V(z') := \begin{pmatrix} x^{\mathsf{T}} \\ y^{\mathsf{T}} \end{pmatrix}$. Further denote the real-imaginary correlation of V_J by

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

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 $\{\Xi_j\}_{j\in[m]}$. The estimate on Lévy concentration function of $\sum_{j=1}^m \Xi_j z_j$ then yields probability bounds on the ℓ_2-norm of $(B^D z)_i$ to be small, for each $i \in [n-1]$. However, we remind the reader that the diagonal of $\bar{A}_n$ is zero. This implies that the probability of having a small value of $\|(B^D z)_i\|_2$ depends on the LCD of $z_{J\setminus\{i\}}/\|z_{J\setminus\{i\}}\|_2$ instead of $z_J/\|z_J\|_2$. Hence, we need to modify the definition of the LCDs so that the modified LCDs of $z_{J\setminus\{i\}}/\|z_{J\setminus\{i\}}\|_2$ can be related to the LCDs of $z_J/\|z_J\|_2$, defined in Definition 2.9.

Definition 5.3. Fixing $L \geq 1$, for a non-zero vector $x \in \mathbb{R}^m$, we denote

$$\widetilde{D}_1(x) := \inf\left\{\theta > 0 : \ \operatorname{dist}(\theta x, \mathbb{Z}^m) < 2^4 L \sqrt{\log_1 \frac{\|\theta x\|_2}{2^6 L}}\right\}. \tag{5.3}$$

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

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

Similarly as before, for $z \in \mathbb{C}^m$, we denote

$$\widetilde{D}_2(z) := \widetilde{D}_2(V(z)) \qquad \text{and} \qquad \widetilde{D}_1(z) := \widetilde{D}_1(\tilde{z}(z)),$$

where $V(z)$ and $\tilde{z}(z)$ are as in Definition 2.9.

Equipped with these new definitions of LCDs we have the following estimate, which is a direct corollary of [33, Theorem 7.5].

Proposition 5.4. *Fix any positive integer m and let $\Xi := (\Xi_1, \ldots, \Xi_m) \in \mathbb{R}^m$, $\Xi_j := \delta_j \xi_j$, $j = 1, \ldots, m$, where $\delta_1, \ldots, \delta_m$ are i.i.d.* $\text{Ber}(p)$, *and ξ_j are i.i.d. random variables satisfying*

$$\mathcal{L}(\xi_j, 1) \leq 1 - c_1 \quad \textit{and} \quad \mathbb{P}(|\xi_j| > C_1) \leq c_1/2 \tag{5.4}$$

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

$$\mathcal{L}(V\Xi, \varepsilon\sqrt{p}) \le \frac{C_{5.4}}{d(V_J)}\left(\varepsilon + \frac{1}{\sqrt{p}\widetilde{D}_2(V_J)}\right)^2 \tag{5.5}$$

and

$$\mathcal{L}(V\Xi, \varepsilon\sqrt{p}) \le \overline{C}_{5.4}\left(\varepsilon + \frac{1}{\sqrt{p}\widetilde{D}_1(x_J)}\right), \tag{5.6}$$

where $V := V(z)$, $C_{5.4}$ and $\overline{C}_{5.4}$ are some constants, depending only on c_1 and C_1, and $x_J = \mathrm{Re}\,(z_J)$.

Remark 5.5. We point out to the reader that the definition of LCD in [33] and that of ours are slightly different from each other. For example, to define LCD in [33] 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 5.3 with [33, Definition 7.1]). However, upon investigating the proof of [33, Theorem 7.5] it becomes evident that the same proof can be carried out for the LCDs $\widetilde{D}_2(\cdot)$ and $\widetilde{D}_1(\cdot)$ to obtain the same estimates on the Lévy concentration function. It only changes the constant that appears in [33, Eqn. (7.3)]. Below we apply this version of [33, Theorem 7.5].

Proof of Proposition 5.4. As mentioned above the proof is a straightforward application of [33, 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 (5.4) implies that

$$\mathcal{L}(\Xi_j, 1) \le 1 - pc_1 \quad \text{and} \quad \mathbb{P}(|\Xi_j| > C_1) \le pc_1/2. \tag{5.7}$$

Since $L = (\delta_0 p)^{-1/2}$ (see Remark 2.11), shrinking δ_0 if necessary, the inequality (5.5) follows directly from [33, Theorem 7.5], applied with $m = 2$. To prove (5.6), 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 [33, Theorem 7.5], with $m = 1$ we obtain (5.6). □

Applying Proposition 5.4 and standard tensorization techniques we obtain the following result.

Proposition 5.6. *Let B^D be as in Proposition 3.1. Fix any $z \in S_{\mathbb{C}}^{n-1}$ such that $z \notin Dom(M, \alpha_\star)$ for some $M, \alpha_\star$ satisfying $\alpha_\star\sqrt{M} > 2$. Let $J := \mathrm{small}(z)$. Then for any $\varepsilon > 0$, we have*

$$\mathcal{L}(B^D z, \varepsilon\sqrt{p(n-1)}\inf_{j\in[n]}\left\|z_{J\setminus\{j\}}\right\|_2) \le \left[\frac{C_{5.6}}{\inf_{j\in[n]} d(z_{J\setminus\{j\}}/\left\|z_{J\setminus\{j\}}\right\|_2)}\left(\varepsilon + \frac{1}{\sqrt{p}D_2(z_J/\left\|z_J\right\|_2)}\right)^2\right]^{n-1}, \tag{5.8}$$

and

$$\mathcal{L}(B^D z, \varepsilon\sqrt{p(n-1)}\inf_{j\in[n]}\left\|z_{J\setminus\{j\}}\right\|_2) \le \left[\overline{C}_{5.6}\left(\varepsilon + \frac{1}{\sqrt{p}D_1(\mathrm{Re}\,(z_J)/\left\|z_J\right\|_2)}\right)\right]^{n-1}, \tag{5.9}$$

for some constants $C_{5.6}$, and $\overline{C}_{5.6}$, depending only on $\mathbb{E}|\xi_{ij}|$ and $\mathbb{E}(\xi_{ij}^4)$.

Proof. Since $\{\xi_{ij}\}$ have finite fourth moment one can easily check that the assumption (5.4) holds. Also note that the diagonal entries of B^D are non-random, for $i \in [n-1]$. Since

$$\mathcal{L}\left((B^D z)_i, \sqrt{p}\varepsilon\inf_{j\in[n]}\left\|z_{J\setminus\{j\}}\right\|_2\right) \le \mathcal{L}\left((B^D|_{J\setminus\{i\}}z_{J\setminus\{i\}})_i, \sqrt{p}\varepsilon\left\|z_{J\setminus\{i\}}\right\|_2\right),$$

from (5.5) we obtain that

$$\mathcal{L}\left((B^D z)_i, \sqrt{p}\varepsilon \inf_{j\in[n]} \left\|z_{J\setminus\{j\}}\right\|_2\right)$$
$$\le \frac{C_{5.4}}{\inf_{j\in[n]} d(z_{J\setminus\{j\}}/\left\|z_{J\setminus\{j\}}\right\|_2)}\left(\varepsilon + \frac{1}{\sqrt{p}\inf_{j\in[n]} \widetilde{D}_2(z_{J\setminus\{j\}}/\left\|z_{J\setminus\{j\}}\right\|_2)}\right)^2.$$

We claim that

$$D_2(z_J/\left\|z_J\right\|_2) \le 4 \inf_{j\in[n]} \widetilde{D}_2(z_{J\setminus\{j\}}/\left\|z_{J\setminus\{j\}}\right\|_2). \tag{5.10}$$

Equipped with the claim (5.10), the inequality in (5.8) follows directly from a standard tensorization argument (see [38, Remark 3.5]).

Turning to prove (5.10), from the definition of $\widetilde{D}_2(z_{J\setminus\{j\}}/\|z_{J\setminus\{j\}}\|_2)$ it follows that there exists $\theta \in \mathbb{R}^2$ with $\|\theta\|_2 \le 2\widetilde{D}_2(z_{J\setminus\{j\}}/\|z_{J\setminus\{j\}}\|_2)$ such that

$$\operatorname{dist}(V_{-j}^{\mathsf{T}}\theta, \mathbb{Z}^m) < \frac{L}{2}\sqrt{\log_1 \frac{\|V_{-j}^{\mathsf{T}}\theta\|_2}{2^8 L}}, \tag{5.11}$$

where $m = n - |J\setminus\{j\}|$ and $V_{-j} := V(z_{J\setminus\{j\}}/\|z_{J\setminus\{j\}}\|_2)$. If $j \notin J$, we see that there is nothing to be proved. So we only focus on the case $j \in J$. Note that $z \notin \operatorname{Dom}(M, \alpha_\star)$ implies that

$$\|z_J\|_2 \ge \alpha_\star\sqrt{M}\|z_J\|_\infty > 2\|z_J\|_\infty,$$

which further implies that $\|z_{J\setminus\{j\}}\|_2 \ge \frac{1}{2}\|z_J\|_2$. Therefore, denoting $V_J = V(z_J/\|z_J\|_2)$, $\theta' = (\|z_J\|_2/\|z_{J\setminus\{j\}}\|_2)\cdot\theta$, and noting that $m = n - M - 1$ we see that

$$\operatorname{dist}(V_J^{\mathsf{T}}\theta', \mathbb{Z}^{n-M}) \le \operatorname{dist}(V_{-j}^{\mathsf{T}}\theta, \mathbb{Z}^m) + 1,$$

and $\|\theta'\|_2 \le 2\|\theta\|_2$. From the definition of $\log_1(\cdot)$ it follows that the RHS of (5.11) is at least $L/2 = \frac{1}{2}(\delta_0 p)^{-1/2}$. Thus, shrinking δ_0 again (if needed) and noting that $\|V_{-j}^{\mathsf{T}}\theta\|_2 \le \|V_J^{\mathsf{T}}\theta'\|_2$ the claim (5.10) follows. This completes the proof of (5.8). To prove (5.9), arguing similarly as in (5.10), we obtain

$$D_1(\operatorname{Re}(z_J)/\left\|z_J\right\|_2) \le 4 \inf_{j\in[n]} \widetilde{D}_1(\operatorname{Re}(z_{J\setminus\{j\}})/\left\|z_{J\setminus\{j\}}\right\|_2), \tag{5.12}$$

from which the proof follows upon using [38, Remark 3.5] and Proposition 5.4. We omit further details. □

Remark 5.7. The inequality (5.9) 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 non-dominated vectors. Its proof follows from [33, Proposition 7.4] and the definition of dominated vectors.

Lemma 5.8. *For any $z \notin \operatorname{Dom}(\mathrm{M}, \alpha_\star)$ we have*

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

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 $\mathcal{Z}(D,\Delta,d)$ appearing in Section 4. Indeed, using Lemma 5.8 we note that for any $z \in \mathcal{Z}(s)$ we have $D_2(z_{\text{small}}/\|z_{\text{small}}\|_2) \ge \frac{1}{2}\alpha_\star\sqrt{M} \ge \frac{\alpha_\star r_\star}{2}p^{-1/2}$ for some $\alpha_\star > 0$. Since $L = (\delta_0 p)^{-1/2}$, choosing $r_\star$ sufficiently large, we therefore obtain that

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

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

$$d(z_{\text{small}}/\left\|z_{\text{small}}\right\|_2) \ge \frac{4L}{D_2(z_{\text{small}}/\left\|z_{\text{small}}\right\|_2)}\sqrt{\log_1 \frac{\Delta(z_{\text{small}}/\left\|z_{\text{small}}\right\|_2)}{2^7 L}}. \tag{5.13}$$

If $z \in \mathcal{Z}(D,\Delta,d)$ we further have that $D_2(z_{\text{small}}/\|z_{\text{small}}\|_2) \le 2D$, $d(z_{\text{small}}/\|z_{\text{small}}\|_2) \le 2d$, and $\Delta(z_{\text{small}}/\|z_{\text{small}}\|_2) \ge \Delta$. Therefore, it follows from (5.13) that

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

So it allows us to use Proposition 4.3. Also, since $D \ge \frac{\alpha_\star r_\star}{2}p^{-1/2}$, and $L = (\delta_0 p)^{-1/2}$, choosing a sufficiently large $r_\star$, we may assume that

$$C_{4.3}\frac{\alpha}{D} \le C_{4.3}\frac{L}{D}\sqrt{\log_1 \frac{D}{2^7 L}} \le dD \le \frac{1}{8}. \tag{5.14}$$

Next recalling that $M \ge r_\star^2 p^{-1} \ge p^{-1}$, we see that the number of different values of 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

$$\mathbb{P}\left(\exists z \in \mathcal{Z}(D,\Delta,d) \cap \mathcal{Z}(s):\ B^D z = 0,\ \text{and}\ \left\|B^D\right\| \le (K+R)\sqrt{pn}\right) \le e^{-2n}, \tag{5.15}$$

for each such tripple (D,Δ,d).

