ON THE CONDITION NUMBER OF THE SHIFTED REAL GINIBRE ENSEMBLE

GIORGIO CIPOLLONI AND LÁSZLÓ ERDŐS

IST Austria, Am Campus 1, 3400 Klosterneuburg, Austria

DOMINIK SCHRÖDER*

Institute for Theoretical Studies, ETH Zurich, Clausiusstr. 47, 8092 Zurich, Switzerland

Abstract. We derive an accurate lower tail estimate on the lowest singular value $\sigma_1(X-z)$ of a *real* Gaussian (Ginibre) random matrix X shifted by a *complex* parameter z. Such shift effectively changes the upper tail behaviour of the condition number $\kappa(X-z)$ from the slower $\mathbf{P}(\kappa(X-z)\ge t)\lesssim 1/t$ decay typical for real Ginibre matrices to the faster $1/t^2$ decay seen for complex Ginibre matrices as long as z is away from the real axis. This sharpens and resolves a recent conjecture in [8] on the regularizing effect of the real Ginibre ensemble with a genuinely complex shift. As a consequence we obtain an improved upper bound on the eigenvalue condition numbers (known also as the eigenvector overlaps) for real Ginibre matrices. The main technical tool is a rigorous supersymmetric analysis from our earlier work [19].

1. Introduction

The condition number $\kappa(X) = \|X\|\|X^{-1}\|$ of large $N\times N$ random matrices X has been a central object in numerical linear algebra at least since the pioneering work of Goldstine and von-Neumann [28], and Demmel [21]. Demmel showed that for a large class of *complex* random matrices X the probability that $\kappa(X)$ is larger than a threshold $t\gg 1$ decays as $1/t^2$, while for *real* matrices the decay rate is slower, of order $1/t$. While the dependence on N was not optimal in Demmel's work, for the specific Gaussian case much more precise results are available. Gaussian random matrices have frequently been used as a test case since often explicit formulas are available for their spectral distribution.

The simplest non-Hermitian random matrix model is the *real* or *complex Ginibre ensemble*, consisting of matrices with independent identically distributed (i.i.d) Gaussian matrix elements. We fix the customary normalization, $\mathbf{E}\,x_{ab}=0$, $\mathbf{E}\,|x_{ab}|^2=N^{-1}$ that guarantees that the density of eigenvalues of X converges to the uniform measure on the complex unit disk (known as the *Circular law*) and that the spectral radius of X converges to 1 with very high probability (these results also hold for non-Gaussian matrix elements, see e.g. [27, 4, 39, 26, 5, 10, 11]), c.f. Fig. 1. Edelman in [22] gave an exact formula for the distribution of the lowest singular value of a Ginibre matrix in both symmetry classes and derived precise large N asymptotics for the condition number, confirming Demmel's upper tail decay on the distribution of $\kappa(X)$ uniformly in the dimension. Non-asymptotic upper and lower bounds with good explicit constants were obtained in [3] for the real case and later extended to rectangular matrices [23, 15] in both symmetry classes.

In more recent applications Ginibre matrices arise as additive perturbations of a deterministic matrix A. The prominent example is the concept of *smoothed analysis* (originally introduced in [38] in the context of the simplex algorithm), where Sankar, Spielman and Teng [37] considered the Gaussian elimination algorithm without pivoting for solving large dimensional linear systems of equations $A\mathbf{x}=\mathbf{b}$. The *bit-complexity* of Gaussian elimination, i.e. the computational cost of achieving a desired output accuracy, depends primarily on $\kappa(A)$ and its upper tail is mainly determined by the lower tail behaviour of $\sigma_1(A)$, the lowest singular value of A (note that $\sigma_1(A)^2=\lambda_1(AA^*)$, the lowest eigenvalue of

E-mail addresses: giorgio.cipolloni@ist.ac.at, lerdos@ist.ac.at, dschroeder@ethz.ch.
Date: November 2, 2022.
2020 *Mathematics Subject Classification.* 60B20, 68W40, 65F35.
Key words and phrases. Smoothed Analysis, Supersymmetric formalism, Circular Law.
* Supported by Dr. Max Rössler, the Walter Haefner Foundation and the ETH Zürich Foundation.

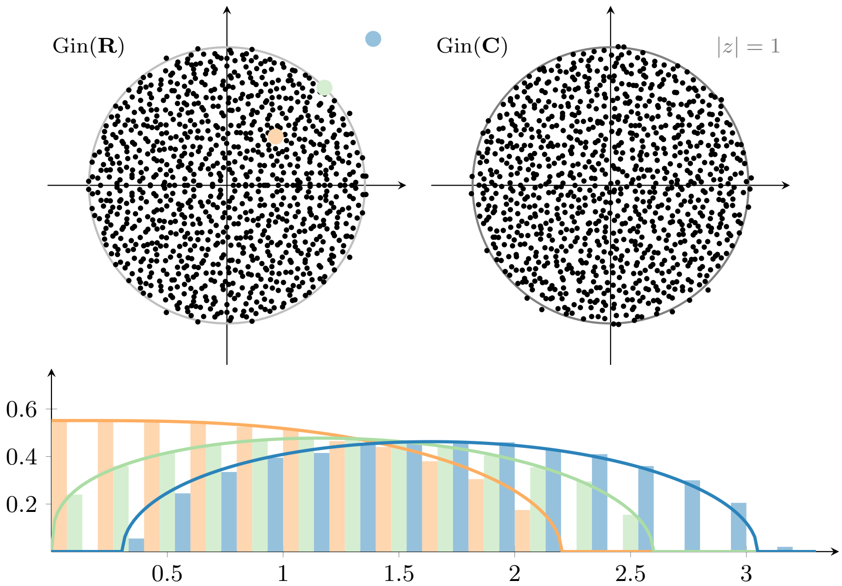

FIGURE 1. The top figure shows the eigenvalues of a single real and complex Ginibre matrix. Note that the eigenvalues of the real Ginibre matrix are symmetric with respect to the real axis, and that some (in fact $\sim \sqrt{N}$) eigenvalues are on the axis itself. The bottom figure shows the singular values of $X - z$ for three different values of z in histogram form (for the single matrix whose eigenvalues are displayed in the top-left figure), together with their theoretic density (solid lines). The singular value density depends only on the absolute value $|z|$ and not on the phase of z since the effect of the real axis is only visible in the density of the smallest singular values, and not in the global density of all singular values. We note that the singular value density is strictly positive at 0 whenever $|z| < 1$.

AA^*). In order to obtain a bound on the *real world accuracy loss* of the problem $A\mathbf{x} = \mathbf{b}$, rather than the *averaged* or *worst case accuracy loss*, the main result of [37] is an estimate on the smoothed value of $\kappa(A+\gamma X)$ for small γ and Ginibre-distributed X. In practice γ is then optimized to balance between the gain in bit-complexity versus the loss in precision.

More recently smoothed analysis has been applied to the problem of finding eigenvalue/eigenvector pairs [2] and to full matrix diagonalization [7, 6] by Banks, Vargas, Kulkarni and Srivastava. This required to develop the ideas of smoothed analysis for the *eigenvector condition number* (see (16) later) in [7] and then in [6] further to the minimal eigenvalue gap (see (29) later).

In [37], the authors proved[1] the following lower tail bound on the (square of the) lowest singular value of the regularised matrix $A + \gamma X$:

$$\mathbf{P}\left(\sigma_1(A+\gamma X) \le \frac{\sqrt{x}}{N}\right) \le C\sqrt{\frac{x}{\gamma^2}}, \qquad x > 0, \tag{1}$$

for a real Ginibre matrix X, and

$$\mathbf{P}\left(\sigma_1(A+\gamma X) \le \frac{\sqrt{x}}{N}\right) \le C\frac{x}{\gamma^2}, \qquad x > 0, \tag{2}$$

for a complex Ginibre matrix X. The constant C is universal, the estimates are uniform in A and γ. The N^{-2} scaling naturally comes from the typical $1/N$ spacing between the eigenvalues of the corresponding Hermitized matrix

$$H^A := \begin{pmatrix} 0 & A+\gamma X \\ (A+\gamma X)^* & 0 \end{pmatrix} \tag{3}$$

[1]The paper [37] states the result only for the real case, but the complex case easily follows by the same proof.

in its bulk spectrum. Comparing the bounds (1) and (2) in the small x regime, note that the regularizing effect of a complex Ginibre matrix is much stronger. Can one achieve the same effect with real Ginibre matrices?

