

RANDOM MATRICES: UNIVERSALITY OF LOCAL SPECTRAL STATISTICS OF NON-HERMITIAN MATRICES

BY TERENCE TAO¹ AND VAN VU²

UCLA and Yale University

It is a classical result of Ginibre that the normalized bulk k -point correlation functions of a complex $n \times n$ Gaussian matrix with independent entries of mean zero and unit variance are asymptotically given by the determinantal point process on \mathbb{C} with kernel $K_\infty(z, w) := \frac{1}{\pi} e^{-|z|^2/2 - |w|^2/2 + z\bar{w}}$ in the limit $n \rightarrow \infty$. In this paper, we show that this asymptotic law is universal among all random $n \times n$ matrices M_n whose entries are jointly independent, exponentially decaying, have independent real and imaginary parts and whose moments match that of the complex Gaussian ensemble to fourth order. Analogous results at the edge of the spectrum are also obtained. As an application, we extend a central limit theorem for the number of eigenvalues of complex Gaussian matrices in a small disk to these more general ensembles.

These results are non-Hermitian analogues of some recent universality results for Hermitian Wigner matrices. However, a key new difficulty arises in the non-Hermitian case, due to the instability of the spectrum for such matrices. To resolve this issue, we need to work with the log-determinants $\log |\det(M_n - z_0)|$ rather than with the Stieltjes transform $\frac{1}{n} \text{tr}(M_n - z_0)^{-1}$, in order to exploit Girko's Hermitization method. Our main tools are a four moment theorem for these log-determinants, together with a strong concentration result for the log-determinants in the Gaussian case. The latter is established by studying the solutions of a certain nonlinear stochastic difference equation.

With some extra consideration, we can extend our arguments to the real case, proving universality for correlation functions of real matrices which match the real Gaussian ensemble to the fourth order. As an application, we show that a real $n \times n$ matrix whose entries are jointly independent, exponentially decaying and whose moments match the real Gaussian ensemble to fourth order has $\sqrt{\frac{2n}{\pi}} + o(\sqrt{n})$ real eigenvalues asymptotically almost surely.

1. Introduction. Let M_n be a random $n \times n$ matrix with complex entries, which is not necessarily assumed to be Hermitian, and can be either a continuous or discrete ensemble of matrices. Then, counting multiplicities, there are n complex

Received November 2012; revised May 2013.

¹Supported by NSF Grant DMS-06-49473.

²Supported by research Grants DMS-09-01216 and AFOSAR-FA-9550-09-1-0167.

MSC2010 subject classifications. 15A52.

Key words and phrases. Universality, random matrices, log-determinant, non-Hermitian matrices, circular law.

(algebraic) eigenvalues, which we enumerate in an arbitrary fashion as

$$\lambda_1(M_n), \dots, \lambda_n(M_n) \in \mathbb{C}.$$

One can then define, for each $1 \leq k \leq n$, the k -point correlation function

$$\rho_n^{(k)} = \rho_n^{(k)}[M_n]: \mathbb{C}^k \rightarrow \mathbb{R}^+$$

of the random matrix ensemble M_n by requiring that

$$(1.1) \quad \int_{\mathbb{C}^k} F(z_1, \dots, z_k) \rho_n^{(k)}(z_1, \dots, z_k) dz_1 \cdots dz_k = \mathbf{E} \sum_{1 \leq i_1, \dots, i_k \leq n, \text{ distinct}} F(\lambda_{i_1}(M_n), \dots, \lambda_{i_k}(M_n))$$

for all continuous, compactly supported test functions F , where dz denotes Lebesgue measure on the complex plane \mathbb{C} . Note that this definition does not depend on the exact order in which the eigenvalues of M_n are enumerated.

If M_n is an absolutely continuous matrix ensemble with a continuous density function, then $\rho^{(k)}$ is a continuous function; but if M_n is a discrete ensemble then $\rho^{(k)}$ is merely a nonnegative measure.³ In the absolutely continuous case with a continuous density function, one can equivalently define $\rho_n^{(k)}(z_1, \dots, z_k)$ for distinct z_1, \dots, z_k to be the quantity such that the probability that there is an eigenvalue of M_n in each of the disks $\{z : |z - z_i| \leq \varepsilon\}$ for $i = 1, \dots, k$ is asymptotically $(\rho_n^{(k)}(z_1, \dots, z_k) + o(1))(\pi \varepsilon^2)^k$ in the limit $\varepsilon \rightarrow 0^+$.

We note two model cases of continuous matrix ensembles that are of interest. The first is the *real Gaussian matrix ensemble*,⁴ in which coefficients ξ_{ij} are independent and identically distributed (or *i.i.d.* for short) and have the distribution $N(0, 1)_{\mathbb{R}}$ of the real Gaussian with mean zero and variance one. We will discuss this case in more detail later, but for now we will focus instead on the simpler and better understood case of the *complex Gaussian matrix ensemble*, in which the ξ_{ij} are *i.i.d.* with the distribution of a complex Gaussian $N(0, 1)_{\mathbb{C}}$ with mean zero and variance one [or in other words, the probability distribution of each ξ_{ij} is $\frac{1}{\pi} e^{-|z|^2} dz$, and the real and imaginary parts of ξ_{ij} independently have the distribution $N(0, 1/2)_{\mathbb{R}}$]. As is well known, the correlation functions of a complex Gaussian matrix are given by the explicit *Ginibre formula* [26]

$$(1.2) \quad \rho_n^{(k)}(z_1, \dots, z_k) = \det(K_n(z_i, z_j))_{1 \leq i, j \leq k},$$

³Here, we have abused notation by identifying a measure $\rho_n^{(k)}(z_1, \dots, z_k) dz_1 \cdots dz_k$ with its density $\rho_n^{(k)}$.

⁴Strictly speaking, the real Gaussian matrix ensemble is only absolutely continuous with respect to Lebesgue measure on the space of *real* $n \times n$ matrices, rather than on the space of *complex* $n \times n$ matrices. However, both ensembles are still continuous in the sense that any individual matrix occurs in the ensemble with probability zero.

where $K_n : \mathbb{C} \times \mathbb{C} \rightarrow \mathbb{C}$ is the kernel

$$(1.3) \quad K_n(z, w) := \frac{1}{\pi} e^{-(|z|^2+|w|^2)/2} \sum_{j=0}^{n-1} \frac{(z\bar{w})^j}{j!}.$$

In particular, one has

$$(1.4) \quad \rho_n^{(1)}(z) = K_n(z, z) = \frac{1}{\pi} e^{-|z|^2} \sum_{j=0}^{n-1} \frac{|z|^{2j}}{j!}$$

and thus (by Taylor expansion of $e^{-|z|^2}$) one has the asymptotic

$$\rho_n^{(1)}(\sqrt{n}z) \rightarrow \frac{1}{\pi} 1_{|z| \leq 1}$$

as $n \rightarrow \infty$ for almost every $z \in \mathbb{C}$. This gives the well-known *circular law* for complex Gaussian matrices, namely that the empirical spectral distribution of $\frac{1}{\sqrt{n}}M_n$ converges (in expectation, at least) to the circular measure $\frac{1}{\pi}1_{B(0,1)} dz$, where we use $B(z_0, r) := \{z \in \mathbb{C} : |z - z_0| < r\}$ to denote an open disk in the complex plane. Informally, this means that the eigenvalues of M_n are asymptotically uniformly distributed on the disk $B(0, \sqrt{n})$. The circular law is also known to hold for many other ensembles of matrices, and for several modes of convergence. In particular, it holds (both in probability and in the almost sure sense) for random matrices with i.i.d. entries having mean 0 and variance 1; see the surveys [5, 52] for further discussion of this and related results. Figures 2 and 3 later in this paper illustrate the circular law for two model instances of i.i.d. ensembles, namely the real Gaussian and real Bernoulli ensembles.

We also remark that from the obvious inequality

$$\sum_{j=0}^{n-1} \frac{|z|^{2j}}{j!} \leq \sum_{j=0}^{\infty} \frac{|z|^{2j}}{j!} = e^{|z|^2}$$

and (1.4) we have the uniform bound

$$|K_n(z, z)| \leq \frac{1}{\pi}$$

for all z , and hence by positivity of $\rho_n^{(2)}(z, w) = K_n(z, z)K_n(w, w) - |K_n(z, w)|^2$ we also have

$$(1.5) \quad |K_n(z, w)| \leq \frac{1}{\pi}$$

for all z, w . In particular, from (1.2) one has

$$(1.6) \quad 0 \leq \rho_{n,z_1, \dots, z_k}^{(k)}(w_1, \dots, w_k) \leq C_k$$

in the case of the complex Gaussian ensemble for all $w_1, \dots, w_k \in \mathbb{C}$, all n , and some constant C_k depending only on k . (Indeed, from the Hadamard inequality one can take $C_k = \pi^{-k} k^{k/2}$, e.g.) This uniform bound will be technically convenient for some of our applications. We will also need an analogous bound for the real Gaussian ensemble; see Lemma 11 below.

Our first main result is to show a universality result of the k -point correlation functions $\rho_{n, z_1, \dots, z_k}^{(k)}(w_1, \dots, w_k)$, in the spirit of the “four moment theorems” for Wigner matrices that first appeared in [55]. Very roughly speaking, the result is that (when measured in the vague topology), the asymptotic behavior of these correlation functions for matrices with independent entries depend only on the first four moments of the entries, though due to our reliance on the Lindeberg exchange method, we will also need to require these matrices to match moments with the complex Gaussian ensemble. To make this statement more precise, we will need some further notation.

DEFINITION 1 (Independent-entry matrices). An *independent-entry matrix ensemble* is an ensemble of random $n \times n$ matrices $M_n = (\xi_{ij})_{1 \leq i, j \leq n}$, where the ξ_{ij} are independent and complex random variables, each with mean zero and variance one; we call the ξ_{ij} the *atom distributions* of M_n . We say that the independent-entry matrix has *independent real and imaginary parts* if for each $1 \leq i, j \leq n$, $\text{Re}(\xi_{ij}), \text{Im}(\xi_{ij})$ are independent. We say that the matrix obeys condition (C1) if one has

$$\mathbf{P}(|\xi_{ij}| \geq t) \leq C \exp(-t^c)$$

for some fixed $C, c > 0$ (independent of n) and all i, j .

If $k \geq 0$, we say that two independent-entry matrix ensembles $M_n = (\xi_{ij})_{1 \leq i, j \leq n}$ and $M'_n = (\xi'_{ij})_{1 \leq i, j \leq n}$ have *matching moments to order k* if one has

$$(1.7) \quad \mathbf{E} \text{Re}(\xi_{ij})^a \text{Im}(\xi_{ij})^b = \mathbf{E} \text{Re}(\xi'_{ij})^a \text{Im}(\xi'_{ij})^b,$$

whenever $1 \leq i, j \leq n, a, b \geq 0$ and $a + b \leq k$.

Our first main result is then as follows.

THEOREM 2 (Four moment theorem for complex matrices). Let M_n, \widetilde{M}_n be independent-entry matrix ensembles with independent real and imaginary parts, obeying condition (C1), such that M_n and \widetilde{M}_n both match moments with the complex Gaussian matrix ensemble to third order, and match moments with each other to fourth order. Let $k \geq 1$ be a fixed integer, let $z_1, \dots, z_k \in \mathbb{C}$ be bounded (thus $|z_i| \leq C$ for all $i = 1, \dots, k$ and some fixed $C > 0$), and let $F: \mathbb{C}^k \rightarrow \mathbb{C}$ be a smooth function, which admits a decomposition of the form

$$(1.8) \quad F(w_1, \dots, w_k) = \sum_{i=1}^m F_{i,1}(w_1) \cdots F_{i,k}(w_k)$$

for some fixed m and some smooth functions $F_{i,j} : \mathbb{C} \rightarrow \mathbb{C}$ for $i = 1, \dots, m$ and $j = 1, \dots, k$ supported on the disk $\{w : |w| \leq C\}$ obeying the derivative bounds⁵

$$(1.9) \quad |\nabla^a F_{i,j}(w)| \leq C$$

for all $0 \leq a \leq 5$, $i = 1, \dots, m$, $j = 1, \dots, k$ and $w \in \mathbb{C}$, and some fixed C . Let $\rho_n^{(k)}, \tilde{\rho}_n^{(k)}$ be the correlation functions for M_n, \tilde{M}_n , respectively. Then

$$\begin{aligned} & \int_{\mathbb{C}^k} F(w_1, \dots, w_k) \rho_n^{(k)}(\sqrt{n}z_1 + w_1, \dots, \sqrt{n}z_k + w_k) dw_1 \cdots dw_k \\ &= \int_{\mathbb{C}^k} F(w_1, \dots, w_k) \tilde{\rho}_n^{(k)}(\sqrt{n}z_1 + w_1, \dots, \sqrt{n}z_k + w_k) dw_1 \cdots dw_k \\ & \quad + O(n^{-c}) \end{aligned}$$

for some absolute constant $c > 0$ (independent of k). Furthermore, the implicit constant in the $O(n^{-c})$ notation is uniform over all z_1, \dots, z_k in the bounded region $\{z : |z| \leq C\}$.

REMARK 3. The regularity hypotheses on the test function F here are somewhat technical, but they are needed to obtain the uniform polynomial decay $O(n^{-c})$ in the conclusion, which is useful for several applications. Note that by rescaling one could allow the bound C in (1.9) to be enlarged somewhat, to $Cn^{c/2k}$, without impacting the conclusion [other than to degrade the $O(n^{-c})$ error slightly to $O(n^{-c/2})$]. If one is only seeking a qualitative error term of $o(1)$, then by applying the Stone–Weierstrass theorem, one only needs F to be continuous and compactly supported, instead of having a smooth factorization of the form (1.8); see the proof of Corollary 7 below. Also, if F is smooth and compactly supported, then by using a partial Fourier expansion one can again obtain a polynomial decay rate $O(n^{-c})$ (with the implied constant depending on the bounds on finitely many derivatives of F). It is possible to improve the value of c somewhat by adding additional matching moment hypotheses, but then one also requires the derivative bounds (1.9) for a larger range of exponents a ; we will not quantify this variant of Theorem 2 here. The requirement that M_n, M'_n match the complex Gaussian ensemble to third order can be removed if z_1, \dots, z_k stays a bounded distance away from the origin, using an extremely recent result of Bourgade, Yau, and Yin [8]; see Remark 22.

Theorem 2 is motivated by the phenomenon, first observed in [55], that the asymptotic local statistics of the spectrum of a random Hermitian matrix of Wigner type typically depend only on the first four moments of the entries; formalizations

⁵See Section 3 for the definition of the a -fold gradient $\nabla^a F_{i,j}$.

of this phenomenon are known as *four moment theorems*. In particular, Corollary 7 is analogous⁶ to the four moment theorems in [55], Theorems 11 and 38.

REMARK 4. The hypothesis of independent real and imaginary parts is primarily for reasons of notational convenience, and it is likely that this hypothesis could be dropped from our results. Note that when M_n and M'_n have independent real and imaginary parts, the moment matching condition (1.7) simplifies to

$$\mathbf{E} \operatorname{Re}(\xi_{ij})^a = \mathbf{E} \operatorname{Re}(\xi'_{ij})^a$$

and

$$\mathbf{E} \operatorname{Im}(\xi_{ij})^b = \mathbf{E} \operatorname{Im}(\xi'_{ij})^b$$

for $1 \leq i, j \leq n$ and $0 \leq a, b \leq k$.

It is also likely that the exponential decay condition in condition (C1) could be replaced with a bound on a sufficiently high moment of the entries. We will however not pursue these refinements here. The vague convergence in the conclusion is natural given that the ensemble M_n is permitted to be discrete (so that $\rho_n^{(k)}$ could be a discrete measure, rather than a continuous function). In analogy with the Hermitian theory (see, e.g., [56]), it is reasonable to conjecture that stronger modes of convergence become available if some additional regularity hypotheses are placed on the entries, but we will not pursue such matters here.

We now discuss some applications of Theorem 2. The first application concerns the asymptotic behavior of the k -point correlation functions as $n \rightarrow \infty$. In the case when M_n is drawn from the complex Gaussian ensemble, these asymptotics have been well understood since the work of Ginibre [26]. To recall these asymptotics, we introduce the following functions.

DEFINITION 5 (Asymptotic kernel). For complex numbers z_1, z_2, w_1, w_2 , define the kernel $K_{\infty, z_1, z_2}(w_1, w_2)$ by the following rules:

- (i) If $z_1 \neq z_2$, then $K_{\infty, z_1, z_2}(w_1, w_2) := 0$.
- (ii) If $z_1 = z_2$ and $|z_1| > 1$, then $K_{\infty, z_1, z_2}(w_1, w_2) := 0$.
- (iii) If $z_1 = z_2$ and $|z_1| < 1$, then $K_{\infty, z_1, z_2}(w_1, w_2) := \frac{1}{\pi} e^{-|w_1|^2/2 - |w_2|^2/2 + w_1 \bar{w}_2}$.

⁶Thanks to more recent results by many authors [16, 20–22, 53, 56], these results are no longer the sharpest results available in the Wigner setting, as the moment matching conditions have now largely been removed, the exponential decay condition relaxed to a finite moment condition, and the bulk results extended to the edge; see the discussion in [56] or the surveys [15, 28, 43, 49] for surveys for more details. In view of these results, it is reasonable to conjecture the moment matching assumptions in Theorem 2 or Corollary 7 may be relaxed; see Remark 22 for some very recent developments in this direction.

(iv) If $z_1 = z_2$ and $|z_1| = 1$, then

$$K_{\infty, z_1, z_2}(w_1, w_2) := \frac{1}{\pi} e^{-|w_1|^2/2 - |w_2|^2/2 + w_1 \bar{w}_2} \left(\frac{1}{2} + \frac{1}{2} \operatorname{erf}(-\sqrt{2}(z_1 \bar{w}_2 + w_1 \bar{z}_2)) \right).$$

Here,

$$\operatorname{erf}(z) := \frac{2}{\sqrt{\pi}} \int_0^z e^{-t^2} dt$$

is the usual error function, defined for all complex z , where the integral is over an arbitrary contour from 0 to z . For complex numbers $z_1, \dots, z_k, w_1, \dots, w_k$, define the correlation function

$$\rho_{\infty, z_1, \dots, z_k}^{(k)}(w_1, \dots, w_k) := \det(K_{\infty, z_i, z_j}(w_i, w_j))_{1 \leq i, j \leq k}.$$

In the model case when z_1, \dots, z_k all avoid the unit circle $\{z \in \mathbb{C} : |z| = 1\}$, the kernel simplifies to

$$K_{\infty, z_i, z_j}(w_i, w_j) = 1_{z_i = z_j} 1_{|z_i| < 1} K_{\infty}(w_i, w_j),$$

where

$$K_{\infty}(z, w) := \frac{1}{\pi} e^{-|z|^2/2 - |w|^2/2 + z \bar{w}}.$$

The kernel K_{∞} can also be interpreted as the reproducing kernel for the orthogonal projection in $L^2(\mathbb{C})$ to (the closure of) the space of functions $f(z)$ that become holomorphic after multiplication by $e^{|z|^2/2}$, or equivalently to the closed span of $z^k e^{-|z|^2/2}$ for $k = 0, 1, \dots$.

LEMMA 6 (Kernel asymptotics). *Let $z_1, \dots, z_k, w_1, \dots, w_k$ be fixed complex numbers for some fixed k , and let M_n be drawn from the complex Gaussian ensemble. Then we have⁷*

$$(1.10) \quad \rho_n^{(k)}(\sqrt{n}z_1 + w_1, \dots, \sqrt{n}z_k + w_k) = \rho_{\infty, z_1, \dots, z_k}^{(k)}(w_1, \dots, w_k) + o(1).$$

If none of the z_1, \dots, z_k lie on the unit circle, then we may improve the error term $o(1)$ to $O(\exp(-\delta n))$ for some fixed $\delta > 0$.

Now suppose that $z_1, \dots, z_k, w_1, \dots, w_k$ are allowed to vary in n , but that the $z_1, \dots, w_1, \dots, w_k$ remain bounded (i.e., $|z_i|, |w_i| \leq C$ for some fixed C and all $1 \leq i \leq k$) and the z_1, \dots, z_k stay bounded away from the unit circle (i.e., $||z_i| - 1| \geq \varepsilon$ for some fixed $\varepsilon > 0$ and all $1 \leq i \leq k$). Then one still has the asymptotic (1.10). In other words, the decay rate of the error term $o(1)$ in (1.10) is uniform across all choices of $z_1, \dots, z_k, w_1, \dots, w_k$ in the ranges specified above.

⁷See Section 3 for the asymptotic notational conventions we will use in this paper.

PROOF. This is a well-known asymptotic (see, e.g., [35, 37] or [7]). For sake of completeness, we have written a proof of these standard facts at Appendix B of the copy of this paper at [arXiv:1206.1893v3](https://arxiv.org/abs/1206.1893v3). \square

From this lemma, we conclude in particular that $\rho_{\infty, z_1, \dots, z_k}^{(k)}(w_1, \dots, w_k) \geq 0$ for all $k, z_1, \dots, z_k, w_1, \dots, w_k$, which [when combined with (1.5)] yields the uniform bound

$$|K_{\infty, z_1, z_2}(w_1, w_2)| \leq \frac{1}{\pi}$$

for all $z_1, z_2, w_1, w_2 \in \mathbb{C}$. In particular, we have

$$(1.11) \quad 0 \leq \rho_{\infty, z_1, \dots, z_k}^{(k)}(w_1, \dots, w_k) \leq C_k$$

for all $w_1, \dots, w_k \in \mathbb{C}$ and some constant C_k depending only on k .

Using Theorem 2, we may extend the above asymptotics for complex Gaussian matrices to more general ensembles (including some discrete ensembles), as follows.

COROLLARY 7 (Universality for complex matrices). *Let M_n be an independent-entry matrix ensemble with independent real and imaginary parts, obeying condition (C1), and which matches moments with the complex Gaussian matrix ensemble to fourth order. Then for any fixed (i.e., independent of n), fixed $k \geq 1$ and fixed $z_1, \dots, z_k \in \mathbb{C}$, and any fixed continuous, compactly supported function $F : \mathbb{C}^k \rightarrow \mathbb{C}$, one has*

$$\begin{aligned} & \int_{\mathbb{C}^k} F(w_1, \dots, w_k) \rho_n^{(k)}(\sqrt{n}z_1 + w_1, \dots, \sqrt{n}z_k + w_k) dw_1 \cdots dw_k \\ &= \int_{\mathbb{C}^k} F(w_1, \dots, w_k) \rho_{\infty, z_1, \dots, z_k}^{(k)}(w_1, \dots, w_k) dw_1 \cdots dw_k + o(1). \end{aligned}$$

In other words, the asymptotic (1.10) is valid in the vague topology for this ensemble. If F is furthermore assumed to be smooth, then we may improve the $o(1)$ error term here to $O(n^{-c})$ for some fixed $c > 0$.

PROOF. From Theorem 2 and Lemma 6, we obtain Corollary 7 in the case when F admits a decomposition of the form given in Theorem 2 [and in this case the $o(1)$ error can be improved to $O(n^{-c})$]. The more general case of continuous, compactly supported F can then be deduced by using the Stone–Weierstrass theorem to approximate a continuous F by an approximant \tilde{F} of the form (1.8) [and by using a further function of the form in Theorem 2 and (1.11) to upper bound the error]. When F is smooth, one can replace the use of the Stone–Weierstrass theorem by a more quantitative partial Fourier series expansion of F (extended periodically in a suitable fashion), followed by a multiplication by a smooth cutoff function, taking advantage of the rapid decrease of the Fourier coefficients in the smooth case; we omit the standard details. \square

REMARK 8. Note that in contrast to the situation in Theorem 2, the parameters z_1, \dots, z_k in Corollary 7 are required to be fixed in n , as opposed to being allowed to vary in n . Related to this, the error term $o(1)$ in Corollary 7 is not asserted to be uniform in the choice of z_1, \dots, z_k , in contrast to the uniformity in Theorem 2. Indeed, given that the limiting correlation function $\rho_{\infty, z_1, \dots, z_k}^{(k)}$ behaves discontinuously in z_1, \dots, z_k whenever two of the z_i collide, or when one of the z_i crosses the unit circle, one would not expect such uniformity in Corollary 7. Thus, while Corollary 7 describes more explicitly the limiting behavior (in certain regimes) of the correlation functions $\rho^{(k)}$, we regard Theorem 2 as the more precise statement regarding the asymptotics of these functions.

In the Hermitian case, four moment theorems can be used to extend various facts about the asymptotic spectral distribution of special matrix ensembles (such as the Gaussian unitary ensemble) to other matrix ensembles which obey appropriate moment matching conditions. Similarly, by using Theorem 2, one may extend some facts about eigenvalues of complex Gaussian matrices can now be extended to i.i.d. matrix models that match the complex Gaussian ensemble to fourth order, although in some “global” cases the extension is only partial in nature due to the “local” nature of the four moment theorem. Rather than provide an exhaustive list of such applications, we will present just one representative such application, namely that of (partially) extending the following central limit theorem of Rider [39].

THEOREM 9 (Central limit theorem, Gaussian case). *Let M_n be drawn from the complex Gaussian ensemble. Let $r > 0$ be a real number (depending on n) such that $1/r, r/n^{1/2} = o(1)$. Let z_0 be a complex number (also depending on n) such that $|z_0| \leq (1 - \varepsilon)\sqrt{n}$ for some fixed $\varepsilon > 0$. Let $N_{B(z_0, r)}$ be the number of eigenvalues of M_n in the ball $B(z_0, r) := \{z \in \mathbb{C} : |z - z_0| < r\}$. Then we have*

$$\frac{N_{B(z_0, r)} - r^2}{r^{1/2}\pi^{-1/4}} \rightarrow N(0, 1)_{\mathbb{R}}$$

in the sense of distributions. In fact, we have the slightly stronger statement that

$$(1.12) \quad \mathbf{E} \left(\frac{N_{B(z_0, r)} - r^2}{r^{1/2}\pi^{-1/4}} \right)^k \rightarrow \mathbf{E} N(0, 1)_{\mathbb{R}}^k$$

for all fixed natural numbers $k \geq 0$.

PROOF. From the general Costin–Lebowitz central limit theorem for determinantal point processes [12, 46, 47], we know that

$$\frac{N_{B(z_0, r)} - \mathbf{E} N_{B(z_0, r)}}{(\text{Var } N_{B(z_0, r)})^{1/2}} \rightarrow N(0, 1)_{\mathbb{R}}$$

provided that $\text{Var } N_{B(z_0,r)} \rightarrow \infty$; indeed, an inspection of the proof in [46] gives the slightly stronger assertion that

$$\mathbf{E} \left(\frac{N_{B(z_0,r)} - \mathbf{E}N_{B(z_0,r)}}{(\text{Var } N_{B(z_0,r)})^{1/2}} \right)^k \rightarrow \mathbf{E}N(0, 1)_{\mathbb{R}}^k$$

for any fixed $k \geq 0$. Thus, it will suffice to establish the asymptotics

$$\mathbf{E}N_{B(z_0,r)} = (1 + o(1))r^2$$

and

$$\text{Var } N_{B(z_0,r)} = (1 + o(1))\pi^{-1/2}r.$$

Using (1.1), (1.2), one can write the left-hand sides here as

$$\int_{B(z_0,r)} K_n(z, z) dz$$

and

$$\int_{B(z_0,r)} K_n(z, z) dz - \int_{B(z_0,r)} \int_{B(z_0,r)} |K_n(z, w)|^2 dz dw,$$

respectively. By Lemma 6, the former expression converges to $\int_{B(z_0,r)} \frac{1}{\pi} dz = r^2$. Lemma 6 also reveals that the second expression is asymptotically independent of z_0 , and so one may without loss of generality take $z_0 = 0$. But then the required asymptotic follows from [39], Theorem 1.6 (after allowing for the different normalization for M_n in that paper). \square

Using Theorem 2, we may extend this result to more general ensembles, at least in the small radius case.

COROLLARY 10 (Central limit theorem, general case). *Let M_n be an independent-entry matrix ensemble with independent real and imaginary parts, obeying condition (C1), such that M_n matches moments with the complex Gaussian matrix ensemble to fourth order. Then the conclusion of Theorem 9 for M_n holds provided that one has the additional assumption $r \leq n^{o(1)}$.*

We prove this result in Section 6.3. The restriction to small radii $r \leq n^{o(1)}$ appears to be a largely technical restriction, relating to the need to take arbitrarily high moments in order to establish a central limit theorem; see, for instance, Figure 1 for some numerical evidence that the central limit theorem should in fact hold for larger radii as well (and for real matrices as well as complex ones). It seems likely that one can also obtain extensions of many of the other results in [39] (or related papers, such as [32, 38]) on Gaussian fluctuations from the circular law from the complex Gaussian ensemble to other ensembles that match the complex Gaussian ensemble to a sufficiently large number of moments, but we will not pursue such results here. We remark that for macroscopic statistics $\frac{1}{n} \sum_{i=1}^n F(\lambda_i / \sqrt{n})$ with F fixed and analytic, such extensions (without the need for matching moments beyond the second moment) were already established in [40].

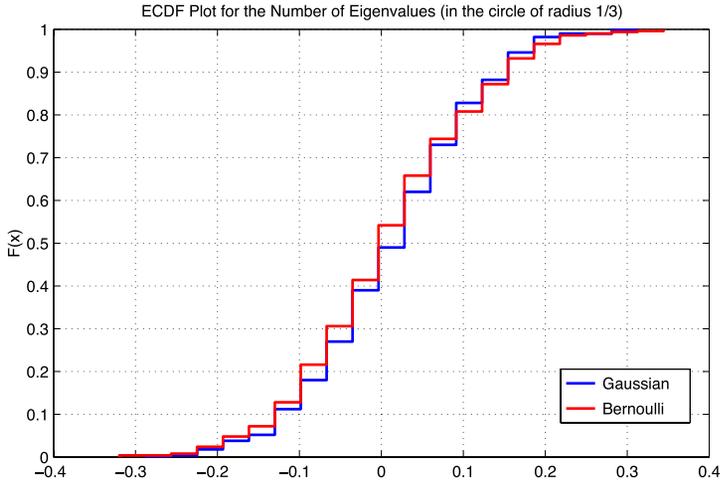


FIG. 1. The cumulative distribution function for the number of eigenvalues in the disk $B(0, \sqrt{n}/3)$ of real Gaussian and real Bernoulli matrices of size $10,000 \times 10,000$, after normalizing the mean by $n/9$ and variance by \sqrt{n} . Thanks to Ke Wang for the data and figure.

1.1. *The real case and applications.* There is a (more complicated) analogue of Theorem 2 in which the complex entries are replaced by real ones. This has the effect of forcing the spectrum $\lambda_1(M_n), \dots, \lambda_n(M_n)$ to split into some number $\lambda_{1,\mathbb{R}}(M_n), \dots, \lambda_{N_{\mathbb{R}}[M_n],\mathbb{R}}(M_n)$ of real eigenvalues, together with some number $\lambda_{1,\mathbb{C}_+}(M_n), \dots, \lambda_{N_{\mathbb{C}_+}[M_n],\mathbb{C}_+}(M_n)$ of complex eigenvalues in the upper half-plane $\mathbb{C}_+ := \{z \in \mathbb{C} : \text{Im}(z) > 0\}$, as well as their complex conjugates $\overline{\lambda_{1,\mathbb{C}_+}(M_n)}, \dots, \overline{\lambda_{N_{\mathbb{C}_+}[M_n],\mathbb{C}_+}(M_n)}$, where $N_{\mathbb{R}}[M_n], N_{\mathbb{C}_+}[M_n]$ denote the number of real eigenvalues of M_n and the number of eigenvalues of M_n in the upper half-plane, respectively (so in particular, $N_{\mathbb{R}}[M_n] + 2N_{\mathbb{C}_+}[M_n] = n$ almost surely). Because of this additional structure of the eigenvalues, it is no longer convenient to consider the correlation functions $\rho_n^{(k)} : \mathbb{C}^k \rightarrow \mathbb{R}^+$ as defined in (1.1), since they become singular when one or more of the variables is real. Instead, it is more convenient to work with the correlation functions $\rho_n^{(k,l)} : \mathbb{R}^k \times \mathbb{C}_+^l \rightarrow \mathbb{R}^+$, defined for $k, l \geq 0$ by the formula

$$\begin{aligned}
 & \int_{\mathbb{R}^k} \int_{\mathbb{C}_+^l} F(x_1, \dots, x_k, z_1, \dots, z_l) \\
 & \quad \times \rho_n^{(k,l)}(x_1, \dots, x_k, z_1, \dots, z_l) dx_1 \cdots dx_k dz_1 \cdots dz_l \\
 (1.13) \quad & = \mathbf{E} \sum_{1 \leq i_1 < \dots < i_k \leq N_{\mathbb{R}}[M_n]} \sum_{1 \leq j_1 < \dots < j_l \leq N_{\mathbb{C}_+}[M_n]} F(\lambda_{i_1,\mathbb{R}}(M_n), \dots, \lambda_{i_k,\mathbb{R}}(M_n), \\
 & \quad \lambda_{j_1,\mathbb{C}_+}(M_n), \dots, \lambda_{j_l,\mathbb{C}_+}(M_n)).
 \end{aligned}$$

Again, the exact ordering of the eigenvalues here is unimportant. When the law of M_n has a continuous density with respect to Lebesgue measure on real matrices (which is, e.g., the case with the real Gaussian ensemble), one can interpret $\rho_n^{(k,l)}(x_1, \dots, x_k, z_1, \dots, z_l)$ for distinct $x_1, \dots, x_k \in \mathbb{R}$ and $z_1, \dots, z_l \in \mathbb{C}_+$ as the unique real number such that, as $\varepsilon \rightarrow 0$, the probability of simultaneously having an eigenvalue of M_n in each of the intervals $(x_i - \varepsilon, x_i + \varepsilon)$ for $i = 1, \dots, k$ and in each of the disks $B(z_j, \varepsilon)$ for $j = 1, \dots, l$ is equal to

$$(1 + o(1))\rho_n^{(k,l)}(x_1, \dots, x_k, z_1, \dots, z_l)(2\varepsilon)^k(\pi\varepsilon^2)^l$$

in the limit as $\varepsilon \rightarrow 0$.