To this end, we note that $\mathcal{Z}(D,\Delta,d)\cap\mathcal{Z}(s)$ admits a net $\mathcal{N}(D,\Delta,d) \subset \mathcal{Z}(D,\Delta,d)\cap\mathcal{Z}(s)$. Since for any $z \in \mathcal{Z}(s)$ we have

$$d(z_{\text{small}}/\left\|z_{\text{small}}\right\|_2) \le 4 \inf_{j\in[n]} d\left(z_{\text{small}\setminus\{j\}}/\|z_{\text{small}\setminus\{j\}}\|_2\right),$$

from Proposition 5.6 it follows that

$$\mathcal{L}(B^D w, \varepsilon\sqrt{p(n-1)} \inf_{j\in[n]} \left\|w_{\text{small}\setminus\{j\}}\right\|_2) \le \left[\frac{4C_{5.6}}{d}\left(\varepsilon + \frac{1}{\sqrt{p}D}\right)^2\right]^{n-1},$$

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

$$\varepsilon_0 := 40 C_{4.3}(K+R)\frac{\alpha}{D}. \tag{5.16}$$

Since $\alpha \ge L = (\delta_0 p)^{-1/2} \ge p^{-1/2}$ and $K,R \ge 1$ we note that $\varepsilon_0 \ge \frac{1}{\sqrt{p}D}$. Recall that $\mathcal{N}(D,\Delta,d) \subset \text{Comp}(M,\rho)^c$. Therefore for any $w \in \mathcal{N}(D,\Delta,d)$ we also have

$$\|w_{\text{small}}\|_2 \le 2 \inf_{j\in[n]} \|w_{\text{small}\setminus\{j\}}\|_2$$

(see also (3.9)). This further implies that

$$\mathbb{P}\left(\left\|B^D w\right\|_2 \le \frac{\varepsilon_0}{3}\|w_{\text{small}}\|_2 \cdot \sqrt{pn}\right) \le \left[\frac{C}{d}\left((K+R)\frac{\alpha}{D}\right)^2\right]^{n-1},$$

for some positive constant C. Hence, by the union bound and applying Proposition 4.3

$$\begin{aligned}&\mathbb{P}\left(\exists w \in \mathcal{N}(D,\Delta,d) \ \ \left\|B^D w\right\|_2 \le \frac{\varepsilon_0}{3}\|w_{\text{small}}\|_2 \cdot \sqrt{pn}\right)\\ &\le |\mathcal{N}(D,\Delta,d)| \cdot \left[\frac{C}{d}\left((K+R)\frac{\alpha}{D}\right)^2\right]^{n-1}\\ &\le \left[\frac{C}{d}\left((K+R)\frac{\alpha}{D}\right)^2\right]^{n-1} \cdot \bar{C}_{4.3}^n \left(\frac{n}{\rho M}\cdot\frac{D}{\alpha}\right)^{5M} \cdot \left(\frac{dD^2}{\alpha}\cdot\left(\frac{1}{\sqrt{n}}+\frac{1}{\Delta}\right)\right)^{n-M}. \qquad (5.17)\end{aligned}$$

Recalling the definition of α and using the inequalities $L \le \alpha \le dD$ and $\Delta \le D$ we note that

$$\frac{1}{d}\cdot\left(\frac{\alpha}{D}\right)^2 \le \frac{\alpha}{D} = \frac{L}{D}\sqrt{\log_1\frac{\Delta}{2^7 L}} \le \frac{L}{D}\sqrt{\log_1\frac{D}{2^7 L}} \le \frac{1}{2^7}\sqrt{\frac{2^7 L}{D}},$$

where the last step follows from the fact that $x^{-1}\log x \to 0$ as $x \to \infty$. As we have already noted that $D \ge C_\star L$, for some large $C_\star$, we can enlarge $C_\star$ further (i.e. we increase $r_\star$) so that $\frac{C}{d}((K+R)\frac{\alpha}{D})^2\bar{C}_{4.3} < 1$. This means that we can drop the term

$$\left[\frac{C}{d}\left((K+R)\frac{\alpha}{D}\right)^2\right]^{M-1} \cdot \bar{C}_{4.3}^M < 1$$

in (5.17). Therefore, from (5.17) we obtain

$$\mathbb{P}\left(\exists w \in \mathcal{N}(D,\Delta,d) \ \ \left\|B^D w\right\|_2 \le \frac{\varepsilon_0}{3}\|w_{\text{small}}\|_2 \cdot \sqrt{pn}\right) \le \exp(-\Gamma n), \qquad (5.18)$$

where

$$\Gamma := -\left(1-\frac{M}{n}\right)\cdot\log\left(\frac{C'(K+R)^2\alpha}{\sqrt{n}} + \frac{C'(K+R)^2\alpha}{\Delta}\right) - \frac{5M}{n}\log\left(\frac{n}{\rho M}\cdot\frac{D}{\alpha}\right),$$

and $C' := C\cdot\bar{C}_{4.3}$. To finish the proof we need to show that $\Gamma \ge 2$.

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

$$\frac{5M}{n}\log\left(\frac{n}{\rho M}\cdot\frac{D}{\alpha}\right) \le \frac{10M}{n}\log\left(\frac{n}{M}\right) + \frac{5M}{n}\log D \le 20.$$

Next recall that $L = (\delta_0 p)^{-1/2}$ and $\Delta \le D \le \exp(c'n/M)$. So

$$\alpha = L\sqrt{\log_1\frac{\Delta}{2^7 L}} \le \sqrt{\frac{c'n}{\delta_0 M p}}.$$

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

$$\Gamma \ge -\frac{1}{2}\cdot\log\left(\frac{C'(K+R)^2\sqrt{c'}}{\sqrt{\delta_0 M p}} + \frac{C'(K+R)^2\alpha}{\Delta}\right) - 20.$$

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

$$\frac{C'(K+R)^2\alpha}{\Delta} < e^{-50}. \qquad (5.19)$$

Thus choosing c' small enough and recalling that $Mp \ge r_\star^2$, from claim (5.19) we see that $\Gamma \ge 2$, providing the required bound for the probability.

Now let us check our claim (5.19). 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

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

for any Δ such that $\Delta \ge s_{5.2}L$. This proves the claim (5.19).

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

$$\mathbb{P}\left(\exists w \in \mathcal{N}(D,\Delta,d)\ \ \left\|B^D w\right\|_2 \le \frac{\varepsilon_0}{3}\|w_{\text{small}}\|_2 \cdot \sqrt{pn}\right) \le \exp(-2n). \tag{5.20}$$

To deduce (5.15) from (5.20) we simply use the property of the net $\mathcal{N}(D,\Delta,d)$. Indeed, let us assume that there exists a $z \in \mathcal{Z}(s_{5.2}) \cap \mathcal{Z}(D,\Delta,d)$ so that $B^D z = 0$. Denoting $J = \text{small}(z)$, using Proposition 4.3, and the triangle inequality we see that there exists a $w \in \mathcal{N}(D,\Delta,d)$ such that

$$\begin{aligned}
\|B^D w\|_2 &= \|B^D(w-z)\|_2 \\
&\le \|B^D\|_2 \cdot \left(\|z_{J^c} - w_{J^c}\|_2 + \|w_J\|_2\left\|\frac{z_J}{\|z_J\|_2} - \frac{w_J}{\|w_J\|_2}\right\|_2 + |\,\|w_J\| - \|z_J\|_2\,|\right) \\
&\le (K+R)\sqrt{np}\cdot\left(2C_{4.3}\frac{\rho\alpha}{D} + C_{4.3}\frac{\alpha}{D}\|w_J\|_2\right) \le 3C_{4.3}(K+R)\frac{\alpha}{D}\|w_J\|_2\sqrt{np},
\end{aligned}$$

where the last inequality follows from the fact that $w \in \mathcal{N}(D,\Delta,d) \subset \mathcal{Z}(s_{5.2}) \subset \text{Comp}(M,\rho)^c$. To complete the proof, let us show that $\|w_{\text{small}}\|_2 \le 4\|w_J\|_2$. Assume for a moment that the opposite inequality holds. Denote $\text{small}(w) = I$. Combining Proposition 4.3 and (5.14), we see that

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

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

Recalling the definition of ε_0 in (5.16), we now deduce (5.15) from (5.20). This finishes the proof. $\square$

Using Proposition 5.2 we now finish 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 [38, Lemma 6.2].

Lemma 5.9. *For any $x \in \mathbb{R}^m$,*

$$D_1(x) \ge \frac{1}{2\|x\|_\infty}.$$

Proof of Theorem 5.1. First we claim that

$$\mathbb{P}\left(\exists z \in \text{Ker}(B^D) \cap S_{\mathbb{C}}^{n-1} : d(z_J/\|z_J\|_2) > 2\inf_{j\in[n]} d(z_{J\setminus\{j\}}/\|z_{J\setminus\{j\}}\|_2)\right) \le 2\exp(-\bar{c}_{3.1}np), \tag{5.21}$$

where $J = \text{small}(z)$. Indeed, recalling that $z_J/\|z_J\|_2 = \phi + \mathrm{i}\psi$ where $\phi \perp \psi$, we note that $d(z_J/\|z_J\|_2) = \|\phi\|_2 \cdot \|\psi\|_2$. On other hand, for $j \in J$, denoting $z_{J\setminus\{j\}}/\|z_{J\setminus\{j\}}\|_2 =: \phi_{-j} + \mathrm{i}\psi_{-j}$ we see that

$$|\langle \phi_{-j}, \psi_{-j}\rangle| \le \frac{\|z_J\|_2^2}{\|z_{J\setminus\{j\}}\|_2^2} \cdot \|\phi\|_\infty \cdot \|\psi\|_\infty.$$

Therefore,

$$\begin{aligned} d^2(z_{J\setminus\{j\}}/\|z_{J\setminus\{j\}}\|_2) &= \|\phi_{-j}\|_2^2 \cdot \|\psi_{-j}\|_2^2 - |\langle \phi_{-j}, \psi_{-j}\rangle|^2 \\ &\ge \left(\frac{\|z_J\|_2}{\|z_{J\setminus\{j\}}\|_2}\right)^4 \cdot \left[(\|\phi\|_2^2 - \|\phi\|_\infty^2)(\|\psi\|_2^2 - \|\psi\|_\infty^2) - \|\phi\|_\infty^2 \cdot \|\psi\|_\infty^2\right], \end{aligned} \tag{5.22}$$

where the last inequality follows from the relations between (ϕ, ψ) and (ϕ_{-j}, ψ_{-j}). Since $z \in \text{Ker}(B^D)$ implies that $\mathrm{i}z \in \text{Ker}(\mathrm{B}^\mathrm{D})$, using the union bound and setting $c_{5.1} < c_{3.1}$, from Proposition 3.1 we have that

$$\mathbb{P}\Bigg(\exists z \in \text{Ker}(B^D) \cap S_{\mathbb{C}}^{n-1} : \left\|\frac{\text{Re}(z_{\text{small}})}{\|\text{Re}(z_{\text{small}})\|_2}\right\|_\infty \bigvee \left\|\frac{\text{Im}(z_{\text{small}})}{\|\text{Im}\,(z_{\text{small}})\|_2}\right\|_\infty \ge \mu^{-1} p^{1/2}, \\ . \qquad \text{and } \left\|\bar{A}_n\right\| \le K\sqrt{np}\Bigg) \le 2\exp(-\bar{c}_{3.1} np),$$

where for two real numbers x and y we denote $x \vee y := \max\{x, y\}$. Therefore, fixing any $\mu > 2$ we have $\max\{\|\phi\|_\infty/\|\phi\|_2, \|\psi\|_\infty/\|\psi\|_2\} \le \frac{1}{2}$, on a set with probability at least $1 - 2\exp(-\bar{c}_{3.1}np)$. This together with (5.22) establishes the claim (5.21). Next note that, if p satisfies the assumption (5.2) for a sufficiently small $c_{5.1}$, then the inequality $r_\star^2 p^{-1} \le M \le \rho n$ holds for $M = C_{3.1}\mu^2\rho^{-4}p^{-1}$, with $2 \le \mu \le \rho^{-1}$. Therefore we can apply Proposition 5.2 and hence from (5.21) it follows that

$$\mathbb{P}\Big(\exists z \in \mathcal{Z}_0(s_{5.2}) \cap \text{Ker}(B^D) : \ D_2(z_{\text{small}}/\left\|z_{\text{small}}\right\|_2) \le \exp(c'_{5.2}\frac{n}{M}), \left\|\bar{A}_n\right\| \le K\sqrt{pn}\Big) \le 3\exp(-\bar{c}_{3.1}np), \tag{5.23}$$

whenever $\mu > 2$, where

$$\mathcal{Z}_0(s) := \left\{z \in \text{Compl}(\mathcal{Z}) : \ \Delta(z_{\text{small}}/\left\|z_{\text{small}}\right\|_2) \ge sL\right\}.$$

To complete the proof we next show that

$$\mathbb{P}\left(\exists z \in \text{Ker}(B^D) \cap S_{\mathbb{C}}^{n-1} : \Delta\left(z_{\text{small}}/\|z_{\text{small}}\|_2\right) \le \frac{\mu}{4\sqrt{p}}\right) \le \exp(-\bar{c}_{3.1}np). \tag{5.24}$$

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 (5.24), we fix $z \in \text{Ker}(B^D) \cap S_{\mathbb{C}}^{n-1}$ and denote $z_{\text{small}}/\left\|z_{\text{small}}\right\|_2 = \phi + \mathrm{i}\psi$, where $\phi, \psi \in \mathbb{R}^J$, $J = \text{small}(z)$. Let $\theta = (\theta_1, \theta_2) \in \mathbb{R}^2$ be such that

$$\text{dist}(\theta_1\phi + \theta_2\psi, \mathbb{Z}^J) < L\sqrt{\log_1 \frac{\|\theta_1\phi + \theta_2\psi\|_2}{2^8 L}} \tag{5.25}$$

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

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

Then $w \in \text{Ker}(B^D) \cap S_{\mathbb{C}}^{n-1}$ and $w_{\text{small}} = z_{\text{small}}$. Therefore (5.25) implies that

$$D_1\left(\text{Re}(w_{\text{small}})/\|\text{Re}(w_{\text{small}})\|_2\right) \le 2\Delta\left(z_{\text{small}}/\|z_{\text{small}}\|_2\right).$$

Upon applying Lemma 5.9 we find that

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

Since $w \in \mathrm{Ker}(B^D) \cap S_{\mathbb{C}}^{n-1}$ the claim (5.24) now follows from another application of Proposition 3.1.

Finally, recalling the fact that $L = (\delta_0 p)^{-1/2}$, and shrinking ρ, if necessary, we choose $2 < \mu \leq \rho^{-1}$ so that

$$s_{5.2} L < \frac{\mu}{4\sqrt{p}}.$$

The desired result then follows from (5.23) and (5.24). □

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

$$\mathrm{Real}(\mathcal{Z}) := \mathcal{Z} \backslash \mathrm{Compl}(\mathcal{Z}). \tag{6.1}$$

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 correlation $d(\cdot)$ is very small which precludes using (5.8). Instead we have to rely on the probability bound obtained in (5.9), 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 $\mathrm{real}(\mathcal{Z})$ would be easier. To construct such a net we will follow the method of [33]. 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 bounds 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.