On one hand, inspecting the proof in [37], the exponents of x in the right hand side of (1) and (2) are direct consequences of the one- vs. two-dimensionality of the support of the real vs. complex random variables x_{ab} and the effect is completely independent of A. On the other hand, quite remarkably, the local *eigenvalue* statistics of the real and complex Ginibre ensemble coincide away from the real axis, see [12, Theorem 11]. Very recently in [16, 18] we showed an analogous phenomenon for the singular values of the *shifted Ginibre matrix*. More precisely, in [16] the density of the low lying singular values of $X - z$ for a real and complex Ginibre X was shown to coincide if the shift parameter z is genuinely complex, $|\Im z| \gg N^{-1/2}$. In the regime $|\Im z| \sim 1$ the same coincidence was proven for all k-point correlations functions [18, Theorem 2.8]. In particular, on the level of the small singular values, the *real* Ginibre matrix with a *complex* shift behaves as a *complex* Ginibre matrix!

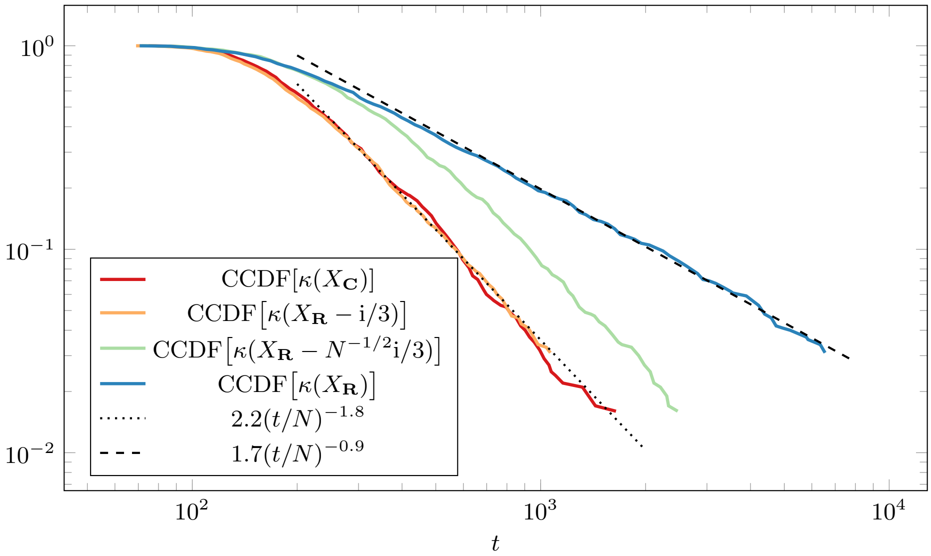

FIGURE 2. The complementary cumulative distribution functions (CCDF) $\mathbf{P}(\kappa \geq t)$ for the condition number $\kappa(X - z)$ of shifted Ginibre matrices obtained from 10 000 random matrices of size 100×100. Away from the real axis the probability of having condition number larger than t decays as t^{-2} for both real and complex Ginibre matrices. On the real axis the real Ginibre ensemble has a slower tail of t^{-1} and exhibits an interpolating behaviour as $\Im z \sim N^{-1/2}$.

For the purpose of the smoothed analysis this indicates the possibility that real Ginibre matrices are as effectively regularizing as the complex ones, at least away from the real axis. To test this hypothesis, we consider the simplest $A = -zI$ case, the shifted Ginibre ensemble. In fact, the following conjecture in this spirit was very recently posed in Section 7 of [8] (with our notations and with $\gamma = 1$ for simplicity):

$$\mathbf{P}\left(\sigma_1(X - z) \leq \frac{\sqrt{x}}{N}\right) \leq C\frac{x}{|\Im z|}, \qquad x > 0. \tag{4}$$

Here and in the sequel we will frequently omit the identity matrix for brevity and write simply $X - z$ for $X - zI$. While (4) highlights the role of $\Im z$, its scaling is far from optimal: by analogy with the eigenvalues, one expects that a real Ginibre matrix X near z behaves essentially as a complex Ginibre matrix as long as $|\Im z| \gg N^{-1/2}$. Indeed, our main result in Theorem 2.1 shows that

$$\mathbf{P}\left(\sigma_1(X - z) \leq \frac{\sqrt{x}}{N}\right) \lesssim (1 + |\log x|)x + e^{-\frac{1}{2}N|\Im z|^2} \min\left\{\sqrt{x}, \frac{x}{\sqrt{N}|\Im z|}\right\}, \quad x > 0, \tag{5}$$

proving that the essentially linear bound (in x) from (2) already dominates the tail behaviour of the lowest singular value for $|\Im z| \gg N^{-1/2}$, while the much larger $\sqrt{x}$ tail prevails in the opposite regime.

Since $\kappa(X-z) \sim \sigma_1(X-z)^{-1}$, we directly obtain the transition from $1/t$ to $1/t^2$ for the upper tail $\mathbf{P}(\kappa(X-z) \geq t)$ as $|\Im z|$ increases well above $N^{-1/2}$, see Fig. 2. A similar behaviour is expected to hold for general matrix A, see Conjecture 2.4.

The bound (5) has several consequences on the *eigenvalue condition number* $\kappa(\lambda) := \|L\|\|R\|$ that determines the stability of the eigenvalue λ against small perturbations, where L, R denote the corresponding left and right eigenvectors with the customary normalization $\langle L, R\rangle = 1$. For complex Ginibre matrices it is known that $\kappa(\lambda)$ is of order $\sqrt{N}$, see [14, 13, 24]. For the real Ginibre case a similar result is obtained in [24] but only for real eigenvalues $\lambda \in \mathbf{R}$. Suboptimal bounds in N have very recently been established in [8] and [33] that also hold for a general matrix A, in particular for any shift $z \in \mathbf{C}$.

The proof of our main estimate (5) uses the supersymmetric (SUSY) approach that is common in the physics literature on random matrices, but is less known in the numerics community. Most of the necessary technical work has already been done in our previous paper [19]; hence the current paper is short and focuses on the results. Our purpose is to demonstrate the power of the SUSY method to obtain very accurate estimates. For example, the exponential suppression factor $\exp(-\frac{1}{2}N|\Im z|^2)$ in (5) expressing the true effect of the non-zero imaginary part of the shift parameter seems very hard to obtain with any other method, while it easily comes out from the SUSY formalism.

Notations and conventions. For positive quantities f, g we write $f \lesssim g$ and $f \sim g$ if $f \leq Cg$ or $cg \leq f \leq Cg$, respectively, for some constants $c, C > 0$ which are independent of N and z. We write $\mathbf{D} \subset \mathbf{C}$ for the open unit disk. We abbreviate the minimum and maximum of real numbers by $a \wedge b := \min\{a, b\}$ and $a \vee b := \max\{a, b\}$.

2. Main results

We consider the ensemble $Y^z := (X-z)(X-z)^*$ with $X \in \mathbf{R}^{N\times N}$ being a *real Ginibre* matrix, i.e. its entries x_{ab} are such that $\sqrt{N}x_{ab}$ are i.i.d. standard real Gaussian random variables, and $z \in \mathbf{C}$ is a fixed complex parameter such that $|z| \leq 1$. Our main results are an optimal lower tail estimate for the smallest singular value of $X-z$ and its consequences for eigenvector overlaps and eigenvalue condition number of real Ginibre matrices.

2.1. Singular value and condition number tail estimates for $X-z$.

The following theorem gives an estimate on the lowest singular value of $X-z$ uniformly in all the relevant parameters and on the optimal scale. Its direct corollary is an analogous estimate on $\kappa(X-z)$.

Theorem 2.1. *Let $\eta := \Im z$, $\delta := 1-|z|^2$ and*

$$c(N,\delta) := \frac{1}{N^{3/2}} \wedge \frac{1}{N^2|\delta|}. \tag{6}$$

Then, denoting by $\lambda_1(Y^z)$ the smallest eigenvalue of Y^z, uniformly in $x \in [0,\infty)$, $\eta \in [-1,1]$, $\delta \in [-10N^{-1/2}, 1]$, it holds

$$\mathbf{P}\left(\lambda_1(Y^z) \leq xc(N,\delta)\right) \leq C_*(1+|\log x|)x + C_* e^{-\frac{1}{2}N\eta^2}\left(\sqrt{x} \wedge \frac{x}{\sqrt{N}|\eta|}\right) \tag{7}$$

where C_* is a universal constant.

Note that $c(N,\delta)$ is the correct scale of the typical size of $\lambda_1(Y^z)$. Indeed the level spacing of the eigenvalues of Y^z close to zero for $|z| < 1$ is given by $N^{-2}\delta^{-1}$ and for $|z| = 1$ by $N^{-3/2}$, see [1, Section 5]. The $N^{-3/2}$ scaling in the edge regime $|z| = 1$ comes from the fact that the density of eigenvalues of the Hermitized matrix

$$H^z := \begin{pmatrix} 0 & X-z \\ (X-z)^* & 0 \end{pmatrix} \tag{8}$$

develops a cubic cusp singularity that has a natural eigenvalue spacing $N^{-3/4}$.