Define $\mathbb{C}_- := \{z \in \mathbb{C} : \text{Im}(z) < 0\}$ and $\mathbb{C}_* := \mathbb{C}_+ \cup \mathbb{C}_- = \mathbb{C} \setminus \mathbb{R}$. We extend the correlation functions $\rho_n^{(k,l)}$ from $\mathbb{R}^k \times \mathbb{C}_+^l$ to $\mathbb{R}^k \times \mathbb{C}_*^l$ by requiring that the functions be invariant with respect to conjugations of any of the l coefficients of \mathbb{C}^l . We then extend $\rho_n^{(k,l)}$ by zero from $\mathbb{R}^k \times \mathbb{C}_*^l$ to $\mathbb{R}^k \times \mathbb{C}^l$.

When M_n is given by the real Gaussian ensemble, the correlation functions $\rho_n^{(k,l)}$ were computed by a variety of methods, for both odd and even n , in [1, 6, 7, 23, 30, 44, 45] (with the $(k, l) = (1, 0), (0, 1)$ cases worked out previously in [13, 14, 34], building in turn on the foundational work of Ginibre [26]). The precise formulae for these correlation functions are somewhat complicated and involve Pfaffians of a certain 2×2 matrix kernel; see Appendix B for the formulae when n is even, and [23, 44] for the case when n is odd. To avoid some technical issues, we shall restrict attention to the case when n is even, although it is virtually certain that the results here should also extend to the odd n case.

For technical reasons, we will need the following variant of (1.6).

LEMMA 11 (Uniform bound on correlation functions). *Let $k, l \geq 0$ be fixed natural numbers, let n be even, and let M_n be drawn from the real Gaussian ensemble. Then for all $x_1, \dots, x_k \in \mathbb{R}$ and $z_1, \dots, z_l \in \mathbb{C}$ one has*

$$0 \leq \rho_n^{(k,l)}(x_1, \dots, x_k, z_1, \dots, z_l) \leq C_{k,l}$$

for some fixed $C_{k,l}$ depending only on k, l .

This lemma follows fairly easily from the computations in [7]; we give the details in Appendix B. We will need this lemma in order to control the event of having two real eigenvalues that are very close to each other, or a complex eigenvalue very close to the real axis, as in those cases, one is close to a transition in which two real eigenvalues become complex or vice versa, creating a potential instability in the correlation functions $\rho_n^{(k,l)}$. One can in fact establish stronger *level repulsion* estimates which provide some decay on $\rho_n^{(k,l)}(x_1, \dots, x_k, z_1, \dots, z_l)$ as two of the $x_1, \dots, x_k, z_1, \dots, z_l$ get close to each other, or as one of the z_i gets close to the real axis, but we will not need such estimates here.

We then have the following analogue of Theorem 2, which is the second main result of this paper:

THEOREM 12 (Four moment theorem for real matrices). *Let M_n, \widetilde{M}_n be independent-entry matrix ensembles with real coefficients, obeying condition (C1), such that M_n and \widetilde{M}_n both match moments with the real Gaussian matrix ensemble to fourth order. Let $k, l \geq 0$ be fixed integers, and let x_1, \dots, x_k and $z_1, \dots, z_l \in \mathbb{C}$ be bounded. Assume that n is even. Let $F: \mathbb{R}^k \times \mathbb{C}^l \rightarrow \mathbb{C}$ be a smooth function which admits a decomposition of the form*

$$F(y_1, \dots, y_k, w_1, \dots, w_l) = \sum_{i=1}^m G_{i,1}(y_1) \cdots G_{i,k}(y_k) F_{i,1}(w_1) \cdots F_{i,l}(w_l)$$

for some fixed m and some smooth functions $G_{i,p}: \mathbb{R} \rightarrow \mathbb{C}$ and $F_{i,j}: \mathbb{C} \rightarrow \mathbb{C}$ for $i = 1, \dots, m, p = 1, \dots, k$ and $j = 1, \dots, l$ supported on the interval $\{y \in \mathbb{R}: |y| \leq C\}$ and disk $\{w \in \mathbb{C}: |w| \leq C\}$, respectively, obeying the derivative bounds

$$|\nabla^a G_{i,p}(y)|, |\nabla^a F_{i,j}(w)| \leq C$$

for all $0 \leq a \leq 5, i = 1, \dots, m, p = 1, \dots, k, j = 1, \dots, l, y \in \mathbb{R}$, and $w \in \mathbb{C}$, and some fixed C . Let $\rho_n^{(k,l)}, \tilde{\rho}_n^{(k,l)}$ be the correlation functions for M_n, \widetilde{M}_n , respectively. Then

$$\begin{aligned} & \int_{\mathbb{R}^k} \int_{\mathbb{C}^l} F(y_1, \dots, y_k, w_1, \dots, w_l) \\ & \quad \times \rho_n^{(k,l)}(\sqrt{n}x_1 + y_1, \dots, \sqrt{n}x_k + y_k, \\ & \quad \quad \sqrt{n}z_1 + w_1, \dots, \sqrt{n}z_l + w_l) dw_1 \cdots dw_l dy_1 \cdots dy_k \\ & = \int_{\mathbb{R}^k} \int_{\mathbb{C}^l} F(y_1, \dots, y_k, w_1, \dots, w_l) \\ & \quad \times \tilde{\rho}_n^{(k,l)}(\sqrt{n}x_1 + y_1, \dots, \sqrt{n}x_k + y_k, \\ & \quad \quad \sqrt{n}z_1 + w_1, \dots, \sqrt{n}z_l + w_l) dw_1 \cdots dw_l dy_1 \cdots dy_k \\ & \quad + O(n^{-c}) \end{aligned}$$

for some absolute constant $c > 0$ (independent of k, l). Furthermore, the implicit constant in the $O(n^{-c})$ notation is uniform over all x_1, \dots, x_k and z_1, \dots, z_l in the bounded regions $\{x \in \mathbb{R}: |x| \leq C\}$ and $\{z \in \mathbb{C}: |z| \leq C\}$, respectively.

As will be seen in Section 6.2, the proof of Theorem 12 proceeds along the same lines as Theorem 2, but with some additional arguments involving Lemma 11 required to prevent pairs of eigenvalues from escaping or entering the real axis due to collisions. It is because of these additional arguments that matching to fourth order, rather than third order, is required. It is, however, expected that the moment conditions should be relaxed; see, for instance, Figures 2 and 3 for the close resemblance in spectral statistics between real Gaussian and Bernoulli matrices, which only match to third order rather than to fourth order.

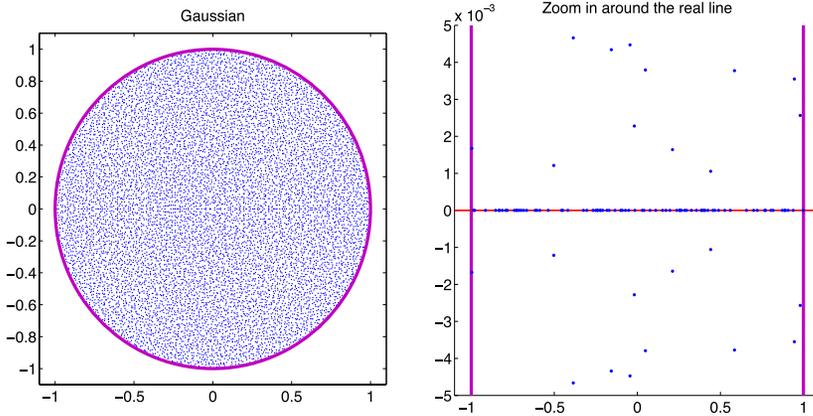


FIG. 2. The spectrum of a random real Gaussian $10,000 \times 10,000$ matrix, with additional detail near the origin to show the concentration on the real axis. Thanks to Ke Wang for the data and figure.

REMARK 13. In [44], some explicit formulae for the correlation functions of real Gaussian matrices in the case of odd n were given, while in [23] a relationship between the correlation functions for odd and even n is established. In principle, one could use either of these two results to extend Lemma 11 to the odd n case. Once the odd case of Lemma 11 is obtained, Theorem 12 extends automatically to this case. Due to space limitations, we do not attempt to execute this calculation here.

We now turn to applications of Theorem 12. In the complex case, the asymptotics for complex Gaussian matrices given in Lemma 6 could be extended to other independent-entry matrices using Theorem 2, yielding Corollary 7. We now de-

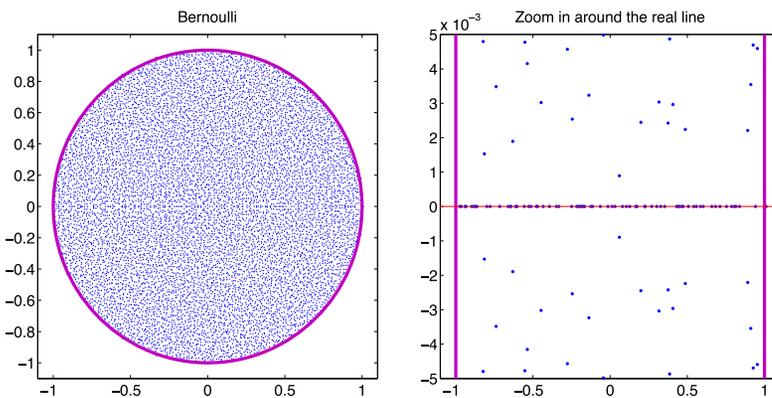


FIG. 3. The spectrum of a random real Bernoulli $10,000 \times 10,000$ matrix, with additional detail near the origin. Thanks to Ke Wang for the data and figure.

velop some analogous results in the real Gaussian case. We first recall the following result of Borodin and Sinclair [7].

LEMMA 14 (Kernel asymptotics, real case). *Let $k, l \geq 0$ be fixed natural numbers, and let z be a fixed complex number. Assume either that $k = 0$, or that z is real. Then there is a function $\rho_{\infty, z}^{(k, l)} : \mathbb{R}^k \times \mathbb{C}^l \rightarrow \mathbb{R}^+$ with the property that one has the pointwise convergence*

$$\begin{aligned} &\rho_n^{(k, l)}(\sqrt{n}z + y_1, \dots, \sqrt{n}z + y_k, \sqrt{n}z + w_1, \dots, \sqrt{n}z + w_l) \\ &\rightarrow \rho_{\infty, z}^{(k, l)}(y_1, \dots, y_k, w_1, \dots, w_l) \end{aligned}$$

as $n \rightarrow \infty$, provided that M_n is drawn from the real Gaussian ensemble and n is restricted to be even.

PROOF. See [6], Section 7, or [7], Section 8. The limit $\rho_{\infty, z}^{(k, l)}$ is explicitly computed in these references, although when z is real the limit is quite complicated, being given in terms of a Pfaffian of a moderately complicated matrix kernel involving the error function erf. However, when z is strictly complex the limit is the same as in the complex Gaussian case, thus $\rho_{\infty, z}^{(0, l)} = \rho_{\infty, z, \dots, z}^{(l)}$; see [7] for further details. It is likely that the same asymptotic also holds for odd n , by using the explicit formulae in [44] or the relation between the odd and even n correlation functions given in [23]; if the restriction to even n is similarly dropped from Lemma 11, then Corollary 15 below can be extended to the odd n case. However, we will not pursue this matter here. \square

We can then obtain the following universality theorem for the correlation functions of real matrices.

COROLLARY 15 (Universality for real matrices). *Let M_n be an independent-entry matrix ensemble with real coefficients obeying condition (C1), and which matches moments with the real Gaussian matrix ensemble to fourth order. Assume n is even. Let $k, l \geq 0$ be fixed natural numbers, and let z be a fixed complex number. Assume either that $k = 0$, or that z is real. Let $F : \mathbb{R}^k \times \mathbb{C}^l \rightarrow \mathbb{R}^+$ be a fixed continuous, compactly supported function. Then*

$$\begin{aligned} &\int_{\mathbb{R}^k} \int_{\mathbb{C}^l_*} F(y_1, \dots, y_k, w_1, \dots, w_l) \\ &\quad \times \rho_n^{(k, l)}(\sqrt{n}z + y_1, \dots, \sqrt{n}z + y_k, \\ &\quad \quad \sqrt{n}z + w_1, \dots, \sqrt{n}z + w_l) dw_1 \cdots dw_l dy_1 \cdots dy_k \\ &\rightarrow \int_{\mathbb{R}^k} \int_{\mathbb{C}^l_*} F(y_1, \dots, y_k, w_1, \dots, w_l) \\ &\quad \times \rho_{\infty, z}^{(k, l)}(y_1, \dots, y_k, w_1, \dots, w_l) dw_1 \cdots dw_l dy_1 \cdots dy_k, \end{aligned}$$

where $\rho_{\infty, x_1, \dots, x_k, z_1, \dots, z_l}^{(k, l)}$ is as in Lemma 14.

PROOF. In the case when M_n is drawn from the real Gaussian ensemble, this follows from Lemmas 14 and 11, and the dominated convergence theorem. The extension to more general independent-entry matrices then follows from Theorem 12 by repeating the arguments used to prove Corollary 7. \square

As in the complex case, Theorem 12 can be used to (partially) extend various-known facts about the distribution of the eigenvalues of a real Gaussian matrices to other real independent-entry matrices. Rather than giving an exhaustive list of such extensions, we illustrate this with two sample applications. Let $N_{\mathbb{R}}(M_n)$ denote the number of real zeroes of a random matrix M_n . Thanks to earlier results [14, 24], we have the following asymptotics.

THEOREM 16 (Real eigenvalues of a real Gaussian matrix). *Let M_n be drawn from the real Gaussian ensemble. Then*

$$\mathbf{E}N_{\mathbb{R}}(M_n) = \sqrt{\frac{2n}{\pi}} + O(1)$$

and

$$\text{Var } N_{\mathbb{R}}(M_n) = (2 - \sqrt{2})\sqrt{\frac{2n}{\pi}} + o(\sqrt{n}).$$

PROOF. The expectation bound was established in [14], and the variance bound in [24]. In fact, more precise asymptotics are available for both the expectation and the variance; we refer the reader to these two papers [14, 24] for further details. \square

By using the above universality results, we may partially extend this result to more general ensembles:

COROLLARY 17 (Real eigenvalues of a real matrix). *Let M_n be an independent-entry matrix ensemble with real coefficients obeying condition (C1), and which matches moments with the real Gaussian matrix ensemble to fourth order. Assume n is even. Then*

$$\mathbf{E}N_{\mathbb{R}}(M_n) = \sqrt{\frac{2n}{\pi}} + O(n^{1/2-c})$$

and

$$\text{Var } N_{\mathbb{R}}(M_n) = O(n^{1-c})$$

for some fixed $c > 0$. In particular, from Chebyshev’s inequality, we have

$$N_{\mathbb{R}}(M_n) = \sqrt{\frac{2n}{\pi}} + O(n^{1/2-c'})$$

with probability $1 - O(n^{-c'})$, for some fixed $c' > 0$.

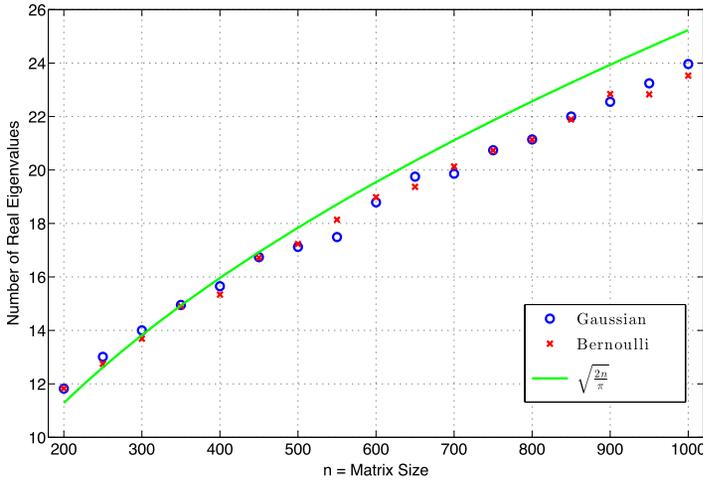


FIG. 4. The empirical average number of real eigenvalues of 200 samples of real Gaussian and real Bernoulli matrices of various sizes, plotted against $\sqrt{\frac{2n}{\pi}}$. Thanks to Ke Wang for the data and figure.

We prove this result in Section 6.3. See Figure 4 for a numerical illustration of this theorem.

As another quick application, we can show for many ensembles that most of the eigenvalues are simple.

COROLLARY 18 (Most eigenvalues simple). *Let M_n be an independent matrix ensemble obeying condition (C1), and which matches moments with the real or complex Gaussian matrix to fourth order. In the real case, assume n is even. Then with probability $1 - O(n^{-c})$, at most $O(n^{1-c})$ of the complex eigenvalues, and $O(n^{1/2-c})$ of the real eigenvalues, are repeated, for some fixed $c > 0$.*

We establish this result in Section 6.3 also. It should in fact be the case that with overwhelming probability, none of the eigenvalues are repeated, but this seems to be beyond the reach of our methods.

2. Key ideas and a sketch of the proof. The proof of the four moment theorem for (Hermitian) Wigner ensembles in [55] is based on the *Lindeberg exchange strategy*, in which one shows that various statistics of ensembles are stable with respect to the swapping of one or two of the coefficients of that ensemble. The original argument in [55] was based on a swapping analysis of individual eigenvalues $\lambda_i(M_n)$, which was somewhat complicated technically; but in [21, 31] it was observed that one could work instead with the simpler swapping analysis of resolvents⁸ (or Green’s functions) $R(z) := (W_n - z)^{-1}$, particularly if one was mainly

⁸Here and in the sequel, we adopt the abbreviation z for the scalar multiple zI of the identity matrix.

focused on obtaining a four moment theorem for correlation functions, rather than for individual eigenvalues (which in any event are not natural to work with in the non-Hermitian case). In all of these arguments for Wigner matrices, a key role was played by the *local semicircle law*, which could in turn be proven by exploiting concentration results for the Stieltjes transform $s(z) := \frac{1}{n} \text{trace}(W_n - z)^{-1}$ of a Wigner matrix. Again, we refer the reader to the preceding surveys for details.

Our strategy of proof of Theorem 2 and Theorem 12 is broadly analogous to that in the Hermitian case, in that it relies on a four moment theorem (Theorem 25 below) and on a local circular law (Theorem 20 below). However, this is highly nontrivial to execute this plan. We are going to need a number of new ideas, coming from different fields of mathematics, and a fair amount of delicate analysis using advanced sharp concentration tools.

To start, there is an essential difference between handling non-Hermitian and Hermitian matrices, namely that the spectrum of a non-Hermitian matrix is highly unstable (see [3] for a discussion). Due to this difficulty, even the (global) circular law, which is the non-Hermitian analogue of Wigner semi-circle law, required several decades of effort to prove, and was solved completely only recently (see the surveys [5, 52] for further discussion). For this reason, it is no longer practical to make the resolvent $(M_n - z)^{-1}$ [and the closely related Stieltjes transform $\frac{1}{n} \text{trace}(M_n - z)^{-1}$] the principal object of study. Instead, following the foundational works of Girko [27] and Brown [10], we shall focus on the *log-determinant*

$$\log|\det(M_n - z)|$$

for a complex number parameter z .

The log-determinant is connected to the eigenvalues of the i.i.d. matrix M_n via the obvious identity

$$(2.1) \quad \log|\det(M_n - z)| = \sum_{i=1}^n \log|\lambda_i(M_n) - z|.$$

In order to restrict to a local region, our idea is to use *Jensen's formula*. Suppose that f is an analytic function in a region in the complex plane which contains the closed disk D of radius r about the origin, a_1, a_2, \dots, a_n are the zeros of f in the interior of D (counting multiplicity), and $f(0) \neq 0$, then

$$\log|f(0)| = \sum_{i=1}^k \log \frac{|a_i|}{r} + \frac{1}{2\pi} \int_0^{2\pi} \log|f(re^{\sqrt{-1}\theta})| d\theta.$$

Applied Jensen's formula to (2.1), we obtain

$$(2.2) \quad \begin{aligned} \log|\det(M_n - z_0)| = & - \sum_{1 \leq i \leq n: \lambda_i(M_n) \in B(z_0, r)} \log \frac{r}{|\lambda_i(M_n) - z_0|} \\ & + \frac{1}{2\pi} \int_0^{2\pi} \log|\det(M_n - z_0 - re^{\sqrt{-1}\theta})| d\theta \end{aligned}$$

for any ball $B(z_0, r)$ (with the convention that both sides are equal to $-\infty$ when z_0 is an eigenvalue of M_n).

From (2.2), we see (in principle, at least) that information on the (joint) distribution of the log-determinants $\log |\det(M_n - z)|$ for various values of z should lead to information on the eigenvalues of M_n , and in particular on the k -point correlation functions $\rho_n^{(k)}$ of M_n . As Jensen formula is a classical tool in complex analysis, this step looks quite robust and would potentially find applications in the study of local properties of many other random processes.

On the other hand, we can also write the log-determinant in terms of the Hermitian $2n \times 2n$ random matrix

$$(2.3) \quad W_{n,z} := \frac{1}{\sqrt{n}} \begin{pmatrix} 0 & M_n - z \\ (M_n - z)^* & 0 \end{pmatrix}$$

via the easily verified identity

$$(2.4) \quad \log |\det(M_n - z)| = \frac{1}{2} \log |\det W_{n,z}| + \frac{1}{2} n \log n.$$

This observation is known as the *Girko Hermitization trick*, and in principle reduces the spectral theory of non-Hermitian matrices to the spectral theory of Hermitian matrices.

The log-determinant of $W_{n,z}$ is in turn related to other spectral information of $W_{n,z}$, such as the *Stieltjes transform*⁹

$$s_{W_{n,z}}(E + \sqrt{-1}\eta) := \frac{1}{2n} \text{trace}((W_{n,z} - E - \sqrt{-1}\eta)^{-1})$$

of $W_{n,z}$, for instance, via the identity

$$(2.5) \quad \log |\det W_{n,z}| = \log |\det(W_{n,z} - \sqrt{-1}T)| - 2n \text{Im} \int_0^T s_{W_{n,z}}(\sqrt{-1}\eta) d\eta,$$

valid for arbitrary $T > 0$. Thus, in principle at least, information on the distribution of the Stieltjes transform $s_{W_{n,z}}$ will imply information on the log-determinant of $W_{n,z}$, and hence on $M_n - z$, which in turn gives information on the eigenvalue distribution of M_n . This is the route taken, for instance, to establish the circular law for i.i.d. matrices; see [5, 52] for further discussion. There is a nontrivial issue with the possible divergence or instability of the integral in (2.5) near $\eta = 0$, but it is now well understood how to control this issue via a regularization or truncation of this integral, provided that one has adequate bounds on the least singular value of $W_{n,z}$; again, see [5, 52] for further discussion. Fortunately, we and many other researchers have proved such bounds in previous papers, using methods from a seemingly unrelated area of Additive Combinatorics (see Proposition 27 below).

⁹We use $\sqrt{-1}$ to denote the standard imaginary unit, in order to free up the symbol i to be an index of summation.

There is a significant technical issue arising from the fact that formulae such as (2.5) or (2.2) require one to control the value of various random functions, such as log-determinants or Stieltjes transforms, for an uncountable number of choices of parameters such as z and η , so that one can no longer directly use union bound to control exceptional events when the expected control on these quantities fails. To overcome this, we appeal to the Monte Carlo method, frequently used in combinatorics and theoretical computer science. This method enables us to use random sampling arguments to replace many of these integral expressions by discrete, random, approximations, to which the union bound can be safely applied (see Section 5).

The application of the Monte Carlo method (Lemma 36), on the other hand, is far from straightforward, since in certain situations (see Section 6), the variance is too high and so the bound implied by Lemma 36 becomes rather weak. We handle this situation by a *variance reduction* argument, exploiting analytical properties of the relevant functions. This step also looks robust and may be useful for practitioners of the Monte Carlo method in other fields.

After these steps, the rest of the proof essentially boils down to error control, in form of a sharp concentration inequality (Theorem 33), which will be done by analyzing a delicate (and rather unstable) random process, using recent martingale inequalities and various ad hoc ideas.

REMARK 19. For Hermitian ensembles, swapping methods (such as the four moment theorem) are not the only way to obtain universality results; there is also an important class of methods (such as the local relaxation flow method) that are based on analysing the effect of a Dyson-type Brownian motion on the spectrum of a random matrix ensemble; see, for example, [15]. However, there is a significant obstruction to adapting such methods to the non-Hermitian setting, because the equations of the analogue to Dyson Brownian motion either¹⁰ couple together the eigenvectors and the eigenvalues in a complicated fashion, or need to be phrased in terms of a triangular form of the matrix, rather than a diagonal one (cf. [35]). We were unable to resolve these difficulties in the non-Hermitian case, and rely solely on swapping methods instead; unfortunately, this then requires us to place moment matching hypotheses on our matrix ensembles. It seems of interest to develop further tools that are able to remove these moment matching hypotheses in non-Hermitian settings.

¹⁰One can explain this by observing that in the Hermitian case, the eigenvalues determine the matrix up to a $U_n(\mathbb{C})$ symmetry, but in the non-Hermitian case the symmetry group is now the much larger group $GL_n(\mathbb{C})$. Dyson Brownian motion is $U_n(\mathbb{C})$ -invariant, but is not $GL_n(\mathbb{C})$ -invariant, which is why this motion can be reduced to dynamics purely on eigenvalues in the Hermitian case but not in the non-Hermitian one.

2.1. *Key propositions.* The proof of Theorem 2 relies on two key facts, both of which may be of independent interest. The first is a “local circular law.” Given a subset Ω of the complex plane, let

$$N_\Omega = N_\Omega[M_n] := |\{1 \leq i \leq n : \lambda_i(M_n) \in \Omega\}|$$

denote the number of eigenvalues of M_n in Ω .

THEOREM 20 (Local circular law). *Let $M_n = (\xi_{ij})_{1 \leq i, j \leq n}$ be an independent-entry matrix with independent real and imaginary parts obeying condition (C1), and which matches either the real or complex Gaussian matrix to third order. Then for any fixed $C > 0$, one has with overwhelming probability¹¹ that*

$$(2.6) \quad N_{B(z_0, r)} = \int_{B(z_0, r)} \frac{1}{\pi} 1_{|z| \leq \sqrt{n}} dz + O(n^{o(1)}r)$$

uniformly for all $z_0 \in B(0, C\sqrt{n})$ and all $r \geq 1$. In particular, we have

$$(2.7) \quad N_{B(z_0, r)} \leq n^{o(1)}r^2$$

with overwhelming probability, uniformly for all $z_0 \in B(0, C\sqrt{n})$ and all $r \geq 1$.

REMARK 21. The bound (2.6) is probably not best possible, even if one ignores the $n^{o(1)}$ term. In the complex Gaussian case, it has been shown [39] that the variance of $N_{B(z_0, r)}$ is actually of order r , suggesting a fluctuation of $O(n^{o(1)}r^{1/2})$ rather than $O(n^{o(1)}r)$; the closely related results in Theorem 9 and Corollary 10 also support this prediction. Also notice that we assume only three matching moments in this theorem, so the statement applies, for instance, to random sign matrices (which match the real Gaussian ensemble to third order). For our applications to Theorems 2, 12, we do not need the full strength (2.6) of the above theorem; the weaker bound (2.7) will suffice.

REMARK 22. Very recently, Bourgade, Yau and Yin [8] have established a variant of Theorem 20 (and also Theorem 25) which does not require matching to third order, but with the disk $B(z_0, r)$ assumed to lie a distance at least $\varepsilon\sqrt{n}$ from the circle $\{|z| = \sqrt{n}\}$ for some fixed $\varepsilon > 0$. By using the main result of [8] as a substitute for Theorem 20 (and also Theorem 25), we may similarly remove the third-order matching hypotheses from Theorem 2, at least in the case when z_1, \dots, z_k stay a distance $\varepsilon\sqrt{n}$ from the circle $\{|z| = \sqrt{n}\}$. Since the initial release of this paper, an alternate proof of Theorem 20 (in the case when one matches the complex Gaussian ensemble to third order, as opposed to the real Gaussian ensemble) which works both in the bulk and in the edge was given in [9].

¹¹See Section 3 for a definition of this term, and for the definition of asymptotic notation such as $o(1)$ and \ll .

The second key fact is a “four moment theorem” for the log-determinants $\log |\det(M_n - z)|$:

THEOREM 23 (Four moment theorem for determinants). *Let $c_0 > 0$ be a sufficiently small absolute constant. Let M_n, M'_n be two independent random matrices with independent real and imaginary parts obeying condition (C1), which match each other to fourth order, and which both match the real Gaussian matrix (or both match the complex Gaussian matrix) to third order. Let $1 \leq k \leq n^{c_0}$, let $C > 0$ be fixed, and let $z_1, \dots, z_k \in B(0, C\sqrt{n})$. Let $G : \mathbb{R}^k \rightarrow \mathbb{C}$ be a smooth function obeying the derivative bounds*

$$|\nabla^j G(x_1, \dots, x_k)| \ll n^{c_0}$$

for all $j = 0, \dots, 5$ and $x_1, \dots, x_k \in \mathbb{R}$, where ∇ denotes the gradient in \mathbb{R}^k . Then we have

$$\begin{aligned} & \mathbf{E}G(\log |\det(M_n - z_1)|, \dots, \log |\det(M_n - z_k)|) \\ &= \mathbf{E}G(\log |\det(M'_n - z_1)|, \dots, \log |\det(M'_n - z_k)|) + O(n^{-c_0}) \end{aligned}$$

with the convention that the expression $G(\log |\det(M_n - z_1)|, \dots, \log |\det(M_n - z_k)|)$ vanishes if one of the z_1, \dots, z_k is an eigenvalue of M_n , and similarly for the expression $G(\log |\det(M'_n - z_1)|, \dots, \log |\det(M'_n - z_k)|)$.

The proof of Theorem 2 follows fairly easily from Theorem 20 [in fact, we will only need the weaker conclusion (2.7)] and Theorem 23 [and (1.10)], using the well-known connection between spectral statistics and the log-determinant which goes back to the work of Girko [27] and Brown [10], and which was mentioned previously in this Introduction; we give this implication in Section 6. A slightly more sophisticated version of the same argument also works to give Theorem 12; we give this implication in Section 6.2.

It remains to establish the local circular law (Theorem 20) and the four moment theorem for log-determinants (Theorem 23). The key lemma in the establishment of the local circular law is the following concentration result for the log-determinant.

DEFINITION 24 (Concentration). Let $n > 1$ be a large parameter, and let X be a real or complex random variable depending on n . We say that X *concentrates around M* for some deterministic scalar M (depending on n) if one has

$$X = M + O(n^{o(1)})$$

with overwhelming probability. Equivalently, for every $\varepsilon, A > 0$ independent of n , one has $X = M + O(n^\varepsilon)$ outside of an event of probability $O(n^{-A})$. We say that X *concentrates* if it concentrates around some M .

THEOREM 25 (Concentration bound on log-determinant). *Let $M_n = (\xi_{ij})_{1 \leq i, j \leq n}$ be an independent-entry matrix obeying condition (C1) and matching the real or complex Gaussian ensemble to third order. Then for any fixed $C > 0$, and any $z_0 \in B(0, C)$, $\log |\det(M_n - z_0 \sqrt{n})|$ concentrates around $\frac{1}{2}n \log n + \frac{1}{2}n(|z_0|^2 - 1)$ for $|z_0| \leq 1$ and around $\frac{1}{2}n \log n + n \log |z_0|$ for $|z_0| \geq 1$, uniformly in z_0 .*

REMARK 26. The reason we require only three moments in this theorem instead of four (as in the previous theorem) is that in this theorem the error in Definition 24 is allowed to be a positive power of n while in the previous one it needs to be a negative power. We remark that this theorem is consistent with (2.1) and the circular law; indeed, the quantity $\int_{B(0,1)} \frac{1}{\pi} \log |z - z_0| dz$ can be computed to be equal to $\frac{1}{2}(|z_0|^2 - 1)$ when $|z_0| \leq 1$ and $\log |z_0|$ when $|z_0| \geq 1$. As in Remark 22, a variant of Theorem 25 without the third order hypothesis, but requiring z_0 bounded away from the circle $\{|z| = 1\}$, was recently given in [8].

We give the derivation of Theorem 20 from Theorem 25 in Section 5. The main tools are Jensen’s formula (2.2) and a random sampling argument to approximate the integral in (2.2) by a Monte Carlo type sum, which can then be estimated by Theorem 25.

It remains to establish Theorem 23 and Theorem 25. For both of these theorems, we will work with the Hermitian matrix $W_{n,z}$ defined in (2.3), taking advantage of the identity (2.4). In order to manipulate quantities such as the log-determinant of $W_{n,z}$ efficiently, we will need some basic estimates on the spectrum of this operator (as well as on related objects, such as resolvent coefficients). We first need a lower bound on the least singular value that is already in the literature.

PROPOSITION 27 (Least singular value). *Let M_n be an independent-entry matrix ensemble with independent real and imaginary parts, obeying condition (C1), and let $z_0 \in B(0, C\sqrt{n})$ for some fixed $C > 0$. Then with overwhelming probability, one has*

$$\inf_{1 \leq i \leq n} |\lambda_i(W_{n,z})| \geq n^{-\log n}.$$

Furthermore, for any fixed $c_0 > 0$ one has

$$\mathbf{P}\left(\inf_{1 \leq i \leq n} |\lambda_i(W_{n,z})| \leq n^{-1/2-c_0}\right) \ll n^{-c_0/2}.$$

The bounds in the tail probability are uniform in z_0 .