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

Proof. Let us denote $J = \mathrm{small}(z)$. Denoting $D = D_2(\phi + \mathrm{i}\psi)$, we see that there exists $\theta = (\theta_1, \theta_2) \in \mathbb{R}^2$ with $\|\theta\|_2 \leq 2D$ and $\|\phi\theta_1 + \psi\theta_2\|_2 \geq \Delta(z_J/\|z_J\|_2)/\sqrt{2}$, 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 $|\theta_2| \leq \|\theta\|_2$, $\|\phi\|_2 \cdot \|\psi\|_2 = d(z_J/\|z_J\|_2)$, and $\|\phi\|_2 \geq 1/2$, we also obtain

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

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

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

(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

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

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 $|\theta_1| \le \|\theta\|_2 \le 2D$, 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 $0 < \widetilde{\alpha} \le D$. Define*

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

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

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

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

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

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

Then, using the triangle inequality,

$$\|q\|_2 < \widetilde{\alpha} + |\theta| \le 4D,$$

and so $q \in \mathbb{Z}_{4D}$. This implies that

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

From the definition of real-imaginary correlation it also follows that

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

Let $\mathcal{N}_1$ be an $(\widetilde{\alpha}/D)$-net in $[-1,1]$ with

$$|\mathcal{N}_1| \le 2\frac{D}{\widetilde{\alpha}}.$$

Define $\mathcal{M}_J^1(D)$ by

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

Then from (6.5)-(6.6) we deduce that $\mathcal{M}_J^1(D)$ is a $(7\widetilde{\alpha}/D)$-net for $S_J(D)$ and $|\mathcal{M}_J^1(D)| = |\mathbb{Z}_{4D}|\cdot|\mathcal{N}_1|$. This in combination with the bound in (4.2) 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 $S_J(D)$. This is done in the same way as in Lemma 4.1. 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 $\widetilde{\mathcal{Z}}(D)$ by

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

where

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

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

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

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

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

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

Proposition 6.3 is derived from Lemma 6.2 in the same way as Proposition 4.3 was derived from Lemma 4.1. 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 $B^D, \bar{A}_n, \rho, K, R, r$, 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*

$$\mathbb{P}\Big(\exists z \in \mathrm{Real}(\mathcal{Z}) \cap \mathrm{Ker}(B^D) : D_2(z_{\mathrm{small}}/\,\|z_{\mathrm{small}}\|_2) \leq \exp(c'_{6.4} n/M) \, and \, \left\|\bar{A}_n\right\| \leq K\sqrt{pn} \Big) \leq e^{-n},$$

where $M = C_{3.1}\mu_{5.1}^2 \rho^{-2} p^{-1}$.

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 $D_1(\phi(z)) \leq \exp(c'n/M)$ for some small constant c', where

$z_{\text{small}}/\|z_{\text{small}}\|_2 =: \phi(z) + \mathrm{i}\psi(z)$ with $\|\phi(z)\|_2 \ge \|\psi(z)\|_2$. We then claim that the subset of $\text{Real}(\mathcal{Z})$ in context can be partitioned into the sets $\widetilde{\mathcal{Z}}(D)$ as follows:

$$\{z \in \text{Real}(\mathcal{Z}) : \; D_1(\phi(z)) \le \exp(c'n/M)\} \subset \bigcup_D \widetilde{\mathcal{Z}}(D), \tag{6.8}$$

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

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

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

(6.9)

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

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 $D_1(\phi(z)) \in [D, 2D]$ then recalling the definition of $\widetilde{\alpha}$, from (6.9) we see that

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

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

$$D_0 := C_0 \mu_{5.1} p^{-1/2},$$

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

$$\|\phi(z)\|_\infty \le \frac{\|z_{\text{small}}\|_\infty}{\|z_{\text{small}}\|_2} \le \frac{1}{\rho\sqrt{M}} = \frac{\sqrt{p}}{\mu_{5.1}\sqrt{C_{3.1}}},$$

where the last step follows from our choice of M. Hence, using Lemma 5.9 we see that for any $z \in \text{Real}(\mathcal{Z})$ we must have $D_1(\phi(z)) \ge D_0$. This establishes that the union in the RHS of (6.8) can be taken over all $D = 2^k, \; D_0 \le D \le \exp(c'n/M)$. So using the union bound, we deduce that it is enough to show that

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

for each such D.

To this end, using Proposition 5.6 we see that for any $w \in \widetilde{\mathcal{N}}(D)$ we have

$$\mathcal{L}(B^D w, \varepsilon\sqrt{p(n-1)} \inf_{j \in [n]} \|w_{\text{small}\setminus\{j\}}\|_2) \le \left[\overline{C}_{5.6}\left(\varepsilon + \frac{1}{\sqrt{p}D}\right)\right]^{n-1}.$$

Since $\widetilde{\mathcal{N}}(D) \subset \text{Comp}(M, \rho)^c$ we have

$$\rho \le \|w_{\text{small}}\|_2 \le 2 \inf_{j \in [n]} \|w_{\text{small}\setminus\{j\}}\|_2.$$

Now set

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

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

$$\mathbb{P}\left(\left\|B^D w\right\|_2 \le \frac{\varepsilon_0}{3}\|w_{\text{small}}\|_2 \cdot \sqrt{pn}\right) \le \left[\overline{C}_{5.6}\left(\widetilde{\varepsilon}_0 + \frac{1}{\sqrt{pD}}\right)\right]^{n-1} \le \left(\widetilde{C}(K+R)\frac{\widetilde{\alpha}}{D}\right)^{n-1},$$

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

$$\begin{aligned}&\mathbb{P}\left(\exists w \in \widetilde{\mathcal{N}}(D): \ \left\|B^D w\right\|_2 \le \frac{\varepsilon_0}{3}\|w_{\text{small}}\|_2 \cdot \sqrt{pn}\right)\\ &\le |\widetilde{\mathcal{N}}(D)| \cdot \left(\widetilde{C}(K+R)\frac{\widetilde{\alpha}}{D}\right)^{n-1} \le \left(C'(K+R)\frac{\widetilde{\alpha}}{D}\right)^{n-1}\left(\frac{n}{\rho M}\cdot\frac{D}{\widetilde{\alpha}}\right)^{4M}\left(\frac{D}{\sqrt{n}}+1\right)^{n-M},\end{aligned}$$

where C' is some large constant. Next recalling the definitions of $\widetilde{\alpha}$ and D_0, using the facts that $D \ge D_0$, $L = (\delta_0 p)^{-1/2}$ and the function $f(x) := x\sqrt{\log_1(1/x)}$ is increasing for $x \in (0, e^{-1})$ we find that

$$\frac{\widetilde{\alpha}}{D} = \frac{2^5 L}{D}\sqrt{\log_1 \frac{D}{2^5 L}} \le \frac{2^5 L}{D_0}\sqrt{\log_1\frac{D_0}{2^5 L}} = f\left(\frac{2^5}{C_0 \mu_{5.1}\delta_0^{1/2}}\right). \tag{6.10}$$

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

$$\mathbb{P}\left(\exists w \in \widetilde{\mathcal{N}}(D): \ \left\|B^D w\right\|_2 \le \frac{\varepsilon_0}{3}\|w_{\text{small}}\|_2 \cdot \sqrt{pn}\right) \le \exp(-\widetilde{\Gamma} n),$$

where

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

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

$$\mathbb{P}\left(\exists w \in \widetilde{\mathcal{N}}(D): \ \left\|B^D w\right\|_2 \le \frac{\varepsilon_0}{3}\|w_{\text{small}}\|_2 \cdot \sqrt{pn}\right) \le \exp(-2n). \tag{6.11}$$

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

$$\frac{4M}{n}\log\left(\frac{n}{\rho M}\cdot\frac{D}{\widetilde{\alpha}}\right) \le 10,$$

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

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

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

$$\frac{\widetilde{\alpha}}{\sqrt{n}} \le \frac{2^5 L\sqrt{\log_1\frac{D}{2^5 L}}}{\sqrt{n}} \le \frac{2^5 L\sqrt{c'\frac{n}{M}}}{\sqrt{n}} = \frac{2^5\sqrt{c'}}{\sqrt{\delta_0 M p}} \le \frac{2^5\rho^2\sqrt{c'}}{\sqrt{C_{3.1}\delta_0}} \le e^{-30},$$

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 $\widetilde{\mathcal{Z}}(D)$ by a point of the set $\widetilde{\mathcal{N}}(D)$, 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 finish the proof of Theorem 2.2. First let us state the following general result from which Theorem 2.2 follows.

Theorem 7.1. *Let $\bar{A}_n$ be an $n \times n$ matrix with zero diagonal and i.i.d. off-diagonal 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 with unit variance and finite fourth moment. Fix $K, R \geq 1$, and $r \in (0,1]$ and let $\Omega_K := \{\|\bar{A}_n\| \leq K\sqrt{np_n}\}$. Assume that D_n is a diagonal matrix such that $\|D_n\| \leq R\sqrt{np_n}$ and $\operatorname{Im}(D_n) = r'\sqrt{np_n}I_n$ 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

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

then

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

(ii) Additionally, if

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

then

$$\mathbb{P}\left(\left\{s_{\min}(\bar{A}_n + D_n) \leq c_{7.1}\varepsilon \exp\left(-C_{7.1}\frac{\log(1/p_n)}{\log(np_n)}\right)\sqrt{\frac{p_n}{n}}\right\}\bigcap \Omega_K\right) \leq \varepsilon + \exp(-c'_{7.1}\sqrt{np_n}).$$

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 $\operatorname{Ker}(B^D)$ with a sub-exponential 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 $\operatorname{Ker}(B^D)$ has an exponential two-dimensional LCD. However, we do not have any control the real-imaginary correlation of this vector. This means that we cannot use the bound (5.8), and have to rely on (5.9). To apply (5.9), we therefore need to show 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 non-zero vector $x \in \mathbb{R}^m$, we set

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

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

Lemma 7.3. *For $z := x + iy \in \mathbb{C}^m$ we have*

$$\widehat{D}_1(x) \geq D_2(z).$$

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

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

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

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

□

Now we are ready to prove Theorem 7.1.

Proof of Theorem 7.1. The proof is similar to that of [8, Theorem 1.1]. We include it for completeness. Note that for any $\vartheta > 0$,

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

where

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

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

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

Therefore it only remains to find an upper bound on the second term in the RHS of (7.2). Applying Lemma 2.5 we see that to find an upper bound of

$$\mathbb{P}\Big(\Big\{\inf_{x\in V} \big\|(\bar{A}_n + D_n)x\big\|_2 \le \varepsilon\rho^2\sqrt{\frac{p}{n}}\Big\} \cap \Omega_K\Big)$$

is enough to find the same for

$$\mathbb{P}\Big(\Big\{\operatorname{dist}(\bar{A}_{n,j}, H_{n,j}) \le \rho\sqrt{p}\varepsilon\Big\} \cap \Omega_K\Big) \text{ for a fixed } j,$$

where $\bar{A}_{n,j}$ are columns of $(\bar{A}_n + D_n)$. 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 $\bar{A}_{n,2}, \ldots, \bar{A}_{n,n}$. Therefore

$$\operatorname{dist}(\bar{A}_{n,1}, H_{n,1}) \ge |\langle v, \bar{A}_{n,1}\rangle|,$$

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

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

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

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

$$\mathbb{P}\left(|\langle \bar{A}_{n,1}, v\rangle| \le \rho\varepsilon\sqrt{p}\right) \le \mathcal{L}\left(\sum_{i\in J} v_i\delta_i\xi_i, \rho\sqrt{p}\varepsilon\right).$$

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

$$\|v_{[c_{3.4}n+1,n]}\|_\infty \le \frac{C_{3.4}(K+R))^4}{\sqrt{c_{3.4}n}} \|v_{[c_{3.4}n+1,n]}\|_2 \quad \text{and} \quad \|v_{[c_{3.4}n+1,n]}\|_2 \ge \rho,$$

from which it easily follows that

$$\|v_J\|_\infty \le \frac{2C_{3.4}(K+R))^4}{\sqrt{c_{3.4}n}} \|v_J\|_2 \quad \text{and} \quad \|v_J\|_2 \ge \rho/2.$$

Therefore the Berry–Esséen Theorem (see [28, Theorem 2.2.17]) yields that

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

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

$$C'' = \frac{2C_{3.4}(K+R))^4}{\sqrt{c_{3.4}}} \cdot C'.$$

Now replacing ε by ε/C the proof of part (i) of the theorem finishes.

It now 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 $M_0 = C_{3.1}\mu_{5.1}^2 \rho^{-2} p^{-1}$ from Proposition 3.4 we find that it is enough to bound

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

where

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

Further denote

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

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

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

for some $\bar{c} > 0$. Since $\text{Ker}(B^D)$ is invariant under rotation, recalling the definition of the set $\mathcal{Z}$ (see (4.1)), we see that it is enough to show that

$$\mathbb{P}\Big(\Big\{\exists v \in \mathcal{Z} \cap \text{Ker}(B^D) : D_2\left(v_{\text{small}}/\|v_{\text{small}}\|_2\right) \le \exp(c'n/M_0)\Big\} \cap \Omega_K\Big) \le \exp(-\bar{c}np).$$

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 now it only remains to find an upper bound on

$$\mathbb{P}\Big(\Big\{\exists v \in \mathcal{Z} \cap \text{Ker}(B^D) : D_2\left(v_{\text{small}}/\|v_{\text{small}}\|_2\right) > \exp(c'n/M_0) \text{ and } |\langle \bar{A}_{n,1}, v\rangle| \le \rho\varepsilon\sqrt{p}\Big\} \cap \Omega_K\Big). \tag{7.7}$$

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

$$D_2\left(v_{\text{small}}/\|v_{\text{small}}\|_2\right) > \exp(c'n/M_0).$$

Lemma 7.3 implies that

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

where we recall that $v_{\text{small}}/\|v_{\text{small}}\|_2 = \phi(v) + \mathrm{i}\psi(v)$. Now repeating the proofs of (5.9) and (5.12), and recalling the definition of M_0 we deduce that

$$\mathbb{P}(|\langle \bar{A}_{n,1}, v\rangle| \le \varepsilon\rho\sqrt{p}) \le \bar{C}\left(\varepsilon + \frac{1}{\sqrt{p}\widehat{D}_1(\phi(v))}\right) \le \bar{C}\left(\varepsilon + \frac{1}{\sqrt{p}}\exp(-c''np\rho^4)\right),$$

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

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

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

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

This completes the proof of the theorem. □

Proof of Theorem 2.2. Proof follows from Theorem 7.1, [8, Theorem 1.7], and the triangle inequality. We omit the details. □

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 [8, Proposition 3.1] we have $\rho = \exp(-C\log(1/p)/\log(np))$, for some large C. If one can improve the conclusion of [8, 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

In this short section, our goal is to prove Theorem 2.12 which shows that there are not too many singular values of the matrix $\frac{1}{\sqrt{np}}A_n - wI_n$ near zero. To prove Theorem 2.12 we employ the same strategy as in [12, 37, 41]. 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. $\mathrm{Ber}(p)$. 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 $c_{8.1}$[2] such that for every sub-space H of $\mathbb{C}^n$ with $1 \le \dim(H) \le n - \psi(n)$, we have*

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

A result similar to Lemma 8.1 was obtained in [37] (see Proposition 5.1 there) for the dense case. Later in [12] (and [41]) it was improved for the sparse case. Our Lemma 8.1 follows from [12, 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 [12, Lemma 3.5].