Remark 2.2. *Introducing the coupling parameter γ and thus replacing $\lambda_1(Y^z)$ by $\lambda_1[(\gamma X-z)(\gamma X-z)^*]$ we then conclude a bound analogous to (7) after replacing x by $x\gamma^{-2}$.*

Remark 2.3. *Theorem 2.1 is proven only for matrices X with Gaussian entries. However, the bound (7) can be extended to matrices X with generic independent identically distributed (i.i.d.) entries at the price of an additional error term. More precisely, for such matrices there exists $\omega > 0$ such that*

$$\mathbf{P}\left(\lambda_1(Y^z) \leq xc(N,\delta)\right) \leq \textit{rhs. of } (7) + CN^{-\omega}, \tag{9}$$

for any $x \geq N^{-\omega}$. Given (7), the bound in (9) is obtained by a standard Green function comparison (GFT) argument (see e.g. [17, Proposition 3]).

For a general deterministic matrix A we can make the following conjecture. Note that (7) proves Conjecture 2.4 for the special case $A = -zI$ up to a logarithmic correction.

Conjecture 2.4. *Let X be an $N \times N$ real Ginibre matrix. There exist constants $c_*, C_* > 0$ such that for any deterministic matrix A and for any $\gamma > 0$ it holds*

$$\mathbf{P}\left(\sigma_1(\gamma X + A) \leq \frac{\sqrt{x}}{N}\right) \leq C_* \frac{x}{\gamma^2} + C_* e^{-c_* \operatorname{Tr}(\Im A)^2} \left(\frac{\sqrt{x}}{\gamma} \wedge \frac{x}{\sqrt{\operatorname{Tr}(\Im A)^2 \gamma^2}}\right) \tag{10}$$

where $\Im A := \frac{1}{2\mathrm{i}}(A - A^*)$.

We conclude this section by remarking that, by $\kappa(X-z) = [\lambda_{\max}(Y^z)/\lambda_1(Y^z)]^{1/2}$, from a lower tail estimate (7) on $\lambda_1(Y^z)$ we immediately obtain an upper tail bound on $\kappa(X-z)$ for $|z| < 99/100$ (the complementary bound in the edge regime $|z| \approx 1$ follows similarly)

$$\mathbf{P}\left(\kappa(X-z) \geq t\right) \lesssim |\log t| \left(\frac{N}{t}\right)^2 + e^{-\frac{1}{2}N\eta^2}\left(\frac{N}{t} \wedge \frac{N^{3/2}}{|\eta| t^2}\right) + e^{-N}. \tag{11}$$

Here we used that the largest eigenvalue $\lambda_{\max}(Y^z)$ can be controlled by the large deviation bound $\mathbf{P}(\|X\| \geq K) \leq e^{-\alpha K^2 N}$ for some small α and any large K. The bound (11) shows Demmel's transition between the $1/t$ and $1/t^2$ tail behaviour up to an exponentially small additive error.

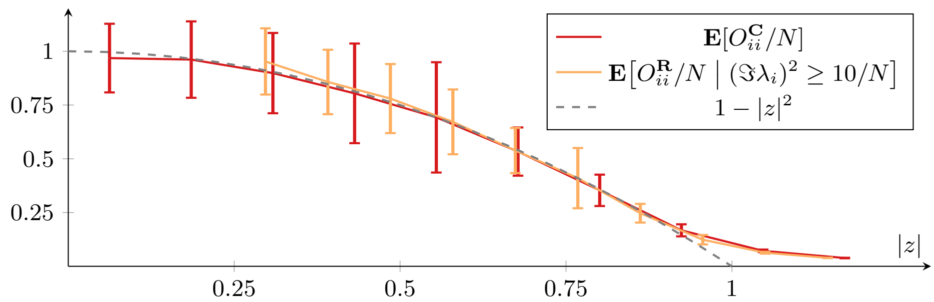

FIGURE 3. The figure shows the empirical averaged overlap conditioned on $|\lambda_i| = |z|$, as well as $(\Im\lambda_i)^2 \geq 10/N$ in the real case, together 95% confidence intervals, obtained from computing the eigenvalues for 10,000 Ginibre matrices of size 100×100. For the complex case, the corresponding behaviour $\mathbf{E}\left[O_{ii}^{\mathbf{C}}/N \middle| \lambda_i = z\right] \approx 1 - |z|^2$ in the large N limit has been established by Chalker and Mehlig [36, 14]. The figure above suggests that the same relation holds true for real Ginibre matrices sufficiently far away from the real axis.

2.2. **Overlaps and condition numbers for X.** Our main result on the tail of the smallest singular value from (7) directly translates into optimal (up to logarithmic corrections) bounds on the *eigenvector overlaps* and *eigenvector condition number*. We denote the left- and right eigenvectors of X corresponding to an eigenvalue λ_i by L_i, R_i so that $X = \sum \lambda_i L_i R_i^*$ with the normalization $\langle L_i, R_i \rangle = 1$ customary in the theory of non-normal matrices. The *diagonal eigenvector overlap* O_{ii} or *eigenvalue condition number* $\kappa(\lambda_i)$ are defined as

$$O_{ii} := \|L_i\|^2 \|R_i\|^2 =: \kappa(\lambda_i)^2. \tag{12}$$

The overlap O_{ii} is directly related to the stability of the eigenvalue $\lambda_i = \lambda_i(X)$ under perturbations in the sense that

$$\lim_{\epsilon \to 0} \sup_{\|E\| \leq 1} \frac{|\lambda_i(X + \epsilon E) - \lambda_i|}{\epsilon} = \sqrt{O_{ii}}. \tag{13}$$

Theorem 2.5. *Let X be a real Ginibre matrix with left and right eigenvectors L_i, R_i corresponding to the eigenvalue λ_i. Then for any open set $\Omega \subset \mathbf{D}$ and any $K > 0$ we have*

$$\mathbf{P}\left(\sum_{i:\lambda_i \in \Omega} \|L_i\|^2 \|R_i\|^2 \geq t(\log N) N^2 \int_{\Omega + B(0, N^{-1/2})} (1 - |z|^2)_+ \,\mathrm{d}^2 z\right) \leq CKt^{-1} + \frac{C_K}{N^K} \tag{14}$$

for any $t > 0$, with some universal constant C and K-dependent constants C_K. Here $\Omega + B(0, N^{-1/2})$ denotes the Minkowski sum of Ω with the ball of radius $N^{-1/2}$.

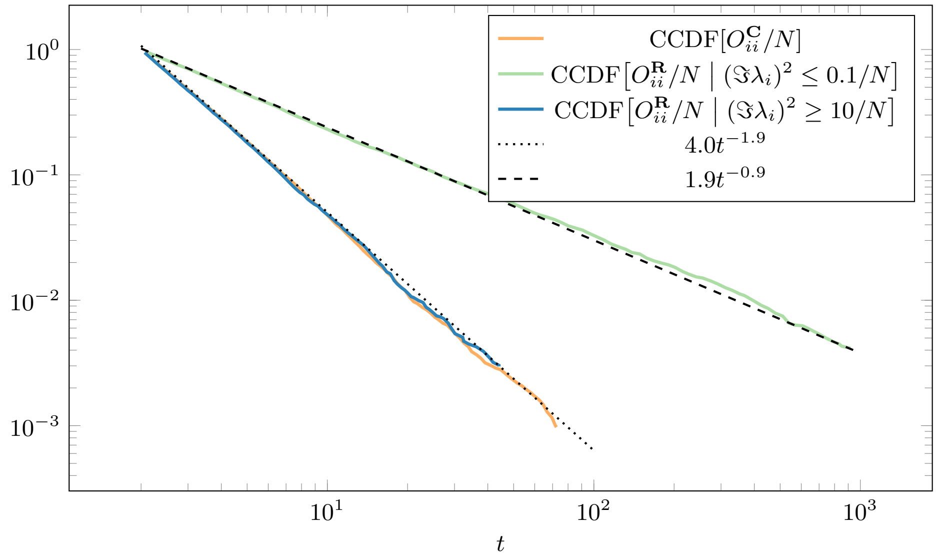

FIGURE 4. The complementary cumulative distribution functions $\mathbf{P}(O_{ii}/N \ge t)$ for the eigenvector overlaps of real and complex Ginibre matrices obtained from 10,000 matrices of size 100×100. The complex overlaps as well as the real overlaps away from the real axis share the same decay exponent of $1.9 \approx 2$, consistent with the fact [13] that in the complex case the overlaps are $1/\gamma_2$-distributed. The fatter tail of the real overlap close to the real axis is responsible for the fact [24] that $\mathbf{E}[O_{ii}^{\mathbf{R}} | \lambda_i \in \mathbf{R}] = \infty$.