PROOF. Note from (2.3) that $\inf_{1 \leq i \leq n} |\lambda_i(W_{n,z})|$ is the least singular value of $\frac{1}{\sqrt{n}}(M_n - z)$. The first bound then follows from [51], Theorem 2.5 (and can also be deduced from the second bound). The lower bound $n^{-\log n}$ can be improved to

any bound decaying faster than a polynomial, but for our applications any lower bound of the form $\exp(-n^{o(1)})$ will suffice. The second bound follows from [54], Theorem 3.2 (and can also be essentially derived from the results in [41], after adapting those results to the case of random matrices whose entries are uncentered [i.e., can have nonzero mean]). We remark that in the z_0 case, significantly sharper bounds can be obtained; see [41] for details. \square

REMARK 28. The proof of this bound relies heavily on the so-called *inverse Littlewood–Offord theory* introduced by the authors in [59], which was motivated by Additive Combinatorics (see [50], Chapter 7), a seemingly unrelated area. Interestingly, this is, at this point, the only way to obtain good lower bound on the least singular values of random matrices when the ensemble is discrete (see also [41, 42, 52] for more results and discussion).

Next, we establish some bounds on the counting function

$$N_I := |\{1 \leq i \leq n : \lambda_i(W_{n,z}) \in I\}|$$

and on coefficients $R(\sqrt{-1}\eta)_{ij}$ of the resolvents $R(\sqrt{-1}\eta) := (W_{n,z} - \sqrt{-1}\eta)^{-1}$ on the imaginary axis.

PROPOSITION 29 (Crude upper bound on N_I). *Let M_n be an independent-entry matrix ensemble with independent real and imaginary parts, obeying condition (C1). Let $C > 0$ be fixed, and let $z_0 \in B(0, C\sqrt{n})$. Then with overwhelming probability, one has*

$$N_I \ll n^{o(1)}(1 + n|I|)$$

for all intervals I . The bounds in the tail probability [and in the $o(1)$ exponent] are uniform in z_0 .

REMARK 30. It is likely that one can strengthen Proposition 29 to a “local distorted quarter-circular law” that gives more accurate upper and lower bounds on N_I , analogous to the local semicircular law from [17–19] (or, for that matter, the local circular law given by Theorem 20). However, we will not need such improvements here.

PROPOSITION 31 (Resolvent bounds). *Let M_n be an independent-entry matrix ensemble with independent real and imaginary parts, obeying condition (C1). Let $C > 0$ be fixed, and let $z_0 \in B(0, C\sqrt{n})$. Then with overwhelming probability, one has*

$$|R(\sqrt{-1}\eta)_{ij}| \ll n^{o(1)}\left(1 + \frac{1}{n\eta}\right)$$

for all $\eta > 0$ and all $1 \leq i, j \leq n$.

REMARK 32. One can also establish similar bounds on the resolvent (as well as closely related *delocalization* bounds on eigenvectors) for more general spectral parameters $E + \sqrt{-1}\eta$. However, in our application we will only need the resolvent bounds in the $E = 0$ case.

Propositions 29 and 31 are proven by standard Stieltjes transform techniques, based on analysis of the self-consistent equation of $W_{n,z}$ as studied, for instance, by Bai [3], combined with concentration of measure results on quadratic forms. The arguments are well established in the literature; indeed, the $z = 0$ case of these theorems essentially appeared in [21, 57], while the analogous estimates for Wigner matrices appeared in [17–19, 55]. As the proofs of these results are fairly routine modifications of existing arguments in the literature, we will place the proof of these propositions in Appendix A. We remark that in the very recent paper [8], some stronger eigenvalue rigidity estimates for $W_{n,z}$ are obtained (at least for z staying away from the unit circle $\{|z| = 1\}$), which among other things allows one to prove variants of Theorem 25 and Theorem 20 without the moment matching hypothesis, and without the need to study the Gaussian case separately (see Theorem 33 below).

One can use Propositions 27, 29, 31 to regularize the log-determinant of $W_{n,z}$, and then show that this log-determinant is quite stable with respect to swapping (real and imaginary parts of) individual entries of the $M_{n,z}$, so long as one keeps the matching moments assumption. In particular, one can now establish Theorem 23 without much difficulty, using standard resolvent perturbation arguments; see Section 8. A similar argument, which we give in Section 10, reduces Theorem 25 to the Gaussian case. Thus, after all these works, the remaining task is to prove:

THEOREM 33. *Theorem 25 holds when M_n is drawn from the real or complex Gaussian ensemble.*

We prove this theorem in Section 9. This section is the most technically involved part of the paper. The starting point is to use an idea from our previous paper [58], which studied the limiting distribution of the log-determinant of a shifted GUE matrix. In that paper, the first step was to conjugate the GUE matrix into the Trotter tridiagonal form [60], so that the log-determinant could be computed in terms of the solution to a certain linear stochastic difference equation. In the case in this paper, the analogue of the Trotter tridiagonal form is a Hessenberg matrix form (i.e., a matrix form which vanishes above the upper diagonal), which (after some linear algebraic transformations) can be used to express the log-determinant $\log |\det(M_n - z_0\sqrt{n})|$ in terms of the solution to a certain *nonlinear* stochastic difference equation. This Hessenberg form of the complex Gaussian ensemble was introduced in [33], although the difference equation we derive is different from the one used in that paper. To obtain the desired level of concentration in the log-determinant, the main difficulty is then to satisfactorily control the interplay

between the diffusive components of this stochastic difference equation, and the stable and unstable equilibria of the nonlinearity, and in particular to show that the deviation of the solution from the stable equilibrium behaves like a martingale. This then allows us to deduce the desired concentration from a martingale concentration result (see Proposition 35 below).

3. Notation. Throughout this paper, n is a natural number parameter going off to infinity. A quantity is said to be *fixed* if it does not depend on n . We write $X = O(Y)$, $X \ll Y$, $Y = \Omega(X)$, or $Y \gg X$ if one has $|X| \leq CY$ for some fixed C , and $X = o(Y)$ if one has $X/Y \rightarrow 0$ as $n \rightarrow \infty$. Absolute constants such as C_0 or c_0 are always understood to be fixed.

We say that an event E occurs with *overwhelming probability* if it occurs with probability $1 - O(n^{-A})$ for all fixed $A > 0$. We use 1_E to denote the indicator of E , thus 1_E equals 1 when E is true and 0 when E is false. We also write $1_{\Omega}(x)$ for $1_{x \in \Omega}$.

As we will be using two-dimensional integration on the complex plane $\mathbb{C} := \{z = x + \sqrt{-1}y : x, y \in \mathbb{R}\}$ far more often than we will be using contour integration, we use $dz = dx dy$ to denote Lebesgue measure on the complex numbers, rather than the complex line element $dx + \sqrt{-1} dy$.

We use $N(\mu, \sigma^2)_{\mathbb{R}}$ to denote a real Gaussian distribution of mean μ and variance σ^2 , so that the probability distribution is given by $\frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x-\mu)^2/2\sigma^2} dx$. Similarly, we let $N(\mu, \sigma^2)_{\mathbb{C}}$ denote the complex Gaussian distribution of μ and variance σ^2 , so that the probability distribution is given by $\frac{1}{\pi\sigma^2} e^{-|z-\mu|^2/\sigma^2} dz$. Of course, the two distributions are closely related: the real and imaginary parts of $N(\mu, \sigma^2)_{\mathbb{C}}$ are independent copies of $N(\text{Re } \mu, \sigma^2/2)_{\mathbb{R}}$ and $N(\text{Im } \mu, \sigma^2/2)_{\mathbb{R}}$, respectively. In a similar spirit, for any natural number, we use $\chi_{i, \mathbb{R}}$ to denote the real χ distribution with i degrees of freedom, thus $\chi_{i, \mathbb{R}} \equiv \sqrt{\xi_1^2 + \dots + \xi_i^2}$ for independent copies ξ_1, \dots, ξ_i of $N(0, 1)_{\mathbb{R}}$. Similarly, we use $\chi_{i, \mathbb{C}}$ to denote the complex χ distribution with i degrees of freedom, thus $\chi_{i, \mathbb{C}} \equiv \sqrt{\xi_1^2 + \dots + \xi_i^2}$ for independent copies ξ_1, \dots, ξ_i of $N(0, 1)_{\mathbb{C}}$. Again, the two distributions are closely related: one has $\chi_{i, \mathbb{C}} \equiv \frac{1}{\sqrt{2}} \chi_{2i, \mathbb{R}}$ for all i .

If $F: \mathbb{C}^k \rightarrow \mathbb{C}$ is a smooth function, we use $\nabla F(z_1, \dots, z_k)$ to denote the $2k$ -dimensional vector whose components are the partial derivatives $\frac{\partial F}{\partial \text{Re } z_i}(z_1, \dots, z_k)$, $\frac{\partial F}{\partial \text{Im } z_i}(z_1, \dots, z_k)$ for $i = 1, \dots, k$. Iterating this, we can define $\nabla^a F(z_1, \dots, z_k)$ for any natural number a as a tensor with $(2k)^a$ coefficients, each of which is an a -fold partial derivative of F at z_1, \dots, z_k . The magnitude $|\nabla^a F(z_1, \dots, z_k)|$ is then defined as the ℓ^2 norm of these coefficients; similarly for functions defined on \mathbb{R}^k instead of \mathbb{C}^k .

4. A concentration inequality. In this section, we recall a martingale type concentration inequality which will be useful in our arguments. Let $Y =$

$Y(\xi_1, \dots, \xi_n)$ be a random variable depending on independent atom variables $\xi_i \in \mathbb{C}$. For $1 \leq i \leq n$ and $\xi = (\xi_1, \dots, \xi_n) \in \mathbb{C}^n$, define the martingale differences

$$(4.1) \quad C_i(\xi) := |\mathbf{E}(Y|\xi_1, \dots, \xi_i) - \mathbf{E}(Y|\xi_1, \dots, \xi_{i-1})|.$$

The classical Azuma’s inequality (see, e.g., [2]) states that if $C_i \leq \alpha_i$ with probability one, then

$$\mathbf{P}\left(|Y - \mathbf{E}Y| \geq \lambda \sqrt{\sum_{i=1}^n \alpha_i^2}\right) = O(\exp(-\Omega(\lambda^2))).$$

In applications, the assumption that $C_i \leq \alpha_i$ with probability one sometimes fails. However, we can overcome this using a trick from [61]. In particular, the following is a simple variant of [61], Lemma 3.1.

PROPOSITION 34. *For any $\alpha_i \geq 0$, we have the inequality*

$$\mathbf{P}\left(|Y - \mathbf{E}Y| \geq \lambda \sqrt{\sum_{i=1}^n \alpha_i^2}\right) = O(\exp(-\Omega(\lambda^2))) + \sum_{i=1}^n \mathbf{P}(C_i(\xi) \geq \alpha_i).$$

PROOF. For each ξ , let i_ξ be the first index where $C_i(\xi) \geq \alpha_i$. Thus, the sets $B_i := \{\xi | i_\xi = i\}$ are disjoint. Define a function $Y'(\xi)$ of ξ which agrees with $Y(\xi)$ for ξ in the complement of $\cup_i B_i$, with $Y'(\xi) := \mathbf{E}_{B_i} Y$ if $\xi \in B_i$. It is clear that Y' and Y has the same mean and

$$\mathbf{P}(Y \neq Y') \leq \sum_{i=1}^n \mathbf{P}(C_i(\xi) \geq \alpha_i).$$

Moreover, Y' satisfies the condition of Azuma’s inequality, so

$$\mathbf{P}\left(|Y' - \mathbf{E}Y'| \geq \lambda \sqrt{\sum_{i=1}^n \alpha_i^2}\right) \ll \exp(-\Omega(\lambda^2))$$

and the bound follows. \square

We have the following useful corollary.

PROPOSITION 35 (Martingale concentration). *Let ξ_1, \dots, ξ_n be independent complex random variables of mean zero and $|\xi_i| = n^{o(1)}$ with overwhelming probability for all i . Let $\alpha_1, \dots, \alpha_n > 0$ be positive real numbers, and for each $i = 1, \dots, n$, let $c_i(\xi_1, \dots, \xi_{i-1})$ be a complex random variable depending only on ξ_1, \dots, ξ_{i-1} obeying the bound*

$$|c_i(\xi_1, \dots, \xi_{i-1})| \leq \alpha_i$$

with overwhelming probability. Define $Y := \sum_{i=1}^n c_i(\xi_1, \dots, \xi_{i-1})\xi_i$. Then

$$|Y| \ll n^{o(1)} \left(\sum_{i=1}^n \alpha_i^2 \right)^{1/2}$$

with overwhelming probability.

PROOF. Let $C_i(\xi)$ be the martingale difference (4.1). It is easy to see that $C_i(\xi) = |c_i(\xi_1, \dots, \xi_{i-1})\xi_i|$. By the assumptions, $C_i(\xi) \leq n^{o(1)}\alpha_i$ with overwhelming probability. Now apply Proposition 34 with a suitable choice of parameter $\lambda = n^{o(1)}$. \square

5. From log-determinant concentration to the local circular law. In this section, we prove Theorem 20 using Theorem 25. The first step is to deduce the crude bound (2.7) from Theorem 25. We first make some basic reductions. By a covering argument and the union bound it suffices to establish the claim for $r = 1$ and for a fixed $z_0 \in B(0, 2C\sqrt{n})$.

The main tool will be Jensen’s formula (2.2). Applying this to the disk $B(z_0, 2)$, we see in particular that

$$(5.1) \quad N_{B(z_0,1)} \ll \frac{1}{2\pi} \int_0^{2\pi} (\log|\det(M_n - z_0 - 2e^{\sqrt{-1}\theta})| - \log|\det(M_n - z_0)|) d\theta.$$

Let $A \geq 1$ be an arbitrary fixed quantity. In view of (5.1), it suffices to show that

$$\frac{1}{2\pi} \int_0^{2\pi} (\log|\det(M_n - z_0 - 2e^{\sqrt{-1}\theta})| - \log|\det(M_n - z_0)|) d\theta = O(n^\varepsilon)$$

with probability $1 - O(n^{-A})$.

We will control this integral¹² by a Monte Carlo sum, using the following standard sampling lemma.

LEMMA 36 (Monte Carlo sampling lemma). *Let (X, μ) be a probability space, and let $F : X \rightarrow \mathbb{C}$ be a square-integrable function. Let $m \geq 1$, let x_1, \dots, x_m be drawn independently at random from X with distribution μ , and let S be the empirical average*

$$S := \frac{1}{m} (F(x_1) + \dots + F(x_m)).$$

¹²One can also control this integral by a Riemann sum, using an argument similar to that used to prove Theorem 20 below. On the other hand, we will use Lemma 36 again in Section 6, and one can view the arguments below as a simplified warmup for the more complicated arguments in that section.

Then S has mean $\int_X F d\mu$ and variance $\int_X (F - \int_X F d\mu)^2 d\mu$. In particular, by Chebyshev's inequality, one has

$$\mathbf{P}\left(\left|S - \int_X F d\mu\right| \geq \lambda\right) \leq \frac{1}{m\lambda^2} \int_X \left(F - \int_X F d\mu\right)^2 d\mu$$

for any $\lambda > 0$, or equivalently, for any $\delta > 0$ one has with probability at least $1 - \delta$ that

$$\left|S - \int_X F d\mu\right| \leq \frac{1}{\sqrt{m\delta}} \left(\int_X \left(F - \int_X F d\mu\right)^2 d\mu\right)^{1/2}.$$

PROOF. The random variables $F(x_i)$ for $i = 1, \dots, m$ are jointly independent with mean $\int_X F d\mu$ and variance $\frac{1}{m} \int_X (F - \int_X F d\mu)^2 d\mu$. Averaging these variables, we obtain the claim. \square

We apply this lemma to the probability space $X := [0, 2\pi]$ with uniform measure $\frac{1}{2\pi} d\theta$, and to the function

$$F(\theta) := \log|\det(M_n - z_0 - 2e^{\sqrt{-1}\theta})| - \log|\det(M_n - z_0)|.$$

Observe that for any complex number z , the function $\log|z - 2e^{\sqrt{-1}\theta}|$ has an $L^2(X)$ norm of $O(1)$. Thus, by the triangle inequality and (2.1), we have the crude bound

$$\int_X \left(F - \int_X F d\mu\right)^2 d\mu \ll n^2.$$

We set $\delta := n^{-A}$ and $m := n^{A+2}$. Let $\theta_1, \dots, \theta_m$ be drawn independently uniformly at random from X (and independently of M_n) and set $\Theta := (\theta_1, \dots, \theta_m)$. Let \mathcal{E}_1 denote the event that the inequality

$$\left|S - \int_X F d\mu\right| \leq \frac{1}{\sqrt{m\delta}} \left(\int_X \left(F - \int_X F d\mu\right)^2 d\mu\right)^{1/2}$$

holds, and let \mathcal{E}_2 denote the event that the inequality

$$|\log|\det(M - z_0 - 2e^{\sqrt{-1}\theta_j})| - \log|\det(M_n - z_0)|| \leq n^\epsilon$$

holds for all $j = 1, \dots, m$. Call a pair (M, Θ) is good if \mathcal{E}_1 and \mathcal{E}_2 both hold. It suffices to show that the probability that a pair (M, Θ) (with $M = M_n$) is good is $1 - O(n^{-A})$.

By Lemma 36, for each fixed M , the probability that \mathcal{E}_1 fails is at most $\delta = n^{-A}$. Moreover, by Theorem 25, we see that for each fixed θ_i , the probability that $|\log|\det(M - z_0 - 2e^{\sqrt{-1}\theta_j})| - \log|\det(M_n - z_0)|| \leq n^\epsilon$ fails is less than $O(n^{-2A-2})$. Thus, by the union bound, the probability that (M, Θ) is not good (over the product space $M_n \times X^m$) is at most

$$n^{-A} + m \times O(n^{-2A-2}) = O(n^{-A}),$$

concluding the proof of (2.7).

Now we are ready to prove Theorem 20. We assume $r \geq 10$ as the claim follows trivially from Theorem 25 otherwise. Consider the circle $C_{z_0,r} := \{z \in \mathbb{C} : |z - z_0| = r\}$. By the pigeonhole principle, there is some $0 \leq j \leq n$ such that the $\frac{1}{n^3}$ -neighborhood of the circle $C_j := C_{z_0,r_j}$ with $r_j := r - \frac{j}{n^2}$ contains no eigenvalues of M_n (notice that these neighborhoods are disjoint). If j is such an index, we see from (2.1) that the function

$$F(\theta) := \log|\det(M_n - z_0 - r_j e^{-\sqrt{-1}\theta})| - \log|\det(M_n - z_0)|$$

then has a Lipschitz norm of $O(n^{O(1)})$ on $[0, 2\pi]$. Setting $m := n^{A+2}$ for a sufficiently large constant A , we then see from quadrature that the Riemann sum $\frac{1}{m} \sum_{k=1}^m F(2\pi k/m)$ approximates the integral $\frac{1}{2\pi} \int_0^{2\pi} F(\theta) d\theta$ within an additive error at most $n^{o(1)}$. By (2.2), we conclude that

$$\sum_{|\lambda_i - z_0| < r_j} \log \frac{r_j}{|\lambda_i - z_0|} = \frac{1}{m} \sum_{k=1}^m F(k/m) + O(n^{o(1)}).$$

On the other hand, from Theorem 25 (after applying rescaling by \sqrt{n}) and the union bound we see that with overwhelming probability, we have

$$F(k/m) = G(z_0 + r_j e^{\sqrt{-1}2\pi k/m}) - G(z_0) + O(n^{o(1)})$$

for all $1 \leq k \leq m$, where $G(z)$ is defined as $\frac{1}{2}(|z|^2 - n)$ for $|z| \leq \sqrt{n}$, and $n \log \frac{|z|}{\sqrt{n}}$ for $|z| \geq \sqrt{n}$. Applying quadrature again, we conclude (for A large enough) that

$$G(z_0) = - \sum_{|\lambda_i - z_0| < r_j} \log \frac{r_j}{|\lambda_i - z_0|} + \frac{1}{2\pi} \int_0^{2\pi} G(z_0 + r_j e^{\sqrt{-1}\theta}) d\theta + O(n^{o(1)}).$$

A similar argument (replacing r by $r - 1$) shows that with overwhelming probability, there exists $0 \leq j' \leq n$ such that

$$\begin{aligned} G(z_0) = & - \sum_{|\lambda_i - z_0| < r_{j'} - 1} \log \frac{r_{j'} - 1}{|\lambda_i - z_0|} \\ & + \frac{1}{2\pi} \int_0^{2\pi} G(z_0 + (r_{j'} - 1)e^{\sqrt{-1}\theta}) d\theta + O(n^{o(1)}). \end{aligned}$$

Also, from (2.7) and a simple covering argument, we know that with overwhelming probability, there are at most $O(n^{o(1)}r)$ eigenvalues in the annular region between $C_{z_0,r_{j'}-1}$ and $C_{z_0,r}$, and in this region, the quantities $\log \frac{r_j}{|\lambda_i - z_0|}$ and $\log \frac{r_{j'}-1}{|\lambda_i - z_0|}$ have magnitude $O(1/r)$. We may thus subtract the above two estimates and conclude that

$$\begin{aligned} 0 = & -N(z_0, r) \log \frac{r_j}{r_{j'} - 1} + \frac{1}{2\pi} \int_0^{2\pi} G(z_0 + r_j e^{\sqrt{-1}\theta}) d\theta \\ (5.2) \quad & - \frac{1}{2\pi} \int_0^{2\pi} G(z_0 + (r_{j'} - 1)e^{\sqrt{-1}\theta}) d\theta + O(n^{o(1)}). \end{aligned}$$

On the other hand, from applying Green’s theorem¹³

$$\int_{\Omega} F(z)\Delta G(z) - \Delta G(z)F(z) dz = \int_{\partial\Omega} F(z)\frac{\partial}{\partial n}G(z) - \frac{\partial}{\partial n}F(z)G(z)$$

to the domain $\Omega := B(z_0, r_j) \setminus B(z_0, \varepsilon)$ with $F(z) := \log \frac{r_j}{|z-z_0|}$, and then sending $\varepsilon \rightarrow 0$, one sees that

$$G(z_0) = -\frac{1}{2\pi} \int_{B(z_0, r_j)} \Delta G(z) \log \frac{r_j}{|z-z_0|} dz + \frac{1}{2\pi} \int_0^{2\pi} G(z_0 + r_j e^{\sqrt{-1}\theta}) d\theta,$$

where Δ is the usual Laplacian on \mathbb{C} ; one easily computes that $\Delta G(z) = 21_{|z|\leq\sqrt{n}}$, and thus

$$G(z_0) = -\frac{1}{\pi} \int_{B(z_0, r_j)} 1_{|z|\leq\sqrt{n}} \log \frac{r_j}{|z-z_0|} dz + \frac{1}{2\pi} \int_0^{2\pi} G(z_0 + r_j e^{\sqrt{-1}\theta}) d\theta.$$

Similarly, one has

$$\begin{aligned} G(z_0) &= -\frac{1}{\pi} \int_{B(z_0, r_{j'}-1)} 1_{|z|\leq\sqrt{n}} \log \frac{r_{j'}-1}{|z-z_0|} dz \\ &\quad + \frac{1}{2\pi} \int_0^{2\pi} G(z_0 + (r_{j'}-1)e^{\sqrt{-1}\theta}) d\theta. \end{aligned}$$

Subtracting, and observing that the integrands $1_{|z|\leq\sqrt{n}} \log \frac{r_j}{|z-z_0|}$, $1_{|z|\leq\sqrt{n}} \log \frac{r_{j'}-1}{|z-z_0|}$ have magnitude $O(1/r)$ in the annular region between $C_{z_0, r_{j'}-1}$ and $C_{z_0, r}$, we conclude that

$$\begin{aligned} 0 &= -\int_{B(z_0, r)} \frac{1}{\pi} 1_{|z|\leq\sqrt{n}} dz \times \log \frac{r_j}{r_{j'}-1} + \frac{1}{2\pi} \int_0^{2\pi} G(z_0 + r_j e^{\sqrt{-1}\theta}) d\theta \\ &\quad - \frac{1}{2\pi} \int_0^{2\pi} G(z_0 + (r_{j'}-1)e^{\sqrt{-1}\theta}) d\theta + O(n^{o(1)}). \end{aligned}$$

Comparing this with (5.2), we conclude with overwhelming probability that

$$\left(N_{B(z_0, r)} - \int_{B(z_0, r)} \frac{1}{\pi} 1_{|z|\leq\sqrt{n}} dz \right) \times \log \frac{r_j}{r_{j'}-1} = O(n^{o(1)}).$$

Since $\log \frac{r_j}{r_{j'}-1}$ is comparable to $1/r$, we obtain (2.6) as desired.

6. Reduction to the four moment theorem and log-determinant concentration. We now begin the task of proving Theorem 2 and Theorem 12, by reducing it the four moment theorem for determinants (Theorem 23) and the local circular law (Proposition 20). In the preceding section, of course, the local circular law has been reduced in turn to the concentration of the log-determinant (Theorem 25).

¹³The function G has a mild singularity on the circle $|z| = \sqrt{n}$, but one can verify that as the first derivatives of G remain continuous across this circle, there is no difficulty in applying Green’s theorem even when $B(z_0, r_j)$ crosses this circle.

6.1. *The complex case.* We begin with Theorem 2, deferring the slightly more complicated argument for Theorem 12 to the end of this section.

Let M_n, \widetilde{M}_n be as in Theorem 2. Call a statistic $S(M_n)$ of (the law of) a random matrix M_n asymptotically (M_n, \widetilde{M}_n) insensitive, or insensitive for short, if we have

$$S(M_n) - S(\widetilde{M}_n) = O(n^{-c})$$

for some fixed $c > 0$. Our objective is then to show that the statistic

$$(6.1) \quad \int_{\mathbb{C}^k} F(w_1, \dots, w_k) \rho_n^{(k)}(\sqrt{n}z_1 + w_1, \dots, \sqrt{n}z_k + w_k) dw_1 \cdots dw_k$$

is insensitive for all fixed $k \geq 1$ and all F of the form (1.8) for some fixed $m \geq 1$.

Fix k ; we may assume inductively that the claim has already been proven for all smaller k . By linearity we may take $m = 1$, thus we may assume that F takes the tensor product form

$$(6.2) \quad F(w_1, \dots, w_k) = F_1(w_1) \cdots F_k(w_k)$$

for some smooth, compactly supported $F_1, \dots, F_k : \mathbb{C} \rightarrow \mathbb{C}$ supported on a fixed ball, with bounds on derivatives up to second order.

Henceforth, we assume that F is in tensor product form (6.2). By (1.1) and the inclusion–exclusion formula, we may thus write (6.1) in this case as

$$(6.3) \quad \mathbf{E} \prod_{j=1}^k X_{z_j, F_j}$$

plus a fixed finite number of lower order terms that are of the form (6.1) for a smaller value of k (and a different choice of F_j), where X_{z_j, F_j} is the linear statistic

$$X_{z_j, F_j} := \sum_{i=1}^n F_j(\lambda_i(M_n) - \sqrt{n}z_j).$$

By the induction hypothesis, it thus suffices to show that the expression (6.3) is insensitive.

Using the local circular law (Proposition 20), we see that for any $1 \leq j \leq k$, one has $X_{z_j, F_j} = O(n^{o(1)})$ with overwhelming probability. Thus, one can truncate the product function $\zeta_1, \dots, \zeta_k \mapsto \zeta_1 \cdots \zeta_k$ and write

$$\mathbf{E} \prod_{j=1}^k X_{z_j, F_j} = \mathbf{E} G(X_{z_1, F_1}, \dots, X_{z_k, F_k}) + O(n^{-B})$$

for any fixed B , where G is a smooth truncation of the product function $\zeta_1, \dots, \zeta_k \mapsto \zeta_1 \cdots \zeta_k$ to the region $\zeta_1, \dots, \zeta_k = n^{o(1)}$. Thus, it suffices to show that the quantity

$$(6.4) \quad \mathbf{E} G(X_{z_1, F_1}, \dots, X_{z_k, F_k})$$

is insensitive whenever $G : \mathbb{C}^k \rightarrow \mathbb{C}$ is a smooth function obeying the bounds

$$(6.5) \quad |\nabla^j G(\zeta_1, \dots, \zeta_k)| \leq n^{o(1)}$$

for all fixed j and all $\zeta_1, \dots, \zeta_k \in \mathbb{C}$.

Fix G . As is standard in the spectral theory of random non-Hermitian matrices (cf. [10, 27]), we now express the linear statistics X_{z_j, F_j} in terms of the log-determinant (2.1). By Green’s theorem, we have

$$(6.6) \quad X_{z_j, F_j} = \int_{\mathbb{C}} \log |\det(M_n - z)| H_j(z) dz,$$

where $H_j : \mathbb{C} \rightarrow \mathbb{C}$ is the function

$$H_j(z) := -\frac{1}{2\pi} \Delta F_j(z - \sqrt{n}z_j)$$

and Δ is the Laplacian on \mathbb{C} . From the derivative and support bounds on F_j , we see that H_j is supported on $B(\sqrt{n}z_j, C)$ and is bounded.

Naively, to control (6.6), one would apply Lemma 36 with the function $\log |\det(M_n - z)| H_j(z)$. Unfortunately, the variance of this expression is too large, due to the contributions of the eigenvalues far away from $\sqrt{n}z_j$. To cancel¹⁴ off these contributions, we exploit the fact that $H_j(z)$, being the Laplacian of a smooth compactly supported function, is orthogonal to all harmonic functions, and in particular to all (real-)linear functions:

$$\int_{\mathbb{C}} (a + b \operatorname{Re}(z) + c \operatorname{Im}(z)) H_j(z) dz = 0.$$

(Recall that we use dz to denote Lebesgue measure on \mathbb{C} .) We will need a reference element $w_{j,0}$ drawn uniformly at random from $B(\sqrt{n}z_j, 1)$ (independently of M_n and the $w_{j,i}$), and let $L(z) = L_j(z)$ denote the random linear function which equals $\log |\det(M_n - z)|$ for $z = w_{j,0}, w_{j,0} + 1, w_{j,0} + \sqrt{-1}$. More explicitly, one has

$$(6.7) \quad \begin{aligned} L(z) := & \log |\det(M_n - w_{j,0})| \\ & + (\log |\det(M_n - w_{j,0} - 1)| - \log |\det(M_n - w_{j,0})|) \operatorname{Re}(z - w_{j,0}) \\ & + (\log |\det(M_n - w_{j,0} - \sqrt{-1})| - \log |\det(M_n - w_{j,0})|) \\ & \times \operatorname{Im}(z - w_{j,0}). \end{aligned}$$

REMARK 37. There is some freedom in how to select $L(z)$; for instance, it is arguably more natural to replace the coefficients $\log |\det(M_n - w_{j,0} - 1)| - \log |\det(M_n - w_{j,0})|$ and $\log |\det(M_n - w_{j,0} - \sqrt{-1})| - \log |\det(M_n - w_{j,0})|$ in the above formula by the Taylor coefficients $\frac{d}{dt} \log |\det(M_n - w_{j,0} - t)| |_{t=0}$

¹⁴It is natural to expect that these nonlocal contributions can be canceled, since the statistics X_{z_i, F_i} are clearly local in nature.

and $\frac{d}{dt} \log |\det(M_n - w_{j,0} - \sqrt{-1}t)|_{t=0}$ instead. However, this would require extending the four moment theorem for log-determinants to derivatives of log-determinants, which can be done but will not be pursued here.

Subtracting off $L(z)$, we have

$$(6.8) \quad X_{z_j, F_j} = \int_{\mathbb{C}} K_j(z) dz,$$

where $K_j : \mathbb{C} \rightarrow \mathbb{C}$ is the random function

$$(6.9) \quad K_j(z) := (\log |\det(M_n - z)| - L(z)) H_j(z).$$

Let us control the L^2 norm

$$\|K_j\|_{L^2} := \left(\int_{\mathbb{C}} |K_j(z)|^2 dz \right)^{1/2}$$

of this quantity.

LEMMA 38. *For any $\varepsilon > 0$, one has*

$$(6.10) \quad \|K_j\|_{L^2} \ll n^{\varepsilon+o(1)}$$

with probability $1 - O(n^{-\varepsilon})$ and all $1 \leq j \leq k$.

PROOF. By the union bound, it suffices to prove the claim for a single k . We can split $K_j = \sum_{i=1}^n K_{j,i}(z)$, where

$$K_{j,i}(z) := (\log |\lambda_i(M_n) - z| - L_i(z)) H_j(z)$$

and $L_i : \mathbb{C} \rightarrow \mathbb{C}$ is the random linear function that equals $\log |\lambda_i(M_n) - z|$ when $z = w_{j,0}, w_{j,0} + 1, w_{j,0} + \sqrt{-1}$. By the triangle inequality, we thus have

$$\|K_j\|_{L^2} \leq \sum_{i=1}^n \|K_{j,i}\|_{L^2}.$$

Thanks to Proposition 20, we know with overwhelming probability that one has

$$(6.11) \quad N_{B(z_j \sqrt{n}, r)} \ll n^{o(1)} r^2$$

for all r . Let us condition on the event that this holds, and then freeze M_n (so that the only remaining source of randomness is $w_{j,0}$). In particular, the eigenvalues $\lambda_i(M_n)$ are now deterministic.