We now complete the proof of Theorem 2.12 using Lemma 8.1. We employ same strategy as in [37, pp. 2055-2056] (see also the proof of [12, Lemma 3.14]).

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

Proof of Theorem 2.12. To lighten the notation, let us denote by $s_1 \ge s_2 \ge \cdots \ge s_n$ the singular values of $(A_n - \sqrt{np} w I_n)$. Fix i such that $3\psi(n) \le i \le n-1$ and denote $A_n^{m,w}$ to be the sub-matrix formed by first m rows of the matrix $(A_n - \sqrt{np} w I_n)$, where $m = n - \lceil i/2 \rceil$. Further denote $s_1' \ge s_2' \ge \cdots \ge s_m'$ to be 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 [37, Lemma A.4] it follows that

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

where $\text{dist}_j' := \text{dist}(\boldsymbol{a}_j - w\sqrt{np} e_j, H_{j,n}^{m,w})$, $\boldsymbol{a}_j^{\mathsf{T}}$ is the j-th 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 j-th row, and e_j is the j-th canonical basis. We also note that $\text{dist}_j \le \text{dist}_j'$, where $\text{dist}_j := \text{dist}(\boldsymbol{a}_j, \text{span}(H_{j,n}^{m,w}, e_j))$. Thus from (8.1)-(8.2) we deduce

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

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

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

where we used the fact that $n - \dim(\text{span}(H_{j,n}^{m,w}, e_j)) \ge n - (m+1) \ge i/3$ and chose $C_{2.12} \ge 4c_{8.1}^{-1}$. Hence, from (8.3) we see that

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

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

9. Weak Convergence

Here our goal is to prove Theorem 2.13. 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 2.13 holds for $\{\xi_{i,j}\}_{i,j=1}^n$ bounded then it continues to hold without the boundedness assumption.*

To prove Lemma 9.1 we need to use the notion of bounded Lipschitz metric between two probability measures.

Definition 9.2. Let μ and ν be two probability measures on $\mathbb{R}$. The bounded Lipschitz metric is defined as

$$d_{BL}(\mu, \nu) := \sup\left\{\int f d\mu - \int f d\nu : \|f\|_\infty + \|f\|_L \le 1\right\},$$

where $\|f\|_\infty := \sup_x |f(x)|$ and $\|f\|_L$ denotes the Lipschitz norm of f.

In the context of Lemma 9.1, an important property of the bounded Lipschitz metric is that it characterizes the weak convergence. That is, a sequence of probability measures $\{\mu_n\}$ converges weakly to μ_∞ if and only if $d_{BL}(\mu_n, \mu_\infty) \to 0$ as $n \to \infty$ (see [3, Theorem C.8]). We will exploit this property of $d_{BL}(\cdot,\cdot)$ to prove Lemma 9.1. The approach of our proof is similar to that of [15, Proposition 4.1].

Proof of Lemma 9.1. Define

$$\xi_{ij}^K := \frac{\xi_{i,j}\mathbb{I}(|\xi_{i,j}| \le K) - \mu(K)}{\sigma(K)},$$

where

$$\mu(K) := \mathbb{E}[\xi_{i,j}\mathbb{I}(|\xi_{i,j}| \le K)] \qquad \text{and} \qquad \sigma^2(K) := \mathrm{Var}(\xi_{i,j}\mathbb{I}(|\xi_{i,j}| \le K)).$$

Using dominated convergence theorem we note that $\mu(K) \to 0$ and $\sigma(K) \to 1$ as $K \to \infty$.

Let A_n^K be the matrix whose (i,j)-th entry is $\xi_{ij,}^K \cdot \delta_{i,j}$. Denote $\nu_n^{w,K}$ to be the ESD of

$$\begin{bmatrix} 0 & \frac{1}{\sqrt{np}} A_n^K - wI_n \\ \frac{1}{\sqrt{np}}(A_n^K)^* - \bar{w}I_n & 0 \end{bmatrix}.$$

Then by the assumption of this lemma $\nu_n^{w,K}$ converges weakly to ν_∞^w for every $K > 0$. This in particular implies that $d_{BL}(\nu_n^{w,K}, \nu_\infty^w) \to 0$ as $n \to \infty$, for all $K > 0$. From the definition of bounded Lipschitz metric it is easy to see that for any two $2n \times 2n$ matrices Hermitian matrices B_1 and B_2

$$d_{BL}(\mu^{B_1}, \mu^{B_2}) \le \sup\left\{ \frac{1}{2n}\sum_{j=1}^{2n} |f(\lambda_j(B_1)) - f(\lambda_j(B_2))|, \, \|f\|_L \le 1 \right\} \le \frac{1}{2n}\sum_{j=1}^{2n} |\lambda_j(B_1) - \lambda_j(B_2)|,$$

where μ^{B_1} and μ^{B_2} denote the ESD of B_1 and B_2 respectively, and $\{\lambda_j(B_1)\}_{j=1}^{2n}$ and $\{\lambda_j(B_2)\}_{j=1}^{2n}$ denote the eigenvalues of B_1 snd B_2, arranged in non-decreasing order.

Therefore using Hoffman-Wielandt inequality (see [3, Lemma 2.1.19]) we see that

$$d_{BL}^2(\nu_n^{K,w}, \nu_n^w) \le \frac{1}{n^2 p} \mathrm{Tr}[(A_n - A_n^K)(A_n - A_n^K)^*] = \frac{\sum_{i,j=1}^n (\tilde{\xi}_{i,j}^K)^2 \delta_{ij}}{n^2 p},$$

where $\tilde{\xi}_{i,j}^K := \xi_{i,j} - \xi_{i,j}^K$. Using Chernoff's inequality, followed by an application of Borel-Cantelli lemma we note that

$$\sum_{i,j=1}^n \delta_{i,j} \le Cn^2 p, \qquad \text{almost surely },$$

for some absolute constant C. Therefore

$$\mathbb{P}\left(d_{BL}^2(\nu_n^{K,w}, \nu_n^w) \ge 2C\mathbb{E}[(\xi_{i,j}^K)^2]\right) \le \mathbb{P}\left(\sum_{\ell=1}^{Cn^2 p} (\xi_\ell^K)^2 \ge 2Cn^2 p \mathbb{E}[(\xi_{i,j}^K)^2] \right), \tag{9.1}$$

where $\{\xi_\ell^K\}$ are i.i.d. copies of $\{\xi_{i,j}^K\}$. One can check that $\mathbb{E}[(\xi_\ell^K)^2] \to 0$ as $K \to \infty$. Thus applying the weak law of large numbers we obtain that for any $\varepsilon > 0$ there exists $K_0(\varepsilon)$ such that for all $K \ge K_0(\varepsilon)$,

$$\lim_{n\to\infty} \mathbb{P}(d_{BL}(\mu_n^w, \nu_n^{K,w}) \ge \varepsilon/2) = 0.$$

Hence using triangle inequality we find

$$\lim_{n\to\infty} \mathbb{P}(d_{BL}(\mu_n^w, \nu_\infty^w) \ge \varepsilon) = 0.$$

Since $\varepsilon > 0$ is arbitrary we obtain that the conclusion of part (i) of Theorem 2.13 holds without the boundedness assumption. To prove that part (ii) of Theorem 2.13 also holds without the boundedness assumption we proceed from (9.1) by applying Markov's inequality, followed by another application of the Borel-Cantelli lemma. We omit the details. This finishes the proof of the lemma. ☐

Equipped with Lemma 9.1, hereafter we assume that $\{\xi_{i,j}\}_{i,j=1}^n$ are bounded, and under that assumption we show that the Stieltjes transform of the ESD of $\boldsymbol{A}_n^w$ converges, i.e. we prove Theorem 2.15 (recall the definition of $\boldsymbol{A}_n^w$ from (2.7)). We remind the reader that the limit of $m_n(\zeta)$ (Stieltjes transform of the ESD of $\boldsymbol{A}_n^w$), denoted by $m_\infty(\zeta)$, satisfy the equation

$$P(m) := m(m+\zeta)^2 + m(1-|w|^2) + \zeta = 0.$$

Our strategy for proving Theorem 2.15 is to show that $P(m_n(\zeta))$ is small for large n. Namely, we establish the following result:

Theorem 9.3. *Fix $w \in B_{\mathbb{C}}(0,1)$ and a constant $C_0 \geq 2$. Fixing any arbitrary positive constant $c_0 < 1/2$ define*

$$\mathscr{S}_{c_0} := \left\{\zeta \in \mathbb{C}^+ \cap B_{\mathbb{C}}(0, 2c_0^{-1}) : \operatorname{Im}\zeta \geq c_0\right\}.$$

Assume $\{\xi_{i,j}\}$ are bounded by K for some $K \geq 1$ and $p = \omega(\frac{\log n}{n})$. Then there exist a constant $C_{9.3}$, depending only on c_0, C_0, K, and R, and an absolute constant $\overline{C}_{9.3}$ such that for any $\zeta \in \mathscr{S}_{c_0}$

$$\mathbb{P}\left(|P(m_n(\zeta))| \geq C_{9.3}\sqrt{\frac{\log n}{np}}\right) \leq n\overline{C}_{9.3}\exp(-C_0\log n).$$

Since there are three roots of the cubic equation $P(m) = 0$, the fact that $P(m_n(\zeta))$ is small does not automatically guarantee that $m_n(\zeta)$ is close to $m_\infty(\zeta)$. We will see later (see Lemma 9.10) that only one of the three roots of the cubic equation $P(m) = 0$ can be the Stieltjes transform of a probability measure. This fact together with Theorem 9.3 finishes the proof of Theorem 2.15.

Now we turn our attention to the proof of Theorem 9.3. A key tool to show that $P(m_n(\zeta))$ will the formula for the inverse of a block matrix.

Lemma 9.4 (Inverse of a block matrix)**.**

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix}^{-1} = \begin{bmatrix} (A-BD^{-1}C)^{-1} & -A^{-1}B(D-CA^{-1}B)^{-1} \\ -D^{-1}C(A-BD^{-1}C)^{-1} & (D-CA^{-1}B)^{-1} \end{bmatrix}.$$

Using Lemma 9.4, one can obtain equations involving the entries of the inverse of $\boldsymbol{A}_n^w - \zeta I_{2n}$ and that of a sub-matrix of $\boldsymbol{A}_n^w - \zeta I_{2n}$. For the ease of writing, we introduce the following notation.

Definition 9.5. Denote

$$\boldsymbol{A}_n^w(\zeta) := \begin{bmatrix} -\zeta I_n & \frac{1}{\sqrt{np}}A_n - wI_n \\ \frac{1}{\sqrt{np}}A_n^* - \bar{w}I_n & -\zeta I_n \end{bmatrix}.$$

Fix a subset $\mathbb{T} \subset [2n]$ and let $\boldsymbol{A}_n^w(\zeta)^{(\mathbb{T})}$ be the sub matrix of $\boldsymbol{A}_n^w(\zeta)$ obtained after removing the columns and the rows of $\boldsymbol{A}_n^w(\zeta)$ indexed by $\mathbb{T}$. Further denote $G_n^{(\mathbb{T})}(\zeta) := (\boldsymbol{A}_n^w(\zeta)^{(\mathbb{T})})^{-1}$, where we suppress the dependence on w. When $\mathbb{T} = \varnothing$, we write $G_n(\zeta)$ instead of $G_n^{(\varnothing)}(\zeta)$. With this notation, we note that $m_n(\zeta) = \frac{1}{2n}\operatorname{Tr} G_n(\zeta)$. For the ease of writing, we abbreviate $G_n^{(\mathbb{T}\cup\{k\})}(\zeta)$

by $G_n^{(k\mathbb{T})}(\zeta)$ for any $k \in [2n]$. Similarly we write $G_n^{(i)}(\zeta)$ and $G_n^{(ij)}(\zeta)$ when $\mathbb{T} = \{i\}$ and $\{i, j\}$ respectively.

Equipped with the above notations, we now state an easy consequence of Lemma 9.4 that allows us to relate the entries of $G_n^{(\mathbb{T})}(\zeta)$ and $G_n^{(k\mathbb{T})}(\zeta)$. Its proof can be found in [19, Lemma 4.2].

Lemma 9.6. *Fix $\mathbb{T} \subset [2n]$ and $k \in [2n]$. For $i, j \neq k$ and $i, j, k \notin \mathbb{T}$ we have*

$$G_n^{(\mathbb{T})}(\zeta)_{i,j} = G_n^{(k\mathbb{T})}(\zeta)_{i,j} + \frac{G_n^{(\mathbb{T})}(\zeta)_{i,k} G_n^{(\mathbb{T})}(\zeta)_{k,j}}{G_n^{(\mathbb{T})}(\zeta)_{k,k}}.$$

In the proof of Theorem 9.3, using Lemma 9.6, we obtain equations involving entries of $G_n^{(\mathbb{T})}(\zeta)$ for certain choices of the index set $\mathbb{T}$. From those equations we then need to identify the negligible and the non-negligible expressions. To do this we use the following concentration inequality.