Remark 2.6.

(i) *For complex Ginibre matrices Chalker and Mehlig [36, 14] computed the expected overlap (rigorously only for $z = 0$) and showed for its conditional expectation that*

$$\mathbf{E}\Big[O_{ii}^{\mathbf{C}} \,\Big|\, \lambda_i = z\Big] = \mathbf{E}\big[\|L_i\|^2 \|R_i\|^2 \,\big|\, \lambda_i = z\big] \sim N(1 - |z|^2). \tag{15}$$

See Fig. 3 for a comparison between the real Ginibre overlaps far away from the real axis and the complex Ginibre overlaps as a function of $|z|$.

(ii) *In the case of complex Ginibre matrices the distribution of individual overlaps $\|L_i\|^2 \|R_i\|^2$ has been identified in [13, 24], showing that*

$$N^{-1}(1 - |\lambda_i|^2)^{-1} \|L_i\|^2 \|R_i\|^2$$

converges in distribution to an inverse γ_2 random variable.

(iii) *Similarly, in the case of real Ginibre matrices the joint distribution of overlaps and their corresponding real eigenvalues has been identified in [24] via supersymmetric techniques. See Fig. 4 for the tail decay of the CCDF of the overlaps corresponding to real and complex eigenvalues in the real Ginibre ensemble. Our numerics reproduce the t^{-1} decay for the CCDF of overlaps corresponding to real eigenvalues as shown in [24, Eq. (2.2)].*

(iv) *Recently, suboptimal versions of (14) could be established for the more general case of non-Hermitian random matrices X with i.i.d. entries with deterministic additive deformation $X + A$ in [8] and [33], resulting in bounds of order N^5 and N^3, respectively.*

(v) *Assuming Conjecture 2.4 holds we expect that near-optimal versions of (14) can be established for general perturbations $A + X$ with real Ginibre matrices X, potentially extending previous results [7] with complex Ginibre regularisation.*

Eigenvector overlaps can be used to estimate the *eigenvector condition number*[2]

$$\kappa_v(X) := \inf_{VXV^{-1}=D} \|V\|\|V^{-1}\| \le \sqrt{N}\sqrt{\sum_{i=1}^{N} \|L_i\|^2\|R_i\|^2}, \tag{16}$$

where D is the diagonal matrix of eigenvalues. Thus we immediately obtain the following corollary.

Corollary 2.7. *For real Ginibre matrices X we have*

$$\mathbf{P}\left(\kappa_v(X) \ge tN^{3/2}\sqrt{\log N}\right) \le CKt^{-2} + \frac{C_K}{N^K}, \qquad t > 0, \tag{17}$$

for any $K > 0$ with some universal constant C and K-dependent constants C_K.

We note that it is generally expected that for dense random matrices $\kappa_v(X)$ scales linearly in N, c.f. [40, Page 338]. Therefore our Corollary 2.7 is still an overestimate by a factor $N^{1/2}\sqrt{\log N}$, even though the estimate in (14) seems optimal. This is essentially due to the fact the ultimate inequality in (16) loses a factor of $\sqrt{N}$ by estimating operator norms by Frobenius norms.

2.3. **Implications for numerical analysis.** Our results on the condition number have direct implications for the running time of various algorithms from numerical analysis. The conjugate gradient (CG) algorithm is commonly used for solving positive-definite linear systems [32]. Given an iid. random matrix X and a random vector $\boldsymbol{b}$ the running time of the CG algorithm for solving the positive definite linear system $X^*X\boldsymbol{x} = \boldsymbol{b}$ has empirically been shown to be universal [20] with respect to the distribution of the input data X (for rectangular matrices with real Gaussian and Bernoulli distributed entries). Theoretical bounds for the convergence rate of the CG algorithm for solving $A\boldsymbol{x} = \boldsymbol{b}$, $A > 0$ in terms of the condition number are given by [31]

$$\|\boldsymbol{x}_k - \boldsymbol{x}\| \le 2\Big(\frac{\sqrt{\kappa(A)} - 1}{\sqrt{\kappa(A) + 1}}\Big)^k \|\boldsymbol{x}_0 - \boldsymbol{x}\|, \tag{18}$$

where $\boldsymbol{x}_k$ is the k-th iterate of the CG algorithm. In Figs. 5 and 6 we study the running time of the CG algorithm for random matrices of the form $(X-z)^*(X-z)$ with real or complex shift z, and X with real/complex Gaussian or Bernoulli entries. In agreement to our results on the condition number we find that real matrices with real shift lead to longer running times compared with real matrices with complex shift, or complex matrices with any shift.

3. Proof of Theorem 2.1

The proof of Theorem 2.1 relies on the following triple integral representation for the expected trace of the resolvent $[Y^z + E]^{-1} = [Y^z + E \cdot I]^{-1}$ at any positive number $E > 0$ (cf. [19, Eqs. (34)-(37)]):

$$\mathbf{E}\,\mathrm{Tr}[Y^z + E]^{-1} = \frac{N}{4\pi\mathrm{i}} \oint_\Gamma \mathrm{d}\xi \int_0^{+\infty} \mathrm{d}a \int_0^1 \mathrm{d}\tau \frac{\xi^2 a}{\tau^{1/2}} e^{N[f(\xi) - g(a,\tau,\eta)]} G_N(a,\tau,\xi,z), \tag{19}$$

where Γ any counter-clockwise contour around zero not crossing or encircling -1. Here the functions f and g are given explicitly as

$$f(\xi) = f(\xi, E) := E\xi + \log(1+\xi) - \log\xi - \frac{|z|^2}{1+\xi}, \tag{20}$$

$$\begin{aligned} g(a,\tau,\eta) = g(a,\tau,\eta,E) := Ea + \frac{1}{2}\log[1 + 2a + a^2\tau] - \log a - \frac{1}{2}\log\tau \\ - \frac{|z|^2(1+a) - 2\eta^2 a^2(1-\tau)}{1 + 2a + a^2\tau}, \end{aligned} \tag{21}$$

where we denoted $\eta := \Im z$. Furthermore, the function $G_N = G_N(a,\tau,\xi,\eta)$ is given by

$$\begin{aligned} G_N := \Big(N^2 \frac{p_{2,0,0}}{a^2\xi^2(\xi+1)^2\tau} - N\frac{p_{1,0,0}}{a^2\xi^2(\xi+1)\tau} + \delta N^2 \frac{p_{2,0,1}}{a\xi(\xi+1)^2\tau} - N\delta\frac{p_{1,0,1}}{a\xi(\xi+1)\tau} \\ + N^2\delta^2\frac{p_{2,0,2}}{(\xi+1)^2} + N^2\eta^2\frac{p_{2,2,0}}{a\xi(\xi+1)^3\tau} - N\eta^2\frac{p_{1,2,0}}{a\xi\tau} + N^2\eta^2\delta\frac{p_{2,2,1}}{(\xi+1)}\Big) \\ \times \Big((a^2\tau + 2a + 1)^2(\xi+1)^2\Big)^{-1}, \end{aligned} \tag{22}$$

[2]The proof of the simple inequality can be found e.g. in [7, Lemma 3.1]

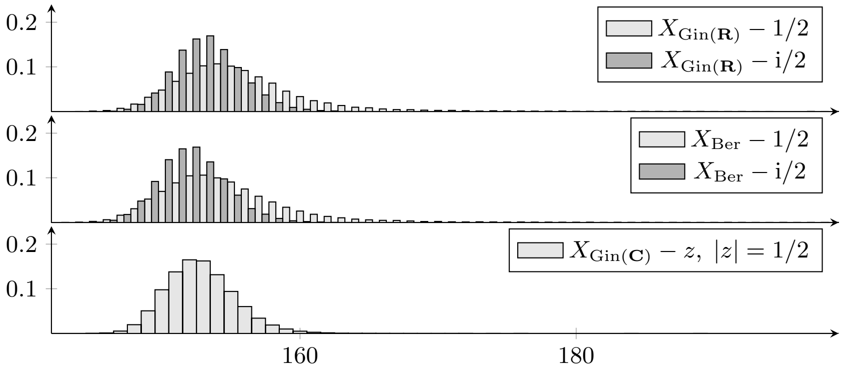

Figure 5. The figure shows the distribution of the running time of the CG algorithm for solving the linear system $(X - z)(X - z)^*\boldsymbol{x} = \boldsymbol{b}$ for different choices of z and distributions of X. The distributions have been obtained by sampling 32 000 random matrices X of size 100×100 with independent (a) real Gaussian, (b) Bernoulli, and (c) complex Gaussian entries, and random vectors $\boldsymbol{b}$ with iid uniformly distributed entries. The random matrices have been scaled such that their empirical spectrum is approximately uniformly distributed in the unit disc, and the random vectors $\boldsymbol{b}$ have been normalised by their Euclidean norm. The horizontal axis shows the number of steps in the CG algorithm to reach a tolerance of 10^{-8}. While for complex random matrices X we observe no difference between real and complex shift, we find that for real random matrices both mean and fluctuation of the running time are smaller for the complex shift $X - \mathrm{i}/2$ than for the real shift $X - 1/2$. We note that the observed runtimes seem to be significantly influenced by rounding errors as the CG algorithm in exact arithmetic is guaranteed terminate in at most $N = 100$ steps [29, 30].