Let $C_0 > 1$ be such that H_j is supported in $B(z_0 \sqrt{n}, C_0)$. If $1 \leq i \leq n$ is such that $\lambda_i(M_n) \in B(z_j \sqrt{n}, 2C_0)$, then a short computation (based on the square-integrability of the logarithm function) shows that the expected value of $\|K_{j,i}\|_{L^2}$ (averaged over all choices of $w_{j,0}$) is $O(1)$. On the other hand, if $\lambda_i(M_n) \notin B(z_j \sqrt{n}, 2C_0)$, then the second derivatives of $\log |\lambda_i(M_n) - z|$ has size $O(1/|\lambda_i(M_n) - z_j \sqrt{n}|^2)$ on $B(z_j \sqrt{n}, 2C_0)$. From this and Taylor expansion, one

sees that the function $\log |\lambda_i(M_n) - z| - L_i(z)$ has magnitude $O(1/|\lambda_i(M_n) - z_j\sqrt{n}|^2)$ on this ball, and so $\|K_{j,i}\|_{L^2}$ has this size as well. Summing, we conclude that the (conditional) expected value of $\|K_j\|_{L^2}$ is at most

$$(6.12) \quad \ll \sum_{i=1}^n \frac{1}{1 + |\lambda_i(M_n) - z_j\sqrt{n}|^2}.$$

We claim that the summation in (6.12) has magnitude $O(n^{o(1)})$ with overwhelming probability, which will give the claim from Markov’s inequality. To see this, first observe that the eigenvalues $\lambda_i(M_n)$ with $|\lambda_i(M_n) - z_j\sqrt{n}| \geq \sqrt{n}$ certainly contribute at most $O(1)$ in total to the above sum. Next, from (6.11) we see that with overwhelming probability that there are only $O(n^{o(1)})$ eigenvalues with $|\lambda_i(M_n) - z_j\sqrt{n}| \leq 1$, giving another contribution of $O(n^{o(1)})$ to the above sum. Similarly, for any 2^k between 1 and \sqrt{n} , another application of (6.11) reveals that the eigenvalues with $2^k \leq |\lambda_i(M_n) - z_j\sqrt{n}| < 2^{k+1}$ contribute another term of $O(n^{o(1)})$ to the above sum with overwhelming probability. As there are only $O(\log \sqrt{n}) = O(n^{o(1)})$ possible choices for k , the claim then follows by summing all the contributions estimated above. \square

Now let $\varepsilon > 0$ be a sufficiently small fixed constant that will be chosen later. Set $m := \lfloor n^{10\varepsilon} \rfloor$, and for each $1 \leq j \leq k$ let $w_{j,1}, \dots, w_{j,m}$ be drawn uniformly at random from $B(\sqrt{n}z_j, C_0)$ (independently of M_n and $w_{j,0}$). By (6.10), (6.8) and Lemma 36, we see that with probability $1 - O(n^{-\varepsilon})$, one has

$$X_{z_j, F_j} = \frac{\pi C_0^2}{m} \sum_{i=1}^m K_j(w_{j,i}) + O(n^{-3\varepsilon+o(1)}).$$

In particular, from (6.5) we see that with probability $1 - O(n^{-\varepsilon})$, one has

$$G(X_{z_1, F_1}, \dots, X_{z_k, F_k}) = G\left(\left(\frac{\pi C_0^2}{m} \sum_{i=1}^m K_j(w_{j,i})\right)_{1 \leq j \leq k}\right) + O(n^{-3\varepsilon+o(1)})$$

and hence

$$\mathbf{E}G(X_{z_1, F_1}, \dots, X_{z_k, F_k}) = \mathbf{E}G\left(\left(\frac{\pi C_0^2}{m} \sum_{i=1}^m K_j(w_{j,i})\right)_{1 \leq j \leq k}\right) + O(n^{-\varepsilon+o(1)}).$$

Thus, to show that (6.4) is insensitive, it suffices to show that

$$\mathbf{E}G\left(\left(\frac{\pi C_0^2}{m} \sum_{i=1}^m K_j(w_{j,i})\right)_{1 \leq j \leq k}\right)$$

is insensitive, uniformly for all deterministic choices of $w_{j,0} \in B(\sqrt{n}z_j, 1)$ and $w_{j,i} \in B(\sqrt{n}z_j, C_0)$ for $1 \leq j \leq k$ and $1 \leq i \leq m$. But this follows from the four moment theorem (Theorem 23), if ε is small enough; indeed, once the $w_{j,0}, w_{j,i}$

are conditioned to be deterministic, we see from (6.9), (6.7) that the quantities $K_j(w_{j,i})$ can be expressed as deterministic linear combinations of a bounded number of log-determinants $\log |\det(M_n - z)|$, with coefficients uniformly bounded in n [recall that $w_{j,i} - w_{j,0} = O(C_0)$ and that the H_j are uniformly bounded]. This concludes the derivation of Theorem 2 from Theorem 23 and Proposition 20.

6.2. *The real case.* We now turn to the proof of Theorem 12. Let M_n be as in Theorem 12, and let \tilde{M}_n be a real Gaussian matrix. Our task is to show that that the quantity

$$(6.13) \quad \int_{\mathbb{R}^k} \int_{\mathbb{C}^m} F(y_1, \dots, y_k, w_1, \dots, w_l) \times \rho_n^{(k,l)}(\sqrt{n}x_1 + y_1, \dots, \sqrt{n}x_k + y_k, \sqrt{n}z_1 + w_1, \dots, \sqrt{n}z_l + w_l) dw_1 \cdots dw_l dy_1 \cdots dy_k$$

is insensitive whenever $k, l \geq 0$ are fixed, $x_1, \dots, x_k \in \mathbb{R}$ and $z_1, \dots, z_l \in \mathbb{C}$ are bounded, and F decomposes as in Theorem 12.

By induction on $k + l$, much as in the complex case, and separating the spectrum into contributions from $\mathbb{R}, \mathbb{C}_+, \mathbb{C}_-$, it thus suffices to show that the quantity

$$(6.14) \quad \mathbf{E} \left(\prod_{i=1}^k X_{x_i, F_i, \mathbb{R}} \right) \left(\prod_{j=1}^l X_{z_j, G_j, \mathbb{C}_+} \right) \left(\prod_{j'=1}^{l'} X_{z'_{j'}, G'_{j'}, \mathbb{C}_-} \right)$$

is insensitive, where k, l, l' are fixed, $x_1, \dots, x_k \in \mathbb{R}$ and $z_1, \dots, z_l, z'_1, \dots, z'_{l'} \in \mathbb{C}$ are bounded,

$$X_{x, F, \mathbb{R}} := \sum_{1 \leq i \leq n : \lambda_i(M_n) \in \mathbb{R}} F(\lambda_i(M_n) - \sqrt{n}x)$$

and

$$X_{z, G, \mathbb{C}_\pm} := \sum_{1 \leq i \leq n : \lambda_i(M_n) \in \mathbb{C}_\pm} G(\lambda_i(M_n) - \sqrt{n}z)$$

and the $F_i : \mathbb{R} \rightarrow \mathbb{C}, G_j : \mathbb{C} \rightarrow \mathbb{C}, G'_{j'} : \mathbb{C} \rightarrow \mathbb{C}$ are smooth functions supported on bounded sets obeying the bounds

$$|\nabla^a F_i(x)|, |\nabla^a G_j(z)|, |\nabla^a G'_{j'}(z)| \leq C$$

for all $0 \leq a \leq 5, x \in \mathbb{R}, z \in \mathbb{C}$. Indeed, one can express any statistic of the form (6.13) as a linear combination of a bounded number of statistics of the form (6.14), plus a bounded number of additional statistics of the form (6.13) with smaller values of $k + l$.

As the spectrum is symmetric around the real axis, one has

$$X_{z, G, \mathbb{C}_-} = X_{\bar{z}, \tilde{G}, \mathbb{C}_+},$$

where $\tilde{G}(z) := G(\bar{z})$. Thus, we may concatenate the G_j with the G'_j , and assume without loss of generality that $l' = 0$, thus we are now seeking to establish the insensitivity of

$$(6.15) \quad \mathbf{E} \left(\prod_{i=1}^k X_{x_i, F_i, \mathbb{R}} \right) \left(\prod_{j=1}^l X_{z_j, G_j, \mathbb{C}_+} \right).$$

On the other hand, by repeating the remainder of the arguments for the complex case with essentially no changes, we can show that the quantity

$$(6.16) \quad \mathbf{E} \prod_{p=1}^m X_{z_p, H_p}$$

is insensitive for any fixed m , any bounded complex numbers z_1, \dots, z_m , and any smooth $H_p : \mathbb{C} \rightarrow \mathbb{C}$ supported in a bounded set and obeying the bounds

$$|\nabla^a H_p(z)| \leq C$$

for all $0 \leq a \leq 5$ and $z \in \mathbb{C}$, where

$$X_{z, H} := \sum_{1 \leq i \leq n} H(\lambda_i(M_n) - z).$$

Thus, the remaining task is to deduce the insensitivity of (6.15) from the insensitivity of (6.16).

Specializing (6.16) to the case when $z_p = z$ is independent of p , and $H_p = H$ is real-valued, we see that

$$\mathbf{E} X_{z, H}^m$$

is insensitive for any m . In particular, we see from (the smooth version of) Urysohn’s lemma and Lemma 11 that we have the bound

$$(6.17) \quad \mathbf{E} N_{B(z\sqrt{n}, C)}^m \ll 1$$

for any fixed radius C and any bounded complex number z , where $N_\Omega = N_\Omega[M_n]$ denotes the number of eigenvalues of M_n in Ω . Among other things, this implies that

$$(6.18) \quad \mathbf{E} |X_{x_i, F_i, \mathbb{R}}|^A, \mathbf{E} |X_{y_j, G_j, \mathbb{C}_+}|^A \ll 1$$

for any fixed A and all i, j .

To proceed further, we need a level repulsion result.

LEMMA 39 (Weak level repulsion). *Let $C > 0$ be fixed, $x \in \mathbb{R}$ be bounded, and ε be such that $n^{-c_0} \leq \varepsilon \leq C$ for a sufficiently small fixed $c_0 > 0$, and let $E_{x, C, \varepsilon}$ be the event that there are two eigenvalues $\lambda_i(M_n), \lambda_j(M_n)$ in the strip $S_{x, C, \varepsilon} := \{z \in B(x\sqrt{n}, C) : \text{Im}(z) \leq \varepsilon\}$ with $i \neq j$ such that $|\lambda_i(M_n) - \lambda_j(M_n)| \leq 2\varepsilon$. Then $\mathbf{P}(E_{x, C, \varepsilon}) \ll \varepsilon$, where the implied constant in the \ll notation is independent of ε .*

PROOF. In this proof, all implied constants in the \ll notation are understood to be independent of ε . By a covering argument, it suffices to show that

$$\mathbf{P}(N_{B(x\sqrt{n}+t, 10\varepsilon)} \geq 2) \ll \varepsilon^2$$

uniformly for all $t = O(1)$.

Let H be a nonnegative bump function supported on $B(t, 20\varepsilon)$ that equals one on $B(t, 10\varepsilon)$. Then the expression $X_{x,H}^2 - X_{x,H^2}$ is nonnegative, and is at least 2 when $N_{B(x\sqrt{n}+t, 10\varepsilon)} \geq 2$. Thus, by Markov's inequality it suffices to show that

$$\mathbf{E}X_{x,H}^2 - X_{x,H^2} \ll \varepsilon^2.$$

By the insensitivity of (6.16) and the lower bound on ε , it suffices to verify the claim when M_n is drawn from the real Gaussian distribution. [Note that the derivatives of H, H^2 can be as large as $O(\varepsilon^{-O(1)})$, causing additional factors of $O(\varepsilon^{-O(1)})$ to appear in the error term created when swapping M_n with the real Gaussian ensemble, but the n^{-c} gain coming from the insensitivity will counteract this if c_0 is small enough.]

We split

$$X_{x,H} = X_{x,H,\mathbb{R}} + 2X_{x,H,\mathbb{C}_+}$$

and similarly for H^2 . It will suffice to establish the estimates

$$(6.19) \quad \mathbf{E}X_{x,H,\mathbb{R}}^2 - X_{x,H^2,\mathbb{R}} \ll \varepsilon^2,$$

$$(6.20) \quad \mathbf{E}X_{x,H,\mathbb{R}}X_{x,H,\mathbb{C}_+} \ll \varepsilon^2$$

and

$$(6.21) \quad \mathbf{E}X_{x,H,\mathbb{C}_+}^2 \ll \varepsilon^2.$$

The left-hand sides of (6.19), (6.20), (6.21) may be expanded as

$$\int_{\mathbb{R}} \int_{\mathbb{R}} \rho_n^{(2,0)}(x\sqrt{n} + y, x\sqrt{n} + y') H(y) H(y') dy dy',$$

$$\int_{\mathbb{R}} \int_{\mathbb{C}_+} \rho_n^{(1,1)}(x\sqrt{n} + y, x\sqrt{n} + z) H(y) H(z) dy dz$$

and

$$\int_{\mathbb{C}_+} \rho_n^{(0,1)}(x\sqrt{n} + z) H^2(z) dz$$

$$+ 2 \int_{\mathbb{C}_+} \int_{\mathbb{C}_+} \rho_n^{(0,2)}(x\sqrt{n} + z, x\sqrt{n} + w) H(z) H(w) dz dw,$$

respectively. Using Lemma 11, we see that these expressions are $O(\varepsilon^2)$ as required. \square

REMARK 40. In fact, a closer inspection of the explicit form of the correlation functions reveals that one can gain some additional powers of ε here, giving a stronger amount of level repulsion, but for our purposes any bound that goes to zero as $\varepsilon \rightarrow 0$ will suffice.

From the symmetry of the spectrum, we observe that if $E_{x,C,\varepsilon}$ does not hold, then there cannot be any strictly complex eigenvalue $\lambda_i(M_n)$ in the strip $S_{x,C,\varepsilon}$, since in that case $\overline{\lambda_i(M_n)}$ would be distinct eigenvalue in the strip at a distance at most 2ε from $\lambda_i(M_n)$. In particular, we see that

$$(6.22) \quad \mathbf{P}(N_{S_{x,C,\varepsilon} \setminus [x\sqrt{n}-C, x\sqrt{n}+C]} = 0) = 1 - O(\varepsilon).$$

Informally, this estimate tells us that we can usually thicken the interval $[x\sqrt{n} - C, x\sqrt{n} + C]$ to the strip $S_{x,C,\varepsilon}$ without encountering any additional spectrum.

Fix $\varepsilon := n^{-c_0}$ for some sufficiently small fixed $c_0 > 0$. We can use (6.22) to simplify the expression (6.15) in two ways. First, thanks to (6.22), (6.18) and Hölder’s inequality, we may replace each of the G_j in (6.14) with a function \tilde{G}_j that vanishes on the strip $\{z - z_j : |\text{Im}(z)| \leq \varepsilon\}$, while only picking up an error of $O(\varepsilon^c)$ for some fixed $c > 0$, which will be acceptable from the choice of ε . By discarding the component of \tilde{G}_j below the strip, we may then assume \tilde{G}_j is supported on the half-space $\mathbb{C}_+ - z_j$. In particular, we have

$$X_{z_j, \tilde{G}_j, \mathbb{C}_+} = X_{z_j, \tilde{G}_j}.$$

Also, by performing a smooth truncation, we see that we have the derivative bounds $\nabla^a \tilde{G}_j = O(\varepsilon^{-O(1)})$ for all $0 \leq a \leq 5$.

Second, by another application of (6.22), (6.18), and Hölder’s inequality, we may “thicken” each factor $X_{x_i, F_i, \mathbb{R}}$ by replacing it with X_{x_i, \tilde{F}_i} , where $\tilde{F}_i : \mathbb{C} \rightarrow \mathbb{C}$ is a smooth extension of F_i that is supported on the strip $\{z : |\text{Im}(z)| \leq \varepsilon\}$, while only acquiring an error of $O(\varepsilon^c)$ for some fixed $c > 0$. Again, we have the derivative bounds $\nabla^a \tilde{F}_i = O(\varepsilon^{-O(1)})$ for $0 \leq a \leq 5$. From the insensitivity of (6.16) [and using the n^{-c} gain coming from insensitivity to absorb all $O(\varepsilon^{-O(1)})$ losses from the derivative bounds], we see that

$$(6.23) \quad \mathbf{E} \left(\prod_{i=1}^k X_{x_i, \tilde{F}_i} \right) \left(\prod_{j=1}^l X_{z_j, \tilde{G}_j} \right)$$

is insensitive, which by the preceding discussion yields (for c_0 small enough) that (6.15) is insensitive also, as required. This concludes the derivation of Theorem 12 from Theorem 23 and Proposition 20.

6.3. *Quick applications.* As quick consequences of Theorems 2 and 12, we now prove Corollaries 10, 17 and 18.

We first prove we prove Corollary 18. Let M_n be as in that theorem. Set $\varepsilon := n^{-c_0}$ for some sufficiently small $c_0 > 0$. A routine modification of the proof of

Lemma 39 (or, alternatively, Theorem 12 combined with Lemma 11) shows that for any $z \in B(0, O(\sqrt{n}))$, one has

$$\mathbf{E} \left(\begin{matrix} 0 \\ N_{B(z, \varepsilon)} \end{matrix} \right)^2 \ll \varepsilon^4,$$

when $|\operatorname{Im} z| \geq \varepsilon$, if c_0 is small enough; in particular, the expected number of eigenvalues in $B(z, \varepsilon)$ which are repeated is $O(\varepsilon^4)$. We then cover $B(0, 3\sqrt{n})$ by $O(n/\varepsilon^2)$ balls $B(z, \varepsilon)$ with $|\operatorname{Im} z| \geq \varepsilon$, together with the strip $\{z : |\operatorname{Im} z| \leq \varepsilon\}$. By (6.22) (or Theorem 12 and Lemma 11) and linearity of expectation, the strip contains $O(\varepsilon\sqrt{n})$ eigenvalues. By [4, 25], the spectral radius of M_n is known to equal $(1 + o(1))\sqrt{n}$ with overwhelming probability.¹⁵ We conclude that the expected number of repeated complex eigenvalues is at most

$$O(n/\varepsilon^2) \times O(\varepsilon^4) + O(\varepsilon\sqrt{n}) + O(n^{-100}),$$

which becomes $O(n^{1-c})$ for some fixed $c > 0$; a similar argument gives a bound of $O(n^{1/2-c})$ for the expected number of repeated real eigenvalues. The claim now follows from Markov’s inequality.

Now we prove Corollary 17. Let M_n be as in that theorem. As mentioned previously, the spectral radius of M_n is known to equal $(1 + o(1))\sqrt{n}$ with overwhelming probability. In particular, we have

$$\mathbf{E}N_{\mathbb{R}}(M_n) = \mathbf{E}N_{[-3\sqrt{n}, 3\sqrt{n}]}(M_n) + O(n^{-100})$$

(say). By the smooth form of Urysohn’s lemma, we can select fixed smooth, non-negative functions F_-, F_+ such that we have the pointwise bounds

$$1_{[-2, 2]} \leq F_- \leq 1_{[-3, 3]} \leq F_+ \leq 1_{[-4, 4]}.$$

By definition of $\rho^{(1,0)}$, we observe that

$$\begin{aligned} \mathbf{E}N_{[-2\sqrt{n}, 2\sqrt{n}]}(M_n) &\leq \int_{\mathbb{R}} \rho^{(1,0)}(x) F_-(x/\sqrt{n}) dx \\ &\leq \mathbf{E}N_{[-3\sqrt{n}, 3\sqrt{n}]}(M_n) \\ &\leq \int_{\mathbb{R}} \rho^{(1,0)}(x) F_+(x/\sqrt{n}) dx \\ &\leq \mathbf{E}N_{[-4\sqrt{n}, 4\sqrt{n}]}(M_n). \end{aligned}$$

By smoothly partitioning $F_{\pm}(x/\sqrt{n})$ into $O(\sqrt{n})$ pieces supported on intervals of size $O(1)$, and applying Theorem 12 to each piece, we see upon summing that the two integrals above are only modified by $O(n^{1/2-c})$ for some fixed $c > 0$ if we

¹⁵Actually, for this argument, the easier bound of $O(1)$ would suffice, which can be obtained by a variety of methods, for example, by an epsilon net argument or by Talagrand’s inequality [48].

replace M_n with a real Gaussian matrix M'_n . On the other hand, when M'_n is real Gaussian we see from Theorem 16 (and the spectral radius bound) that

$$\mathbf{E}N_{[-2\sqrt{n}, 2\sqrt{n}]}(M'_n), \quad \mathbf{E}N_{[-4\sqrt{n}, 4\sqrt{n}]}(M'_n) = \sqrt{\frac{2n}{\pi}} + O(1).$$

Putting these bounds together, we obtain the expectation claim of Corollary 17. The variance claim is similar. Indeed, we have

$$\mathbf{E}N_{\mathbb{R}}(M_n)^2 = \mathbf{E}N_{[-3\sqrt{n}, 3\sqrt{n}]}(M_n)^2 + O(n^{-90})$$

(say) and

$$\begin{aligned} \mathbf{E}N_{[-2\sqrt{n}, 2\sqrt{n}]}(M_n)^2 &\leq \int_{\mathbb{R}} \rho^{(1,0)}(x) F_-(x/\sqrt{n})^2 dx \\ &\quad + \int_{\mathbb{R}} \int_{\mathbb{R}} \rho^{(2,0)}(x, y) F_-(x/\sqrt{n}) F_-(y/\sqrt{n}) dx dy \\ &\leq \mathbf{E}N_{[-3\sqrt{n}, 3\sqrt{n}]}(M_n)^2 \\ &\leq \int_{\mathbb{R}} \rho^{(1,0)}(x) F_+(x/\sqrt{n})^2 dx \\ &\quad + \int_{\mathbb{R}} \int_{\mathbb{R}} \rho^{(2,0)}(x, y) F_+(x/\sqrt{n}) F_+(y/\sqrt{n}) dx dy \\ &\leq \mathbf{E}N_{[-4\sqrt{n}, 4\sqrt{n}]}(M_n)^2. \end{aligned}$$

From Theorem 12 and smooth decomposition, we see that all of the above integrals vary by $O(n^{1-c})$ at most for some fixed $c > 0$ if M_n is replaced with a real Gaussian matrix, and then the variance claim can be deduced from Theorem 16 and the spectral radius bound as before.

REMARK 41. A similar argument shows that in the complex case, the expected number of real eigenvalues is $O(n^{1/2-c})$, which can be improved to $O(n^{-A})$ for any $A > 0$ if one assumes sufficiently many matching moments depending on A . Of course, one expects typically in this case that there are no real eigenvalues whatsoever (and this is almost surely the case when the matrix ensemble is continuous), but this is beyond the ability of our current methods to establish in the case of discrete complex matrices.

Finally, we prove Corollary 10. Let M_n, z_0, r be as in that theorem, and let \tilde{M}_n be drawn from the complex Gaussian matrix ensemble. Let $\varepsilon = o(1)$ be a slowly decaying function of n to be chosen later. Let R be any rectangle in $B(0, 100\sqrt{n})$ of side length $1 \times n^{-\varepsilon}$, and let $3R$ be the rectangle with the same center as R but three times the side lengths. By the smooth form of Urysohn’s lemma, we can construct a smooth function $F : \mathbb{C} \rightarrow \mathbb{R}^+$ with the pointwise bounds

$$1_R \leq F \leq 1_{3R}$$

such that $|\nabla^j F| \ll n^{j\varepsilon}$ for all $0 \leq j \leq 5$. Applying Corollary 15 (to $n^{-5\varepsilon} F$), we conclude that

$$\int_{\mathbb{C}} F(z)\rho_n^{(1)}(z) dz = \int_{\mathbb{C}} F(z)\tilde{\rho}_n^{(1)}(z) dz + O(n^{-c+5\varepsilon})$$

for some absolute constant c . On the other hand, from (1.5) we see that $\int_{\mathbb{C}} F(z)\tilde{\rho}_n^{(1)}(z) dz \ll n^{-\varepsilon}$, since $3R$ has area $O(n^{-\varepsilon})$. Since $\varepsilon = o(1)$, we conclude that

$$\int_{\mathbb{C}} F(z)\rho_n^{(1)}(z) dz \ll n^{-\varepsilon}$$

and in particular that

$$(6.24) \quad \mathbf{E}N_R(M_n) \ll n^{-\varepsilon}.$$

A similar argument (with larger values of k) gives

$$(6.25) \quad \mathbf{E}N_{R_1}(M_n) \cdots N_{R_k}(M_n) \ll n^{-k\varepsilon},$$

whenever k is fixed and R_1, \dots, R_k are $1 \times n^{-\varepsilon}$ rectangles (possibly overlapping) in $B(0, 100\sqrt{n})$.

Now let $G : \mathbb{C} \rightarrow \mathbb{R}^+$ be a smooth function supported on $B(z_0, r + n^{-\varepsilon})$ which equals 1 on $B(z_0, r)$ and has the derivative bounds $|\nabla^j G| \ll n^{j\varepsilon}$ for all $0 \leq j \leq 5$. By covering the annulus $B(z_0, r + n^{-\varepsilon}) \setminus B(z_0, r)$ by $O(r)$ rectangles of dimension $1 \times n^{-\varepsilon}$, we see from (6.24) that

$$\mathbf{E}N_{B(z_0, r+n^{-\varepsilon}) \setminus B(z_0, r)}(M_n) \ll rn^{-\varepsilon}$$

and similarly from (6.25) one has

$$\mathbf{E}N_{B(z_0, r+n^{-\varepsilon}) \setminus B(z_0, r)}(M_n)^k \ll r^k n^{-k\varepsilon}$$

for any fixed k . Since we are assuming $r \leq n^{o(1)}$, we conclude (if ε decays to zero sufficiently slowly) that

$$\mathbf{E}N_{B(z_0, r+n^{-\varepsilon}) \setminus B(z_0, r)}(M_n)^k = o(1)$$

for all k . In particular, if we introduce the linear statistic

$$(6.26) \quad X := \frac{\sum_{i=1}^n G(\lambda_i(M_n)) - r^2}{r^{1/2}\pi^{-1/4}}$$

we see from the triangle inequality that the asymptotics

$$\mathbf{E}\left(\frac{N_{B(z_0, r)} - r^2}{r^{1/2}\pi^{-1/4}}\right)^k \rightarrow \mathbf{E}N(0, 1)_{\mathbb{R}}^k$$

for all fixed $k \geq 0$ are equivalent to the asymptotics

$$\mathbf{E}X^k \rightarrow \mathbf{E}N(0, 1)_{\mathbb{R}}^k.$$

Let \tilde{X} be the analogue of X for \tilde{M}_n . From Theorem 9 and the preceding arguments, we have

$$\mathbf{E}\tilde{X}^k \rightarrow \mathbf{E}N(0, 1)_{\mathbb{R}}^k$$

and so it will suffice to show that

$$\mathbf{E}X^k - \mathbf{E}\tilde{X}^k = o(1)$$

for all fixed $k \geq 1$. By (6.26) and the hypotheses that $1 \leq r \leq n^{o(1)}$ and $\varepsilon = o(1)$, it will suffice to show that

$$\mathbf{E}\left(\sum_{i=1}^n G(\lambda_i(M_n))\right)^k - \mathbf{E}\left(\sum_{i=1}^n G(\lambda_i(\tilde{M}_n))\right)^k = O(r^{O(k)}n^{-c+O(k\varepsilon)})$$

for all fixed $k \geq 0$ and some fixed $c > 0$ (which will in fact turn out to be uniform in k , although we will not need this fact). Expanding out the k th powers and collecting terms¹⁶ depending on the multiplicities of the i indices, we see that it suffices to show that

$$\begin{aligned} &\mathbf{E} \sum_{1 \leq i_1 < \dots < i_{k'} \leq n} G^{a_1}(\lambda_{i_1}(M_n)) \dots G^{a_{k'}}(\lambda_{i_{k'}}(M_n)) \\ &\quad - G^{a_1}(\lambda_{i_1}(\tilde{M}_n)) \dots G^{a_{k'}}(\lambda_{i_{k'}}(\tilde{M}_n)) \\ &= O(r^{O(k)}n^{-c+O(k\varepsilon)}) \end{aligned}$$

for all fixed $k', a_1, \dots, a_{k'} \geq 1$ and some fixed $c > 0$, where $k := a_1 + \dots + a_{k'}$. But the left-hand side can be rewritten using (1.1) as

$$\int_{\mathbb{C}^k} \left(\prod_{j=1}^k G(z_j)^{a_j}\right) (\rho_n^{(k)}(z_1, \dots, z_k) - \tilde{\rho}_n^{(k)}(z_1, \dots, z_k)) dz_1 \dots dz_k.$$

One can smoothly decompose $(\prod_{j=1}^k G(z_j)^{a_j})$ as the sum of $O(r^{O(k)}n^{O(\varepsilon)})$ smooth functions supported on balls of bounded radius, whose derivatives up to fifth order are all uniformly bounded. Applying Theorem 2 to each such function and summing, one obtains the claim.

REMARK 42. The main reason why the radius r was restricted to be $O(n^{o(1)})$ was because of the need to obtain asymptotics for k th moments for arbitrary fixed k . For any given k , the above arguments show that one obtains the right asymptotics for all $r \leq n^{c/k}$ for some absolute constant $c > 0$. If one increases the number of matching moment assumptions, one can increase the value of k , but we were unable to find an argument that allowed one to take r as large as n^α for some fixed $\alpha > 0$ independent of k , even after assuming a large number of matching moments.

¹⁶The observant reader will note that this step is inverting one of the first steps in the proof of Theorem 2 given previously, and one could shorten the total length of the argument here if desired by skipping directly to that point of the proof of Theorem 2 and continuing onward from there.

7. Resolvent swapping. In this section, we recall some facts about the stability of the resolvent of Hermitian matrices with respect to permutation in just one or two entries, in order to perform swapping arguments. Such swapping arguments were introduced to random matrix theory in [11], and first applied to establish universality results for local spectral statistics in [55]. In [21], it was observed that the stability analysis of such swapping was particularly simple if one worked with the resolvents (or Green’s function) rather than with individual eigenvalues. Our formalization of this analysis here is drawn from [58]. We will use this resolvent swapping analysis twice in this paper; once to establish the four moment theorem for the determinant (Theorem 23) in Section 8, and once to deduce concentration of the log-determinant for i.i.d. matrices (Theorem 25) from concentration for Gaussian matrices (Theorem 33) in Section 10.

We will need the matrix norm

$$\|A\|_{(\infty,1)} = \sup_{1 \leq i, j \leq n} |a_{ij}|$$

and the following definition.

DEFINITION 43 (Elementary matrix). An *elementary matrix* is a matrix which has one of the following forms:

$$(7.1) \quad V = e_a e_a^*, e_a e_b^* + e_b e_a^*, \sqrt{-1} e_a e_b^* - \sqrt{-1} e_b e_a^*$$

with $1 \leq a, b \leq n$ distinct, where e_1, \dots, e_n is the standard basis of \mathbb{C}^n .

Let M_0 be a Hermitian matrix, let $z = E + i\eta$ be a complex number, and let V be an elementary matrix. We then introduce, for each $t \in \mathbb{R}$, the Hermitian matrices

$$M_t := M_0 + \frac{1}{\sqrt{n}} t V,$$

the resolvents

$$(7.2) \quad R_t = R_t(E + i\eta) := (M_t - E - i\eta)^{-1}$$

and the Stieltjes transform

$$s_t := s_t(E + i\eta) := \frac{1}{n} \text{trace } R_t(E + i\eta).$$

We have the following Neumann series expansion.

LEMMA 44 (Neumann series). Let M_0 be a Hermitian $n \times n$ matrix, let $E \in \mathbb{R}$, $\eta > 0$, and $t \in \mathbb{R}$, and let V be an elementary matrix. Suppose one has

$$(7.3) \quad |t| \|R_0\|_{(\infty,1)} = o(\sqrt{n}).$$

Then one has the Neumann series formula

$$(7.4) \quad R_t = R_0 + \sum_{j=1}^{\infty} \left(-\frac{t}{\sqrt{n}}\right)^j (R_0 V)^j R_0$$

with the right-hand side being absolutely convergent, where R_t is defined by (7.2). Furthermore, we have

$$(7.5) \quad \|R_t\|_{(\infty,1)} \leq (1 + o(1)) \|R_0\|_{(\infty,1)}.$$

In practice, we will have $t = n^{O(c_0)}$ (from a decay hypothesis on the atom distribution) and $\|R_0\|_{(\infty,1)} = n^{O(c_0)}$ (from eigenvector delocalization and a level repulsion hypothesis), where $c_0 > 0$ is a small constant, so (7.3) is quite a mild condition.

PROOF OF LEMMA 44. See [58], Lemma 12. \square

We now can describe the dependence of s_t on t :

PROPOSITION 45 (Taylor expansion of s_t). *Let the notation be as above, and suppose that (7.3) holds. Let $k \geq 1$ be fixed. Then one has*

$$(7.6) \quad s_t = s_0 + \sum_{j=1}^k n^{-j/2} c_j t^j + O\left(n^{-(k+1)/2} |t|^{k+1} \|R_0\|_{(\infty,1)}^{k+1} \min\left(\|R_0\|_{(\infty,1)}, \frac{1}{n\eta}\right)\right),$$

where the coefficients c_j are independent of t and obey the bounds

$$(7.7) \quad |c_j| \ll \|R_0\|_{(\infty,1)}^j \min\left(\|R_0\|_{(\infty,1)}, \frac{1}{n\eta}\right)$$

for all $1 \leq j \leq k$.

PROOF. See [58], Proposition 13. \square

8. Proof of the four moment theorem. We now prove Theorem 23.

We begin with some simple reductions. Observe that each entry ξ_{ij} of M_n has size at most $O(n^{o(1)})$ with overwhelming probability. Thus, by modifying the distributions of the ξ_{ij} slightly (taking care to retain the moment matching property¹⁷) and assume that all entries *surely* have size $O(n^{o(1)})$. Thus,

$$(8.1) \quad \|M_n\|_{(\infty,1)}, \|M'_n\|_{(\infty,1)} \ll n^{o(1)}.$$

¹⁷Alternatively, one can allow the moments to deviate from each other by, say, $O(n^{-100})$, which one can verify will not affect the argument. See [3], Chapter 2, or [36], Appendix A, for details.

We may also assume that G is bounded by 1 rather than by n^{c_0} , since the general claim then follows by normalizing G and shrinking c_0 as necessary; thus,

$$(8.2) \quad |G(x_1, \dots, x_k)| \leq 1$$

for all $x_1, \dots, x_k \in \mathbb{R}$.

Fix M_n, M'_n . Recall that a statistic S is *asymptotically (M_n, M'_n) -insensitive*, or *insensitive* for short, if one has

$$|S(M_n) - S(M'_n)| \ll n^{-c}$$

for some fixed $c > 0$. By shrinking c_0 if necessary, our task is thus to show that the quantity

$$EG(\log|\det(M_n - z_1)|, \dots, \log|\det(M_n - z_k)|)$$

is insensitive.