Lemma 9.7. *(i) Let $\boldsymbol{a} := (\xi_i \delta_i)_{i=1}^n$ be an n-dimensional random vector where $\{\xi_i\}_{i=1}^n$ are i.i.d. with zero mean and unit variance and $\{\delta_i\}_{i=1}^n$ are i.i.d.* $\text{Ber}(p)$*, where $p = \omega(\frac{\log n}{n})$. Also assume that $\{\xi_i\}_{i=1}^n$ are bounded by some constant $K \geq 1$. Fix a matrix $\mathfrak{R}$ and a vector ϑ such that $\|\vartheta\|_2, \|\mathfrak{R}\| \leq R$ for some $R \geq 1$. Then for every $C_0 \geq 1$ there exist a constant $C_{9.7}$, depending only on C_0, K, and R, and an absolute constant $\overline{C}_{9.7}$ such that we have*

$$\mathbb{P}\left(\left|\boldsymbol{a}^* \mathfrak{R} \boldsymbol{a} - p \sum_{i=1}^n \mathfrak{R}_{i,i}\right| \geq C_{9.7} \sqrt{np \log n}\right) \leq \overline{C}_{9.7} \exp(-C_0 \log n) \tag{9.2}$$

and

$$\mathbb{P}\left(|\vartheta^* \boldsymbol{a}| \geq C_{9.7} \sqrt{\log n}\right) \leq \overline{C}_{9.7} \exp(-C_0 \log n). \tag{9.3}$$

(ii) Let $\boldsymbol{a} := (\delta_i - \mathbb{E}\delta_i)_{i=1}^n$ where $\{\delta_i\}_{i=1}^n$ are i.i.d. $\text{Ber}(p)$*. Further let $p, \mathfrak{R}, \vartheta$, and R be as in part (i). Then for every positive constant $C_0 \geq 1$ there exist a constant $C'_{9.7}$, depending only on C_0 and R, and an absolute constant $\overline{C}_{9.7}$ such that*

$$\mathbb{P}\left(\left|\boldsymbol{a}^* \mathfrak{R} \boldsymbol{a} - p(1-p) \sum_{i=1}^n \mathfrak{R}_{i,i}\right| \geq C'_{9.7} \sqrt{np \log n}\right) \leq \overline{C}_{9.7} \exp(-C_0 \log n) \tag{9.4}$$

and

$$\mathbb{P}\left(|\vartheta^* \boldsymbol{a}| \geq C'_{9.7} \sqrt{\log n}\right) \leq \overline{C}_{9.7} \exp(-C_0 \log n). \tag{9.5}$$

The proof of Lemma 9.7 has been motivated from the arguments in [35, pp. 175-177]. Although the arguments there are for the dense case, we show below that the same approach can be taken in the sparse case. The part (ii) of Lemma 9.7 will be used in Section 11 to prove Theorem 1.8.

Before proceeding with the proof, we should mention that Lemma 9.7 continues to hold for any $n' \times n'$ matrix $\mathfrak{R}$ and n'-dimensional vector ϑ as long as $n' \geq c'n$ for some absolute constant c'. This modification only worsens the constants appear in Lemma 9.7. During the proof of Theorem 9.3 we use Lemma 9.7 for $n' = n - 1$ and $n - 2$. In the proof of Lemma 9.7 we stick to the case $n' = n$ for simplicity.

Proof of Lemma 9.7. First let us prove part (i). Then we will outline the required changes to prove part (ii). To this end, using Chernoff's inequality we see that there exist absolute constants C and c such that

$$\mathbb{P}\left(\sum_{i=1}^n \delta_i \geq Cnp\right) \leq \exp(-cnp). \tag{9.6}$$

Therefore using the fact that $\{\xi_i\}_{i=1}^n$ is a sequence of bounded i.i.d. random variables and applying Hoeffding's inequality, we obtain

$$\mathbb{P}\left(\|\boldsymbol{a}\|_2^2 \geq 2Cnp\right) \leq \mathbb{P}\left(\sum_{i=1}^{Cnp} \xi_i^2 \geq 2Cnp\right) + \exp(-cnp) \leq \exp(-Cnp/K^2) + \exp(-cnp).$$

Hence for any non-negative definite matrix $\mathfrak{R}$

$$\mathbb{P}\left(\|\mathfrak{R}^{1/2}\boldsymbol{a}\|_2 \geq \sqrt{2C\,\|\mathfrak{R}\|\,np}\right) \leq \exp(-Cnp/K^2) + \exp(-cnp). \tag{9.7}$$

Note that the above inequality in particular implies that $M(\|\mathfrak{R}^{1/2}\boldsymbol{a}\|_2) \leq \sqrt{2C\|\mathfrak{R}\|np}$, where for any random variable X the notation $M(X)$ denotes the median of its distribution. Now applying Talagrand's concentration inequality for convex Lipschitz functions (see [35, Theorem 2.1.13]) we deduce that for every $\varepsilon > 0$,

$$\mathbb{P}\left(\left|\|\mathfrak{R}^{1/2}\boldsymbol{a}\|_2 - M(\|\mathfrak{R}^{1/2}\boldsymbol{a}\|_2)\right| \geq \|\mathfrak{R}\|^{1/2}K\varepsilon\right) \leq C'\exp(-c'\varepsilon^2), \tag{9.8}$$

where c' and C' are absolute constants. Upon using integration by parts, the above inequality also yields that $|M(\|\mathfrak{R}^{1/2}\boldsymbol{a}\|_2) - \mathbb{E}\|\mathfrak{R}^{1/2}\boldsymbol{a}\|_2| \leq C''$ for some constant C'', depending only on K and R, where we recall that $\|\mathfrak{R}\| \leq R$. Therefore an application of the triangle inequality further shows that

$$\mathbb{P}\left(\left|\|\mathfrak{R}^{1/2}\boldsymbol{a}\|_2 - \mathbb{E}\|\mathfrak{R}^{1/2}\boldsymbol{a}\|_2\right| \geq 2\|\mathfrak{R}\|^{1/2}K\varepsilon\right) \leq C'\exp(-c'\varepsilon^2), \tag{9.9}$$

for all $\varepsilon \geq C''$. Using integration by parts once more we obtain that $\text{Var}(\|\mathfrak{R}^{1/2}\boldsymbol{a}\|_2) \leq 2(C'')^2$ where we enlarge C'', if necessary. Thus

$$\begin{aligned}
&\left|\mathbb{E}(\|\mathfrak{R}^{1/2}\boldsymbol{a}\|_2^2) - M(\|\mathfrak{R}^{1/2}\boldsymbol{a}\|_2^2)\right| \\
&\leq \text{Var}(\|\mathfrak{R}^{1/2}\boldsymbol{a}\|_2) + \left|\left(\mathbb{E}\|\mathfrak{R}^{1/2}\boldsymbol{a}\|_2\right)^2 - \left(M(\|\mathfrak{R}^{1/2}\boldsymbol{a}\|_2)\right)^2\right| \leq 3C''\sqrt{2C\|\mathfrak{R}\|np},
\end{aligned} \tag{9.10}$$

for all large n. Since for any non-negative definite matrix $\mathfrak{R}$ we have $\boldsymbol{a}^*\mathfrak{R}\boldsymbol{a} = \left\|\mathfrak{R}^{1/2}\boldsymbol{a}\right\|_2^2$, from (9.7)-(9.8), and (9.10), upon applying triangle inequality, we further deduce that

$$\begin{aligned}
&\mathbb{P}\left(|\boldsymbol{a}^*\mathfrak{R}\boldsymbol{a} - \mathbb{E}(\boldsymbol{a}^*\mathfrak{R}\boldsymbol{a})| \geq 3\sqrt{2CC_0c'^{-1}}K\|\mathfrak{R}\|\sqrt{np\log n}\right) \\
&\quad\leq \mathbb{P}\left(|\boldsymbol{a}^*\mathfrak{R}\boldsymbol{a} - M(\boldsymbol{a}^*\mathfrak{R}\boldsymbol{a})| \geq 2\sqrt{2CC_0c'^{-1}}K\|\mathfrak{R}\|\sqrt{np\log n}\right) \\
&\quad\leq \mathbb{P}\left(\|\mathfrak{R}^{1/2}\boldsymbol{a}\|_2 + M\left(\|\mathfrak{R}^{1/2}\boldsymbol{a}\|_2\right) \geq 2\sqrt{2C\,\|\mathfrak{R}\|\,np}\right) \\
&\qquad + \mathbb{P}\left(\left|\|\mathfrak{R}^{1/2}\boldsymbol{a}\|_2 - M(\|\mathfrak{R}^{1/2}\boldsymbol{a}\|_2)\right| \geq K\sqrt{\|\mathfrak{R}\|C_0c'^{-1}\log n}\right) \\
&\quad\leq C'\exp(-C_0\log n) + \exp(-Cnp/K^2) + \exp(-cnp),
\end{aligned} \tag{9.11}$$

for all large n. Now recalling that $\mathbb{E}[\xi_i] = 0$, $\mathbb{E}[\xi_i^2] = 1$, $\mathbb{E}[\delta_i] = p$, and the fact that $\{\xi_i\delta_i\}_{i=1}^n$ are i.i.d. we see that $\mathbb{E}[\boldsymbol{a}^*\mathfrak{R}\boldsymbol{a}] = p\sum_{i=1}^n \mathfrak{R}_{i,i}$. Thus using the fact $np = \omega(\log n)$, when $\mathfrak{R}$ is a non-negative definite matrix, the inequality (9.2) follows from (9.11). Then applications of triangle inequality extend (9.2) for the Hermitian case first and then for any $\mathfrak{R}$ with bounded norm. This completes the proof of (9.2). To prove of (9.3) we follow the same lines of approach as above and note that $\mathbb{E}(\vartheta^* \boldsymbol{a}) = 0$. This finishes the proof of part (i).

Next we turn our attention to the proof of part (ii). Since $\|\boldsymbol{a}\|_2^2 \le 2\sum_{i=1}^n (\delta_i^2 + p^2) \le 2\sum_{i=1}^n \delta_i + 2np$. Thus using (9.6) we note that

$$\mathbb{P}\left(\|\boldsymbol{a}\|_2^2 \ge 4Cnp \right) \le \exp(-cnp).$$

Now repeating the remaining steps of the proofs of (9.2) the proof of (9.4) completes. We omit the details. The proof of (9.5) is exactly same as (9.3). ☐

During the proof of Theorem 9.3, we apply Lemma 9.7 to the matrix $\mathfrak{R} = G_n^{(\mathbb{T})}(\zeta)$, with some suitably chosen $\mathbb{T}$, while its rows (columns) are playing the roles of ϑ. So we need bounds on their norms, which are derived in the next two lemmas.

Lemma 9.8. *For any $\mathbb{T} \subset [2n]$ and $k \in [2n]\backslash\mathbb{T}$ both $|G_n^{(\mathbb{T})}(\zeta)_{k,k}|$ and $\|G_n^{(\mathbb{T})}(\zeta)\|$ are bounded by $1/\mathrm{Im}\,\zeta$.*

The proof follows by simply observing that

$$G_n^{(\mathbb{T})}(\zeta) = \sum_{\ell \in [2n]\backslash\mathbb{T}} \frac{1}{\lambda_\ell - \zeta} \boldsymbol{u}_\ell \boldsymbol{u}_\ell^*,$$

where $\{\lambda_\ell\}$, and $\{\boldsymbol{u}_\ell\}$ are the eigenvalues and eigenvectors of $\boldsymbol{A}_n^w(\zeta)^{(\mathbb{T})}$. The next result provides bound on the Euclidean norm of a row (column) of $G_n^{(\mathbb{T})}(\zeta)$.

Lemma 9.9 (Ward identity)**.** *Fix any $\mathbb{T} \in [2n]$. Then for any $i \in [2n]\backslash\mathbb{T}$*

$$\sum_{k \in [2n]\backslash\mathbb{T}} \left| G_n^{(\mathbb{T})}(\zeta)_{k,i} \right|^2 = \sum_{k \in [2n]\backslash\mathbb{T}} \left| G_n^{(\mathbb{T})}(\zeta)_{i,k} \right|^2 = \frac{\mathrm{Im}\,[G_n^{(\mathbb{T})}(\zeta)_{i,i}]}{\mathrm{Im}\,\zeta}.$$

The proof of Ward identity follows from the resolvent identity

$$C^{-1} - D^{-1} = C^{-1}(D - C)D^{-1},$$

applied with $C = \boldsymbol{A}_n^w(\zeta)^{(\mathbb{T})}$ and $D = C^*$. Now we are ready to prove Theorem 9.3.

Proof of Theorem 9.3. Fix any $i \in [n]$. Using Lemma 9.4 we note that

$$G_n(\zeta)_{i,i} = -(\zeta + Z_i)^{-1}, \tag{9.12}$$

where

$$Z_i := \sum_{k,\ell \neq i} a_{i,k} G_n^{(i)}(\zeta)_{k,\ell} a_{\ell,i}$$

and $a_{k,\ell}$ is the (k,ℓ)-th entry of $A_n^w(\zeta)$. We see that Z_i is a quadratic function in the variables $\{a_{i,k}\}_{k \in [2n]\backslash\{i\}}$. Hence, one would like to appeal to Lemma 9.7(i) and show that Z_i must be close to its expectation. However, we cannot directly apply Lemma 9.7(i) because all the entries of $A_n^w(\zeta)$ are not of the form $\xi \cdot \delta$. In particular, for every $i \in [n]$ the entry $a_{i,n+i}$ does not have that product structure, so we need to separate this case. We carry out the details below.

To this end, we first note that for any $i \in [n]$ we have that $a_{i,k} = 0$ for any $k \in [n]$. Therefore the expression of Z_i simplifies to

$$Z_i = \sum_{k,\ell \in [2n]\setminus[n]} a_{i,k} G_n^{(i)}(\zeta)_{k,\ell} a_{\ell,i}. \tag{9.13}$$

For the ease of writing, denote $[n]^{(i)} := [2n]\setminus([n]\cup\{n+i\})$. Recalling that $\|G_n^{(i)}(\zeta)\| \le 1/\mathrm{Im}\,\zeta \le c_0^{-1}$, setting $R = c_0^{-1}$, and applying Lemma 9.7(i), we see that

$$\mathbb{P}\left(\left|\sum_{k,\ell\in[n]^{(i)}} a_{i,k} G_n^{(i)}(\zeta)_{k,\ell} a_{\ell,i} - \frac{1}{n}\sum_{k\in[n]^{(i)}} G_n^{(i)}(\zeta)_{k,k}\right| \ge C_{9.7}\sqrt{\frac{\log n}{np}}\right) \le \overline{C}_{9.7}\exp(-C_0 \log n). \tag{9.14}$$

Using Lemma 9.6 we note

$$\frac{1}{n}\sum_{k\in[n]^{(i)}} G_n^{(i)}(\zeta)_{k,k} = \frac{1}{n}\sum_{k\in[n]^{(i)}} G_n(\zeta)_{k,k} - \frac{1}{n}\sum_{k\in[n]^{(i)}} \frac{G_n(\zeta)_{k,i} G_n(\zeta)_{i,k}}{G_n(\zeta)_{i,i}}.$$

Therefore an application of Ward identity (see Lemma 9.9) and Cauchy-Schwarz inequality yields

$$\left|\frac{1}{n}\sum_{k\in[n]^{(i)}} G_n^{(i)}(\zeta)_{k,k} - \frac{1}{n}\sum_{k\in[n]^{(i)}} G_n(\zeta)_{k,k}\right| \le \frac{1}{n}\cdot\frac{\mathrm{Im}\, G_n(\zeta)_{i,i}}{|G_n(\zeta)_{i,i}|\mathrm{Im}\,\zeta} \le \frac{1}{n\mathrm{Im}\,\zeta}. \tag{9.15}$$

From Lemma 9.4 an easy computation also shows that

$$\frac{1}{n}\sum_{k\in[n]} G_n(\zeta)_{k,k} = \frac{1}{n}\sum_{k\in[2n]\setminus[n]} G_n(\zeta)_{k,k}, \tag{9.16}$$

and therefore recalling that $|G_n(\zeta)_{n+i,n+i}| \le 1/\mathrm{Im}\,\zeta$, from (9.15) and (9.14) we deduce that

$$\mathbb{P}\left(\left|\sum_{k,\ell\in[n]^{(i)}} a_{i,k} G_n^{(i)}(\zeta)_{k,\ell} a_{\ell,i} - m_n(\zeta)\right| \ge 2C_{9.7}\sqrt{\frac{\log n}{np}}\right) \le \overline{C}_{9.7}\exp(-C_0 \log n), \tag{9.17}$$

for all large n.