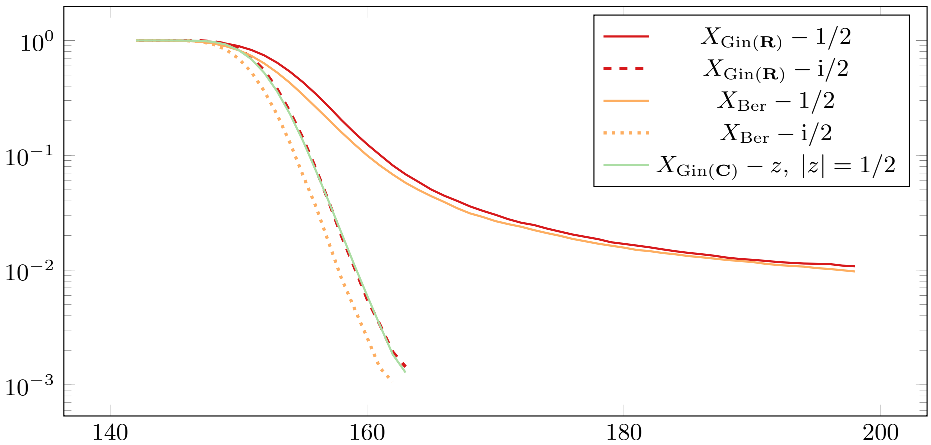

Figure 6. The figure shows the tail behaviour of the CCDF of the running times from Fig. 5. It is evident that the running times for real matrices with real shift has a significantly heavier tail than those for real matrices with complex shift, or complex matrices with any shift.

where $\delta := 1 - |z|^2$ and $p_{i,j,k} = p_{i,j,k}(a, \tau, \xi)$ are explicit polynomials in a, τ, ξ whose precise form is not particularly relevant so we defer listing them to Appendix B. The indices i, j, k in the definition of $p_{i,j,k}$ indicate the N, η and δ powers, respectively.

The formula (19) looks somewhat complicated, but it is especially well suited for an accurate asymptotic analysis in the large N regime via contour deformations and Laplace asymptotics. Note a remarkable reduction in the number of integration variables: while the expectation in the lhs. of (19) involves N^2 real integrations, the rhs. is a three-fold integral.

We derived this formula in [19, Section 3] using supersymmetric (SUSY) methods. We will not repeat here the entire derivation of (19), but we explain the main steps by giving a very short glimpse into SUSY. The interested reader can find the detailed and self-contained proof of (19) in [19, Section 3.4].

A fundamental identity on which SUSY methods rely is the following integral representation for the trace of the resolvent of any $N \times N$ Hermitian matrix H for any spectral parameter[3] $w \in \mathbf{C}$ with $\Im w > 0$:

$$\operatorname{Tr}(H-w)^{-1} = \mathrm{i}\int \langle \chi, \chi\rangle e^{-\mathrm{i}\operatorname{Tr}[H-w](ss^*+\chi\chi^*)}, \qquad \int := \int_{\mathbf{C}^N} \mathrm{d}s\, \partial_\chi. \tag{23}$$

We now explain the individual components in this formula. Here $s \in \mathbf{C}^N$ is a standard complex vector and $\int_{\mathbf{C}^N} \mathrm{d}s$ is the usual N-fold complex area integral of the entries of s, e.g. $\mathrm{d}s_1 = \pi^{-1}\,\mathrm{d}\Re s_1\,\mathrm{d}\Im s_1$. The letter χ denotes the column vector with entries $\chi_1, \dots, \chi_N$, while χ^* denotes the row vector with entries $\overline{\chi_1}, \dots, \overline{\chi_N}$, where the collection of $2N$ symbols $\{\chi_i, \overline{\chi_i} \,:\, i = 1, 2, \dots, N\}$ consists of independent Grassmannian variables, i.e. they are $2N$ non-commuting algebraic variables satisfying

$$\chi_i\chi_j = -\chi_j\chi_i, \qquad \chi_i\overline{\chi_j} = -\overline{\chi_j}\chi_i, \qquad \overline{\chi_i}\,\overline{\chi_j} = -\overline{\chi_j}\,\overline{\chi_i}$$

(the overline does not indicate any complex conjugation). These variables naturally generate a 2^{2N}-dimensional algebra over the complex scalar field. For Grassmannian vectors χ, ϕ we define the "scalar product"

$$\langle \chi, \phi\rangle := \sum_{i=1}^N \overline{\chi_i}\phi_i.$$

Beyond polynomials, one may define analytic functions of Grassmannian variables via power series, but notice that any polynomial of degree higher than $2N$ vanishes since $\chi_i^2 = 0$, hence in practice any analytic function is a polynomial. For example

$$\exp(2\chi_1 + 3\chi_2\chi_3) = 1 + 2\chi_1 + 3\chi_2\chi_3 + \frac{1}{2}(2\chi_1 + 3\chi_2\chi_3)^2 = 1 + 2\chi_1 + 3\chi_2\chi_3 + 6\chi_1\chi_2\chi_3.$$

One may also define a concept of integration (in the sense of Berezin [9]) over Grassmannian variables, which may be expressed equivalently via introducing the (formal) derivatives

$$\partial_{\chi_i}\chi_i = \partial_{\overline{\chi_i}}\overline{\chi_i} = 1, \qquad \partial_{\chi_i}1 = \partial_{\overline{\chi_i}}1 = 0, \qquad \partial_\chi := \partial_{\chi_1}\partial_{\overline{\chi_1}}\dots\partial_{\chi_N}\partial_{\overline{\chi_N}}$$

and extend them by multilinearity to all finite combinations of monomials in Grassmannians.

The integral representation in (23) consists of N complex integrals and $2N$ Grassmann integrals. To heavily reduce the number of integrals, and obtain (19), we rely on another key ingredients of SUSY calculus, the *superbosonization formula* [34, Eq. (1.13)]. We will not write it up here in full generality, but only in the form we need.

We now choose $H = Y^z$ in (23) and define the $N \times 4$ matrix $\Psi := (s, \overline{s}, \chi, \overline{\chi})$. The main advantage of the r.h.s. of (23) over its l.h.s. is that Gaussian expectation of a quadratic function (as $\operatorname{Tr} Y^z$) in the exponent can be directly computed. Taking the expectation of (23), and performing the Gaussian integration for the entries of Y^z, we are left with an integrand that turns out to be a meromorphic function depending on the variables only via $\Psi^*\Psi$, which is the 4×4 matrix consisting of all the scalar products of the four vectors $s, \overline{s}, \chi, \overline{\chi}$. This special form of the resulting function allows us to use the superbosonization formula:

$$\begin{aligned} &\int F(\Psi^*\Psi) = \int_Q \operatorname{SDet}^{N/2}(Q) F(Q), \qquad Q := \begin{pmatrix} x & \sigma \\ \tau & y \end{pmatrix} \\ &\int_Q := \frac{1}{(2\pi)^2\mathrm{i}} \int \mathrm{d}x \oint \mathrm{d}y \partial_\sigma \left(\frac{\det(y)}{\det(x)}\right)^{1/2} \det\left(1 - \frac{x^{-1}}{y}\sigma\tau\right)^{1/2}, \end{aligned} \tag{24}$$

for any meromorphic F (see [19, Appendix A] for the extension of the superbosonization formula to meromorphic functions). Here SDet denotes the *superdeterminant*, which is defined by

$$\operatorname{SDet}\begin{pmatrix} x & \sigma \\ \tau & y \end{pmatrix} := \frac{\det(x)}{\det(y - \tau x^{-1}\sigma)}.$$

[3]The imaginary part $\Im w > 0$ of the spectral parameter is a regularisation which can be removed later

The matrix Q in (24) is a 4×4 matrix written as a 2×2 block matrix of 2×2 blocks: x is non-negative Hermitian with $x_{11}=x_{22}$, y is a scalar multiple of the identity matrix. The off-diagonal block σ is a 2×2 matrix with fresh Grassmannian entries and τ is given by

$$\tau := -\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \sigma^t \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

In (24) $\int \mathrm{d}x$ denotes the integral over the Lebesgue measure on non-negative Hermitian matrices with the additional constraint $x_{11}=x_{22}$, $\oint \mathrm{d}y$ denotes the contour integral over $|y|=1$ in a counter-clockwise direction (as before, we identified the scalar multiple of the identity matrix y with the corresponding scalar) and $\partial_\sigma := \partial_{\sigma_{11}}\partial_{\sigma_{22}}\partial_{\sigma_{21}}\partial_{\sigma_{12}}$ denote the Grassmann derivatives.