The next step is to use (2.4) to replace the log-determinants $\log|\det(M_n - z)|$ with the log-determinants $\log|\det W_{n,z}|$, where the $W_{n,z}$ are defined by (2.3). After translating and rescaling the function G , we thus see that it suffices to show that

$$EG(\log|\det(W_{n,z_1})|, \dots, \log|\det(W_{n,z_k})|)$$

is insensitive.

We observe the identity

$$\log|\det(W_{n,z_j})| = \log|\det(W_{n,z_j} - \sqrt{-1}T)| - n \operatorname{Im} \int_0^T s_j(\sqrt{-1}\eta) d\eta$$

for any $T > 0$ for all $1 \leq j \leq k$, where $s_j(z) := \frac{1}{n} \operatorname{trace}(W_{n,z_j} - z)^{-1}$ is the Stieltjes transform, as can be seen by writing everything in terms of the eigenvalues of W_{n,z_j} . If we set $T := n^{100}$, then we see that

$$\begin{aligned} \log|\det(W_{n,z_j} - \sqrt{-1}T)| &= n \log T + \log|\det(1 - n^{-100}W_{n,z_j})| \\ &= n \log T + O(n^{-10}) \end{aligned}$$

(say), thanks to (8.1) and the hypothesis that z_j lies in $B(0, (1 - \delta)\sqrt{n})$. Thus, by translating G again, it suffices to show that the quantity

$$EG\left(\left(n \operatorname{Im} \int_0^{n^{100}} s_j(\sqrt{-1}\eta) d\eta\right)_{j=1}^k\right)$$

is insensitive.

We need to truncate away from the event that W_{n,z_j} has an eigenvalue too close to zero. Let $\chi : \mathbb{R} \rightarrow \mathbb{R}$ be a smooth cutoff to the region $|x| \leq n^{3c_0}$ that equals 1 for $|x| \leq n^{3c_0}/2$. From Proposition 27 and the union bound, we have with probability $1 - O(n^{-c_0+o(1)})$ that there are no eigenvalues of W_{n,z_j} in the interval $[-n^{1-2c_0}, n^{-1-2c_0}]$ for all $1 \leq j \leq k$. Combining this with Proposition 29 and a

dyadic decomposition, we conclude that with probability $1 - O(n^{-c_0+o(1)})$ one has

$$|\operatorname{Im} s_j(\sqrt{-1}n^{-1-4c_0})| \ll n^{2c_0+o(1)}$$

for all $1 \leq j \leq k$. In particular, one has

$$\chi(\operatorname{Im} s_j(\sqrt{-1}n^{-1-4c_0})) = 1$$

with overwhelming probability.

In view of this fact and (8.2), it suffices to show that the quantity

$$(8.3) \quad \mathbf{E}G\left(n \operatorname{Im} \int_0^{n^{100}} s_j(\sqrt{-1}\eta) d\eta\right) \chi\left(\left(\operatorname{Im} s_j(\sqrt{-1}n^{-1-4c_0})\right)_{j=1}^k\right)$$

is insensitive.

Call a statistic S *very highly insensitive* if one has

$$|S(M_n) - S(M'_n)| \ll n^{-2-c}$$

for some fixed $c > 0$. By swapping the real and imaginary parts of the components of M_n with those of M'_n one at a time, we see from telescoping series that it will suffice to show that (8.3) is very highly insensitive whenever M_n and M'_n are identical in all but one entry, and in that entry either the real parts are identical, or the imaginary parts are identical.

Fix M_n, M'_n as indicated. Then for each $1 \leq j \leq k$, one has

$$\begin{aligned} W_{n,z_j} &= W_{n,z_j,0} + \frac{1}{\sqrt{n}}\xi V, \\ W'_{n,z_j} &= W_{n,z_j,0} + \frac{1}{\sqrt{n}}\xi' V, \end{aligned}$$

where ξ, ξ' are real random variables that match to order 4 and have the magnitude bound

$$(8.4) \quad |\xi|, |\xi'| \ll n^{o(1)},$$

V is an elementary matrix, and $W_{n,z_j,0}$ is a random Hermitian matrix independent of both ξ and ξ' . To emphasize this representation, and to bring the notation closer to that of the preceding section, we rewrite s_j as $s_\xi^{(j)}$, where

$$s_t^{(j)}(z) := \frac{1}{2n} \operatorname{trace} R_t^{(j)}(z)$$

and

$$R_t^{(j)}(z) := \left(W_{n,z_j,0} + \frac{1}{\sqrt{n}}tV - z\right)^{-1}.$$

Our task is now to show that the quantity

$$(8.5) \quad \mathbf{E}G\left(n \operatorname{Im} \int_0^{n^{100}} s_\xi^{(j)}(\sqrt{-1}\eta) d\eta\right) \chi\left(\left(\operatorname{Im} s_\xi^{(j)}(\sqrt{-1}n^{-1-4c_0})\right)_{j=1}^k\right)$$

only changes by $O(n^{-2-c})$ when ξ is replaced by ξ' .

We now place some bounds on $R_t^{(j)}(z)$.

LEMMA 46 (Eigenvector delocalization). *Let $1 \leq j \leq k$, and suppose that we are in the event that $\chi(\operatorname{Im} s_j(\sqrt{-1}n^{-1-4c_0}))$ is nonzero. Then with overwhelming probability, one has*

$$(8.6) \quad \sup_{\eta>0} \|R_\xi^{(j)}(\sqrt{-1}\eta)\|_{(\infty,1)} \ll n^{O(c_0)}$$

and hence [by Lemma 44 and (8.4), swapping the roles of ξ and 0]

$$(8.7) \quad \sup_{\eta>0} \|R_0^{(j)}(\sqrt{-1}\eta)\|_{(\infty,1)} \ll n^{O(c_0)}.$$

The bounds in the above lemma are similar to those from Proposition 31 (and Proposition 31 will be used in the proof of the lemma), but the point here is that the bounds remain uniform in the limit $\eta \rightarrow 0$, whereas the bounds in Proposition 31 blow up at that limit.

PROOF OF LEMMA 46. By hypothesis and the support of χ , one has

$$|\operatorname{Im} s_\xi^{(j)}(\sqrt{-1}n^{-1-4c_0})| \ll n^{-3c_0}.$$

The left-hand side can be expanded as

$$n^{-2-4c_0} \sum_{i=1}^n \frac{1}{\lambda_i(W_{n,z_j})^2 + n^{-2-8c_0}}$$

and so we obtain the lower bound

$$(8.8) \quad \lambda_i(W_{n,z_j}) \gg n^{-1-c_0/2}$$

for all i .

From Proposition 31, one already has

$$\sup_{\eta>1/n} \|R_\xi^{(j)}(\sqrt{-1}\eta)\|_{(\infty,1)} \ll n^{o(1)}$$

with overwhelming probability. In particular, for each $1 \leq j \leq k$ and $\eta > 1/n$, one has

$$\frac{\eta}{n} \sum_{i=1}^n \frac{|e_j^* u_i|^2}{\lambda_i(W_{n,z_j})^2 + \eta^2} \ll n^{o(1)}.$$

Combining this with (8.8), we see that

$$\frac{\eta}{n} \sum_{i=1}^n \frac{|e_l^* u_i|^2}{\lambda_i(W_{n,z_j})^2 + \eta^2} \ll n^{O(c_0)}$$

for all $\eta > 0$, $1 \leq j \leq k$, and $1 \leq l \leq n$. By dyadic summation, we conclude that

$$\sum_{i=1}^n \frac{|e_l^* u_i|^2}{(\lambda_i(W_{n,z_j})^2 + \eta^2)^{1/2}} \ll n^{O(c_0)}$$

for all $\eta > 1/n$, and thus by Cauchy–Schwarz one has

$$\left| \frac{1}{n} \sum_{i=1}^n \frac{(e_l^* u_i) \overline{(e_m^* u_i)}}{\lambda_i(W_{n,z_j}) - \sqrt{-1}\eta} \right| \ll n^{O(c_0)}$$

for all $\eta > 0$ and $1 \leq j \leq k$ and $1 \leq l, m \leq n$. But the left-hand side is the lm coefficient of $R_\xi^{(j)}(\sqrt{-1}\eta)$, and the claim follows. \square

We now condition to the event that (8.7) holds for all $1 \leq j \leq k$; Lemma 46 ensures us that the error in doing so is $O_A(n^{-A})$ for any A . Then by Proposition 45, we have

$$s_\xi^{(j)}(\sqrt{-1}\eta) = s_0^{(j)}(\sqrt{-1}\eta) + \sum_{i=1}^4 \xi^i n^{-i/2} c_i^{(j)}(\eta) + O(n^{-5/2+O(c_0)}) \min\left(1, \frac{1}{n\eta}\right)$$

for each j and all $\eta > 0$, and similarly with ξ replaced by $\tilde{\xi}$, where the coefficients $c_i^{(j)}$ enjoy the bounds

$$|c_i^{(j)}| \ll n^{O(c_0)} \min\left(1, \frac{1}{n\eta}\right).$$

From this and Taylor expansion, we see that the expression

$$G\left(n \operatorname{Im} \int_0^{n^{100}} s_\xi(\sqrt{-1}\eta) d\eta\right) \chi(\operatorname{Im} s_\xi(E + \sqrt{-1}n^{-1-4c_0}))$$

is equal to a polynomial of degree at most 4 in η with coefficients independent of η , plus an error of $O(n^{-5/2+O(c_0)})$, which gives the claim for c_0 small enough.

REMARK 47. If one assumes more than four matching moments, one can improve the final constant c in the conclusion of Theorem 23. However, it appears that one cannot make c arbitrarily large with this method, basically because the Taylor expansion becomes unfavorable when c_0 is too large.

9. Concentration of log-determinant for Gaussian matrices. In this section, we establish Theorem 33. Fix $z_0 \in B(0, C)$; all our implied constants will be uniform in z_0 . Define α to be the quantity $\alpha := \frac{1}{2}(|z_0|^2 - 1)$ if $|z_0| \leq 1$, and $\alpha := \log |z_0|$ if $|z_0| \geq 1$. Our task is to show that $\log |\det(M_n - z_0\sqrt{n})|$ concentrates around $\frac{1}{2}n \log n + \alpha n$.

9.1. *The upper bound.* In this section, we prove that with overwhelming probability

$$\log|\det(M_n - z_0\sqrt{n})| \leq \frac{1}{2}n \log n + \alpha n + n^{o(1)},$$

which is the upper bound of what we need. In fact, the statement (which is based on the second moment method) holds for general random matrices with non-Gaussian entries.

PROPOSITION 48 (Upper bound on log-determinant). *Let $M_n = (\xi_{ij})_{1 \leq i, j \leq n}$ be a random matrix with independent entries having mean zero and variance one. Then for any $z_0 \in \mathbb{C}$, one has*

$$\log|\det(M_n - z_0\sqrt{n})| \leq \frac{1}{2}n \log n + \alpha n + O(n^{o(1)})$$

with overwhelming probability.

The key is the following lemma.

LEMMA 49. *Let $M_n = (\xi_{ij})_{1 \leq i, j \leq n}$ be a random matrix as above. Then for any $z_0 \in \mathbb{C}$, one has*

$$(9.1) \quad \mathbf{E}|\det(M_n - z_0\sqrt{n})|^2 \leq n! \exp(|z_0|^2 n)$$

for all z_0 . When $|z_0| \geq 1$, we have the variant bound

$$(9.2) \quad \mathbf{E}|\det(M_n - z_0\sqrt{n})|^2 \leq n^{n+1} |z_0|^{2n}.$$

PROOF. By cofactor expansion, one has

$$\det(M_n - z_0\sqrt{n}) = \sum_{\sigma \in S_n} \text{sgn}(\sigma) \prod_{i=1}^n (\xi_{i\sigma(i)} - z_0\sqrt{n}1_{\sigma(i)=i}),$$

where S_n is the set of permutations on $\{1, \dots, n\}$. We can rewrite this expression as

$$\sum_{A \subset \{1, \dots, n\}} \sum_{\sigma \in S_{n,A}} F_{A,\sigma},$$

where $S_{n,A}$ is the set of permutations $\sigma \in S_n$ that fix A , thus $\sigma(i) = i$ for all $i \in A$, and

$$F_{A,\sigma} := (-z_0\sqrt{n})^{|A|} \prod_{i \notin A} \xi_{i\sigma(i)}.$$

As the ξ_{ij} are jointly independent and have mean zero, we see that $\mathbf{E}F_{A,\sigma} \overline{F_{A',\sigma'}} = 0$ whenever $(A, \sigma) \neq (A', \sigma')$. Also, as the ξ_{ij} also have unit variance, we have $\mathbf{E}|F_{A,\sigma}|^2 = |z_0|^{2|A|} n^{|A|}$. We conclude that

$$\mathbf{E}|\det(M_n - z_0\sqrt{n})|^2 = \sum_{A \subset \{1, \dots, n\}} \sum_{\sigma \in S_{n,A}} |z_0|^{2|A|} n^{|A|}.$$

Write $j = |A|$. For each choice of $j = 0, \dots, n$, there are $\frac{n!}{j!(n-j)!}$ choices for A , and $(n - j)!$ choices for σ . We conclude that

$$\mathbf{E}|\det(M_n - z_0\sqrt{n})|^2 = n! \sum_{j=0}^n \frac{|z_0|^{2j} n^j}{j!}.$$

(This formula is well known in the literature; see, e.g., [13], Theorem 3.1.) Since

$$\sum_{j=0}^{\infty} \frac{|z_0|^{2j} n^j}{j!} = \exp(|z_0|^2 n)$$

we obtain (9.1).

Now suppose that $|z_0| \geq 1$, then the terms $\frac{|z_0|^{2j} n^j}{j!}$ are nondecreasing in j , and are thus each bounded by $|z_0|^{2n} n^n / n!$, and (9.2) follows. \square

From Lemma 49 and Stirling’s formula, we see that

$$\mathbf{E}|\det(M_n - z_0\sqrt{n})|^2 \leq \exp(n \log n + 2\alpha n + O(n^{o(1)}))$$

and thus by Markov’s inequality we see that

$$|\det(M_n - z_0\sqrt{n})|^2 \leq \exp(n \log n + 2\alpha n + O(n^{o(1)}))$$

with overwhelming probability, which gives Proposition 48 as desired.

9.2. *Hessenberg form.* To complete the proof of Theorem 33, we need to show the lower bound

$$\log|\det(M_n - z_0\sqrt{n})| \geq \frac{1}{2}n \log n + \alpha n - O(n^{o(1)})$$

with overwhelming probability. As we shall see later, the fact that we only seek a one-sided bound now instead of a two-sided one will lead to some convenient simplifications to the argument.¹⁸

Now we will make essential use of the fact that the entries are Gaussian. The first step is to conjugate a complex Gaussian matrix into an almost lower-triangular form first observed in [33], in the spirit of the tridiagonalization of GUE matrices first observed by Trotter [60], as follows.

¹⁸If one really wished, one could adapt the arguments below to also give the upper bound, giving an alternate proof of Proposition 48, but this argument would be more complicated than the proof given in the previous section, and we will not pursue it here.

PROPOSITION 50 (Hessenberg matrix form). [33] *Let M_n be a complex Gaussian matrix, and let M'_n be the random matrix*

$$M'_n = \begin{pmatrix} \xi_{11} & \chi_{n-1,\mathbb{C}} & 0 & 0 & \cdots & 0 \\ \xi_{21} & \xi_{22} & \chi_{n-2,\mathbb{C}} & 0 & \cdots & 0 \\ \xi_{31} & \xi_{32} & \xi_{33} & \chi_{n-3,\mathbb{C}} & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \xi_{(n-1)1} & \xi_{(n-1)2} & \xi_{(n-1)3} & \xi_{(n-1)4} & \cdots & \chi_{1,\mathbb{C}} \\ \xi_{n1} & \xi_{n2} & \xi_{n3} & \xi_{n4} & \cdots & \xi_{nn} \end{pmatrix},$$

where ξ_{ij} for $1 \leq j \leq i \leq n$ are i.i.d. copies of the complex Gaussian $N(0, 1)_{\mathbb{C}}$, and for each $1 \leq i \leq n - 1$, $\chi_{i,\mathbb{C}}$ is a complex χ distribution of i degrees of freedom (see Section 3 for definitions), with the ξ_{ij} and $\chi_{i,\mathbb{C}}$ being jointly independent. Then the spectrum of M_n has the same distribution as the spectrum of M'_n .

The same result holds when M_n is a real Gaussian matrix, except that ξ_{ij} are now i.i.d. copies of the real Gaussian $N(0, 1)_{\mathbb{R}}$, and the $\chi_{i,\mathbb{C}}$ are replaced with real χ distributions $\chi_{i,\mathbb{R}}$ with i degrees of freedom.

PROOF. This result appears in [33], Section 2, but for the convenience of the reader we supply a proof here. We establish the complex case only, as the real case is similar, making the obvious changes (such as replacing the unitary matrices in the argument below by orthogonal matrices instead).

The idea will be to exploit the unitary invariance of complex Gaussian vectors by taking a complex Gaussian matrix M_n and conjugating it by unitary matrices (which will depend on M_n) until one arrives at a matrix with the distribution of M'_n .

Write the first row of M_n as $(\xi_{11}, \dots, \xi_{1n})$. Then there is a unitary transformation U_1 that preserves the first basis vector e_1 , and maps $(\xi_{11}, \dots, \xi_{1n})$ to $(\xi_{11}, \chi_{n-1,\mathbb{C}}, 0, \dots, 0)$, where $\chi_{n-1,\mathbb{C}}$ is a complex χ distribution with $n - 1$ degrees of freedom. If we then conjugate M_n by U_1 , and use the fact that the conjugate of a Gaussian vector by a unitary matrix that is independent of that vector, remains distributed as a Gaussian vector, we see that the conjugate $U_1 M_n U_1^*$ to a matrix takes the form

$$\begin{pmatrix} \xi_{11} & \chi_{n-1,\mathbb{C}} & 0 & \cdots & 0 \\ \xi_{21} & \xi_{22} & \xi_{23} & \cdots & \xi_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \xi_{n1} & \xi_{n2} & \xi_{n3} & \cdots & \xi_{nn} \end{pmatrix},$$

where the ξ_{ij} coefficients appearing in this matrix are i.i.d. copies of $N(0, 1)_{\mathbb{C}}$ (and are not necessarily equal to the corresponding coefficients of M_n), and $\chi_{n-1,\mathbb{C}}$ is independent of all of the ξ_{ij} .

We may then find another unitary transformation U_2 that preserves e_1 and e_2 , and maps the second row $(\xi_{21}, \dots, \xi_{2n})$ of $U_1 M_n U_1^*$ to $(\xi_{21}, \xi_{22}, \chi_{n-2,\mathbb{C}}, 0, \dots, 0)$,

where $\chi_{n-2,\mathbb{C}}$ is distributed by the complex χ distribution with $n - 2$ degrees of freedom. Conjugating $U_1 M_n U_1^*$ by U_2 , we arrive at a matrix of the form

$$\begin{pmatrix} \xi_{11} & \chi_{n-1,\mathbb{C}} & 0 & 0 & \cdots & 0 \\ \xi_{21} & \xi_{22} & \chi_{n-2,\mathbb{C}} & 0 & \cdots & 0 \\ \xi_{31} & \xi_{32} & \xi_{33} & \xi_{34} & \cdots & \xi_{3n} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \xi_{n1} & \xi_{n2} & \xi_{n3} & \cdots & \xi_{nn} \end{pmatrix},$$

where the ξ_{ij} coefficients appearing in this matrix are again i.i.d. copies of $N(0, 1)_{\mathbb{C}}$ (though they are not necessarily identical to their counterparts in the previous matrix $U_1 M_n U_1^*$), and $\chi_{n-1,\mathbb{C}}$ and $\chi_{n-2,\mathbb{C}}$ are independent of each other and of the ξ_{ij} . Iterating this procedure a total of $n - 1$ times, we obtain the claim. \square

We now use this conjugated form of the complex Gaussian matrix M_n to describe the characteristic polynomial $\det(M_n - z_0\sqrt{n})$.

PROPOSITION 51. *Let z_0 be a complex number, and let M_n be a complex Gaussian matrix. Let $\chi_{1,\mathbb{C}}, \dots, \chi_{n-1,\mathbb{C}}$ be a sequence of independent random variables distributed according to the complex χ distributions with $1, \dots, n - 1$ degrees of freedom, respectively. Let ξ_1, \dots, ξ_n be another sequence of independent random variables distributed according to the complex Gaussian $N(0, 1)_{\mathbb{C}}$, and independent of the χ_i . Define the sequence a_1, \dots, a_n of complex random variables recursively by setting*

$$(9.3) \quad a_1 := \xi_1 - z_0\sqrt{n}$$

and

$$(9.4) \quad a_{i+1} := \frac{-z_0\sqrt{n}a_i}{\sqrt{|a_i|^2 + \chi_{n-i,\mathbb{C}}^2}} + \xi_{i+1}$$

for $i = 1, \dots, n - 1$. (Note that the a_i are almost surely well defined.) Then the random variable

$$\left(\prod_{i=1}^{n-1} \sqrt{|a_i|^2 + \chi_{n-i,\mathbb{C}}^2} \right) a_n$$

has the same distribution as $\det(M_n - z_0\sqrt{n})$.

The same conclusions hold when M_n is a real Gaussian matrix, after replacing ξ_i with copies of the real Gaussian $N(0, 1)_{\mathbb{C}}$, and replacing $\chi_{i,\mathbb{C}}$ with a real χ distribution $\chi_{i,\mathbb{R}}$ with i degrees of freedom.

We remark that in [33] a slightly different stochastic equation (a Hilbert space variant of the Pólya urn process) for the determinants $\det(M_n - z_0\sqrt{n})$ were given, in which the value of each determinant was influenced by a Gaussian variable whose variance depended on all of the determinants of the top left $k \times k$ minors for $k = 1, \dots, n - 1$. In contrast, the recurrence here is more explicitly Markovian in the sense that the state a_{i+1} of the recursion at time $i + 1$ only depends (stochastically) on the state a_i at the immediately preceding time. We will rely heavily on the Markovian nature of the process in the subsequent analysis.

PROOF OF PROPOSITION 51. Again, we argue for the complex Gaussian case only, as the real Gaussian case proceeds similarly with the obvious modifications.

By Proposition 50, $\det(M_n - z_0\sqrt{n})$ has the same distribution as $\det(M'_n - z_0\sqrt{n})$. The strategy is then to manipulate $M'_n - z_0\sqrt{n}$ by elementary column operations that preserve the determinant, until it becomes a lower triangular matrix whose diagonal entries have the joint distribution of $(\sqrt{|a_i|^2 + \chi_{n-i,\mathbb{C}}^2})_{i=1}^{n-1}, a_n$, at which point the claim follows.

We turn to the details. Writing $\xi_1 := \xi_{11}$, we see that $M'_n - z_0\sqrt{n}$ can be written as

$$\begin{pmatrix} a_1 & \chi_{n-1,\mathbb{C}} & 0 & 0 & \cdots & 0 \\ \xi_{21} & \xi_{22} - z_0\sqrt{n} & \chi_{n-2,\mathbb{C}} & 0 & \cdots & 0 \\ \xi_{31} & \xi_{32} & \xi_{33} - z_0\sqrt{n} & \chi_{n-3,\mathbb{C}} & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \xi_{(n-1)1} & \xi_{(n-1)2} & \xi_{(n-1)3} & \xi_{(n-1)4} & \cdots & \chi_{1,\mathbb{C}} \\ \xi_{n1} & \xi_{n2} & \xi_{n3} & \xi_{n4} & \cdots & \xi_{nn} - z_0\sqrt{n} \end{pmatrix}.$$

Note that there is a unitary matrix U_1 whose action on row vectors (multiplying on the right) maps $(a_1, \chi_{n-1,\mathbb{C}}, 0, \dots, 0)$ to $(\sqrt{|a_1|^2 + \chi_{n-1,\mathbb{C}}^2}, 0, \dots, 0)$, and which only modifies the first two coefficients of a row vector. This corresponds to a column operation that modifies the first two columns of a matrix in a unitary fashion (by multiplying that matrix on the right by U_1). Because complex Gaussian vectors remain Gaussian after unitary transformations, we see (after a brief computation) that this transformation maps the second row $(\xi_{21}, \xi_{22} - z_0\sqrt{n}, \chi_{n-2,\mathbb{C}}, 0, \dots, 0)$ of the above matrix to a vector of the form

$$\left(*, \frac{-z_0\sqrt{n}a_1}{\sqrt{|a_1|^2 + \chi_{n-1,\mathbb{C}}^2}} + \xi_2, \chi_{n-2,\mathbb{C}}, \dots, 0 \right),$$

where ξ_2 is a complex Gaussian (formed by some combination of ξ_{21} and ξ_{22}) and $*$ is a quantity whose exact value will not be relevant for us. By (9.4), we may denote the second coefficient of this vector by a_2 . The remaining rows of the matrix have their distribution unchanged by the unitary matrix U_1 , because

their first two entries form a complex Gaussian vector. Thus, after applying the U_1 column operation to the above matrix, we arrive at a matrix with the distribution

$$\begin{pmatrix} \sqrt{|a_1|^2 + \chi_{n-1, \mathbb{C}}^2} & 0 & 0 & 0 & \cdots & 0 \\ * & a_2 & \chi_{n-2, \mathbb{C}} & 0 & \cdots & 0 \\ \xi_{31} & \xi_{32} & \xi_{33} - z_0\sqrt{n} & \chi_{n-3, \mathbb{C}} & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \xi_{(n-1)1} & \xi_{(n-1)2} & \xi_{(n-1)3} & \xi_{(n-1)4} & \cdots & \chi_{1, \mathbb{C}} \\ \xi_{n1} & \xi_{n2} & \xi_{n3} & \xi_{n4} & \cdots & \xi_{nn} - z_0\sqrt{n} \end{pmatrix},$$

where the ξ_{ij} here are i.i.d. copies of $N(0, 1)_{\mathbb{C}}$ that are independent of a_1, a_2 and the $\chi_{i, \mathbb{C}}$ (and which are not necessarily identical to their counterparts in the previous matrix under consideration). Of course, the determinant of this matrix has the same distribution as the determinant of the preceding matrix.

In a similar fashion, we may find a unitary matrix U_2 whose action on row vectors maps $(*, a_1, \chi_{n-2, \mathbb{C}}, 0, \dots, 0)$ to $(*, \sqrt{|a_2|^2 + \chi_{n-2, \mathbb{C}}^2}, 0, \dots, 0)$, and which only modifies the second and third coefficients of a row vector. Applying the associated column operation, and arguing as before, we arrive at a matrix with the distribution

$$\begin{pmatrix} \sqrt{|a_1|^2 + \chi_{n-1, \mathbb{C}}^2} & 0 & 0 & 0 & \cdots & 0 \\ * & \sqrt{|a_2|^2 + \chi_{n-2, \mathbb{C}}^2} & 0 & 0 & \cdots & 0 \\ * & * & a_3 & \chi_{n-3, \mathbb{C}} & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \xi_{(n-1)1} & \xi_{(n-1)2} & \xi_{(n-1)3} & \xi_{(n-1)4} & \cdots & \chi_{1, \mathbb{C}} \\ \xi_{n1} & \xi_{n2} & \xi_{n3} & \xi_{n4} & \cdots & \xi_{nn} - z_0\sqrt{n} \end{pmatrix},$$

where again the values of the entries marked $*$ are not relevant for us. Iterating this procedure a total of $n - 1$ times, we finally arrive at a lower triangular matrix whose diagonal entries have the distribution of

$$\left(\sqrt{|a_1|^2 + \chi_{n-1, \mathbb{C}}^2}, \sqrt{|a_2|^2 + \chi_{n-2, \mathbb{C}}^2}, \dots, \sqrt{|a_{n-1}|^2 + \chi_{1, \mathbb{C}}^2}, a_n \right)$$

and whose determinant has the same distribution as that of $M'_n - z_0\sqrt{n}$ or $M_n - z_0\sqrt{n}$. The claim follows. \square

9.3. *A nonlinear stochastic difference equation.* For the sake of exposition, we now specialize to the complex Gaussian case; the case when M_n is a real Gaussian is similar and we will indicate at various junctures what changes need to be made.

From Proposition 51, we see that $\log |\det(M_n - z_0\sqrt{n})|$ has the same distribution as

$$(9.5) \quad \frac{1}{2} \sum_{i=1}^{n-1} \log(|a_i|^2 + \chi_{n-i, \mathbb{C}}^2) + \log |a_n|.$$

It thus suffices to establish the lower bound

$$(9.6) \quad \frac{1}{2} \sum_{i=1}^{n-1} \log(|a_i|^2 + \chi_{n-i,\mathbb{C}}^2) + \log |a_n| \geq \frac{1}{2}n \log n + \alpha n - n^{o(1)}$$

with overwhelming probability.

We first note that as the distribution of $\log |\det(M_n - z_0\sqrt{n})|$ is invariant with respect to phase rotation $z_0 \mapsto z_0 e^{\sqrt{-1}\theta}$, we may assume without loss of generality that z_0 is real and nonpositive, thus

$$(9.7) \quad a_{i+1} := \frac{|z_0|\sqrt{n}a_i}{\sqrt{|a_i|^2 + \chi_{n-i,\mathbb{C}}^2}} + \xi_{i+1}.$$

REMARK 52. In the real Gaussian case, one does not have phase rotation invariance. However, by making the change of variables $a'_i := a_i e^{-\sqrt{-1}i\theta}$ one can obtain the variant

$$(9.8) \quad a'_{i+1} := \frac{|z_0|\sqrt{n}a'_i}{\sqrt{|a_i|^2 + \chi_{n-i,\mathbb{R}}^2}} + \xi'_{i+1}$$

to (9.7), where $\xi'_{i+1} := e^{-\sqrt{-1}i\theta} \xi_{i+1}$. It will turn out that this recurrence is similar enough to (9.7) that the arguments below used to study (9.7) can be adapted to (9.8); the ξ'_i are no longer identically distributed, but they still have mean zero, variance one, and are jointly independent, and this is all that is needed in the arguments that follow.

The random variable $\chi_{n-i,\mathbb{C}}^2$ has mean $n - i$ and variance $n - i$. As such, it is natural to make the change of variables

$$\chi_{n-i,\mathbb{C}} =: n - i + \sqrt{n - i} \eta_{n-i},$$

where the $\eta_1, \dots, \eta_{n-1}$ have mean zero, variance one and are independent of each other and of the ξ_i .

REMARK 53. For real Gaussian matrices, the situation is very similar, except that the error terms η_{n-i} now have variance two instead of one. However, this will not significantly affect the concentration results for the log-determinant in this paper. (This will however presumably affect any central limit theorems one could establish for the log-determinant, in analogy with [58], though we will not pursue such theorems here.)

We now pause to perform a technical truncation. As the ξ_i are distributed in a Gaussian fashion, we know that

$$(9.9) \quad \sup_{1 \leq i \leq n} |\xi_i| \leq n^{o(1)}$$

with overwhelming probability. Similarly, standard asymptotics for chi-square distributions also give the bound

$$(9.10) \quad \sup_{1 \leq i < n} |\eta_i| \leq n^{o(1)}$$

with overwhelming probability (this bound also follows from Proposition 35).

We may now condition on the event that (9.9), (9.10) hold [for a suitable choice of the $o(1)$ decay exponent]. Importantly, the joint independence of the $\xi_1, \dots, \xi_n, \eta_1, \dots, \eta_{n-1}$ remain unchanged by this conditioning. Of course, the distribution of the ξ_i and η_i will be slightly distorted by this conditioning, but this will not cause a difficulty in practice, as the mean, variances, and higher moments of these variables are only modified by $O(n^{-100})$ (say) at most, and also we will at key junctures in the proof be able to undo the conditioning (after accepting an event of negligible probability) in order to restore the original distributions of ξ_i and η_i if needed.

We return to the task of proving (9.6). We write (9.7) as

$$(9.11) \quad a_{i+1} := \frac{|z_0| \sqrt{na_i}}{\sqrt{|a_i|^2 + n - i + \sqrt{n - i} \eta_{n-i}}} + \xi_{i+1}.$$

We will treat this as a nonlinear stochastic difference equation in the a_i . If we ignore the diffusion terms η_{n-i}, ξ_{i+1} , we see that (9.11) is governed by the dynamics of the maps

$$(9.12) \quad a \mapsto \frac{|z_0| \sqrt{na}}{\sqrt{|a|^2 + n - i}}$$

as i increases from 1 to $n - 1$. In the regime $i < (1 - |z_0|^2)n$, we see that this map has a stable fixed point at zero, while in the regime $i > (1 - |z_0|^2)n$, this map has an unstable fixed point at zero and a fixed circle at $|a| = \sqrt{|z_0|^2 n - (n - i)}$. This suggests that $|a_i|$ should concentrate somehow around 0 for $i \leq (1 - |z_0|^2)n$ and around $\sqrt{|z_0|^2 n - (n - i)}$ for $i \geq (1 - |z_0|^2)n$. In particular, this leads to the heuristic

$$|a_i|^2 + \chi_{n-i, \mathbb{C}}^2 \approx \max(n - i, |z_0|^2 n).$$

Note from the integral test that

$$(9.13) \quad \begin{aligned} & \frac{1}{2} \sum_{i=1}^{n-1} \log \max(n - i, |z_0|^2 n) \\ &= \frac{1}{2} \int_1^n \log \max(n - t, |z_0|^2 n) dt + O(n^{o(1)}) \\ &= \frac{1}{2} n \log n + \alpha n + O(n^{o(1)}), \end{aligned}$$

where the second identity follows from a routine integration (treating the cases $|z_0| \leq 1$ and $|z_0| \geq 1$ separately). This gives heuristic support for the desired bound (9.6).

We now make the above analysis rigorous. Because we are only seeking a lower bound (9.6), the main task will be to obtain lower bounds that are roughly of the form

$$|a_i|^2 + \chi_{n-i, \mathbb{C}}^2 \gtrsim \max(n - i, |z_0|^2 n)$$

with overwhelming probability. In the “early regime” $i \leq (1 - |z_0|^2)n$, we will be able to achieve this easily from the trivial bound $|a_i| \geq 0$. In the “late regime” $i \geq (1 - |z_0|^2)n$, the main difficulty is then to show (with overwhelming probability) that a_i avoids the unstable fixed point at zero, and instead is essentially at least as far away from the origin as the fixed circle $|a| = \sqrt{|z_0|^2 n - (n - i)}$.