Next we need to analyze the remaining terms of Z_i. Since $w \in B_{\mathbb{C}}(0,1)$ using the triangle inequality we see that $|a_{i,n+i}| \le 2$ for all large n. Now applying Ward identity again we see

$$\sum_{\ell\in[n]^{(i)}} \left|G_n^{(i)}(\zeta)_{n+i,\ell}\right|^2 \le \frac{\mathrm{Im}\, G_n^{(i)}(\zeta)_{n+i,n+i}}{\mathrm{Im}\,\zeta} \le \frac{1}{(\mathrm{Im}\,\zeta)^2}.$$

So from Lemma 9.7(i) it follows that

$$\mathbb{P}\left(\left|\sum_{\ell\in[n]^{(i)}} a_{i,n+i} G_n^{(i)}(\zeta)_{n+i,\ell} a_{\ell,i}\right| \ge 2C_{9.7}\sqrt{\frac{\log n}{np}}\right) \le \overline{C}_{9.7}\exp(-C_0 \log n). \tag{9.18}$$

A similar argument as above shows that

$$\mathbb{P}\left(\left|\sum_{k\in[n]^{(i)}} a_{i,k} G_n^{(i)}(\zeta)_{k,n+i} a_{n+i,i}\right| \ge 2C_{9.7}\sqrt{\frac{\log n}{np}}\right) \le \overline{C}_{9.7}\exp(-C_0 \log n). \tag{9.19}$$

Using the fact that $\{\xi_{i,j}\}_{i,j=1}^n$ are bounded and recalling that

$$a_{i,n+i} = \frac{1}{\sqrt{np}}\xi_{i,n+i}\delta_{i,n+i} - w \qquad \text{and} \qquad a_{n+i,i} = \frac{1}{\sqrt{np}}\xi_{n+i,i}\delta_{n+i,i} - \bar{w},$$

we note

$$\left| a_{i,n+i}G_n^{(i)}(\zeta)_{n+i,n+i}a_{n+i,i} - |w|^2 G_n^{(i)}(\zeta)_{n+i,n+i} \right| \le \sqrt{\frac{\log n}{np}},$$

for all large n. Thus denoting

$$Z_i^0 := Z_i - m_n(\zeta) - |w|^2 G_n^{(i)}(\zeta)_{n+i,n+i},$$

from (9.12)-(9.13) and (9.17)-(9.19) we deduce

$$G_n(\zeta)_{i,i} = -(\zeta + m_n(\zeta) + |w|^2 G_n^{(i)}(\zeta)_{n+i,n+i} + Z_i^0)^{-1}, \tag{9.20}$$

where

$$\mathbb{P}\left(|Z_i^0| \ge 7C_{9.7}\sqrt{\frac{\log n}{np}} \right) \le 3\overline{C}_{9.7}\exp(-C_0 \log n), \tag{9.21}$$

for all large n.

Equation (9.20) together with inequality (9.21) is the first step towards obtaining an approximate fixed point equation for $m_n(\zeta)$. Next our goal would be to show that $G_n^{(i)}(\zeta)_{n+i,n+i}$ is close to a certain function of $m_n(\zeta)$ with high probability. This would enable us to approximate $G_n^{(i)}(\zeta)_{n+i,n+i}$ by that function of $m_n(\zeta)$ and derive the desired equation for $m_n(\zeta)$.

To control the behavior of $G_n^{(i)}(\zeta)_{n+i,n+i}$, we apply the formula for the inverse of the block matrix again. This yields

$$G_n^{(i)}(\zeta)_{n+i,n+i} = -(\zeta + \tilde{Z}_i)^{-1},$$

where

$$\tilde{Z}_i := \sum_{k,\ell \ne n+i,i} a_{n+i,k}G_n^{(i,n+i)}(\zeta)_{k,\ell}a_{\ell,n+i} = \sum_{k,\ell \in [n]\setminus\{i\}} a_{n+i,k}G_n^{(i,n+i)}(\zeta)_{k,\ell}a_{\ell,n+i},$$

and in the last step above we have again used the fact that $a_{k,n+i} = a_{n+i,k} = 0$ for $k \in [2n]\setminus[n]$. Note that for every $k \in [n]\setminus\{i\}$ the random variable $a_{n+i,k}$ has the desired product structure. Hence, an application of Lemma 9.7(i) now shows that

$$\mathbb{P}\left(\left| \tilde{Z}_i - \frac{1}{n}\sum_{k\in[n]\setminus\{i\}} G_n^{(i,n+i)}(\zeta)_{k,k} \right| \ge C_{9.7}\sqrt{\frac{\log n}{np}} \right) \le \overline{C}_{9.7}\exp(-C_0 \log n).$$

Therefore, arguing similarly as in (9.15)-(9.16) and denoting

$$\tilde{Z}_i^0 := \tilde{Z}_i - m_n(\zeta),$$

we conclude

$$G_n^{(i)}(\zeta)_{n+i,n+i} = -(\zeta + m_n(\zeta) + \tilde{Z}_i^0)^{-1}, \tag{9.22}$$

where

$$\mathbb{P}\left(|\tilde{Z}_i^0| \ge 2C_{9.7}\sqrt{\frac{\log n}{np}} \right) \le \overline{C}_{9.7}\exp(-C_0 \log n), \tag{9.23}$$

for all large n. So combining (9.20) and (9.22) we see that

$$G_n(\zeta)_{i,i} = -\frac{1}{\zeta + m_n(\zeta) - \frac{|w|^2}{\zeta+m_n(\zeta)+\tilde{Z}_i^0} + Z_i^0}. \tag{9.24}$$

Since $\operatorname{Im} m_n(\zeta) \geq 0$ for any $\zeta \in \mathbb{C}^+$ we have $|\zeta + m_n(\zeta)| \geq \operatorname{Im} \zeta \geq c_0$. Thus for any $i \in [n]$

$$G_n(\zeta)_{i,i} = -\frac{1}{\zeta + m_n(\zeta) - \frac{|w|^2}{\zeta+m_n(\zeta)} + \hat{Z}_i^0}, \tag{9.25}$$

for some $\hat{Z}_i^0$ such that

$$\mathbb{P}\left(|\hat{Z}_i^0| \geq C\sqrt{\frac{\log n}{np}}\right) \leq 4\overline{C}_{9.7} \exp(-C_0 \log n), \tag{9.26}$$

for some large constant C. By a similar argument (9.25)-(9.26) holds for any $i \in [2n]\backslash[n]$. We omit the details.

Next we claim that

$$\left|\zeta + m_n(\zeta) - \frac{|w|^2}{\zeta + m_n(\zeta)}\right| \geq c_0/2.$$

If not, then together with (9.25)-(9.26) it implies that $|G_n(\zeta)_{i,i}| \geq \frac{3}{2}c_0^{-1}$, for all large n, on a set with high probability. This yields a contradiction to Lemma 9.8. Therefore, from (9.25)-(9.26) again, we deduce that for every $i \in [2n]$,

$$G_n(\zeta)_{i,i} = -\frac{1}{\zeta + m_n(\zeta) - \frac{|w|^2}{\zeta+m_n(\zeta)}} + \mathcal{E}_i^0, \tag{9.27}$$

for some $\mathcal{E}_i^0$ such that

$$\mathbb{P}\left(|\mathcal{E}_i^0| \geq C'\sqrt{\frac{\log n}{np}}\right) \leq 4\overline{C}_{9.7} \exp(-C_0 \log n),$$

for some other constant C'. Next recalling that $w \in B_{\mathbb{C}}(0,1)$ and $|\zeta + m_n(\zeta)|$ is bounded on $\mathscr{S}_{c_0}$, the proof of the theorem finally completes from (9.27) by taking an average over $i \in [2n]$ and rearranging the terms. ☐

As we have already mentioned Theorem 9.3 alone cannot prove Theorem 2.15. To complete the proof of Theorem 2.15 we need the following uniqueness property of the limit $m_\infty(\zeta)$, which is borrowed from [7].

Lemma 9.10 ([7, Lemma 8.5])**.** *There exists an absolute constant $C_{9.10}$ such that for every $\xi \in \mathbb{C}^+$ with $\operatorname{Im} \zeta \geq C_{9.10}$ there is one and only one root $m(\zeta) := m(\zeta, z)$ of the cubic equation $P(m) = 0$ such that $|m(\zeta)| \leq 2/\operatorname{Im} \zeta$.*

We are now ready to finish the proof of Theorem 2.15.

Proof of Theorem 2.15. Recall that the limit $m_\infty(\zeta)$ is a root of the equation $P(m) = 0$. Let $m_1(\zeta)$ and $m_2(\zeta)$ denote the other two roots of the same equation. Therefore

$$P(m_n(\zeta)) = (m_n(\zeta) - m_\infty(\zeta))(m_n(\zeta) - m_1(\zeta))(m_n(\zeta) - m_2(\zeta)).$$

From Lemma 9.8 and Lemma 9.10, and the triangle inequality, we note

$$|m_n(\zeta) - m_j(\zeta)| \ge 1/\mathrm{Im}\,\zeta, \qquad \text{for } j = 1, 2.$$

Hence, upon choosing c_0 in Theorem 9.3 such that $c_0^{-1} \ge 2C_{9.10}$, from Theorem 9.3 it follows that

$$\mathbb{P}\left(|(m_n(\zeta) - m_\infty(\zeta)| \ge C\sqrt{\frac{\log n}{np}} \right) \le n\overline{C}_{9.3} \exp(-C_0 \log n),$$

for any $\zeta \in \mathscr{T}_{C_{9.10}}$, where C is some large constant. Choosing $C_0 \ge 3$ the proof completes by an application of Borel-Cantelli lemma. □

Proof of Theorem 2.13 is now immediate.

Proof of Theorem 2.13. When $\{\xi_{i,j}\}$ are bounded, the weak convergence of ν_n^w to ν_∞^w follows from Theorem 2.15 and Proposition 2.16. Then the boundedness assumption is removed using Lemma 9.1. □

10. Proof of Theorem 1.3

In this section we combine Theorem 2.2, Theorem 2.12, and Theorem 2.13 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 $\mathbb{D}_r := \{w \in B_{\mathbb{C}}(0, 1-r) : |\mathrm{Im}\, w| \ge r\}$. 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 $\frac{1}{\sqrt{np}} A_n$. Afterwards 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 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*

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

(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 $w \in B_{\mathbb{C}}(0, 1-r) \backslash B_{\mathbb{C}}(0, r)$ we have

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

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

Proof. In [4, 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 $t > 0$, $\widetilde{\nu}_\infty^w\left((0,t^2)\right) = \nu_\infty^w((-t,t))$. 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 $0 < \tau < 1$,

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

Using [7, Lemma 8.4(i)] and [7, Lemma 10.3] we see that for any $t \in (0,1)$,

$$\nu_\infty^w((0,t)) \le \nu_\infty^w((-t,t)) \le 2t \cdot (\operatorname{Im} m_\infty(it)) \le 2Ct,$$

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 $\mathbb{D}_r := \{w \in B_{\mathbb{C}}(0, 1-r) : |\operatorname{Im} w| \ge r\}$. Let us also fix $w \in \mathbb{D}_r$. Define

$$\Omega_n' := \left\{ s_{\min}\left(\frac{A_n}{\sqrt{np}} - wI_n\right) \ge c_n \right\}, \qquad \text{where } c_n := c_{2.2} \exp\left(-C_{2.2}\frac{\log(1/p)}{\log(np)}\right)\frac{1}{n\sqrt{np}}.$$

Setting $D_n := \operatorname{diag}(A_n) - w\sqrt{np}I_n$, conditioning on $\operatorname{diag}(A_n)$, the diagonal of A_n, applying Theorem 2.2, and then taking an average over $\operatorname{diag}(A_n)$ we deduce that

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

Fix any $\delta \in (0,1)$ and let $\tau := \tau(\delta) := c_{2.12}\delta$. 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 notations we recall the definition of ν_n^w to see that

$$\int_{-\tau}^{\tau} |\log(|x|)| d\nu_n^w(x) = \frac{1}{n}\sum_{i=1}^{n-3\psi(n)} |\log(s_i)|\mathbb{I}(s_i \le \tau) + \frac{1}{n}\sum_{i=n-3\psi(n)+1}^{n} |\log(s_i)|\mathbb{I}(s_i \le \tau). \tag{10.3}$$

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

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

We next consider the first term of (10.3). Since $\min\{p\psi(n), \psi^2(n)/n\} \ge C_{2.12}\log n$ we therefore deduce from Theorem 2.12 that on an event Ω_n'' we have

$$\begin{aligned}
\frac{1}{n}\sum_{i=1}^{n-3\psi(n)} |\log(s_i)|\mathbb{I}(s_i \le \tau) &= \frac{1}{n}\sum_{i=3\psi(n)}^{n-1} |\log(s_{n-i})|\mathbb{I}(s_{n-i} \le \tau) \\
&\le \frac{\log(1/c_{2.12})}{n}\sum_{i=3\psi(n)}^{n-1} \mathbb{I}(s_{n-i} \le \tau) + \frac{1}{n}\sum_{i=3\psi(n)}^{n-1} \log\left(\frac{n}{i}\right)\mathbb{I}(s_{n-i} \le \tau). \qquad (10.5)
\end{aligned}$$

where $\mathbb{P}((\Omega_n'')^c) \le 2/n^2$. Recalling the definition of τ, from Theorem 2.12 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

$$\frac{1}{n}\sum_{i=1}^{n-3\psi(n)} |\log(s_i)|\mathbb{I}(s_i \le \tau) \le \delta \cdot \log(1/c_{2.12}) + \frac{1}{n}\sum_{i=3\psi(n)}^{\delta n} \log\left(\frac{n}{i}\right)$$

$$\le \delta \cdot \log(1/c_{2.12}) - 2\int_0^\delta \log x\, dx, \tag{10.6}$$

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

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

for all large n, on the event Ω_n, where $\kappa(\delta) := 2\delta\cdot\log(1/c_{2.12}) - 2\int_0^\delta \log x\, dx$. 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

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

Next noting that $\log(\cdot)$ is a bounded function on a compact interval that is bounded away from zero, we apply Theorem 2.13 to deduce that

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

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, using that $\log|x| \le |x|^2$ for $|x| > e^{1/2}$ and choosing $R > e^{1/2}$, we have

$$\int_{(-R,R)^c} |\log|x|| d\nu_n^w(x) \le \frac{\log R}{R^2}\int_{(-R,R)^c} |x|^2 d\nu_n^w(x) \le \frac{2\log R}{R^2}\left[\frac{\sum_{i,j=1}^n a_{i,j}^2}{n^2 p} + |w|^2\right].$$

The weak law of large numbers for a triangular array (see [17, Theorem 2.2.6]) implies that $\sum_{i,j=1}^n a_{i,j}^2/(n^2p) \to 1$ in probability.