Notice that the superbosonisation formula (24) entails a drastic reduction in the number of integration variables; the l.h.s. involves N complex integration, the r.h.s. has only four one-dimensional integrals after all Grassmannians are eliminated. In our concrete application one of these four integrals can be performed trivially, yielding eventually the three-fold integration in (19). For more details on SUSY calculus and for the complete proof of (19), see the proof of [19, Eqs. (34)-(37)].

Having explained our key formula (19), we now conclude this section with the proof of Theorem 2.1.

Proof of Theorem 2.1. The proof of this theorem is completely analogous to the proof of [19, Theorem 2.3] after replacing the bound in [19, Lemma 6.4] with the improved bound in Lemma 3.1 below. Furthermore, two technical estimates from [19, Lemma 5.2, Lemma 6.3] are used throughout the proof. These are straightforward but tedious bounds on explicit integrals of the form

$$\int_{\Gamma\setminus\widetilde{\Gamma}} \frac{e^{Nf(\xi)}}{\xi^b}\,\mathrm{d}\xi \qquad \text{and} \qquad \int_{N^\rho}^{\infty}\int_{N^{\rho/2}a^{-1}}^{1} \frac{e^{-Ng(a,\tau,\eta)}}{a^{\alpha-1}\tau^{\gamma-1/2}}\,\mathrm{d}\tau\,\mathrm{d}a,$$

respectively. Here Γ is any contour around zero not crossing or encircling -1,

$$\widetilde{\Gamma} := \{\xi\in\Gamma : |\xi|\le N^\rho\},$$

with some small fixed constant $\rho>0$ and $b\ge 0$, $\alpha\ge 2$, $1\le\gamma\le\alpha$ are additional parameters. We will use these technical lemmas also in the proof of Theorem 2.1, but we will not repeat them here, just refer to [19]. To make the presentation cleaner we only present the proof in the more critical case $\delta\in[0,1]$, the case $\delta\in[-10N^{-1/2},0)$ is analogous and is omitted (see e.g. [19, Section 6.2]).

First we show that the regime where at least one among a, τ and ξ is small gives a negligible contribution to the triple integral. The proof of this lemma is postponed to Appendix A.

Lemma 3.1. *Fix a small $\rho>0$. Let $\delta\in[0,1]$, $I=I_a:=[0,N^{\rho/2}a^{-1}]$, and let $c(N,\delta)$ be defined as in (6). Then, for any positive $E\lesssim c(N,\delta)$ it holds that*

$$\begin{aligned}
&\int_\Gamma \mathrm{d}\xi \int_0^{+\infty}\mathrm{d}a\int_0^1 \mathrm{d}\tau \frac{\xi^2 a}{\tau^{1/2}} e^{N[f(\xi)-g(a,\tau,\eta)]} G_N(a,\tau,\xi,z) \\
&\quad = \int_{\Gamma\setminus\widetilde{\Gamma}} \mathrm{d}\xi \int_{N^\rho}^{+\infty}\mathrm{d}a\int_{[0,1]\setminus I} \mathrm{d}\tau \frac{\xi^2 a}{\tau^{1/2}} e^{N[f(\xi)-g(a,\tau,\eta)]} G_N(a,\tau,\xi,z) + \mathcal{O}(\mathcal{E}) \\
&\mathcal{E} := N^{5/2+\rho} e^{-\frac{1}{2}N^{1-2\rho}} \left(\frac{c(N,\delta)^{-1/2} e^{-(N\eta^2)/2}}{E^{1/2}\vee [c(N,\delta)^{1/2}N^{1/2}|\eta|]} + \frac{1+|\log(NE^{2/3})|}{c(N,\delta)} \right).
\end{aligned} \tag{25}$$

Using Lemma 3.1, by (19), we conclude that

$$\begin{aligned}
&\mathbf{E}\,\mathrm{Tr}[Y^z+E]^{-1} \\
&= \frac{N}{4\pi\mathrm{i}} \int_{\Gamma\setminus\widetilde{\Gamma}} \mathrm{d}\xi \int_{N^\rho}^{+\infty}\mathrm{d}a\int_{[0,1]\setminus I} \mathrm{d}\tau \frac{\xi^2 a}{\tau^{1/2}} e^{N[f(\xi)-g(a,\tau,\eta)]} G_N(a,\tau,\xi,z) + \mathcal{O}(N\mathcal{E}),
\end{aligned} \tag{26}$$

with $\mathcal{E}$ defined in (25). In particular, we are left only with the regime where all a, τ, ξ are large (in absolute value). In this regime we can use Taylor and Laurent expansions for the functions f, g, G_N; these are listed in [19, Eqs. (75)-(77)]. Then we use the bounds from [19, Lemma 5.2] to estimate the regime $|\xi|\ge N^\rho$ and the ones from [19, Lemma 6.3] to estimate the regime $(a,\tau)\in[N^\rho,+\infty)\times([0,1]\setminus I)$, which together imply

$$\begin{aligned}\left|\frac{N}{4\pi \mathrm{i}}\int_{\Gamma\setminus\widetilde{\Gamma}}\mathrm{d}\xi\int_{N^\rho}^{+\infty}\mathrm{d}a\int_{[0,1]\setminus I}\mathrm{d}\tau\,\frac{\xi^2 a}{\tau^{1/2}}e^{N[f(\xi)-g(a,\tau,\eta)]}G_N(a,\tau,\xi,z)\right| \\ \lesssim\frac{e^{-\frac{1}{2}N\eta^2}c(N,\delta)^{-1/2}}{\sqrt{E}\vee[c(N,\delta)^{1/2}N^{1/2}|\eta|]}+c(N,\delta)^{-1}\big[1+|\log(NE^{2/3})|\big].\end{aligned} \tag{27}$$

Finally, combining (26) and (27), we readily conclude

$$|\mathbf{E}\,\mathrm{Tr}[Y^z+E]^{-1}|\lesssim\frac{e^{-\frac{1}{2}N\eta^2}c(N,\delta)^{-1/2}}{\sqrt{E}\vee[c(N,\delta)^{1/2}N^{1/2}|\eta|]}+c(N,\delta)^{-1}\big[1+|\log(NE^{2/3})|\big], \tag{28}$$

where we used the $N\mathcal{E}$ is always smaller than the r.h.s. of (28). Given (28), choosing $E=xc(N,\delta)$, the bound in (7) follows by a simple Markov inequality. □

4. Proof of Theorem 2.5

We begin with an a priori bound on the minimal eigenvalue gap

$$\Delta:=\min_{i\neq j}|\lambda_i-\lambda_j| \tag{29}$$

which follows directly from estimating the smallest two singular values of $X-z$, see e.g. [25, Theorem 3.1.1] or [35, Theorem 1.9].

Lemma 4.1. *There exist constants $c,C>0$ such that for Ginibre matrices X and for any $K\geq 100$ we have*

$$\Omega_\Delta:=\{\Delta>N^{-3K}\},\qquad \mathbf{P}(\Omega_\Delta^c)\leq\frac{C}{N^{2K-5}}+Ce^{-cN}. \tag{30}$$

Together with the singular value bound from Theorem 2.1 we obtain the following a priori bound on overlaps.

Lemma 4.2. *There exists a constant $C>0$ such that for Ginibre matrices X the event $\Omega_O:=\{\max_i O_{ii}<N^{12K}\}$ satisfies*

$$\mathbf{P}(\Omega_O^c\cap\Omega_\Delta)\leq\frac{C}{N^K}, \tag{31}$$

for any $K\geq 100$.