We turn to the details. We begin with a crude bound on the magnitude of the quantities a_i .

LEMMA 54 (Crude lower bound). *Almost surely [after conditioning to (9.9) and (9.10)], one has*

$$(9.14) \quad \sup_{1 \leq i \leq n} |a_i| \leq (1 + |z_0|)\sqrt{n}$$

and with overwhelming probability

$$(9.15) \quad \inf_{1 \leq i \leq n} |a_i| \geq \exp(-n^{o(1)}).$$

PROOF. From (9.3), (9.9), we see that we have

$$|a_1| \leq 2\sqrt{n}.$$

From (9.7) (trivially bounding χ_{n-i} from below by zero), we have

$$|a_{i+1}| \leq |z_0|\sqrt{n} + |\xi_{i+1}|$$

and so the bound (9.14) follows from (9.9) and the assumption that $|z_0| \leq 1$.

Now we prove (9.15). Let $A \geq 0$ be fixed. Observe that ξ_1 has a bounded density function [even after conditioning on (9.9)], so from (9.3) we have

$$|a_1| \geq n^{-A}$$

with probability¹⁹ $1 - O(n^{-2A})$. In a similar spirit, for any $i = 1, \dots, n - 1$, ξ_{i+1} has a bounded density function, so from (9.7) or (9.11) (after temporarily conditioning a_i and η_{n-i} to be fixed) that

$$|a_{i+1}| \geq n^{-A}$$

¹⁹In the real Gaussian case, the n^{-2A} factor worsens to n^{-A} , but this does not impact the final conclusion.

with probability $1 - O(n^{-2A})$. By the union bound, we conclude that

$$\inf_{1 \leq i \leq n} |a_i| \geq n^{-A}$$

with probability $1 - O(n^{-2A+1})$. Diagonalizing in A , we obtain the claim. \square

From this lemma, we conclude that

$$(9.16) \quad \log |a_i| = n^{o(1)}$$

with overwhelming probability for each $1 \leq i \leq n$. To show (9.6), it thus suffices to establish, for each fixed $\varepsilon > 0$, that

$$\frac{1}{2} \sum_{i=1}^{n-1} \log(|a_i|^2 + \chi_{n-i, \mathbb{C}}^2) \geq \frac{1}{2} n \log n + \alpha n - O(n^{O(\varepsilon)})$$

with overwhelming probability, where the implied constant in the $O(\varepsilon)$ notation is understood to be independent of ε of course.

In view of (9.13), it will suffice to show that

$$(9.17) \quad \sum_{n^\varepsilon < i \leq n-n^\varepsilon} (\log(|a_i|^2 + \chi_{n-i, \mathbb{C}}^2) - \log \max(n-i, |z_0|^2 n)) \geq -O(n^{O(\varepsilon)})$$

with overwhelming probability, as the contributions of the i within n^ε of 1 or n can be controlled by $O(n^{\varepsilon+o(1)})$ thanks to Lemma 54.

9.4. *Lower bound at early times.* We partition $\sum_{n^\varepsilon < i \leq n-n^\varepsilon} (\log(|a_i|^2 + \chi_{n-i, \mathbb{C}}^2) - \log \max(n-i, |z_0|^2 n))$ into two parts, according to the heuristics following (9.12). The following simple lemma handles the first part of the partition.

LEMMA 55 (Concentration at early times). *One has*

$$\begin{aligned} & \sum_{n^\varepsilon < i \leq \min((1-|z_0|^2)n + |z_0|n^{1/2+\varepsilon}, n-n^\varepsilon)} \log(|a_i|^2 + \chi_{n-i, \mathbb{C}}^2) - \log \max(n-i, |z_0|^2 n) \\ & \geq -O(n^{O(\varepsilon)}) \end{aligned}$$

with overwhelming probability.

PROOF. We abbreviate the summation as \sum_i . The key observation here is that we need only a lower bound, so we can use the trivial inequality

$$\log(|a_i|^2 + \chi_{n-i, \mathbb{C}}) \geq \log \chi_{n-i, \mathbb{C}}.$$

It suffices to show that

$$(9.18) \quad \sum_i |\log(n-i) - \log \max(n-i, |z_0|^2 n)| = O(n^{O(\varepsilon)})$$

and

$$(9.19) \quad \sum_i \log \chi_{n-i, \mathbb{C}}^2 - \log(n - i) = O(n^{O(\varepsilon)})$$

with overwhelming probability.

We first verify (9.18). The summand is only nonzero when $i = (1 - |z_0|^2)n + j$ for some $0 < j \leq \min(|z_0|n^{1/2+\varepsilon}, |z_0|^2n - n^\varepsilon)$, and so one can bound the left-hand side of (9.18) by

$$\sum_{0 < j \leq \min(|z_0|n^{1/2+\varepsilon}, |z_0|^2n - n^\varepsilon)} |\log(|z_0|^2n - j) - \log(|z_0|^2n)|.$$

When $j \leq |z_0|^2n - n^\varepsilon$, we may bound

$$|\log(|z_0|^2n - j) - \log(|z_0|^2n)| \ll n^{o(1)} \frac{j}{|z_0|^2n}$$

and the claim then follows by summing over all $0 < j \leq |z_0|n^{1/2+\varepsilon}$.

Now we verify (9.19), which is quite standard. Writing $\chi_{n-i, \mathbb{C}}^2 = n - i + \sqrt{n - i}\eta_{n-i}$, we can write the left-hand side of (9.19) as

$$\sum_i \log\left(1 + \frac{\eta_{n-i}}{\sqrt{n - i}}\right).$$

From Taylor expansion and (9.10) we then have

$$\log\left(1 + \frac{\eta_{n-i}}{\sqrt{n - i}}\right) = \frac{\eta_{n-i}}{\sqrt{n - i}} + O\left(\frac{n^{o(1)}}{n - i}\right).$$

The sum of the error term is acceptable, so it suffices to show that

$$\sum_i \frac{\eta_{n-i}}{\sqrt{n - i}} = O(n^{O(\varepsilon)})$$

with overwhelming probability. But this follows²⁰ from Proposition 35. \square

REMARK 56. Following the heuristics after (9.12), it would be more natural to consider $n^\varepsilon \leq i \leq (1 - |z_0|^2)n$. The extra term $|z_0|n^{1/2+\varepsilon}$ in the upper bound of i is needed for a technical reason which will be clear in the analysis of larger i (see Lemma 58).

²⁰Strictly speaking, Proposition 35 does not apply directly because the mean of the random variables η_{n-i} deviates very slightly from zero when the conditioning (9.10) is applied. However, one can first apply Proposition 35 to the unconditioned variables η_{n-i} , and then apply the conditioning (9.10) that is in force elsewhere in this argument, noting that such conditioning does not affect the property of an event occurring with overwhelming probability.

9.5. *Concentration at late times.* Define

$$(9.20) \quad i_0 := \max(n^\varepsilon, (1 - |z_0|^2)n + |z_0|n^{1/2+\varepsilon}).$$

In view of Lemma 55, we see that to prove (9.17) it now suffices to establish the lower bound

$$(9.21) \quad \sum_{i_0 < i \leq n-n^\varepsilon} \log(|a_i|^2 + \chi_{n-i, \mathbb{C}}^2) - \log(|z_0|^2 n) = O(n^{O(\varepsilon)})$$

with overwhelming probability. In fact, we only need the lower bound from (9.21), but the argument given here gives the matching upper bound as well with no additional effort.

Let us first deal with the easy case when

$$(9.22) \quad |z_0| < n^{-1/2+400\varepsilon}$$

(say). In this case, there are only $O(n^{800\varepsilon})$ terms in the sum, and from Lemma 54 (discarding the nonnegative $\chi_{n-i, \mathbb{C}}^2$ term) each term is at least $-O(n^{o(1)})$, so the claim (9.21) follows immediately. [Note that the summation is in fact empty unless $|z_0| \geq n^{-1/2+\varepsilon/2}$, so the $\log(|z_0|^2 n)$ term is $O(n^{o(1)})$.] Thus, in the arguments below we can assume that

$$(9.23) \quad |z_0| \geq n^{-1/2+400\varepsilon}.$$

Observe from (9.7) that

$$\log(|a_i|^2 + \chi_{n-i, \mathbb{C}}^2) - \log(|z_0|^2 n) = \log \frac{|a_{i+1} - \xi_{i+1}|^2}{|a_i|^2}.$$

From telescoping series and (9.16) we have

$$\sum_{i_0 < i \leq n-n^\varepsilon} \log \frac{|a_{i+1}|^2}{|a_i|^2} = O(n^{o(1)})$$

with overwhelming probability, so by the triangle inequality it suffices to show that

$$\sum_{i_0 < i \leq n-n^\varepsilon} \log \frac{|a_{i+1} - \xi_{i+1}|^2}{|a_{i+1}|^2} = O(n^{O(\varepsilon)})$$

with overwhelming probability. We can rewrite

$$\frac{|a_{i+1} - \xi_{i+1}|^2}{|a_{i+1}|^2} = \left| 1 + \frac{\xi_{i+1}}{a'_i} \right|^{-2},$$

where

$$(9.24) \quad a'_i := a_{i+1} - \xi_{i+1} = \frac{|z_0| \sqrt{n} a_i}{\sqrt{|a_i|^2 + \chi_{n-1, \mathbb{C}}}}.$$

It suffices to show that

$$\sum_{i_0 < i \leq n - n^\varepsilon} \log \left| 1 + \frac{\xi_{i+1}}{a'_i} \right| = O(n^{O(\varepsilon)})$$

with overwhelming probability.

The heart of the matter will be the following lemma.

LEMMA 57. *With overwhelming probability*

$$(9.25) \quad |a'_i| \gg n^{-100\varepsilon} (i - (1 - |z_0|^2)n)^{1/2}$$

holds for all $i_0 < i \leq n - n^\varepsilon$.

Assuming this lemma for the moment, we can then use it to conclude the proof as follows. For any $i_0 < i \leq n - n^\varepsilon$, one has

$$(9.26) \quad (i - (1 - |z_0|^2)n)^{1/2} > (i_0 - (1 - |z_0|^2)n)^{1/2} \geq (|z_0|n^{1/2+\varepsilon})^{1/2} \geq n^{200\varepsilon}$$

by (9.20) and (9.23), and thus by Lemma 57

$$|a'_i| \gg n^{100\varepsilon}$$

with overwhelming probability. From this and (9.9), we see that

$$\left| \frac{\xi_{i+1}}{a'_i} \right| = o(1);$$

indeed, the same argument gives the more precise bound

$$\left| \frac{\xi_{i+1}}{a'_i} \right| \ll n^{O(\varepsilon)} (i - (1 - |z_0|^2)n)^{-1/2}.$$

Performing a Taylor expansion (up to the second-order term), we conclude that

$$\log \left| 1 + \frac{\xi_{i+1}}{a'_i} \right| = \operatorname{Re} \xi_{i+1}/a'_i + O(n^{O(\varepsilon)} (i - (1 - |z_0|^2)n)^{-1})$$

with overwhelming probability.

The error terms $O(n^{O(\varepsilon)} (i - (1 - |z_0|^2)n)^{-1})$ sum to $O(n^{O(\varepsilon)})$, so it suffices to show that

$$(9.27) \quad \sum_{i_0 < i \leq n - n^\varepsilon} \frac{\xi_{i+1}}{a'_i} = O(n^{O(\varepsilon)})$$

with overwhelming probability. But from (9.25), one has

$$\frac{1}{a'_i} = O(n^{O(\varepsilon)} (i - (1 - |z_0|^2)n)^{-1/2})$$

with overwhelming probability. Also, the coefficient $\frac{1}{a_i}$ depends on ξ_1, \dots, ξ_i and $\chi_{1,\mathbb{C}}, \dots, \chi_{n,\mathbb{C}}$ and is independent of ξ_{i+1}, \dots, ξ_n , so the sum in (9.27) becomes a martingale sum.²¹ The claim then follows from Proposition 35.

It remains to prove (9.25). From (9.7), (9.24), (9.9), we have

$$a'_i = a_{i+1} - \xi_{i+1} = a_{i+1} + O(n^{o(1)})$$

and so by (9.26) it will suffice to establish the bound

$$(9.28) \quad |a_i| \gg n^{-99\varepsilon} (i - (1 - |z_0|^2)n)^{1/2}$$

with overwhelming probability for each $i_0 < i \leq n - n^\varepsilon + 1$.

In order to prove (9.28), let us first establish a preliminary largeness result on a_i , which uses the diffusive term ξ_{i+1} in (9.7) to push this random variable away from the unstable equilibrium 0 of the map (9.12).

LEMMA 58 (Initial largeness). *With overwhelming probability, one has*

$$(9.29) \quad \sup_{\max(i_0 - (1/2)|z_0|n^{1/2+\varepsilon}, 0) \leq i \leq i_0} |a_i| > A,$$

where A is the quantity

$$A := |z_0|^{1/2} n^{1/4+\varepsilon/10}.$$

PROOF. Suppose first that

$$i_0 - \frac{1}{2}|z_0|n^{1/2+\varepsilon} \leq 0.$$

By (9.20), this implies that $|z_0| \gg 1$, and then from (9.3), (9.9) we have $|a_1| \gg n^{1/2}$, which certainly gives (9.29) in this case. Thus, we may assume that

$$i_0 - \frac{1}{2}|z_0|n^{1/2+\varepsilon} > 0.$$

It will suffice to show that, for each integer

$$i_0 - \frac{1}{2}|z_0|n^{1/2+\varepsilon} \leq i_1 \leq i_0$$

and each fixed (i.e., conditioned) choice of ξ_1, \dots, ξ_{i_1} and $\chi_{n-1,\mathbb{C}}, \dots, \chi_{n-i_1}$, one has

$$(9.30) \quad \sup_{i_1 \leq i \leq i_1 + |z_0|n^{1/2+\varepsilon/2}} |a_i| > A$$

with conditional probability at least q for some fixed $q > 0$. Indeed, we can choose in the interval $[i_0 - \frac{1}{2}|z_0|n^{1/2+\varepsilon}, i_0 - |z_0|n^{1/2+\varepsilon/2}]$ at least $\frac{n^{\varepsilon/2}}{100}$ initial points

²¹Again, strictly speaking one should apply Proposition 35 to the unconditioned variables and then apply the conditioning (9.9), (9.10), as in Lemma 55.

i_1, \dots, i_m so that the distance between any two of them is at least $|z_0|n^{1/2+\varepsilon/2}$. If we let E_j for $j = 1, \dots, m$ be the event that (9.30) holds with i_1 replaced by i_j , then the above claim asserts that after conditioning on the failure of the events E_1, \dots, E_{j-1} , the event E_j holds with conditional probability at least q . Multiplying the conditional probabilities together, we then obtain (9.29) with a failure probability of at most

$$(1 - q)^{n^{\varepsilon/2}/4},$$

which is $O(n^{-A})$ for any fixed $A > 0$ as required.

Fix $i_0 - \frac{1}{2}|z_0|n^{1/2+\varepsilon} \leq i_1 \leq i_0$ and ξ_1, \dots, ξ_{i_1} and $\chi_{n-1, \mathbb{C}}, \dots, \chi_{n-i_1, \mathbb{C}}$; all probabilities in this argument are now understood to be conditioned on these choices. The quantity a_{i_1} is now deterministic, and we may of course assume that

$$(9.31) \quad |a_{i_1}| \leq A$$

as the claim is trivial otherwise. We may also condition on the event that (9.10) hold. Let $i_2 := \lfloor i_1 + |z_0|n^{1/2+\varepsilon/2} \rfloor$. Our goal is to show that

$$\mathbf{P}\left(\sup_{i_1 \leq i \leq i_2} |a_i| > A\right) \gg 1.$$

For technical reasons [having to do with the contractive nature of the recursion (9.7) when a_i becomes large], it will be convenient to replace the random process a_i by a slightly truncated random process \tilde{a}_i for $i_0 \leq i \leq i_1$, which is defined by setting $\tilde{a}_{i_1} := a_{i_1}$ and

$$(9.32) \quad \tilde{a}_{i+1} := \frac{|z_0|\sqrt{n}\tilde{a}_i}{\sqrt{\min(|\tilde{a}_i|, A)^2 + \chi_{n-i, \mathbb{C}}^2}} + \xi_{i+1}$$

for $i_1 \leq i < i_2$. From an induction on the upper range i_2 of the i parameter, we see that

$$\sup_{i_1 \leq i \leq i_2} |a_i| \leq A \iff \sup_{i_1 \leq i \leq i_2} |\tilde{a}_i| \leq A$$

and in particular

$$|\tilde{a}_{i_2}| > A \implies \sup_{i_1 \leq i \leq i_2} |a_i| > A.$$

Thus, it will suffice to show that

$$(9.33) \quad \mathbf{P}(|\tilde{a}_{i_2}| > A) \gg 1.$$

By a standard Paley–Zygmund type argument, it will suffice to obtain the lower bound

$$(9.34) \quad \mathbf{E}|\tilde{a}_{i_2}|^2 \gg |z_0|n^{1/2+\varepsilon/2}$$

on the second moment, and the upper bound

$$(9.35) \quad \mathbf{E}|\tilde{a}_{i_2}|^4 \ll |z_0|^2 n^{1+\varepsilon} + |z_0| n^{1/2+\varepsilon/2} \mathbf{E}|\tilde{a}_{i_2}|^2$$

on the fourth moment. Indeed, if p denotes the probability in (9.33), then from Hölder’s inequality one has

$$\mathbf{E}|\tilde{a}_{i_2}|^2 \ll A^2 + p^{1/2} (\mathbf{E}|\tilde{a}_{i_2}|^4)^{1/2}$$

and then from (9.35) and (9.34) (and the definition of A) we obtain $p \gg 1$ as required.

It remains to establish (9.34) and (9.35). For this, we will use (9.32) to track the growth of the moments $\mathbf{E}|\tilde{a}_i|^2, \mathbf{E}|\tilde{a}_i|^4$ as i increases from i_1 to i_2 .

Let $i_1 \leq i < i_2$. From (9.32), we thus have

$$\mathbf{E}|\tilde{a}_{i+1}|^2 = \mathbf{E} \left| \frac{|z_0| \sqrt{n} \tilde{a}_i}{\sqrt{\min(|\tilde{a}_i|, A)^2 + n - i + \sqrt{n - i} \eta_{n-i}}} + \xi_{i+1} \right|^2$$

The quantity ξ_{i+1} has mean $O(n^{-100})$, variance $1 + O(n^{-100})$ [the $O(n^{-100})$ errors arising from our conditioning to (9.9)], and is independent of the other random variables on the right-hand side. Thus [using (9.14)], we have

$$\mathbf{E}|\tilde{a}_{i+1}|^2 = \mathbf{E} \left| \frac{|z_0| \sqrt{n} \tilde{a}_i}{\sqrt{\min(|\tilde{a}_i|, A)^2 + n - i + \sqrt{n - i} \eta_{n-i}}} \right|^2 + 1 + O(n^{-90}).$$

Upper bounding $\min(|\tilde{a}_i|, A)$ by A and $n - i$ by $|z_0|^2 \sqrt{n} - |z_0| n^{1/2+\varepsilon}/2$, and using (9.10) (which we recall that we have conditioned on), we conclude that

$$\min(|\tilde{a}_i|, A)^2 + n - i + \sqrt{n - i} \eta_{n-i} \leq |z_0|^2 n.$$

This implies that

$$(9.36) \quad \mathbf{E}|\tilde{a}_{i+1}|^2 \geq \mathbf{E}|\tilde{a}_i|^2 + 1 + O(n^{-90}).$$

Iterating this $\gg |z_0| n^{1/2+\varepsilon/2}$ times, we obtain (9.34) as required.

Now we turn to (9.35). Again, we let $i_1 \leq i < i_2$. From (9.32), we have

$$\mathbf{E}|\tilde{a}_{i+1}|^4 = \mathbf{E} \left| \frac{|z_0| \sqrt{n} \tilde{a}_i}{\sqrt{\min(|\tilde{a}_i|, A)^2 + n - i + \sqrt{n - i} \eta_{n-i}}} + \xi_{i+1} \right|^4.$$

Expanding out the left-hand side using the independence and moment properties of ξ_{i+1} , we can estimate the above expression as

$$\begin{aligned} & \mathbf{E} \left| \frac{|z_0| \sqrt{n} \tilde{a}_i}{\sqrt{\min(|\tilde{a}_i|, A)^2 + n - i + \sqrt{n - i} \eta_{n-i}}} \right|^4 \\ & + O \left(\mathbf{E} \left| \frac{|z_0| \sqrt{n} \tilde{a}_i}{\sqrt{\min(|\tilde{a}_i|, A)^2 + n - i + \sqrt{n - i} \eta_{n-i}}} \right|^2 + 1 \right). \end{aligned}$$

Using (9.9), (9.10) and the bound $n - i \geq |z_0|^2 n - O(|z_0| n^{1/2+\varepsilon})$, and discarding the nonnegative $\min(|\tilde{a}_i|, A)^2$ term, we then obtain the upper bound

$$(9.37) \quad \mathbf{E}|\tilde{a}_{i+1}|^4 \leq (1 + O(|z_0|^{-1} n^{-1/2+\varepsilon})) \mathbf{E}|\tilde{a}_i|^4 + O(\mathbf{E}|\tilde{a}_i|^2 + 1),$$

via a routine calculation. From (9.36), we have

$$\mathbf{E}|\tilde{a}_i|^2 \ll \mathbf{E}|\tilde{a}_{i_2}|^2.$$

From (9.31), we also have

$$\mathbf{E}|\tilde{a}_{i_1}|^4 \ll |z_0|^2 n^{1+\varepsilon};$$

if we then iterate (9.37) $O(|z_0| n^{1/2+\varepsilon/2})$ times, we obtain (9.35) as desired. \square

Now we need to use the repulsive properties of (9.12) near the origin to propagate this initial largeness to later values of i . The key proposition is the following.

PROPOSITION 59. *Let $i_0 \leq i_1 \leq i_2 \leq n - n^\varepsilon/2$. Let E_{i_1, i_2} be the event that $|a_i| \leq \frac{1}{2} \sqrt{i - (1 - |z_0|^2)n}$ for all $i_1 \leq i \leq i_2$. Then we have with overwhelming probability that*

$$|a_{i_2}| 1_{E_{i_1, i_2}} \geq \left(1 + \frac{ci_1 - (1 - |z_0|^2)n}{|z_0|^2 n}\right)^{i_2 - i_1} (|a_{i_1}| + O(n^{o(1)} \sqrt{i - i_1})) 1_{E_{i_1, i_2}}$$

for some constant $c > 0$.

PROOF. The probability in question will be computed over the product space generated by ξ_i, η_i with $i_1 < i \leq i_2$, conditioning all the other ξ_i, η_i to be fixed. In particular, a_{i_1} is now deterministic.

For any $i_1 \leq i < i_2$, we see from (9.11) that

$$(9.38) \quad a_{i+1} = \beta_i a_i + \xi_{i+1},$$

where β_i is the positive real number

$$\beta_i := \frac{|z_0| \sqrt{n}}{\sqrt{|a_i|^2 + n - i + \sqrt{n - i} \eta_{n-i}}}.$$

Next, from iterating (9.38) we have

$$a_{i_2} = \gamma_{i_1, i_2} \left(a_{i_1} + \sum_{i_1 \leq i < i_2} \delta_{i_1, i} \xi_{i+1} \right),$$

where $\gamma_{i_1, i_2} := \beta_{i_1} \cdots \beta_{i_2-1}$ and $\delta_{i_1, i} := \beta_{i_1}^{-1} \cdots \beta_i^{-1}$.

As the event E_{i_1, i_2} contains $E_{i_1, i}$ for $i_1 \leq i < i_2$, we have

$$(9.39) \quad a_{i_2} 1_{E_{i_1, i_2}} = \gamma_{i_1, i_2} 1_{E_{i_1, i_2}} \left(a_{i_1} + \sum_{i_1 \leq i < i_2} \delta_{i_1, i} \xi_{i+1} 1_{E_{i_1, i}} \right).$$

Notice that if $E_{i_1,i}$ holds, then

$$|a_i|^2 \leq \frac{1}{4}(i - (1 - |z_0|^2)n),$$

which is equivalent to

$$|a_i|^2 + n - i \leq |z_0|^2 n - \frac{3}{4}(i - (1 - |z_0|^2)n).$$

On the other hand, since

$$i - (1 - |z_0|^2)n \geq i_1 - (1 - |z_0|^2)n \geq |z_0|n^{1/2+\varepsilon}/2$$

and $n - i \leq |z_0|^2 n$, we deduce from (9.10) that

$$|a_i|^2 + n - i + \sqrt{n-i}\eta_{n-i} \leq |z_0|^2 n - \frac{1}{2}(i - (1 - |z_0|^2)n)$$

(say) if n is large enough. This gives a bound of the form

$$\beta_i \geq 1 + c \frac{i - (1 - |z_0|^2)n}{|z_0|^2 n} \geq 1 + c \frac{i_1 - (1 - |z_0|^2)n}{|z_0|^2 n}$$

for some absolute constants $c > 0$.

From the definition of γ_i , we conclude the lower bound

$$(9.40) \quad |\gamma_{i_1,i_2}| 1_{E_{i_1,i_2}} \geq \left(1 + c \frac{i_1 - (1 - |z_0|^2)n}{|z_0|^2 n}\right)^{i_2-i_1} 1_{E_{i_1,i_2}}$$

and the upper bound

$$(9.41) \quad |\delta_{i_1,i}| 1_{E_{i_1,i}} \leq 1_{E_{i_1,i}} \leq 1.$$

Let us now make a critical observation that the random variable $\delta_{i_1,i} 1_{E_{i_1,i}}$ depends on ξ_2, \dots, ξ_i (and on the $\chi_{1,\mathbb{C}}, \dots, \chi_{n-1,\mathbb{C}}$) but is independent of ξ_{i+1}, \dots, ξ_n . This enables us to apply Proposition 35, from which we can conclude that with overwhelming probability

$$(9.42) \quad \sum_{i_1 \leq i < i_2} \delta_{i_1,i} 1_{E_{i_1,i}} \xi_{i+1} = O(n^{o(1)} |i_2 - i_1|^{1/2}) = O(n^{o(1)} \sqrt{i_2 - i_1}),$$

concluding the proof. \square

COROLLARY 60. *Assume that $|a_{i_1}| \geq n^{\varepsilon/100} T^{1/2}$ where*

$$T := \left\lfloor \frac{|z_0|^2 n}{i_1 - (1 - |z_0|^2)n} \log^2 n \right\rfloor.$$

Then $1_{E_{i_1,i_1+T}} = 0$ holds with overwhelming probability.

PROOF. Assume, for contradiction, that there is a fixed A such that $\mathbf{P}(1_{E_T}) \geq n^{-A}$. By the previous lemma, we can assume that

$$|a_{i_1+T}|1_{E_{i_1,i_1+T}} \geq \left(1 + \frac{ci_1 - (1 - |z_0|^2)n}{|z_0|^2n}\right)^T (|a_{i_1}| + O(n^{o(1)}\sqrt{T})1_{E_{i_1,i_1+T}})$$

holds with probability at least $1 - n^{-2A}$. Taking expectations, we conclude

$$\begin{aligned} \mathbf{E}|a_{i_1+T}| &\geq \mathbf{E}|a_{i_1+T}|1_{E_{i_1,i_1+T}} \\ &\geq \left(1 + \frac{ci_1 - (1 - |z_0|^2)n}{|z_0|^2n}\right)^T (\mathbf{E}|a_{i_1}| + O(n^{o(1)}\sqrt{T}))(n^{-A} - n^{-2A}). \end{aligned}$$

Since $|a_{i_1}| \geq n^{\varepsilon/100}T^{1/2}$ and $(1 + \frac{ci_1 - (1 - |z_0|^2)n}{|z_0|^2n})^T \geq \exp(c \log^2 n)$ for some fixed $c > 0$ by the definition of T , the RHS is bounded from below by

$$n^{-A} \exp(c \log^2 n) \gg n.$$

On the other hand, from Lemma 54 we have that

$$\mathbf{E}|a_{i_1+T}| \leq (1 + |z_0|)\sqrt{n} \ll \sqrt{n},$$

yielding the desired contradiction. \square

Next, we observe that a_i cannot drop in magnitude too quickly once it is somewhat small [assuming the hypotheses (9.9), (9.10), of course].

LEMMA 61. *If $|a_i| \leq \frac{1}{2}\sqrt{i - (1 - |z_0|^2)n}$ then $|a_i| \geq |a_{i-1}| - n^{o(1)}$.*

PROOF. From (9.7), we have

$$a_i - \xi_i = \frac{|z_0|\sqrt{n}}{\sqrt{|a_{i-1}|^2 + \chi_{n-i+1,\mathbb{C}}}}a_{i-1}$$

and hence

$$\frac{|z_0|^2n}{|a_{i-1}|^2 + \chi_{n-i+1,\mathbb{C}}} |a_{i-1}|^2 = |a_i - \xi_i|^2.$$

We can rearrange this as

$$|a_{i-1}|^2 = \frac{\chi_{n-i+1,\mathbb{C}}}{|z_0|^2n - |a_i - \xi_i|^2} |a_i - \xi_i|^2.$$

By (9.10), we have

$$\chi_{n-i+1,\mathbb{C}} = n - i + O(\sqrt{n - in^{o(1)}}) = n - i + O(n^{o(1)}|z_0|\sqrt{n}),$$

using the fact that in this range $n - i \leq |z_0|^2n$.

From the assumption of the lemma, we have that

$$|a_i - \xi_i|^2 \leq \frac{1}{4}(i - (1 - |z_0|^2)n) + O(n^{o(1)}\sqrt{i - (1 - |z_0|^2)n})$$

and thus

$$\begin{aligned} &\chi_{n-i+1, \mathbb{C}} - |z_0|^2n + |a_i - \xi_i|^2 \\ &\leq -\frac{3}{4}(i - (1 - |z_0|^2)n) + O(n^{o(1)}|z_0|\sqrt{n}) + O(n^{o(1)}\sqrt{i - (1 - |z_0|^2)n}). \end{aligned}$$

As $i - (1 - |z_0|^2)n \geq |z_0|n^{1/2+\varepsilon}$, we see that the right-hand side is negative for n large enough, thus

$$\frac{\chi_{n-i+1, \mathbb{C}}}{|z_0|^2n - |a_i - \xi_i|^2} \leq 1.$$

We thus have

$$|a_{i-1}| \leq |a_i - \xi_i|,$$

which implies from (9.9) that $|a_i| \geq |a_{i-1}| - n^{o(1)}$ as desired. \square

We can now prove the lower bound (9.28) with overwhelming probability as follows. We first condition on the event that the conclusion of Lemma 58 holds. Now assume that there is some $i_0 < i \leq n - n^\varepsilon$ such that

$$|a_i| \leq \frac{1}{3}\sqrt{i - (1 - |z_0|^2)n}.$$

Let i_2 be the first such index. In particular,

$$(9.43) \quad |a_{i_2}| \leq \frac{1}{3}\sqrt{i_2 - (1 - |z_0|^2)n} \leq \frac{1}{2}\sqrt{i_2 - (1 - |z_0|^2)n}.$$

By Lemma 58, we can then locate an index $\max(i_0 - \frac{1}{2}|z_0|n^{1/2+\varepsilon}, 0) + 1 \leq i_1 < i_2$ such that $|a_i| \leq \frac{1}{2}\sqrt{i - (1 - |z_0|^2)n}$ for all $i_1 \leq i \leq i_2$ (or in other words, E_{i_1, i_2} holds) and

$$|a_{i_1-1}| > \frac{1}{2}\sqrt{i_1 - 1 - (1 - |z_0|^2)n}.$$

From Lemma 61, this implies in particular that

$$(9.44) \quad |a_{i_1}| \geq 0.499\sqrt{i_1 - (1 - |z_0|^2)n}.$$

From the above discussion and the union bound, it thus suffices to show that for any given $i_0 \leq i_1 < i_2 \leq n - n^\varepsilon$, the event that (9.43) and (9.44) and E_{i_1, i_2} all simultaneously hold, is false with overwhelming probability.

Fix i_1, i_2 . If $i_2 - i_1 > T$ then by Corollary 60, $1_{E_{i_1, i_2}} = 0$ with overwhelming probability and we are done. In the other case $i_2 - i_1 \leq T$, by Proposition 59, we have with overwhelming probability

$$(9.45) \quad \begin{aligned} & |a_{i_2}| 1_{E_{i_1, i_2}} \\ & \geq \left(1 + \frac{ci_1 - (1 - |z_0|^2)n}{|z_0|^2n}\right)^{i_2 - i_1} (|a_{i_1}| + O(n^{o(1)}\sqrt{i - i_1})) 1_{E_{i_1, i_2}}. \end{aligned}$$

It now suffices to verify that if $|a_{i_1}| \geq 0.499\sqrt{i_1 - (1 - |z_0|^2)n}$, E_{i_1, i_2} holds, and $|a_{i_2}| \leq \frac{1}{3}\sqrt{i_2 - (1 - |z_0|^2)n}$, then the above inequality is violated. Notice that since $i_2 - i_1 \leq T = \frac{|z_0|^2n}{i_1 - (1 - |z_0|^2)n} \log^2 n$ and $i_1 - (1 - |z_0|^2)n \gg |z_0|n^{1/2 + \varepsilon}$, we have

$$\begin{aligned} |a_{i_1}| + O(n^{o(1)}\sqrt{i_2 - i_1}) & \geq 0.499\sqrt{i_1 - (1 - |z_0|^2)n} - O(n^{o(1)}T^{1/2}) \\ & \geq \frac{5}{12}\sqrt{i_1 - (1 - |z_0|^2)n}. \end{aligned}$$

As E_{i_1, i_2} holds, it follows that the RHS of (9.45) is at least

$$\frac{5}{12}\sqrt{i_1 - (1 - |z_0|^2)n} > \frac{1}{3}\sqrt{i_2 - (1 - |z_0|^2)n}$$

again thanks to the fact that $i_2 - i_1 \leq T = o(i_1 - (1 - |z_0|^2)n)$. Our proof is complete.