Therefore, given any $\kappa > 0$ there exists $R_\kappa := R(\kappa)$ sufficiently large such that

$$\limsup_{n\to\infty} \mathbb{P}\left(\left|\int_{(-R_\kappa,R_\kappa)^c} |\log|x|| d\nu_n^w(x) - \int_{(-R_\kappa,R_\kappa)^c} |\log|x|| d\nu_\infty^z(x)\right| > \kappa\right) = 0 \tag{10.9}$$

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

$$\frac{1}{n}\log|\det(A_n/\sqrt{np} - wI_n)| = \int_{-\infty}^{\infty} \log|x| d\nu_n^w(x) \to \int_{-\infty}^{\infty} \log|x| d\nu_\infty^w(x), \text{ in probability}. \tag{10.11}$$

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,

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

For example, one can obtain a proof of (10.12) using [13, Lemma 4.11, Lemma 4.12], [14, Theorem 3.4], and [30, 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. [37, Theorem 1.13]), we obtain that for every $r > 0$ and every $f_r \in C_c^2(\mathbb{C})$, supported on $\mathbb{D}_r$,

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

To finish 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.

To this end, fix any arbitrary sequence $\{r_m\}$ such that $r_m \downarrow 0$ as $m \to \infty$. Also fix a $f \in C_c^2(\mathbb{C})$. Define a function $i_r \in C_c^2(\mathbb{C})$ such that i_r is supported on $\mathbb{D}_r$ and $i_r \equiv 1$ on $\mathbb{D}_{2r}$. Denote $f_{r_m} := f i_{r_m}$.

Recall that if a sequence of random variables converges in probability then given any subsequence there exists a further subsequence such that almost sure convergence hold along the latter subsequence. Therefore we can find a subsequence $\{n_k^{(1)}\}$ along which (10.13) holds almost surely for f_{r_1} and i_{r_1}. This, in particular implies that

$$\begin{aligned} \limsup_{k\to\infty} \int (1 - i_{r_1}(w)) dL_{n_k^{(1)}}(w) = \int (1 - i_{r_1}(w)) d\mathfrak{m}(w) \; &\le \; 1 - \frac{1}{\pi} \int_{\mathbb{D}_{2r_1}} d\mathfrak{m}(w) \\ &\le \; 4r_1(1 - r_1) + \frac{8r_1}{\pi} \le 8r_1, \end{aligned} \tag{10.14}$$

almost surely. Thus writing $f = f_{r_1} + (1 - i_{r_1}) f$ and assuming that $\|f\|_\infty \le K$, by the triangle inequality we also obtain

$$\limsup_{k\to\infty} \Big| \int f(w) dL_{n_k^{(1)}}(w) - \frac{1}{\pi} \int_{B_{\mathbb{C}}(0,1)} f(w) d\mathfrak{m}(w) \Big| \le 16 K r_1, \text{ almost surely.} \tag{10.15}$$

Repeating the same argument we can find a further subsequence $\{n_k^{(2)}\} \subset \{n_k^{(1)}\}$ such that (10.15) holds for the subsequence $\{n_k^{(2)}\}$ and $r = r_2$. Proceeding by induction we therefore deduce that for any $m > 0$ there exists a subsequence $\{n_k^{(m)}\} \subset \{n_k^{(m-1)}\}$ such that

$$\limsup_{k\to\infty} \Big| \int f(w) dL_{n_k^{(m)}}(w) - \frac{1}{\pi} \int_{B_{\mathbb{C}}(0,1)} f(w) d\mathfrak{m}(w) \Big| \le 16 K r_m, \text{ almost surely.} \tag{10.16}$$

Since $r_m \downarrow 0$ as $m \to \infty$, proceeding along the diagonal subsequence $\{n_m^{(m)}\}$ we further obtain that

$$\lim_{m\to\infty} \Big| \int f(w) dL_{n_m^{(m)}}(w) - \frac{1}{\pi} \int_{B_{\mathbb{C}}(0,1)} f(w) d\mathfrak{m}(w) \Big| = 0, \text{ almost surely.}$$

The above argument also shows that given any subsequence $\{n_k\}$ there exists a further subsequence $\{n_{k_m}\}$ such that $\int f(w) dL_{n_{k_m}}(w) \to \frac{1}{\pi} \int_{B_{\mathbb{C}}(0,1)} f(w) d\mathfrak{m}(w)$, as $m \to \infty$, almost surely, for every $f \in C_c^2(\mathbb{C})$. Hence,

$$\int f(w) dL_n(w) \to \frac{1}{\pi} \int_{B_{\mathbb{C}}(0,1)} f(w) d\mathfrak{m}(w), \qquad \text{in probability.}$$

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

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

where

$$\tilde{\Omega}_n' := \left\{ s_{\min}\left(\frac{A_n}{\sqrt{np}} - wI_n\right) \geq \tilde{c}_n \right\}, \quad \text{and } \tilde{c}_n := c_{2.2} \exp\left(-C_{2.2}\frac{\log(1/p_n)}{\log(np_n)}\right) \cdot n^{-3}.$$

Therefore, proceeding similarly as above, applying Borel-Cantelli lemma, and using Theorem 2.13(ii) we see that the conclusions of (10.7)-(10.8) hold almost surely. To show that (10.9) holds almost surely one needs to use a strong law large number for a triangular array. This can be obtained using Markov inequality and Borel-Cantelli lemma. We omit further details.

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 $f_r \in C_c^2(\mathbb{C})$, supported on $\mathbb{D}_r$,

$$\int f_r(w) dL_n(w) \to \frac{1}{\pi} \int f_r(w) d\mathfrak{m}(w), \text{ almost surely.} \tag{10.17}$$

Hence, arguing similarly as in (10.15) we also deduce that for every $f \in C_c^2(\mathbb{C})$, with $\|f\|_\infty \leq K$, and any $r \in (0, 1)$,

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

Since r is arbitrary the proof follows. This completes the proof of the theorem. □

11. Proof of Theorem 1.8

In this section we provide the proof of Theorem 1.8. It follows from a simple adaptation of the proof of Theorem 1.3. Recall the three key ingredients to the proof of Theorem 1.3 are Theorem 2.2, Theorem 2.12, and Theorem 2.13. So we need to find analogues of these three theorems.

First let us state the analogue of Theorem 2.13.

Theorem 11.1. *Let* Adj_n *be the adjacency matrix of a directed Erdős-Rényi graph with edge connectivity probability* p*. Fix* $w \in B_{\mathbb{C}}(0,1)$ *and denote* $\widehat{\nu}_n^w$ *to be the* ESD *of* $\mathbf{Adj}_n^w$ *where*

$$\mathbf{Adj}_n^w := \begin{bmatrix} 0 & \frac{1}{\sqrt{np(1-p)}}\mathrm{Adj}_n - wI_n \\ \frac{1}{\sqrt{np(1-p)}}\mathrm{Adj}_n^* - \bar{w}I_n & 0 \end{bmatrix}. \tag{11.1}$$

Further denote $\bar{p} := \min\{p, 1-p\}$*. If* $\bar{p} = \omega(\frac{\log n}{n})$ *then* $\widehat{\nu}_n^w$ *converges weakly to* ν_∞^w*, almost surely.*

Proof. Let $\widetilde{\mathrm{Adj}}_n$ denote the matrix of i.i.d. $\mathrm{Ber}(p)$ entries. That is, unlike the matrix Adj_n, the diagonal entries of $\widetilde{\mathrm{Adj}}_n$ are also Bernoulli random variables. Define $\widetilde{\mathbf{Adj}}_n^w$ similarly as in (11.1). Using Hoffman-Wielandt inequality (see [3, Lemma 2.1.19]) we see that it is enough to prove that the ESD of $\widetilde{\mathbf{Adj}}_n^w$ converges weakly to ν_∞^w, almost surely. Hence, for the purpose of this proof we can assume that Adj_n is a matrix of i.i.d. $\mathrm{Ber}(p)$ entries. Also using the rank inequality (see [5, Lemma 2.2]) we see that, without loss of generality, we can assume that $p \leq 1/2$ and so $\bar{p} = p$.

Now recall that the key to the proof of Theorem 2.13 is Theorem 9.3, where the proof of the latter theorem is based on the formula for the inverse of a block matrix and Talagrand's concentration inequality. We already derived relevant concentration inequalities for the Bernoulli set-up in Lemma

9.7(ii). Using Lemma 9.7(ii) one can proceed exactly same as in the proof of Theorem 9.3 to derive an analogue of Theorem 9.3 in the current set-up, from which the proof of the current theorem follows. We omit the tedious details. □

Turning to prove an analogue of Theorem 2.12 for Adj_n we recall that the key to the proof of Theorem 2.12 is Lemma 8.1 where the latter is borrowed from [12]. Though the current set-up of Bernoulli random variables does not directly fit the framework of Lemma 8.1, one can repeat its proof to deduce that

$$\mathbb{P}\left(\operatorname{dist}(\boldsymbol{a}_0, H) \le \frac{1}{2}\sqrt{p(1-p)(n-\dim(H))}\right) \le \exp(-cp(1-p)\psi(n)), \tag{11.2}$$

for any sub-space H with $1 \le \dim(H) \le n-\psi(n)$, where $\boldsymbol{a}_0 := (\delta_i - \mathbb{E}\delta_i)_{i=1}^n$, $\{\delta_i\}_{i=1}^n$ are i.i.d. $\mathrm{Ber}(p)$, and $\psi : \mathbb{N} \mapsto \mathbb{N}$ be such that $\psi(n) \to \infty$ and $C\sqrt{n/(p(1-p))} \le \psi(n) < n$ for some large absolute constant C and c is some small absolute constant.

Next note that $\operatorname{dist}(\boldsymbol{a}, H) \le \operatorname{dist}(\boldsymbol{a}_0, \operatorname{span}\{H, \mathbf{1}\})$ where $\boldsymbol{a} := (\delta_i)_{i=1}^n$ and $\mathbf{1}$ is the vector of ones. Using (11.2) and proceeding same as in the proof of Theorem 2.12 we derive the following result.

Theorem 11.2. *Let Adj_n be the adjacency matrix of a directed Erdős-Rényi graph with edge connectivity probability p. There exist absolute constants $C_{11.2}$ and $c_{11.2}$ such that the following holds: Let $\psi : \mathbb{N} \mapsto \mathbb{N}$ be such that $C_{11.2}\sqrt{n/(p(1-p))} \le \psi(n) < n$. Then for any $w \in B_{\mathbb{C}}(0,1)$ we have*

$$\mathbb{P}\left(\bigcup_{i=3\psi(n)}^{n-1}\left\{s_{n-i}\left(\frac{\mathrm{Adj}_n}{\sqrt{np}} - wI_n\right) \le c_{11.2}\frac{i}{n}\right\}\right) \le \frac{2}{n^2}.$$

Now it remains to find analogue of Theorem 2.2. Similar to [8] we first extend Theorem 2.2 to accommodate non-zero mean. This is done in the result below.

Theorem 11.3. *Let $\bar{A}_n$ be an $n \times n$ matrix with zero diagonal and i.i.d. off-diagonal 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 $p \in (0,1]$, and $\xi_{i,j}$, $i,j \in [n]$ are i.i.d. sub-Gaussian (not necessarily centered) with unit variance. Let $R \ge 1$, $r \in (0,1]$ and let D_n be a diagonal matrix such that $\|D_n\| \le R\sqrt{np}$ and $\mathrm{Im}(D_n) = r'\sqrt{np}I_n$ for some r' with $|r'| \in [r,1]$. Then there exists constants $0 < c_{11.3}, \bar{c}_{11.3}, c'_{11.3}, C_{11.3}, C'_{11.3}, \overline{C}_{11.3} < \infty$, depending only on R, r, and the sub-Gaussian moment of $\{\xi_{i,j}\}$, such that for any $\varepsilon > 0$ we have the following:*

(i) If

$$p \ge \frac{\overline{C}_{11.3}\log n}{n},$$

then

$$\mathbb{P}\left(s_{\min}(\bar{A}_n + D_n) \le c_{11.3}\varepsilon \exp\left(-C_{11.3}\frac{\log(1/p)}{\log(np)}\right)\sqrt{\frac{p}{n}}\right) \le \varepsilon + \frac{C'_{11.3}}{\sqrt{np}}.$$

(ii) Additionally, if

$$\log(1/p) < \bar{c}_{11.3}(\log np)^2,$$

then

$$\mathbb{P}\left(s_{\min}(\bar{A}_n + D_n) \le c_{11.3}\varepsilon \exp\left(-C_{11.3}\frac{\log(1/p)}{\log(np)}\right)\sqrt{\frac{p}{n}}\right) \le \varepsilon + \exp(-c'_{11.3}\sqrt{np}).$$

Proof. The proof follows from a straightforward adaptation of that of Theorem 2.2. We outline the necessary changes.