Proof. We claim that

$$\Omega_O^c\cap\Omega_\Delta\subset\bigcup_z^{N^{8K}}\{\sigma_1(X-z)\leq 2N^{-10K}\}, \tag{32}$$

where the union is taken over z's on an N^{-4K}-grid inside the unit disk. Indeed, for $i:=\arg\max_j O_{jj}$ there exists $z\in B(\lambda_i,N^{-4K})$ on the grid and we now show that $\sigma_1(X-z)\leq 2N^{-10K}$ for this z. We use $\sigma_1(X-z)=\|(X-z)^{-1}\|^{-1}$ and the spectral decomposition

$$\frac{1}{X-z}=\sum_i\frac{R_iL_i^*}{\lambda_i-z}.$$

Since on the event Ω_Δ for each $j\neq i$ we have $|\lambda_j-z|\geq|\lambda_i-\lambda_j|-|\lambda_i-z|\geq N^{-3K}-N^{-4K}\geq N^{-3K}/2$ it follows that

$$\begin{aligned}\frac{1}{\sigma_1(X-z)}&=\|(X-z)^{-1}\|\geq\sqrt{O_{ii}}\Big(\frac{1}{|\lambda_i-z|}-\sum_{j\neq i}\frac{1}{|\lambda_j-z|}\Big)\\&\geq\sqrt{O_{ii}}(N^{4K}-2N^{3K+1})\geq\sqrt{O_{ii}}N^{4K}/2\geq N^{10K}/2,\end{aligned} \tag{33}$$

confirming (32). Thus it follows from Theorem 2.1 and a union bound that

$$\mathbf{P}(\Omega_O^c\cap\Omega_\Delta)\leq CN^{8K-10K+2}\leq\frac{C}{N^K}, \tag{34}$$

concluding the proof of the Lemma. □

Using the a priori bounds from Lemmas 4.1–4.2 we are ready to present the proof of Theorem 2.5. The basic idea is to relate the eigenvalue overlaps to the area of the pseudo-spectrum of X, see e.g. [40, Section 52], [13, Section 3.6], [7, Lemma 3.2] or [33, Lemma 2.3] for a quantitative version.

Proof of Theorem 2.5. We introduce the event $\Omega_{\Delta,O} := \Omega_\Delta \cap \Omega_O$ and claim that on $\Omega_{\Delta,O}$ for $\epsilon := N^{-12K}$ we have

$$\sum_{\lambda_i\in\Omega} O_{ii} \le 4\frac{|\{z\in\Omega+B(0,N^{-6K})\ :\ \sigma_1(X-z)\le\epsilon\}|}{\epsilon^2}. \tag{35}$$

Indeed, first note that $B(\lambda_i, O_{ii}^{1/2}\epsilon/2)\cap B(\lambda_j, O_{jj}^{1/2}\epsilon/2)=\emptyset$ for $i\neq j$ due to

$$|\lambda_i-\lambda_j|\ge\Delta>N^{-3K}>\epsilon(O_{ii}^{1/2}/2+O_{jj}^{1/2}/2).$$

Then, for $z\in B(\lambda_i,O_{ii}^{1/2}\epsilon/2)$ we have

$$\frac{1}{\sigma_1(X-z)}=\|(X-z)^{-1}\|\ge\frac{O_{ii}^{1/2}}{|\lambda_i-z|}-\sum_{j\neq i}\frac{O_{jj}^{1/2}}{|\lambda_j-z|}\ge\frac{2}{\epsilon}-N\frac{N^{6K}}{\Delta-\epsilon O_{ii}^{1/2}/2}\ge\frac{1}{\epsilon}$$

and from this relation it follows that

$$\bigcup_{\lambda_i\in\Omega} B(\lambda_i,O_{ii}^{1/2}\epsilon/2)\subset\Big\{z\in\Omega+B(0,N^{-6K})\ :\ \sigma_1(X-z)\le\epsilon\Big\}.$$

Comparing the volumes of both sides we obtain (35).

Now from (35) and (7) we get

$$\begin{aligned}
&\mathbf{E}\left[\sum_{\lambda_i\in\Omega}O_{ii}\,\middle|\,\Omega_{\Delta,O}\right]\\
&\lesssim\int_{\Omega+B(0,N^{-6K})}\frac{\mathbf{P}[\sigma_1(X-z)\le\epsilon\,|\,\Omega_{\Delta,O}]}{\epsilon^2}\,\mathrm{d}^2z\\
&\lesssim\int_{\Omega+B(0,N^{-1/2})}\Bigg[(N^2\big|1-|z|^2\big|\vee N^{3/2})K\log N\\
&\qquad+e^{-\frac{1}{2}N(\Im z)^2}\left(\frac{N\big|1-|z|^2\big|^{1/2}\vee N^{3/4}}{\epsilon}\wedge\frac{N^{3/2}(1-|z|^2)_+\vee N}{|\Im z|}\right)\Bigg]\,\mathrm{d}^2z\\
&\lesssim K\log N\left(N^2\int_{\Omega+B(0,N^{-1/2})}\big|1-|z|^2\big|\,\mathrm{d}^2z\vee N^{3/2}|\Omega+B(0,N^{-1/2})|\right).
\end{aligned}\tag{36}$$

In the last inequality, in order to estimate the terms multiplied by $e^{-N(\Im z)^2/2}$, we performed the $\mathrm{d}\Re z$ and $\mathrm{d}\Im z$ integrations separately and split the analysis into three regimes: (i) $\big|1-|z|^2\big|\le N^{-1/2}$, (ii) $1-|z|^2>N^{-1/2}$ and $|\Im z|\ge N^{-1/2+\xi}$, for some small $\xi>0$, (iii) $1-|z|^2>N^{-1/2}$ and $|\Im z|\le N^{-1/2+\xi}$. The regime (i) is trivial since the factor $\big|1-|z|^2\big|$ can be neglected; for (ii) we used that $e^{-N(\Im z)^2/2}\le e^{-N^\xi/2}$ and so the contribution of this regime is exponentially small; finally in (iii) we used that $\big|1-|z|^2\big|\le\big|1-|\Re z|^2\big|$ and that

$$\int_{\Pi(\widetilde\Omega)}(1-y^2)\,\mathrm{d}y\lesssim\sqrt{N}\int_{\widetilde\Omega}(1-|z|^2)\,\mathrm{d}^2z,$$

where $\widetilde\Omega:=(\Omega+B(0,N^{-1/2}))\cap\{|\Im z|\le N^{-1/2+\xi}\}\cap\{1-|z|^2>N^{-1/2}\}$ and $\Pi(\widetilde\Omega)\in\mathbf{R}$ is the projection of $\widetilde\Omega$ onto the real axis. Finally, by the estimate (36) used on the event $\Omega_{\Delta,O}$ together with a simple Markov inequality combined with the probability bound on the complement $\mathbf{P}(\Omega^c_{\Delta,O})\le N^{-K}$ from Lemmas 4.1–4.2, we conclude (14). ∎

Appendix A. Proof of Lemma 3.1

The proof of this lemma is very similar to the proof of [19, Lemma 6.4], however we present a detailed proof here since we need the slightly improved bound (25) compared to [19, Eq. (92)]. More precisely, in the current paper we exploit the regularising effect of η in the very small E regime (see Eqs. (42)–(43) below); notice the additional regularisation $c(N,\delta)^{1/2}N^{1/2}|\eta|$ in the denominator of the first error term in (25). The analogous term in [19, Lemma 6.4] had a $E^{-1/2}$ singularity for small E.

Proof of Lemma 3.1. Throughout the proof we choose Γ as in [19, Eq. (48a)] to make the comparison clearer, i.e. $\Gamma = \Gamma_{z_*} := \Gamma_{1,z_*} \cup \Gamma_{2,z_*}$ with

$$\Gamma_{1,z_*} := \left\{ -\frac{2}{3} + \mathrm{i}t : \middle| \, 0 \le |t| \le \sqrt{|z_*|^2 - \frac{4}{9}} \right\}, \qquad \Gamma_{2,z_*} := \left\{ |z_*| e^{\mathrm{i}\psi} : \psi \in [-\psi_{z_*}, \psi_{z_*}] \right\},$$

where $\psi_{z_*} := \arccos[2/(3|z_*|)]$, and $z_* = z_*(E,\delta)$ is defined in [19, Eq. (45)]. We remark that $|z_*| \sim E^{-1/3} \vee \sqrt{\delta E^{-1}}$, and recall that $\widetilde{\Gamma} := \{\xi \in \Gamma : |\xi| \le N^\rho\}$.

Note that $g(a,\tau,\eta) \ge g(a,\tau,0)$ for any a,τ and that the map $\tau \mapsto g(a,\tau,0)$ is decreasing for any $a \in [0,+\infty)$ (see [19, (iii) of Lemma 6.1]). Hence, using that $g(a,\tau,\eta) \ge g(a,1,0) = f(a)$, together with [19, (ii)-(iv) of Lemma 6.1], we conclude that

$$\sup_{\xi \in \widetilde{\Gamma}} \left| e^{Nf(\xi)} \right| + \sup_{a \in [0,N^\rho]} \left| e^{-Ng(a,\tau,\eta)} \right| \lesssim e^{-Nf(N^\rho)} \lesssim e^{-\frac{1}{2}N^{1-2\rho}}, \tag{37}$$

where in the last inequality we also used that $f(N^\rho) = (\delta N^{-\rho} + (2N^{2\rho})^{-1})(1 + \mathcal{O}(N^{-\rho}))$.