REMARK 62. All the above arguments go through without difficulty in the real case, using (9.8) instead of (9.7), replacing $a_i, \xi_i, \chi_{i, \mathbb{C}}$ by $a'_i, \xi'_i, \chi_{i, \mathbb{R}}$, respectively; we leave the details to the interested reader.

10. Concentration of log-determinant for i.i.d. matrices. Now that we have established concentration of the log-determinant in the special case of real and complex Gaussian matrices (Theorem 33), we are now ready to apply the resolvent swapping machinery from Section 7 to obtain concentration for more general i.i.d. matrices (Theorem 25).

Fix δ, z_0 . Let W_{n, z_0} be defined as in (2.3). As in the previous section, set α equal to $\frac{1}{2}(|z_0|^2 - 1)$ if $|z_0| \leq 1$, and $\log |z_0|$ if $|z_0| \geq 1$. It suffices to show that

$$\log |\det(W_{n, z_0})| = 2n\alpha + O(n^{o(1)})$$

with overwhelming probability, uniformly in z_0 . We may assume without loss of generality that all entries of M_n are $O(n^{o(1)})$.

We observe the identity

$$\log |\det(W_{n, z_0})| = \log |\det(W_{n, z_0} - \sqrt{-1}T)| - n \operatorname{Im} \int_0^T s(\sqrt{-1}\eta) d\eta$$

for any $T > 0$, where $s(z) := \frac{1}{n} \text{trace}(W_{n,z_0} - z)^{-1}$ is the Stieltjes transform, as can be seen by writing everything in terms of the eigenvalues of W_{n,z_0} . If we set $T := n^{100}$, then we see that

$$\begin{aligned} \log|\det(W_{n,z_0} - \sqrt{-1}T)| &= n \log T + \log|\det(1 - n^{-100}W_{n,z_0})| \\ &= n \log T + O(n^{-10}) \end{aligned}$$

(say), thanks to (8.1) and the hypothesis that $|z_j| \leq \sqrt{n}$. Thus, it suffices to show that

$$n \operatorname{Im} \int_0^T s(\sqrt{-1}\eta) d\eta = n \log T - 2n\alpha + O(n^{o(1)})$$

with overwhelming probability.

Now we eliminate the contribution of very small η .

LEMMA 63. *One has*

$$n \operatorname{Im} \int_0^{1/n} s(\sqrt{-1}\eta) d\eta = O(n^{o(1)})$$

with overwhelming probability.

PROOF. From Proposition 31, we see with overwhelming probability that

$$|s(\sqrt{-1}\eta)| \ll n^{o(1)} \left(1 + \frac{1}{n\eta}\right)$$

for all $\eta > 0$. This already handles the portion of the integral where $\eta > n^{-2\log n}$ (say). For the remaining portion when $0 < \eta \leq n^{-2\log n}$, we observe from Proposition 27 that with overwhelming probability, all eigenvalues of W_{n,z_0} are at least $n^{-\log n}$ in magnitude, which implies that $s(\sqrt{-1}\eta) = O(n^{1+\log n})$ for all such η , and the claim follows. \square

Set $X := n \operatorname{Im} \int_{1/n}^T s(\sqrt{-1}\eta) d\eta$ and $X_* := n \log T - 2n\alpha$. Fix arbitrary constants $A, \epsilon > 0$. In view of the above lemma, it suffices to show that

$$\mathbf{P}(|X - X_*| \geq n^\epsilon) \ll n^{-A}.$$

By Markov’s inequality, it suffices to show that for $j = 2\lfloor A/\epsilon \rfloor$

$$(10.1) \quad \mathbf{E}(X - X_*)^j = O(n^{j\epsilon/2}).$$

Without loss of generality, we may assume j to be large, for example, $j > 5$. By Theorem 33, we know that a stronger bound

$$(10.2) \quad \mathbf{E}(X' - X_*)^j \leq n^\epsilon$$

holds for the same range of j (for n sufficiently large depending on ε and j), where X' is defined as in X but with M_n replaced by a random real or complex Gaussian matrix M'_n that matches M_n to third order.

We now execute the following swapping process. Start with the random Gaussian matrix M'_n and in each step swap either the real or imaginary part of a Gaussian entry of M'_n to the associated real or imaginary part of the corresponding entry of M_n . The exact order in which we perform this swapping is not important, so long as it is chosen in advance; for instance, one could use lexicographical ordering, swapping the real part and then the imaginary part for each entry in turn. Let $M_n^{[k]}$, $0 \leq k \leq 2n^2$, be the resulting random matrix at time k and define $X^{[k]}$ accordingly. We will show, by induction on k , that

$$(10.3) \quad \mathbf{E}(X^{[k]} - X_*)^j \leq \left(1 + \frac{k}{n^{2+\varepsilon/8j}}\right)n^\varepsilon$$

for n sufficiently large depending on ε and j (but not on k). Note that the base case $k = 0$ of (10.3) holds thanks to (10.2), while the case $k = 2n^2$ implies (10.1) with some room to spare.

For technical reasons, it is convenient to assume that $|\xi|, |\xi'| = n^{o(1)}$ with probability one. This can be done replacing all entries ξ_{ij} by $\xi_{ij} \mathbf{1}_{|\xi_{ij}| \leq \log^B n}$ and ξ'_{ij} by $\xi'_{ij} \mathbf{1}_{|\xi'_{ij}| \leq \log^B n}$, where B is a sufficiently large constant so that with overwhelming probability $|\xi_{ij}| + |\xi'_{ij}| < \log^B n$ for all i, j . It is clear that any event that holds with overwhelming probability in the truncated model also holds with overwhelming probability in the original one. Thus, we can reduce to the truncated case. At this point, we would like to point out that the truncation does change the moments of the entries, but by a very small amount that will only introduce negligible factors such as $O(n^{-100})$ to the swapping argument. Abusing the notion slightly, from now on we still work with ξ and ξ' but under the extra assumption that with probability one $|\xi|, |\xi'| \leq \log^B n = n^{o(1)}$.

Fix a step $0 \leq k < 2n^2$, and consider the difference

$$(10.4) \quad \begin{aligned} D_k &:= \mathbf{E}(X^{[k+1]} - X_*)^j - \mathbf{E}(X^{[k]} - X_*)^j \\ &= \int \mathbf{E}([(X^{[k+1]} - X_*)^j - (X^{[k]} - X_*)^j] | M_0) dM_0, \end{aligned}$$

where M_0 is obtained from $X^{[k+1]}$ by putting 0 at the swapping position (in other words, M_0 is the common part of $M^{[k]}$ and $M^{[k+1]}$), and dM_0 is the law of M_0 . Once conditioned on M_0 , we can simplify the notation by replacing $X^{[k]}$ and $X^{[k+1]}$ by X_ξ and $X_{\xi'}$, respectively.

It is important to notice that since $\eta \geq 1/n$, we can bound $|s_\xi(\sqrt{-1}\eta)|$ crudely by n with probability one (for any matrix $M_n^{[k]}$). As $T = n^{100}$, this implies that $|X^{[k]}| \ll n^{102}$ and

$$(10.5) \quad |(X^{[k]} - X_*)^j - (X^{[k+1]} - X_*)^j| \ll n^{102j}$$

for any j , with probability one.

By Proposition 31, we see with overwhelming probability that

$$\|R_\xi(\sqrt{-1}\eta)\|_{(\infty,1)} \ll n^{o(1)}$$

for all $\eta \geq n^{-1}$. In this case, by Lemma 44 and (8.4)

$$(10.6) \quad \|R_0(\sqrt{-1}\eta)\|_{(\infty,1)} \ll n^{o(1)}$$

for all such η .

If (10.6) holds, we say that M_0 is *good*. The contribution from bad M_0 in the RHS of (10.4) is very small. Indeed, by Proposition 31, we can assume that M_0 is bad with probability at most $n^{-102j-100}$. By the upper bound (10.5), the integral (in D_k) over the bad M_0 is at most

$$(10.7) \quad n^{-102j-100} n^{102j} = n^{-100}.$$

Let us now condition on a good M_0 . By Proposition 45, we have

$$(10.8) \quad s_\xi(\sqrt{-1}\eta) = s_0 + \sum_{i=1}^3 \xi^i n^{-i/2} c_i(\eta) + O\left(n^{-2+o(1)} \frac{1}{n\eta}\right),$$

where the coefficient $c_i(\eta)$ is independent of ξ and enjoys the bound $|c_i(\eta)| \ll n^{o(1)} \frac{1}{n\eta}$.

Multiplying by n and taking the integral over η , we obtain,

$$(10.9) \quad X_\xi = X_0 + P(\xi) + O(n^{-2+o(1)}),$$

where $P = \sum_{i=1}^3 \xi^i n^{-i/2} d_i$ is a polynomial in ξ with coefficients $d_i = O(n^{o(1)})$, and X_0 is a quantity independent of ξ . As $|\xi| = n^{o(1)}$ with probability one, it follows that $|X_\xi - X_0| = n^{-1/2+o(1)}$ with probability one. Furthermore,

$$(10.10) \quad X_\xi - X_* = (X_0 - X_*) + P(\xi) + O(n^{-2+o(1)}).$$

We raise this equation to the power j , focusing on those terms of order ξ^4 or more. As $d_i = O(n^{o(1)})$, using the fact that $|\xi| \leq n^{o(1)}$ with probability one and $j > 5$, we have

$$(10.11) \quad (X_\xi - X_*)^j = P_j(\xi) + O\left(n^{-2+o(1)} \sum_{l=1}^{j-1} |X_0 - X_*|^l + n^{-5/2+o(1)}\right),$$

where P_j is a polynomial of degree at most 3. Therefore,

$$(10.12) \quad \begin{aligned} & \mathbf{E}(X_\xi - X_*)^j \\ &= \mathbf{E}P_j(\xi) + O\left(n^{-2+o(1)} \sum_{k=1}^{j-1} |X_0 - X_*|^k + n^{-5/2+o(1)}\right). \end{aligned}$$

Similarly,

$$(10.13) \quad \begin{aligned} & \mathbf{E}(X_{\xi'} - X_*)^j \\ &= \mathbf{E}P_j(\xi') + O\left(n^{-2+o(1)} \sum_{k=1}^{j-1} |X_0 - X_*|^k + n^{-5/2+o(1)}\right). \end{aligned}$$

Here, the expectations are with respect to ξ and ξ' (as we already conditioned on a good M_0 .) It follows that

$$(10.14) \quad \begin{aligned} & \mathbf{E}(X_\xi - X_*)^j - \mathbf{E}(X_{\xi'} - X_*)^j \\ &= \mathbf{E}(P_j(\xi) - P_j(\xi')) \\ &+ O\left(n^{-2+o(1)} \sum_{k=1}^{j-1} |X_0 - X_*|^k + n^{-5/2+o(1)}\right). \end{aligned}$$

As already pointed out, the first three moments of ξ and ξ' do not entirely match due to the truncation. However, by fixing B large enough, we can assume that the truncation changes each moment by at most n^{-C} for some sufficiently large C [we need C to be larger than the absolute value of the coefficients of P_j , which are of size $O(n^{O(1)})$, again thanks to the fact that $|s_\xi(\sqrt{-1}\eta)| \leq n$ with probability one]. This yields

$$(10.15) \quad \begin{aligned} & \mathbf{E}(X_\xi - X_*)^j - \mathbf{E}(X_{\xi'} - X_*)^j \\ &= O\left(n^{-2+o(1)} \sum_{k=1}^{j-1} |X_0 - X_*|^k + n^{-5/2+o(1)}\right). \end{aligned}$$

But $|X_\xi - X_0| \leq n^{-1/2+o(1)}$ with probability one, so (10.4) implies

$$(10.16) \quad \begin{aligned} & \mathbf{E}(X_\xi - X_*)^j - \mathbf{E}(X_{\xi'} - X_*)^j \\ &= O\left(n^{-2+o(1)} \sum_{k=1}^{j-1} \mathbf{E}|X_\xi - X_*|^k + n^{-5/2+o(1)}\right). \end{aligned}$$

The right-hand side of (10.16) can be bounded as

$$(10.17) \quad O(n^{-2+o(1)} \min\{\mathbf{E}|X_\xi - X_*|^j n^{-\varepsilon/4j}, n^{\varepsilon/2}\}),$$

where the bound comes from considering two cases $\mathbf{E}|X_\xi - X_*|^j$ being not smaller or smaller than $n^{\varepsilon/2}$, and the Holder inequality.

Thus, conditioned on a good M_0 , we have

$$|\mathbf{E}(X_\xi - X_*)^j - \mathbf{E}(X_{\xi'} - X_*)^j| \ll n^{-2+o(1)} \min\{|X_\xi - X_*|^j n^{-\varepsilon/4j}, n^{\varepsilon/2}\}.$$

Taking into account (10.7), we conclude

$$D_k \ll n^{-100} + n^{-2-\varepsilon/4j} \mathbf{E}|X_\xi - X_*|^j + n^{-2+\varepsilon/2+o(1)}$$

and the desired bound (10.3) on $\mathbf{E}(X^{[k+1]} - X_*)^j$ follows easily by the induction hypothesis.

APPENDIX A: SPECTRAL PROPERTIES OF $W_{n,z}$

In this appendix, we prove Propositions 29 and 31. We fix M_n, C, z_0 as in these propositions. By truncation, we may assume that all the coefficients of M_n have magnitude $O(n^{o(1)})$.

A.1. Crude upper bound. We begin with Proposition 29, which we will prove by modifying the argument from [55], Appendix C, and [57], Proposition 28. Write $I = [E - \eta, E + \eta]$. It suffices to establish the claim in the case $1/n \leq \eta \leq 1$, as the general case then follows from this case (and from the trivial bound $N_I \leq 2n$). By rounding η to the nearest integer power of two, and using the union bound, it suffices to establish the claim for a single η in this range, which we now fix. Similarly, we may round E to a multiple of η ; since the claim is easy for (say) $|E| \geq n^{10}$; we see from the union bound that it suffices to establish the claim for a single E , which we now also fix. By symmetry, we may take $E \geq 0$.

By a diagonalization argument, it will suffice to show for each fixed $c > 0$ that one has

$$N_{[E-\eta, E+\eta]} \leq n^{1+c} \eta$$

with overwhelming probability. Accordingly, we assume for contradiction that

$$(A.1) \quad N_{[E-\eta, E+\eta]} > n^{1+c} \eta.$$

We use the Stieltjes transform

$$s(E + \sqrt{-1}\eta) = \frac{1}{2n} \text{trace}(W_{n,z} - E - \sqrt{-1}\eta)^{-1}.$$

Then

$$\text{Im } s(E + \sqrt{-1}\eta) = \frac{1}{2n} \sum_{j=1}^{2n} \frac{\eta}{(\lambda_j(W_{n,z}) - E)^2 + \eta^2}$$

from (A.1) we thus have

$$\text{Im } s(E + \sqrt{-1}\eta) \gg n^c.$$

In particular, since

$$s(E + \sqrt{-1}\eta) = \frac{1}{2n} \sum_{j=1}^{2n} R(E + \sqrt{-1}\eta)_{jj}$$

we see from the pigeonhole principle that we have

$$(A.2) \quad |R(E + \sqrt{-1}\eta)_{jj}| \gg n^c$$

for some $1 \leq j \leq 2n$. By the union bound, it suffices to show that for each j , the hypothesis (A.2) [combined with (A.1)] leads to a contradiction with overwhelming probability.

Fix j ; by symmetry, we may take $j = 2n$, thus

$$(A.3) \quad |R(E + \sqrt{-1}\eta)_{2n,2n}| \gg n^c.$$

We expand $W_{n,z}$ as

$$W_{n,z} = \begin{pmatrix} W'_{n,z} & X \\ X^* & 0 \end{pmatrix},$$

where $W'_{n,z}$ is the $2n - 1 \times 2n - 1$ Hermitian matrix

$$W'_{n,z} := \begin{pmatrix} 0 & 0 & \frac{1}{\sqrt{n}}(M_{n-1} - z) \\ 0 & 0 & Z \\ \frac{1}{\sqrt{n}}(M_{n-1} - z)^* & Z^* & 0 \end{pmatrix},$$

where M_{n-1} is the top left $n - 1 \times n - 1$ minor of M_n , Z is the $n - 1$ -dimensional row vector with entries $\frac{1}{\sqrt{n}}\xi_{nj}$ for $j = 1, \dots, n - 1$, X is the $2n$ -dimensional column vector

$$X := \begin{pmatrix} X' \\ \frac{1}{\sqrt{n}}(\xi_{nn} - z) \\ 0 \end{pmatrix}$$

and X' is the $n - 1$ -dimensional column vector with entries $\frac{1}{\sqrt{n}}\xi_{jn}$ for $j = 1, \dots, n - 1$.

By Schur's complement, the resolvent coefficient $R(E + \sqrt{-1}\eta)_{2n,2n}$ can be expressed as

$$(A.4) \quad R(E + \sqrt{-1}\eta)_{2n,2n} = \frac{1}{-E - \sqrt{-1}\eta - Y_n},$$

where Y_n is the expression

$$Y_n := X^*(W'_{n,z} - E - \sqrt{-1}\eta)^{-1}X.$$

By (A.3), we conclude that

$$|E + \sqrt{-1}\eta + Y_n| \ll n^{-c}$$

as Y_n has a nonnegative imaginary part, we conclude that

$$(A.5) \quad \text{Im } Y_n \ll n^{-c}.$$

Next, we apply the singular value decomposition to the $n \times n - 1$ matrix $\begin{pmatrix} \frac{1}{\sqrt{n}}(M_{n-1} - z) \\ Z \end{pmatrix}$, generating an orthonormal basis of n right singular vectors u_1, \dots, u_n in \mathbb{C}^n , and an orthonormal basis of $n - 1$ left singular vectors in \mathbb{C}^{n-1} ,

associated to singular values $\sigma_1, \dots, \sigma_n$ (with $\sigma_n = 0$). Then $W'_{n,z}$ is conjugate to the direct sum

$$W'_{n,z} \equiv \bigoplus_{j=1}^{n-1} \begin{pmatrix} 0 & \sigma_j \\ \sigma_j & 0 \end{pmatrix} \oplus (0)$$

and thus

$$\begin{aligned} & (W'_{n,z} - E - \sqrt{-1}\eta)^{-1} \\ & \equiv \bigoplus_{j=1}^{n-1} \frac{1}{\sigma_j^2 - (E + \sqrt{-1}\eta)^2} \begin{pmatrix} E + \sqrt{-1}\eta & \sigma_j \\ \sigma_j & E + \sqrt{-1}\eta \end{pmatrix} \oplus \left(\frac{1}{E + \sqrt{-1}\eta} \right) \end{aligned}$$

and thus

$$\begin{aligned} \text{Im } Y_n &= \sum_{j=1}^{n-1} \text{Im} \frac{E + \sqrt{-1}\eta}{\sigma_j^2 - (E + \sqrt{-1}\eta)^2} |\tilde{X}^* u_j|^2 \\ &= \frac{1}{2} \sum_{j=1}^{n-1} \sum_{\epsilon=\pm 1} \frac{1}{\epsilon \sigma_j - (E + \sqrt{-1}\eta)} |\tilde{X}^* u_j|^2 \\ &= \frac{\eta}{2} \sum_{j=1}^{n-1} \sum_{\epsilon=\pm 1} \frac{1}{|E - \epsilon \sigma_j|^2 + \eta^2} |\tilde{X}^* u_j|^2, \end{aligned}$$

where

$$\tilde{X} := \begin{pmatrix} X' \\ \frac{1}{\sqrt{n}}(\xi_{nm} - z) \end{pmatrix}$$

is the top half of X .

By (A.1) and the Cauchy interlacing law, we may find an interval $[j_-, j_+]$ of length $j_+ - j_- \gg n^{1+c}\eta$ such that $|\sigma_j - E| \leq \eta$ for all $j_- \leq j \leq j_+$. We conclude that

$$\sum_{j_- \leq j \leq j_+} |\tilde{X}^* u_j|^2 \ll n^{-c}\eta.$$

At this point, we will follow [19] and invoke a concentration estimate for quadratic forms essentially due to Hanson and Wright [29, 62].

PROPOSITION 64 (Concentration). *Let ξ_1, \dots, ξ_n be i.i.d. complex random variables with mean zero, variance one, and bounded in magnitude by K for some $K \geq 1$. Let $X \in \mathbb{C}^n$ be a random vector of the form $Y + Z$, where*

$$Y := \frac{1}{n^{1/2}} \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix}$$

and Z is a random vector independent of Y . Let $A = (a_{ij})_{1 \leq i, j \leq n}$ be a random complex matrix that is also independent of Y . Then with overwhelming probability one has

$$X^*AX = \frac{1}{n} \text{trace } A + Z^*AZ + O\left(K^2 \log^2 n \left(\frac{1}{n} \|A\|_F + \frac{1}{\sqrt{n}} \|AZ\| + \frac{1}{\sqrt{n}} \|A^*Z\|\right)\right),$$

where $\|A\|_F := (\sum_{1 \leq i, j \leq n} |a_{ij}|^2)^{1/2}$ is the Frobenius norm of A .

We remark that for our applications, one could also use Talagrand’s concentration inequality [48] as a substitute for this concentration inequality, at the cost of a slight degradation in the bounds; see, for example, [55].

PROOF OF PROPOSITION 64. By conditioning, we may assume that Z, A are deterministic (the failure probability in our estimates will be uniform in the choice of Z, A). Let $\tilde{\xi}_i := \xi_i/K$. From [19], Proposition 4.5, we have

$$\sum_{1 \leq i, j \leq n} a_{ij} \tilde{\xi}_i \overline{\tilde{\xi}_j} = \sum_{1 \leq i, j \leq n} a_{ij} \mathbf{E} \tilde{\xi}_i \overline{\tilde{\xi}_j} + O(\|A\|_F \log^2 n)$$

with overwhelming probability. Multiplying by K^2/n and noting that $\mathbf{E} \xi_i \overline{\xi_j} = 1_{i=j}$, we conclude that

$$Y^*AY = \frac{1}{n} \text{trace } A + O\left(\frac{K^2 \log^2 n}{n} \|A\|_F\right)$$

with overwhelming probability. Meanwhile, from the Chernoff inequality we see that

$$Y^*AZ = O\left(\frac{K \log^2 n}{\sqrt{n}} \|AZ\|\right)$$

and similarly

$$Z^*AY = O\left(\frac{K \log^2 n}{\sqrt{n}} \|A^*Z\|\right)$$

with overwhelming probability. The claim follows. \square

Applying Proposition 64 (with A equal to the projection matrix $A := \sum_{j_- \leq j \leq j_+} u_j u_j^*$), one has

$$\begin{aligned} \sum_{j_- \leq j \leq j_+} |\tilde{X}^* u_j|^2 &= \frac{j_+ - j_- + 1}{n} + \left\| \frac{z}{\sqrt{n}} \pi(e_n) \right\|^2 \\ &+ O(n^{-1+o(1)} (j_+ - j_- + 1)^{1/2}) \\ &+ O\left(n^{-1/2+o(1)} \left\| \frac{z}{\sqrt{n}} \pi(e_n) \right\|\right) \end{aligned}$$

with overwhelming probability. By the arithmetic mean–geometric mean inequality, one has $\|\frac{z}{\sqrt{n}}\pi(e_n)\|^2 + O(n^{-1/2+o(1)})\|\frac{z}{\sqrt{n}}\pi(e_n)\| \geq -n^{-1+o(1)}$, and we conclude that

$$\sum_{j_-\leq j\leq j_+} |\tilde{X}^*u_j|^2 \gg n^c\eta$$

with overwhelming probability (conditioning on M_{n-1}, Z). Undoing the conditioning, we thus obtain a contradiction with overwhelming probability, and Proposition 29 follows.

A.2. Resolvent bounds. We now prove Proposition 31, by using a more complicated variant of the arguments above. We first take advantage of the fact that the spectral parameter $\sqrt{-1}\eta$ is on the imaginary axis to make some minor simplifications. Namely, we have

$$\begin{aligned} R(\sqrt{-1}\eta) &= (W_{n,z} - \sqrt{-1}\eta)^{-1} \\ &= W_{n,z}(W_{n,z}^2 + \eta^2)^{-1} + \sqrt{-1}\eta(W_{n,z}^2 + \eta^2)^{-1}. \end{aligned}$$

Note from (2.3) that $W_{n,z}^2 + \eta^2$ is block-diagonal, and thus $W_{n,z}(W_{n,z}^2 + \eta^2)^{-1}$ vanishes on the diagonal. We conclude that $R(\sqrt{-1}\eta)_{jj}$ and $s(\sqrt{-1}\eta)$ are purely imaginary (with nonnegative imaginary part) for $1 \leq j \leq n$, with

$$\begin{aligned} \text{Im } s(\sqrt{-1}\eta) &= \frac{\eta}{2n} \text{trace}(W_{n,z}^2 + \eta^2)^{-1} \\ \text{(A.6)} \quad &= \frac{\eta}{n} \text{trace}((M_n - z)^*(M_n - z) + \eta^2)^{-1}. \end{aligned}$$

Now we observe that it suffices to verify the claim for $\eta \geq n^{-1+c}$ for each fixed c . To see this, observe that

$$\text{Im } R(\sqrt{-1}\eta)_{jj} = \eta \sum_{k=1}^{2n} \frac{|u_{k,j}|^2}{\lambda_k(W_{n,z})^2 + \eta^2}$$

for any $1 \leq j \leq 2n$, where u_1, \dots, u_{2n} are an orthonormal basis of eigenvectors for $W_{n,z}$, and $u_{k,j}$ is the j th coefficient of u_k . Thus, if we can obtain Proposition 31 for $\eta \geq n^{-1+c}$, we conclude with overwhelming probability that

$$\text{(A.7)} \quad \eta \sum_{k=1}^{2n} \frac{|u_{k,j}|^2}{\lambda_k(W_{n,z})^2 + \eta^2} \ll n^{o(1)}$$

for all $\eta \geq n^{-1+c}$, and hence that

$$\sum_{1 \leq k \leq 2n: \lambda_k(W_{n,z}) \leq \eta} |u_{k,j}|^2 \ll n^{o(1)}\eta$$

for all $\eta \geq n^{-1+c}$. This implies that

$$\sum_{1 \leq k \leq 2n : \lambda_k(W_{n,z}) \leq \eta} |u_{k,j}|^2 \ll n^{o(1)}(\eta + n^{-1+c})$$

for all $\eta > 0$. By dyadic summation [using the crude upper bound $\lambda_k(W_{n,z}) = O(n^{O(1)})$], this implies that

$$\sum_{k=1}^{2n} \frac{|u_{k,j}|^2}{(\lambda_k(W_{n,z})^2 + \eta^2)^{1/2}} \ll n^{c+o(1)} \left(1 + \frac{1}{n\eta}\right)$$

for all $\eta > 0$. Similarly, with $u_{k,j}$ replaced by $u_{k,i}$. By Cauchy–Schwarz, we conclude that

$$\left| \sum_{k=1}^{2n} \frac{u_{k,j} \overline{u_{k,i}}}{\lambda_k(W_{n,z}) - \sqrt{-1}\eta} \right| \ll n^{c+o(1)} \left(1 + \frac{1}{n\eta}\right)$$

for any $\eta > 0$. The left-hand side is $R(\sqrt{-1}\eta)_{ij}$. The claim then follows by using a diagonalization argument.

A similar argument reveals that we may assume without loss of generality that η is an integer power of two. Note that the above argument shows that one only needs to verify the diagonal case $i = j$; by symmetry and the union bound we may take $i = j = 2n$. The claim is trivially verified for $\eta \geq n^{10}$ (say), so we may assume that η lies between n^{-1+c} and n^{10} ; by the union bound, we may now consider η as fixed. By diagonalization (and the imaginary nature of the resolvent), it will now suffice to show that

$$(A.8) \quad \text{Im } R(\sqrt{-1}\eta)_{2n,2n} \ll n^{c+o(1)}$$

with overwhelming probability.

From (A.4) [and the fact that $R(\sqrt{-1}\eta)_{2n,2n}$ is imaginary], we have

$$(A.9) \quad \text{Im } R(\sqrt{-1}\eta)_{2n,2n} = \frac{1}{\eta + \text{Im } Y_n},$$

where

$$Y_n := X^*(W'_{n,z} - \sqrt{-1}\eta)^{-1}X.$$

From the block-diagonal nature of $W'_{n,z}$ as before, we see that Y_n is purely imaginary, with nonnegative imaginary part; indeed, we have

$$(A.10) \quad \text{Im } Y_n = \eta \tilde{X}^*(AA^* + \eta^2)^{-1} \tilde{X},$$

where A is the $n \times n - 1$ matrix

$$A := \begin{pmatrix} M_{n-1} - z \\ Y \end{pmatrix}.$$

Thus, we have the crude bound

$$(A.11) \quad \text{Im } R(\sqrt{-1}\eta)_{2n,2n} \leq \frac{1}{\eta},$$

which already takes care of the case when η is large (e.g., $\eta \geq n^{-c}$).

On the other hand, we see from Proposition 64 that with overwhelming probability one has

$$\begin{aligned} \tilde{X}^*(AA^* + \eta^2)^{-1}\tilde{X} &= \frac{1}{n} \text{trace}(AA^* + \eta^2)^{-1} + \frac{|z|^2}{n} e_n^*(AA^* + \eta^2)^{-1} e_n \\ &\quad + O(n^{-1+o(1)} \|(AA^* + \eta^2)^{-1}\|_F) \\ &\quad + O(n^{-1+o(1)} |z| \|(AA^* + \eta^2)^{-1} e_n\|). \end{aligned}$$

From the spectral theorem, one has

$$\|(AA^* + \eta^2)^{-1} e_n\| \leq (e_n^*(AA^* + \eta^2)^{-1} e_n)^{1/2} \eta^{-1}$$

and thus by Young’s inequality (or the arithmetic mean–geometric mean inequality)

$$\begin{aligned} n^{-1+o(1)} |z| \|(AA^* + \eta^2)^{-1} e_n\| &= o\left(\frac{|z|^2}{n} e_n^*(AA^* + \eta^2)^{-1} e_n\right) \\ &\quad + O(n^{-1+o(1)} \eta^{-2}). \end{aligned}$$

Also, we may expand

$$\|(AA^* + \eta^2)^{-1}\|_F = \left(\sum_{j=1}^n \frac{1}{(\sigma_j(A)^2 + \eta^2)^2} \right)^{1/2},$$

where $\sigma_1(A), \dots, \sigma_n(A)$ are the n singular values of A (thus one of these singular values is automatically zero). From Proposition 29 and the Cauchy interlacing law, we see with overwhelming probability that for any interval $[-r, r]$, the number of singular values of A in this interval is $O(n^{o(1)}(1 + nr))$. From dyadic summation, we then see that

$$(A.12) \quad \|(AA^* + \eta^2)^{-1}\|_F \ll n^{o(1)} (n\eta)^{1/2} / \eta^2.$$

Similarly, one has

$$\text{trace}(AA^* + \eta^2)^{-1} = \sum_{j=1}^n \frac{1}{\sigma_j(A)^2 + \eta^2}$$

and thus by interlacing

$$\text{trace}(AA^* + \eta^2)^{-1} = \sum_{j=1}^n \frac{1}{\sigma_j(M_n - z)^2 + \eta^2} + O\left(\frac{1}{\eta^2}\right).$$

But from (A.6) we have

$$\sum_{j=1}^n \frac{1}{\sigma_j(M_n - z)^2 + \eta^2} = \frac{n}{\eta} s(\sqrt{-1}\eta)$$

and thus

$$(A.13) \quad \frac{\eta}{n} \text{trace}(AA^* + \eta^2)^{-1} = s(\sqrt{-1}\eta) + O\left(\frac{1}{n\eta}\right).$$

Putting all this together with (A.10), we see that with overwhelming probability one has

$$\begin{aligned} \text{Im } Y_n &= \text{Im } s(\sqrt{-1}\eta) + (1 + o(1)) \frac{|z|^2}{n} \eta e_n^*(AA^* + \eta^2)^{-1} e_n \\ &\quad + O\left(\frac{n^{o(1)}}{n\eta}\right) + O\left(\frac{n^{o(1)}}{\sqrt{n\eta}}\right), \end{aligned}$$

which, in view of the lower bound $\eta \geq n^{-1+c}$, simplifies to

$$(A.14) \quad \text{Im } Y_n = \text{Im } s(\sqrt{-1}\eta) + (1 + o(1)) \frac{|z|^2}{n} \eta e_n^*(AA^* + \eta^2)^{-1} e_n + o(1).$$

Now we evaluate the expression $e_n^*(AA^* + \eta^2)^{-1} e_n$. Observe that

$$AA^* + \eta^2 = \begin{pmatrix} (M_{n-1} - z)(M_{n-1} - z)^* + \eta^2 & (M_{n-1} - z)Y^* \\ Y(M_{n-1} - z)^* & YY^* + \eta^2 \end{pmatrix}.$$

By Schur’s complement, we thus have

$$\begin{aligned} e_n^*(AA^* + \eta^2)^{-1} e_n &= 1/(YY^* + \eta^2 \\ &\quad - Y(M_{n-1} - z)^*((M_{n-1} - z)(M_{n-1} - z)^* + \eta^2)^{-1}(M_{n-1} - z)Y^*). \end{aligned}$$

One can simplify this using the identity

$$B^*(BB^* + \eta^2)^{-1} B = 1 - \eta^2(B^*B + \eta^2)^{-1},$$

valid for any matrix B [which can be seen either from the singular value decomposition, or by multiplying both sides of the identity by $(B^*B + \eta^2)$] to conclude that

$$\eta e_n^*(AA^* + \eta^2)^{-1} e_n = \frac{1}{\eta + \eta Y((M_{n-1} - z)^*(M_{n-1} - z) + \eta^2)^{-1} Y^*}.$$

Applying Lemma 64, we see with overwhelming probability that

$$\begin{aligned} &\eta Y((M_{n-1} - z)^*(M_{n-1} - z) + \eta^2)^{-1} Y^* \\ &= \frac{\eta}{n} \text{trace}((M_{n-1} - z)^*(M_{n-1} - z) + \eta^2)^{-1} \\ &\quad + O(n^{-1+o(1)} \eta \| (M_{n-1} - z)^*(M_{n-1} - z) + \eta^2 \|_F). \end{aligned}$$

By mimicking the proof of (A.12), one has

$$\|(M_{n-1} - z)^*(M_{n-1} - z) + \eta^2\|_F \ll n^{o(1)}(n\eta)^{1/2}/\eta^2$$

with overwhelming probability. Similarly, by mimicking the proof of (A.13), one has

$$\begin{aligned} & \frac{\eta}{n} \text{trace}((M_{n-1} - z)^*(M_{n-1} - z) + \eta^2)^{-1} \\ &= \text{Im } s(\sqrt{-1}\eta) + O\left(\frac{1}{n\eta}\right). \end{aligned}$$

Putting these bounds together, we conclude that

$$\eta e_n^*(AA^* + \eta^2)^{-1} e_n = \frac{1}{\eta + \text{Im } s(\sqrt{-1}\eta) + o(1)}$$

with overwhelming probability; inserting this back into (A.14) and (A.9) we conclude that

$$\begin{aligned} & \text{Im } R(\sqrt{-1}\eta)_{2n,2n} \\ \text{(A.15)} \quad &= 1 / \left(\eta + \text{Im } s(\sqrt{-1}\eta) \right. \\ & \quad \left. + (1 + o(1)) \frac{|z|^2/n}{\eta + \text{Im } s(\sqrt{-1}\eta) + o(1)} + o(1) \right) \end{aligned}$$

with overwhelming probability.