Recall that to find bounds on $s_{\min}(\cdot)$ we need to treat the compressible and dominated vectors first and then the rest. To find bounds on the infimum over compressible and dominated vectors we adapt the proof of [8, Theorem 7.1]. Note that a key to the proof of of [8, Theorem 7.1] is [8, Proposition 7.3]. One can check that [8, Proposition 7.3] holds also for complex valued diagonal matrix D_n. Therefore, repeating the same steps as in the proof of [8, Theorem 7.1] we deduce that Proposition 3.4 continue to hold for the non-centered case. This takes care of the infimum over compressible and dominated vectors.

Now it remains to obtain bounds on the infimum over vectors that are neither compressible nor dominated. Since (7.4) follows from Berry-Esséen Theorem and the Lévy concentration function is invariant under translation we deduce that (7.4) continues to hold in the non-centered set-up. This finishes the proof of part (i).

To prove part (ii) we need to make appropriate changes to the results of Section 3, Section 5, and Section 6. Beginning with Section 3 we recall that the main result there is Proposition 3.1. Turning to derive an analogue of Proposition 3.1 we see that, as $\{\xi_{i,j}\}$ are non-centered, we do not have $\|\bar{A}_n\| \le K\sqrt{np}$. Instead we have that $\|\bar{A}_n - \mathbb{E}\bar{A}_n\| \le K\sqrt{np}$ with high probability, for large enough K. So in the analogue of Proposition 3.1 we should replace the condition $\|\bar{A}_n\| \le K\sqrt{np}$ by $\|\bar{A}_n - \mathbb{E}\bar{A}_n\| \le K\sqrt{np}$ in (3.2). Hence we need to modify the steps where bounds on $\|\bar{A}_n\|$ have been used.

With the above in mind and noting that $\mathbb{E}(\mathrm{Re}\,(B^D)) = \mathfrak{u}p(\boldsymbol{U}_n - I_n)$ where $\boldsymbol{U}_n$ is the $(n-1)\times n$ matrix of all ones and $\mathfrak{u} = \mathbb{E}[\xi_{i,j}]$, we see that (3.5) changes to

$$\left\|\mathrm{Im}\,(B^D)y - (\mathrm{Re}\,(B^D) - \mathbb{E}(\mathrm{Re}\,(B^D)))|_J x_{[1:M]} - \mathfrak{u}p\boldsymbol{U}_n x\right\|_2 \le 3c\rho\sqrt{np},$$

where we have used the fact that $\frac{1}{\rho} = O(n^{o(1)}) = o(\sqrt{n})$, and $c, M, J, \rho, x,$ and y as in (3.5). Proceeding same as in the proof of (3.6) we obtain that for some $\gamma \in \mathbb{R}$ with $|\gamma| \le \mathfrak{u}/r$,

$$\left\|y|_{[2:n]} - (\mathrm{Re}\,(B^D) - \mathbb{E}(\mathrm{Re}\,(B^D)))|_J x_{[1:M]} - \gamma\mathbf{1}\right\|_2 \le 3r^{-1}c\rho.$$

This necessitates to define $\widetilde{\mathscr{E}}_J := \mathrm{span}\{\mathscr{E}_J, \mathbf{1}_J\}$ where now $\mathscr{E}'_J := \mathrm{span}\{(\mathrm{Re}\,(B^D) - \mathbb{E}(\mathrm{Re}\,(B^D)))|_J\mathbb{R}^J\}$ and $\mathscr{E}_J$ is as in (3.7). Since $\dim(\widetilde{\mathscr{E}}_J) \le \dim(\mathscr{E}_J) + 1$, repeating the rest of the proof of Lemma 3.3 we see that it continues to hold in this set-up.

Now we move on to find an analogue of Proposition 5.2. Recall that the proof of Proposition 5.2 relies on finding bounds on Lévy concentration function, which as already noted is invariant under translation. Thus the proof of Proposition 5.2 extends in the non-centered set-up. We follow the ideas of the proofs of [8, Theorem 7.1] and [8, Proposition 7.2].

Since (5.20) is obtained using bounds on Lévy concentration function and the cardinality of the net of genuinely complex vectors, we see that (5.20) extends to the following:

$$\mathbb{P}\left(\exists w \in \mathcal{N}(D,\Delta,d)\ \ \left\|(B^D - \mathfrak{u}p\boldsymbol{U}_n)w - v\right\|_2 \le \frac{\varepsilon_0}{3}\|w_{\mathrm{small}}\|_2 \cdot \sqrt{pn}\right) \le \exp(-2n),$$

where v is any fixed vector. Recalling the definition of ε_0 (see (5.16)) we see that the set of vectors $\{\gamma\mathbf{1}, |\gamma| \le \mathfrak{u}\sqrt{np}\}$ admits a $\frac{\varepsilon_0\rho\sqrt{np}}{6}$-net of cardinality at most $O(\exp(\sqrt{np}))$. So using an union bound and the triangle inequality we also see that

$$\mathbb{P}\left(\inf_{w\in\mathcal{N}(D,\Delta,d), |\gamma|\le\mathfrak{u}np}\ \left\|(B^D - \mathfrak{u}p\boldsymbol{U}_n)w - \gamma\mathbf{1}\right\|_2 \le \frac{\varepsilon_0}{6}\|w_{\mathrm{small}}\|_2 \cdot \sqrt{pn}\right) \le \exp(-(3/2)n).$$

Since $\|B^D - \mathfrak{u}p\boldsymbol{U}_n\|_2 = O(\sqrt{np})$ and $\boldsymbol{U}_n z = \gamma \mathbf{1}$ for any $z \in S_{\mathbb{C}}^{n-1}$ with $|\gamma| \le \sqrt{n}$, the rest of the argument follows similarly as before. This yields the analogue of Proposition 5.2. The proof of the analogue of Proposition 6.4 is similar. This finishes the proof of Theorem 11.3. ☐

Building on Theorem 11.3 we derive the following result on the smallest singular value of Adj_n.

Theorem 11.4. *Let* Adj_n *be the adjacency matrix of a directed Erdős-Rényi graph, with edge connectivity probability* $p \in (0,1)$. *Denote* $\bar{p} := \min\{p, 1-p\}$. *Fix* $w \in B_{\mathbb{C}}(0,1)$ *such that* $|\mathrm{Im}\, w| \ge r$ *for some* $r \in (0,1)$. *Then there exists constants* $0 < c_{11.4}, \bar{c}_{11.4}, c'_{11.4}, C_{11.4}, C'_{11.4}, \overline{C}_{11.4} < \infty$, *depending only on* r, *such that for any* $\varepsilon > 0$ *we have the following:*

(i) If

$$\bar{p} \ge \frac{\overline{C}_{11.4} \log n}{n},$$

then

$$\mathbb{P}\left(s_{\min}(\mathrm{Adj}_n - w\sqrt{np(1-p)} I_n) \le c_{11.4} \varepsilon \exp\left(-C_{11.4} \frac{\log(1/\bar{p})}{\log(n\bar{p})} \right) \sqrt{\frac{\bar{p}}{n}} \right) \le \varepsilon + \frac{C'_{11.4}}{\sqrt{n\bar{p}}}.$$

(ii) Additionally, if

$$\log(1/\bar{p}) < \bar{c}_{11.4} (\log n\bar{p})^2,$$

then

$$\mathbb{P}\left(s_{\min}(\mathrm{Adj}_n - w\sqrt{np(1-p)} I_n) \le c_{11.4} \varepsilon \exp\left(-C_{11.4} \frac{\log(1/\bar{p})}{\log(n\bar{p})} \right) \sqrt{\frac{\bar{p}}{n}} \right) \le \varepsilon + \exp(-c'_{11.4}\sqrt{n\bar{p}}).$$

Fix $p \in (0, 1/2]$. Then a $\mathrm{Ber}(p)$ can be written as the product of a $\mathrm{Ber}(2p)$ random variable and a $\mathrm{Ber}(1/2)$ random variables. So, Theorem 11.4 is a direct consequence of Theorem 11.3. Proof of the case $p \in (1/2, 1)$ is similar but requires an additional ε-net argument. See [8, Theorem 1.11] for more details.

Now the proof of Theorem 1.8 finishes from Theorem 11.1, Theorem 11.2, and Theorem 11.4.

Acknowledgements. A.B. thanks Amir Dembo and Ofer Zeitouni for suggesting the problem and helpful conversations.

References

[1] R. Adamczak. On the Marchenko-Pastur and circular laws for some classes of random matrices with dependent entries. *Electronic Journal of Probabability*, 16(37), 1068–1095, 2011.

[2] R. Adamczak, D. Chafaï, and P. Wolff. Circular law for random matrices with exchangeable entries. *Random Structures & Algorithms*, 48, 454–479, 2016.

[3] G. W. Anderson, A. Guionnet, and O. Zeitouni. *An introduction to random matrices.* No. 118. Cambridge University Press, 2010.

[4] Z. D. Bai. Circular law. *The Annals of Probability*, 25(1), 494–529, 1997.

[5] Z. D. Bai. Methodologies in spectral analysis of large dimensional random matrices, a review. *Statistica Sinica*, 9, 611–677, 1999.

[6] Z. D. Bai and J. W. Silverstein. *Spectral analysis of large dimensional random matrices*, volume 20. Springer, 2010.

[7] A. Basak, N. Cook, and O. Zeitouni. Circular law for sum of permutations. *Preprint*, 2017. Available at https://arxiv.org/abs/1705.09053.

[8] A. Basak and M. Rudelson. Invertibility of sparse non-Hermitian matrices. *Advances in Mathematics*, 310, 426–483, 2017.

[9] A. Basak and M. Rudelson. The local circular law for sparse non-Hermitian matrices. *In prepration*, 2017.

[10] F. Benaych-Georges and A. Knowles. Lectures on the local semicircle law for Wigner matrices. *To appear in SMF volume Panoramas et Synthéses*, 2016.

[11] C. Bordenave, P. Caputo, and D. Chafaï. Circular Law Theorem for Random Markov Matrices. *Probability Theory and Related Fields*, 152(3-4), 751–779, 2012.

[12] C. Bordenave, P. Caputo, and D. Chafaï. Spectrum of Markov generators on sparse random graphs. *Communications in Pure and Applied Mathematics*, 67(4), 621–669, 2014.

[13] C. Bordenave and D. Chafaï. Around the circular law. *Probability Surveys*, 9(0), 1–89, 2012.

[14] P. Bourgade, H.-T. Yau, and J. Yin. Local circular law for random matrices. *Probability Theory and Related Fields*, 159(3), 545–595, 2014.

[15] W. Bryc, A. Dembo, and T. Jiang. Spectral measure of large random Hankel, Markov and Toeplitz matrices. *Annals of Probability*, 34(1), 1–38, 2006.

[16] N. Cook. The circular law for random regular digraphs. *Preprint*, 2017. Available at https://arxiv.org/abs/1703.05839.

[17] R. Durrett, Rick. *Probability: theory and examples.* Cambridge university press, 2010.

[18] A. Edelman. Eigenvalues and condition numbers of random matrices. *SIAM Journal on Matrix Analysis and Applications*, 9(4), 543–560, 1988.

[19] L. Erdős, H.-T. Yau, and J. Yin. Bulk universality for generalized Wigner matrices. *Probability Theory and Related Fields*, 154, 341-407, 2012.

[20] C. G. Esséen. On the Kolmogorov-Rogozin inequality for the concentration function. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*. 5(3),210–216, 1966.

[21] S. Ge. Eigenvalue spacing of i.i.d. random matrices. *Preprint*, 2017.

[22] V. L. Girko. Circular law. *Teoriya Veroyatnostei i ee Primeneniya*, 29(4), 669–679, 1984.

[23] F. Götze and A. Tikhomirov. The circular law for random matrices. *Annals of Probability*, 38 (4), 1444-1491, 2010.

[24] M. L. Mehta. Random matrices and the statistical theory of energy levels. *Academic Press, New York*, 1967.

[25] H. H. Nguyen. Random doubly stochastic matrices: the circular law. *The Annals of Probability*, 42(3), 1161–1196, 2014.

[26] Hoi H. Nguyen and Van Vu. Circular law for random discrete matrices of given row sum. *Journal of Combinatorics*, 4(1), 1–30, 2013.

[27] L. A. Pastur. On the spectrum of random matrices. *Theoretical and Mathematical Physics*, 10(1), 67–74, 1972.

[28] D. W.Stroock. *Probability theory: An analytic view.* Second edition. Cambridge University Press, Cambridge, 2011.

[29] E. Rebrova and K. Tikhomirov. Coverings of random ellipsoids, and invertibility of matrices with i.i.d. heavy-tailed entries. *Israel Journal of Mathematics*, to appear. Preprint available at https://arxiv.org/abs/1508.06690.

[30] M. Rudelson. Invertibility of random matrices: Norm of the inverse. *Annals of Mathematics*, 168, 575-600, 2008.

[31] M. Rudelson and R. Vershynin. The Littlewood-Offord Problem and invertibility of random matrices. *Advances in Mathematics*, 218 (2), 600–633, 2008.

[32] M. Rudelson and R. Vershynin. Smallest singular value of a random rectangular matrix. *Communications on Pure and Applied Mathematics*, 62, 1707-1739, 2009.

[33] M. Rudelson and R. Vershynin. No-gaps delocalization for general random matrices. *Geometric and Functional Analysis*, 26(6), 1716–1776, 2016.

[34] M. Rudelson and O. Zeitouni. Singular values of Gaussian matrices and permanent estimators. *Random Structures Algorithms* 48(1), 183–212, 2016.

[35] T. Tao. *Topics in random matrix theory.* Vol. 132, American Mathematical Society, 2012.

[36] T. Tao and V. Vu. Random matrices: the circular law. *Communications in Contemporary Mathematics*, 10(2), 261–307, 2008.

[37] T. Tao and V. Vu. Random matrices: universality of the ESDs and the circular law. *Annals of Probability*, 38, 2023-2065, 2010 (with an appendix by M. Krishnapur).

[38] R. Vershynin. Invertibility of symmetric random matrices. *Random Structures & Algorithms*, 44(2), 135–182, 2014.

[39] E. P. Wigner. Characteristic vectors of bordered matrices. *Annals of Mathematics*, 62(3), 548–564, 1955.

[40] E. P. Wigner. On the distribution of the roots of certain symmetric matrices. *Annals of Mathematics*, 67(2), 325–327, 1958.

[41] P. M. Wood. Universality and the circular law for sparse random matrices. *Annals of Applied Probability*, 22(3), 1266–1300, 2012.

*Department of Mathematics, Weizmann Institute of Science, POB 26, Rehovot 76100, Israel.

† Department of Mathematics, University of Michigan, East Hall, 530 Church Street, Ann Arbor, Michigan 48109, USA.