Suppose now that $|z|^2/n \geq 1/2$. Then we have

$$\left| y + \frac{|z|^2/n}{y} \right| \gg 1$$

for any y ; this implies that the denominator in (A.15) has magnitude $\gg 1$, which gives (A.8). Thus, we may assume that $|z|^2/n < 1/2$.

The bound (A.15) similarly with the index $2n$ replaced by any other index. Averaging over these indices, we obtain the *self-consistent equation*

$$\begin{aligned} & \text{Im } s(\sqrt{-1}\eta) \\ \text{(A.16)} \quad &= \frac{1}{2n} \sum_{i=1}^{2n} 1 / \left(\eta + \text{Im } s(\sqrt{-1}\eta) \right. \\ & \quad \left. + (1 + o(1)) \frac{|z|^2/n}{\eta + \text{Im } s(\sqrt{-1}\eta) + o(1)} + o(1) \right) \end{aligned}$$

with overwhelming probability. If we write $x := \eta + \text{Im } s(\sqrt{-1}\eta)$, we thus have

$$x = \frac{1}{2n} \sum_{i=1}^{2n} \frac{1}{x + (1 + o(1))(|z|^2/n)/(x + o(1)) + o(1)} + \eta$$

with overwhelming probability. Note that either $x = o(1)$ or $x + o(1) = (1 + o(1))x$. In the latter case, we can simplify the above equation as

$$x = \frac{1}{2n} \sum_{i=1}^{2n} \frac{1 + o(1)}{x + (|z|^2/n)/x} + \eta$$

and thus

$$x = \frac{(1 + o(1))x}{x^2 + |z|^2/n} + \eta.$$

In particular, this forces $x^2 + |z|^2/n \geq 1 + o(1)$. Since we have assumed that $|z|^2/n \leq 1/2$, we conclude that $x \geq 1/2$ (say). We conclude that for each $n^{-1+c} \leq \eta \leq n^{10}$, we have

$$\text{Im } s(\sqrt{-1}\eta) + \eta = o(1)$$

or

$$\text{Im } s(\sqrt{-1}\eta) + \eta \geq 1/2$$

with overwhelming probability. Rounding η to the nearest multiple of (say) n^{-100} and using the union bound (and crude perturbation theory estimates), we conclude with overwhelming probability that this dichotomy in fact holds for *all* $n^{-1+c} \leq \eta \leq n^{10}$. On the other hand, for $\eta = n^{10}$, one is clearly in the second case of the dichotomy rather than the first. By continuity, we conclude that the second case of this dichotomy in fact holds for all $n^{-1+c} \leq \eta \leq n^{10}$; in particular, we have with overwhelming probability that

$$\text{Im } s(\sqrt{-1}\eta) \gg 1,$$

when $n^{-1+c} \leq \eta \leq n^{-c}$. Inserting this bound into (A.15), we conclude with overwhelming probability that

$$\text{Im } R(\sqrt{-1}\eta)_{2n,2n} \ll 1,$$

when $n^{-1+c} \leq \eta \leq n^{-c}$, which gives Proposition 31 in this case. Finally, the case $\eta > n^{-c}$ can be handled by (A.11).

REMARK 65. A refinement of the above analysis can be used to give more precise control on the Stieltjes transform of $W_{n,z}$, as well as the counting function N_I . See [3] for more details.

APPENDIX B: ASYMPTOTICS FOR THE REAL GAUSSIAN ENSEMBLE

The purpose of this appendix is to establish Lemma 11. Our arguments here will rely heavily on those in [7].

By reflection, we may restrict attention to the case when z_1, \dots, z_l lie in the upper half-plane \mathbb{C}_+ . Our starting point is the explicit formula

$$\rho_n^{(k,l)}(x_1, \dots, x_k, z_1, \dots, z_l) = \text{Pf} \begin{pmatrix} \tilde{K}_n(x_i, x_{i'}) & \tilde{K}_n(x_i, z_{j'}) \\ \tilde{K}_n(z_j, x_{i'}) & \tilde{K}_n(z_j, z_{j'}) \end{pmatrix}_{1 \leq i, i' \leq k; 1 \leq j, j' \leq l}$$

for the correlation functions, where $\tilde{K}_n : (\mathbb{R} \cup \mathbb{C}_+) \times (\mathbb{R} \cup \mathbb{C}_+) \rightarrow M_2(\mathbb{C})$ is a certain explicit 2×2 matrix kernel obeying the antisymmetry law

$$(B.1) \quad \tilde{K}(\zeta, \zeta') = -\tilde{K}(\zeta', \zeta)^T,$$

making the expression inside the Pfaffian Pf an antisymmetric $2(k + l) \times 2(k + l)$ matrix; see [7], Theorem 8. In view of this formula, we see that Lemma 11 will follow if we can establish the uniform bound

$$\tilde{K}_n(\zeta, \zeta') = O(1)$$

for all $\zeta, \zeta' \in \mathbb{R} \cup \mathbb{C}_+$.

To do this, we will need the explicit description of the kernel \tilde{K}_n . Following [7], we will need the partial cosine and exponential functions

$$c_{n/2}(\gamma) := \sum_{m=0}^{n/2-1} \frac{\gamma^{2m}}{(2m)!},$$

$$e_{n/2}(\gamma) := \sum_{m=0}^{n-2} \frac{\gamma^m}{m!}$$

as well as the function

$$r_{n/2}(z, x) := \frac{e^{-z^2/2}}{\sqrt{2\pi}} \sqrt{\text{erfc}(\sqrt{2} \text{Im } z)} \frac{2^{(n-3)/2}}{(n-2)!} \text{sgn}(x) z^{n-1} \gamma \left(\frac{n-1}{2}, \frac{x^2}{2} \right),$$

where $\text{erfc} := 1 - \text{erf}$ is the complementary error function and

$$\gamma(t, x) = \int_0^x y^{t-1} e^{-y} dy$$

is the incomplete gamma function. In [7], Theorem 8, the formula

$$\tilde{K}_n(\gamma, \gamma') := \begin{pmatrix} \tilde{D}S_n(\gamma, \gamma') & \tilde{S}(\gamma, \gamma') \\ -\tilde{S}(\gamma', \gamma) & \tilde{I}SM_n(\gamma, \gamma') + \mathcal{E}(\gamma, \gamma') \end{pmatrix}$$

is given for the kernel \tilde{K}_n , where $\mathcal{E}(\gamma, \gamma')$ is equal to $\frac{1}{2} \text{sgn}(\gamma - \gamma')$ when γ, γ' are real, and equal to 0 otherwise, and the scalar quantities $\tilde{D}S_n(\gamma, \gamma')$, $\tilde{S}(\gamma, \gamma')$, $\tilde{I}SM_n(\gamma, \gamma')$, are defined by the following formulae, depending on whether γ, γ' are real or complex:

(1) (Real–real case). If $x, x' \in \mathbb{R}$, then

$$\begin{aligned} \tilde{S}_n(x, x') &:= \frac{e^{-(x-x')^2/2}}{\sqrt{2\pi}} e^{-xx'} e_{n/2}(xx') + r_{n/2}(x, x'), \\ \widetilde{D}S_n(x, x') &:= \frac{e^{-(x-x')^2/2}}{\sqrt{2\pi}} (x' - x) e^{-xx'} e_{n/2}(xx'), \\ \widetilde{I}S_n(x, x') &:= \frac{e^{-x^2/2}}{2\sqrt{\pi}} \operatorname{sgn}(x') \int_0^{(x')^2/2} \frac{e^{-t}}{\sqrt{t}} c_{n/2}(x\sqrt{2t}) dt \\ &\quad - \frac{e^{-(x')^2/2}}{2\sqrt{\pi}} \operatorname{sgn}(x) \int_0^{x^2/2} \frac{e^{-t}}{\sqrt{t}} c_{n/2}(x'\sqrt{2t}) dt. \end{aligned}$$

(2) (Complex–complex case). If $z, z' \in \mathbb{C}_+$, then

$$\begin{aligned} \tilde{S}_n(z, z') &:= \frac{ie^{-(1/2)(z-\bar{z}')^2}}{\sqrt{2\pi}} (\bar{z}' - z) \sqrt{\operatorname{erfc}(\sqrt{2} \operatorname{Im}(z)) \operatorname{erfc}(\sqrt{2} \operatorname{Im}(z'))} e^{-z\bar{z}'} e_{n/2}(z\bar{z}'), \\ \widetilde{D}S_n(z, z') &:= \frac{e^{-(1/2)(z-z')^2}}{\sqrt{2\pi}} (z' - z) \sqrt{\operatorname{erfc}(\sqrt{2} \operatorname{Im}(z)) \operatorname{erfc}(\sqrt{2} \operatorname{Im}(z'))} e^{-zz'} e_{n/2}(zz'), \\ \widetilde{I}S_n(z, z') &:= -\frac{e^{-(1/2)(\bar{z}-\bar{z}')^2}}{\sqrt{2\pi}} (\bar{z}' - \bar{z}) \sqrt{\operatorname{erfc}(\sqrt{2} \operatorname{Im}(z)) \operatorname{erfc}(\sqrt{2} \operatorname{Im}(z'))} e^{-\bar{z}\bar{z}'} e_{n/2}(\bar{z}\bar{z}'). \end{aligned}$$

(3) (Real–complex case). If $x \in \mathbb{R}$ and $z \in \mathbb{C}_+$, then

$$\begin{aligned} \tilde{S}_n(x, z) &:= \frac{ie^{-(1/2)(x-\bar{z})^2}}{\sqrt{2\pi}} \sqrt{\operatorname{erfc}(\sqrt{2} \operatorname{Im}(z))} e^{-x\bar{z}} e_{n/2}(x\bar{z}), \\ \tilde{S}_n(z, x) &:= \frac{e^{-(1/2)(x-z)^2}}{\sqrt{2\pi}} \sqrt{\operatorname{erfc}(\sqrt{2} \operatorname{Im}(z))} e^{-xz} e_{n/2}(xz) + r_{n/2}(z, x), \\ \widetilde{D}S_n(x, z) &:= \frac{e^{-(1/2)(x-z)^2}}{\sqrt{2\pi}} (z - x) \sqrt{\operatorname{erfc}(\sqrt{2} \operatorname{Im}(z))} e^{-xz} e_{n/2}(xz), \\ \widetilde{I}S_n(x, z) &:= -\frac{ie^{-(1/2)(x-\bar{z})^2}}{\sqrt{2\pi}} \sqrt{\operatorname{erfc}(\sqrt{2} \operatorname{Im}(z))} e^{-x\bar{z}} e_{n/2}(x\bar{z}) - ir_{n/2}(\bar{z}, x). \end{aligned}$$

As $\mathcal{E}(\gamma, \gamma')$ is clearly bounded, it thus suffices [in view of (B.1)] to show that all the expressions $\tilde{S}_n(x, x')$, $\widetilde{D}S_n(x, x')$, $\widetilde{I}S_n(x, x')$, $\tilde{S}_n(z, z')$, $\widetilde{D}S_n(z, z')$, $\widetilde{I}S_n(z, z')$, $\tilde{S}_n(x, z)$, $\tilde{S}_n(z, x)$, $\widetilde{D}S_n(x, z)$, $\widetilde{I}S_n(x, z)$ are all $O(1)$ for $x, x' \in \mathbb{R}$ and $z, z' \in \mathbb{C}_+$.

This will be a variant of the estimates in [7], Section 9, which were concerned with the asymptotic values of these expressions as $n \rightarrow \infty$ rather than uniform bounds.

We first dispose of the $r_{n/2}$ terms. In the proof of [7], Corollary 9, the estimate

$$|r_{n/2}(z, x)| \leq e^{-(1/2)\operatorname{Re}(z^2)} \sqrt{\operatorname{erfc}(\sqrt{2}\operatorname{Im}(z))} \frac{|z|^{n-1}}{2^{n/2}(n/2-1)!}$$

is established for any $x \in \mathbb{R}$ and $z \in \mathbb{C}^+$. Using the standard bound

$$(B.2) \quad \operatorname{erfc}(x) = O\left(\frac{e^{-x^2}}{1+x}\right)$$

for any $x \geq 0$, we thus have

$$|r_{n/2}(z, x)| \ll e^{-|z|^2/2} \frac{|z|^{n-1}}{2^{(n-1)/2}(n/2-1)!}.$$

But $\frac{|z|^{n-1}}{2^{(n-1)/2}(n/2-1)!}$ is one of the Taylor coefficients of $e^{|z|^2/2}$, and so

$$(B.3) \quad r_{n/2}(z, x) = O(1).$$

Thus, we may ignore all terms involving $r_{n/2}$.

Now we handle the real–real case. Recall from the triangle inequality and Taylor expansion that

$$(B.4) \quad |e_{n/2}(z)| \leq e_{n/2}(|z|) \leq \exp(|z|)$$

for any complex number z . Thus, for instance, we have

$$|\tilde{S}_n(x, x')| \ll \exp(-(x-x')^2/2 - xx' + |xx'|) + 1 \ll 1$$

since the expression inside the exponential is either $-(x-x')^2/2$ or $-(x+x')^2/2$.

If one applies the same method to bound $\widetilde{D}S_n(x, x')$, one obtains.

Similarly, one has

$$|\widetilde{D}S_n(x, x')| \ll |x-x'| \exp(-(x-x')^2/2 - xx' + |xx'|).$$

This bound is $O(1)$ when xx' is positive, but can grow linearly when xx' is negative. To deal with this issue, we need an alternate bound to (B.4) that saves an additional polynomial factor in some cases:

LEMMA 66 (Alternate bound). *For any complex number z , one has*

$$|e_{n/2}(z)| \ll \frac{|z|^{1/2}}{||z| - z|} \exp(|z|)$$

with the convention that the right-hand side is infinite when z is a nonnegative real.

PROOF. The claim is trivial for $|z| \leq 1$, so we may assume that $|z| > 1$. Observe that

$$(B.5) \quad (|z| - z)e_{n/2}(z) = \sum_{m=0}^{n/2} \frac{z^m}{m!} (|z| - m) - \frac{z^{n/2+1}}{(n/2)!}.$$

An application of Stirling’s formula reveals that

$$\frac{z^m}{m!} = O\left(\frac{1}{|z|^{1/2}} \exp(|z|)\right)$$

for all m , so the second term on the right-hand side of (B.5) is $O(|z| \frac{1}{|z|^{1/2}} \exp(|z|))$. It thus suffices to show that

$$\sum_{m=0}^{n/2} \frac{z^m}{m!} (|z| - m) = O(|z|^{1/2} \exp(|z|)).$$

By the triangle inequality, the left-hand side can be bounded by

$$\sum_{m \leq |z|} \frac{|z|^m}{m!} (|z| - m) + \sum_{m > |z|} \frac{|z|^m}{m!} (m - |z|).$$

This expression telescopes to

$$2 \frac{|z|^{m+1}}{m!},$$

where $m := \lfloor |z| \rfloor$. By Stirling’s formula, this expression is $O(|z|^{1/2} \exp(|z|))$ as required. \square

Inserting this bound in the case when xx' is negative, we conclude that

$$\begin{aligned} |\widetilde{D}\mathcal{S}_n(x, x')| &\ll |x - x'| \frac{1}{(xx')^{1/2}} \exp(-(x - x')^2/2 - xx' + |xx'|) \\ &= \frac{|x| + |x'|}{|x|^{1/2}|x'|^{1/2}} \exp((|x| - |x'|)^2/2) \end{aligned}$$

and one easily verifies that this expression is $O(1)$.

Finally, to control $\widetilde{I}\mathcal{S}_n(x, x')$, it suffices by symmetry to show that

$$(B.6) \quad \int_0^{(x')^2/2} \frac{e^{-t}}{\sqrt{t}} c_{n/2}(x\sqrt{2t}) dt = O(\exp(x^2/2)).$$

But by Taylor expansion we may bound $c_{n/2}(x\sqrt{2t})$ by $\cosh(x\sqrt{2t})$. Since

$$\int_0^{(x')^2/2} \frac{e^{-t}}{\sqrt{t}} \cosh(x\sqrt{2t}) = \frac{\sqrt{\pi}}{2} e^{(x')^2/2} \left(\operatorname{erf}\left(\frac{|x| + |x'|}{\sqrt{2}}\right) - \operatorname{erf}\left(\frac{|x'| - |x|}{\sqrt{2}}\right) \right),$$

we see from (B.2) that the left-hand side of (B.6) is

$$\ll \exp((x')^2/2) \exp(-\max(|x'| - |x|, 0)^2/2) \leq \exp(x^2/2)$$

as required.

Next we turn to the complex-complex case. From (B.2) and (B.4), we see that

$$|\tilde{S}_n(z, z')| \ll \exp\left(-\frac{1}{2} \operatorname{Re}((z - \bar{z}')^2)\right) |\bar{z}' - z| \frac{\exp(-\operatorname{Im}(z)^2 - \operatorname{Im}(z')^2)}{(1 + \operatorname{Im}(z))^{1/2}(1 + \operatorname{Im}(z'))^{1/2}} \\ \times \exp(|z\bar{z}'| - \operatorname{Re}(z\bar{z}')).$$

After some rearrangement, the right-hand side here becomes

$$\frac{|\bar{z}' - z|}{(1 + \operatorname{Im}(z))^{1/2}(1 + \operatorname{Im}(z'))^{1/2}} \exp\left(-\frac{1}{2}(|z| - |z'|)^2\right).$$

If one uses Lemma 66 instead of (B.4), one gains an additional factor of $\frac{|z|^{1/2}|z'|^{1/2}}{||z||z'| - z\bar{z}'|}$. Thus, it suffices to show that

$$(B.7) \quad \frac{|\bar{z}' - z|}{(1 + \operatorname{Im}(z))^{1/2}(1 + \operatorname{Im}(z'))^{1/2}} \min\left(1, \frac{|z|^{1/2}|z'|^{1/2}}{||z||z'| - z\bar{z}'|}\right) \\ \times \exp\left(-\frac{1}{2}(|z| - |z'|)^2\right) \ll 1.$$

By symmetry, we may assume that $0 < \operatorname{Im}(z) \leq \operatorname{Im}(z')$. We may assume that $|z|$ and $|z'|$ are comparable and larger than 1, since otherwise the claim easily follows from the $\exp(-\frac{1}{2}(|z| - |z'|)^2)$ term.

Let θ denote the angle subtended by z and z' . Observe from the triangle inequality that

$$(B.8) \quad |\bar{z}' - z| \ll ||z| - |z'|| + \operatorname{Im}(z) + |z|\theta$$

and

$$||z||z'| - z\bar{z}'| \gg |z|^2\theta.$$

The first two terms on the right-hand side of (B.8) give an acceptable contribution to (B.7) (bounding the minimum crudely by 1), so it suffices to show that

$$\frac{|z|\theta}{(1 + \operatorname{Im}(z))^{1/2}(1 + \operatorname{Im}(z'))^{1/2}} \min\left(1, \frac{|z|}{|z|^2\theta}\right) \ll 1,$$

but this is clear after discarding the denominator and using the second term in the minimum. This establishes the bound $|\tilde{S}_n(z, z')| \ll 1$. Similar arguments, which we leave to the reader, show that $|\widetilde{DS}_n(z, z')| \ll 1$ and $|\widetilde{IS}_n(z, z')| \ll 1$.

Finally, we turn to the real–complex case. Using (B.4) and (B.2), we can bound

$$|\tilde{S}_n(x, z)| \ll \exp\left(-\frac{1}{2} \operatorname{Re}((x - \bar{z})^2)\right) \frac{\exp(-\operatorname{Im}(z)^2)}{1 + \operatorname{Im}(z)^{1/2}} \exp(-x\bar{z} + |x||z|).$$

The right-hand side simplifies to $\exp(-(x - |z|)^2/2)/(1 + \operatorname{Im}(z)^{1/2})$, which is clearly $O(1)$.

A similar argument [using (B.3)] shows that $\tilde{S}_n(x, z) = O(1)$ and $\tilde{I}\tilde{S}_n(x, z) = O(1)$. The bound $\tilde{D}\tilde{S}_n(x, z) = O(1)$ can be established by the same arguments used to handle the complex–complex case; we leave the details to the reader. This completes the proof of Lemma 11.

Acknowledgments. We thank Anthony Mays and the anonymous referees for corrections and help with the references.

REFERENCES

- [1] AKEMANN, G. and KANZIEPER, E. (2007). Integrable structure of Ginibre’s ensemble of real random matrices and a Pfaffian integration theorem. *J. Stat. Phys.* **129** 1159–1231. [MR2363393](#)
- [2] ALON, N. and SPENCER, J. H. (2008). *The Probabilistic Method*, 3rd ed. Wiley, Hoboken, NJ. [MR2437651](#)
- [3] BAI, Z. D. (1997). Circular law. *Ann. Probab.* **25** 494–529. [MR1428519](#)
- [4] BAI, Z. D. and YIN, Y. Q. (1986). Limiting behavior of the norm of products of random matrices and two problems of Geman–Hwang. *Probab. Theory Related Fields* **73** 555–569. [MR0863545](#)
- [5] BORDENAVE, C. and CHAFAÏ, D. (2012). Around the circular law. *Probab. Surv.* **9** 1–89. [MR2908617](#)
- [6] BORODIN, A. and SINCLAIR, C. D. Correlation functions of ensembles of asymmetric real matrices. Preprint. Available at [arXiv:0706.2670v1](#).
- [7] BORODIN, A. and SINCLAIR, C. D. (2009). The Ginibre ensemble of real random matrices and its scaling limits. *Comm. Math. Phys.* **291** 177–224. [MR2530159](#)
- [8] BOURGADE, P., YAU, H.-T. and YIN, J. (2014). The local circular law II: The edge case. *Probab. Theory Related Fields* **159** 619–660. [MR3230004](#)
- [9] BOURGADE, P., YAU, H.-T. and YIN, J. (2014). Local circular law for random matrices. *Probab. Theory Related Fields* **159** 545–595. [MR3230002](#)
- [10] BROWN, L. G. (1986). Lidskii’s theorem in the type II case. In *Geometric Methods in Operator Algebras (Kyoto, 1983)*. *Pitman Res. Notes Math. Ser.* **123** 1–35. Longman Sci. Tech., Harlow. [MR0866489](#)
- [11] CHATTERJEE, S. (2006). A generalization of the Lindeberg principle. *Ann. Probab.* **34** 2061–2076. [MR2294976](#)
- [12] COSTIN, A. and LEBOWITZ, J. L. (1995). Gaussian fluctuations in random matrices. *Phys. Rev. Lett.* **75** 69–72.
- [13] EDELMAN, A. (1997). The probability that a random real Gaussian matrix has k real eigenvalues, related distributions, and the circular law. *J. Multivariate Anal.* **60** 203–232. [MR1437734](#)
- [14] EDELMAN, A., KOSTLAN, E. and SHUB, M. (1994). How many eigenvalues of a random matrix are real? *J. Amer. Math. Soc.* **7** 247–267. [MR1231689](#)

- [15] ERDŐS, L. (2011). Universality of Wigner random matrices: A survey of recent results. *Russian Math. Surveys* **66** 67–198.
- [16] ERDŐS, L., RAMÍREZ, J., SCHLEIN, B., TAO, T., VU, V. and YAU, H.-T. (2010). Bulk universality for Wigner Hermitian matrices with subexponential decay. *Math. Res. Lett.* **17** 667–674. [MR2661171](#)
- [17] ERDŐS, L., SCHLEIN, B. and YAU, H.-T. (2009). Semicircle law on short scales and delocalization of eigenvectors for Wigner random matrices. *Ann. Probab.* **37** 815–852. [MR2537522](#)
- [18] ERDŐS, L., SCHLEIN, B. and YAU, H.-T. (2009). Local semicircle law and complete delocalization for Wigner random matrices. *Comm. Math. Phys.* **287** 641–655. [MR2481753](#)
- [19] ERDŐS, L., SCHLEIN, B. and YAU, H.-T. (2010). Wegner estimate and level repulsion for Wigner random matrices. *Int. Math. Res. Not. IMRN* **3** 436–479. [MR2587574](#)
- [20] ERDŐS, L., SCHLEIN, B. and YAU, H.-T. (2011). Universality of random matrices and local relaxation flow. *Invent. Math.* **185** 75–119. [MR2810797](#)
- [21] ERDŐS, L., SCHLEIN, B., YAU, H.-T. and YIN, J. (2012). The local relaxation flow approach to universality of the local statistics for random matrices. *Ann. Inst. Henri Poincaré Probab. Stat.* **48** 1–46. [MR2919197](#)
- [22] ERDŐS, L., YAU, H.-T. and YIN, J. (2012). Bulk universality for generalized Wigner matrices. *Probab. Theory Related Fields* **154** 341–407. [MR2981427](#)
- [23] FORRESTER, P. J. and MAYS, A. (2009). A method to calculate correlation functions for $\beta = 1$ random matrices of odd size. *J. Stat. Phys.* **134** 443–462. [MR2485724](#)
- [24] FORRESTER, P. J. and NAGAO, T. (2007). Eigenvalue statistics of the real Ginibre ensemble. *Phys. Rev. Lett.* **99** 050603.
- [25] GEMAN, S. (1986). The spectral radius of large random matrices. *Ann. Probab.* **14** 1318–1328. [MR0866352](#)
- [26] GINIBRE, J. (1965). Statistical ensembles of complex, quaternion, and real matrices. *J. Math. Phys.* **6** 440–449. [MR0173726](#)
- [27] GIRKO, V. L. (1984). The circular law. *Teor. Veroyatnost. i Primenen.* **29** 669–679.
- [28] GUIONNET, A. (2009–2010). Grandes matrices aléatoires et théorèmes d’universalité, *Séminaire BOURBAKI*. Avril 2010. 62ème année, 2009–2010, no 1019.
- [29] HANSON, D. L. and WRIGHT, F. T. (1971). A bound on tail probabilities for quadratic forms in independent random variables. *Ann. Math. Statist.* **42** 1079–1083. [MR0279864](#)
- [30] KANZIEPER, E. and AKEMANN, G. (2005). Statistics of real eigenvalues in Ginibre’s ensemble of random real matrices. *Phys. Rev. Lett.* **95** 230201, 4. [MR2185860](#)
- [31] KNOWLES, A. and YIN, J. (2013). Eigenvector distribution of Wigner matrices. *Probab. Theory Related Fields* **155** 543–582. [MR3034787](#)
- [32] KOSTLAN, E. (1992). On the spectra of Gaussian matrices. *Linear Algebra Appl.* **162/164** 385–388. [MR1148410](#)
- [33] KRISHNAPUR, M. and VIRÁG, B. (2014). The Ginibre ensemble and Gaussian analytic functions. *Int. Math. Res. Not. IMRN* **6** 1441–1464. [MR3180597](#)
- [34] LEHMANN, N. and SOMMERS, H.-J. (1991). Eigenvalue statistics of random real matrices. *Phys. Rev. Lett.* **67** 941–944. [MR1121461](#)
- [35] MEHTA, M. L. (1967). *Random Matrices and the Statistical Theory of Energy Levels*. Academic Press, New York. [MR0220494](#)
- [36] NGUYEN, H. H. and VU, V. (2014). Random matrices: Law of the determinant. *Ann. Probab.* **42** 146–167. [MR3161483](#)
- [37] NOURDIN, I. and PECCATI, G. (2010). Universal Gaussian fluctuations of non-Hermitian matrix ensembles: From weak convergence to almost sure CLTs. *ALEA Lat. Am. J. Probab. Math. Stat.* **7** 341–375. [MR2738319](#)
- [38] RIDER, B. (2003). A limit theorem at the edge of a non-Hermitian random matrix ensemble. *J. Phys. A* **36** 3401–3409. [MR1986426](#)

- [39] RIDER, B. (2004). Deviations from the circular law. *Probab. Theory Related Fields* **130** 337–367. [MR2095933](#)
- [40] RIDER, B. and SILVERSTEIN, J. W. (2006). Gaussian fluctuations for non-Hermitian random matrix ensembles. *Ann. Probab.* **34** 2118–2143. [MR2294978](#)
- [41] RUDELSON, M. and VERSHYNIN, R. (2008). The Littlewood–Offord problem and invertibility of random matrices. *Adv. Math.* **218** 600–633. [MR2407948](#)
- [42] RUDELSON, M. and VERSHYNIN, R. (2010). Non-asymptotic theory of random matrices: Extreme singular values. In *Proceedings of the International Congress of Mathematicians. Volume III* 1576–1602. Hindustan Book Agency, New Delhi. [MR2827856](#)
- [43] SCHLEIN, B. (2011). Spectral properties of Wigner matrices. In *Mathematical Results in Quantum Physics* 79–94. World Sci. Publ., Hackensack, NJ. [MR2885161](#)
- [44] SINCLAIR, C. D. (2009). Correlation functions for $\beta = 1$ ensembles of matrices of odd size. *J. Stat. Phys.* **136** 17–33. [MR2525224](#)
- [45] SOMMERS, H.-J. and WIECZOREK, W. (2008). General eigenvalue correlations for the real Ginibre ensemble. *J. Phys. A* **41** 405003, 24. [MR2439268](#)
- [46] SOSHNIKOV, A. (2002). Gaussian limit for determinantal random point fields. *Ann. Probab.* **30** 171–187. [MR1894104](#)
- [47] SOSHNIKOV, A. B. (2000). Gaussian fluctuation for the number of particles in Airy, Bessel, sine, and other determinantal random point fields. *J. Stat. Phys.* **100** 491–522. [MR1788476](#)
- [48] TALAGRAND, M. (1996). A new look at independence. *Ann. Probab.* **24** 1–34. [MR1387624](#)
- [49] TAO, T. and VU, V. Random matrices: The universality phenomenon for Wigner ensembles. Preprint. Available at [arXiv:1202.0068](#).
- [50] TAO, T. and VU, V. (2006). *Additive Combinatorics*. Cambridge Univ. Press, Cambridge. [MR2289012](#)
- [51] TAO, T. and VU, V. (2007). The condition number of a randomly perturbed matrix. In *STOC'07—Proceedings of the 39th Annual ACM Symposium on Theory of Computing* 248–255. ACM, New York. [MR2402448](#)
- [52] TAO, T. and VU, V. (2009). From the Littlewood–Offord problem to the circular law: Universality of the spectral distribution of random matrices. *Bull. Amer. Math. Soc. (N.S.)* **46** 377–396. [MR2507275](#)
- [53] TAO, T. and VU, V. (2010). Random matrices: Universality of local eigenvalue statistics up to the edge. *Comm. Math. Phys.* **298** 549–572. [MR2669449](#)
- [54] TAO, T. and VU, V. (2010). Smooth analysis of the condition number and the least singular value. *Math. Comp.* **79** 2333–2352. [MR2684367](#)
- [55] TAO, T. and VU, V. (2011). Random matrices: Universality of local eigenvalue statistics. *Acta Math.* **206** 127–204. [MR2784665](#)
- [56] TAO, T. and VU, V. (2011). The Wigner–Dyson–Mehta bulk universality conjecture for Wigner matrices. *Electron. J. Probab.* **16** 2104–2121. [MR2851058](#)
- [57] TAO, T. and VU, V. (2012). Random covariance matrices: Universality of local statistics of eigenvalues. *Ann. Probab.* **40** 1285–1315. [MR2962092](#)
- [58] TAO, T. and VU, V. (2012). A central limit theorem for the determinant of a Wigner matrix. *Adv. Math.* **231** 74–101. [MR2935384](#)
- [59] TAO, T. and VU, V. H. (2009). Inverse Littlewood–Offord theorems and the condition number of random discrete matrices. *Ann. of Math. (2)* **169** 595–632. [MR2480613](#)
- [60] TROTTER, H. F. (1984). Eigenvalue distributions of large Hermitian matrices; Wigner’s semi-circle law and a theorem of Kac, Murdock, and Szegő. *Adv. Math.* **54** 67–82. [MR0761763](#)
- [61] VU, V. H. (2002). Concentration of non-Lipschitz functions and applications. *Random Structures Algorithms* **20** 262–316. [MR1900610](#)

- [62] WRIGHT, F. T. (1973). A bound on tail probabilities for quadratic forms in independent random variables whose distributions are not necessarily symmetric. *Ann. Probab.* **1** 1068–1070.
[MR0353419](#)

DEPARTMENT OF MATHEMATICS
UCLA
LOS ANGELES, CALIFORNIA 90095-1555
USA
E-MAIL: tao@math.ucla.edu

DEPARTMENT OF MATHEMATICS
YALE UNIVERSITY
NEW HAVEN, CONNECTICUT 06520-8283
USA
E-MAIL: van.vu@yale.edu