The proof of (25) is divided into three regimes:

(1) $a \in [0, N^\rho]$, $\xi \in \Gamma$, $\tau \in [0,1]$ (see (38) below),
(2) $a \ge N^\rho$, $\xi \in \Gamma$, $\tau \in I$ (see (42) below),
(3) $a \ge N^\rho$, $\xi \in \widetilde{\Gamma}$, $\tau \in [0,1] \setminus I$ (see (43) below).

We will now prove that all these three regimes give an exponentially small contribution.

To estimate the regime Item (1) we consider two further sub-cases: (i) $(a,\xi) \in [0,N^\rho] \times \widetilde{\Gamma}$, (ii) $(a,\xi) \in [0,N^\rho] \times \Gamma \setminus \widetilde{\Gamma}$. In case (i), we write $Ng(a,\tau,\eta) = (N-2)g(a,\tau,\eta) + 2g(a,\tau,\eta)$. Then we use (37) to estimate $e^{-(N-2)g(a,\tau,\eta)}$ and $e^{Nf(\xi)}$ and we use $e^{-2g(a,\tau,\eta)} \lesssim a^2\tau$ for the remaining part. In this way, together with the bound $|G_N| \lesssim N^{2+5\rho}(a^2|\xi|^2\tau)^{-1}$ which follows by the explicit expression of G_N in (22), we readily get the estimate $N^{2+7\rho}e^{-N^{1-2\rho}/2}$. Here we used that the area of the domain of integration is bounded by $N^{2\rho}$. In case (ii), we first notice that by the explicit form of G_N in (22) and the fact that $|\xi| \ge N^\rho$ we have

$$\left| \int_{\Gamma \setminus \widetilde{\Gamma}} e^{Nf(\xi)} \xi^2 G_N(a,\tau,\xi,z)\, \mathrm{d}\xi \right| \lesssim N^{2+5\rho} \frac{N^{1/2} \vee (N\delta)}{a^2\tau},$$

by [19, Lemma 5.2]. Then, proceeding exactly as in case (i) to estimate the (a,τ)-integral, we conclude that this regime is bounded by $N^{2+5\rho}(N^{1/2} \vee (N\delta))e^{-N^{1-2\rho}/2}$. Combining the bounds in cases (i) and (ii) we conclude that

$$\begin{aligned} &\left| \int_\Gamma \mathrm{d}\xi \int_0^{N^\rho} \mathrm{d}a \int_0^1 \mathrm{d}\tau \frac{\xi^2 a}{\tau^{1/2}} e^{N[f(\xi) - g(a,\tau,\eta)]} G_N(a,\tau,\xi,z) \right| \\ &\quad \lesssim N^{2+7\rho}(N^{1/2} \vee (N\delta)) e^{-\frac{1}{2}N^{1-2\rho}}. \end{aligned} \tag{38}$$

Next we consider the regime Item (2). In this case we will use that $g(a,\tau,0) \ge g(a, N^{\rho/2}a^{-1}, 0)$, by [19, (iii) of Lemma 6.1] and that by explicit computations

$$g(a, N^{\rho/2}a^{-1}, 0) = Ea + \frac{\delta}{N^{\rho/2}} + \frac{1}{2N^\rho} + \mathcal{O}\left(\frac{\delta}{N^\rho} + \frac{1}{N^{3\rho/2}} \right). \tag{39}$$

Then, using (37) to estimate $e^{Nf(\xi)}$ when $\xi \in \widetilde{\Gamma}$, and [19, Lemma 5.2] for the regime $\xi \in \Gamma \setminus \widetilde{\Gamma}$, we get that

$$\left| \int_{\Gamma \setminus \widetilde{\Gamma}} e^{Nf(\xi)} \xi^2 G_N(a,\tau,\xi,z)\, \mathrm{d}\xi \right| \lesssim C(N,\delta,a,\tau), \tag{40}$$

with

$$C(a,\tau) = C(N,\delta,a,\tau) := N^{2+5\rho}(N^{1/2} \vee (N\delta)) \left(1 + \frac{1}{a^2\tau} \right).$$

We now write

$$g(a,\tau,\eta) = g(a,\tau,0) + \frac{2\eta^2 a^2 (1-\tau)}{1 + 2a + a^2\tau} \ge g(a, N^{\rho/2}a^{-1}, 0) + \frac{2\eta^2 a^2(1-\tau)}{1 + 2a + a^2\tau},$$

and so, using that $e^{-2g(a,\tau,\eta)} \lesssim a\tau$ and that

$$\frac{2\eta^2 a^2(1-\tau)}{1 + 2a + a^2\tau} \ge \frac{\eta^2 a}{2},$$

by (39) we get

$$\begin{aligned}&\left|\frac{a}{\tau^{1/2}}e^{-Ng(a,\tau,\eta)}\int_{\Gamma\setminus\widetilde{\Gamma}}e^{Nf(\xi)}\xi^2G_N(a,\tau,\xi,z)\,\mathrm{d}\xi\right|\\&\quad\lesssim C(a,\tau)a^2\tau^{1/2}e^{-(N-2)g(a,\tau,\eta)}\\&\quad\lesssim C(a,\tau)a^2\tau^{1/2}e^{-\frac{1}{2}N^{1-\rho}}e^{-(N-2)[Ea+\eta^2a/2]}.\end{aligned}\tag{41}$$

Computing the (a,τ)-integral of the r.h.s. of (41), we conclude that

$$\begin{aligned}&\left|\int_{\Gamma}\mathrm{d}\xi\int_{N^\rho}^{\infty}\mathrm{d}a\int_I\mathrm{d}\tau\frac{\xi^2a}{\tau^{1/2}}e^{N[f(\xi)-g(a,\tau,\eta)]}G_N(a,\tau,\xi,z)\right|\\&\quad\lesssim N^{3/2+7\rho}\big(N^{1/2}\vee(N\delta)\big)e^{-\frac{1}{2}N^{1-2\rho}}\frac{e^{-(N\eta^2)/2}}{E^{1/2}\vee[c(N,\delta)^{1/2}N^{1/2}|\eta|]}.\end{aligned}\tag{42}$$

Notice that in (42) we estimated the integral of r.h.s. of (41) more precisely than in [19]. This additional improvement, which is relevant only for small E, comes from using the regularising effect of $e^{-(N-2)\eta^2a/2}$.

Finally, in order to conclude the proof of this lemma, we are left with the regime Item (3). In this case, since $|\xi|\le N^\rho$, we use (37) to bound $e^{Nf(\xi)}$ and [19, Lemma 6.3] to estimate the (a,τ)-integral, with again exploiting the regularising effect of η. Hence, we conclude that

$$\begin{aligned}&\left|\int_{\widetilde{\Gamma}}\mathrm{d}\xi\int_{N^\rho}^{\infty}\mathrm{d}a\int_{[0,1]\setminus I}\mathrm{d}\tau\frac{\xi^2a}{\tau^{1/2}}e^{N[f(\xi)-g(a,\tau,\eta)]}G_N(a,\tau,\xi,z)\right|\\&\quad\lesssim N^{5/2+\rho}\big(N^{1/2}\vee(N\delta)\big)e^{-\frac{1}{2}N^{1-2\rho}}\\&\qquad\times\left(\frac{e^{-(N\eta^2)/2}}{E^{1/2}\vee[c(N,\delta)^{1/2}N^{1/2}|\eta|]}+\big[N^{1/2}\vee(N\delta)\big]\cdot\big[1+|\log(NE^{2/3})|\big]\right).\end{aligned}\tag{43}$$

Combining (38), (42), (43) we conclude (25). □

Appendix B. Explicit formulas for the real symmetric integral representation

Here we collect the explicit formulas for the polynomials of a,ξ,τ in the definition of G_N in (19).

$$\begin{aligned}p_{2,0,0}&:=a^4\tau^2+2a^3\xi\tau+4a^3\tau-a^2\xi^2\tau+4a^2\xi^2+8a^2\xi+2a^2\tau\\&\quad+4a^2+2a\xi^3+8a\xi^2+10a\xi+4a+\xi^4+4\xi^3+6\xi^2+4\xi+1,\\p_{1,0,0}&:=-a^4\xi\tau^2+a^4\tau^2-2a^3\xi^2\tau-2a^3\xi\tau+4a^3\tau-a^2\xi^3\tau-3a^2\xi^2\tau\\&\quad-2a^2\xi\tau+4a^2\xi+2a^2\tau+4a^2+2a\xi^2+6a\xi+4a+\xi^3+3\xi^2+3\xi+1,\\p_{2,2,0}&:=4(a+1)\left(a^2\tau+a\xi\tau+2a\tau+\xi^2+2\xi+1\right),\\p_{1,2,0}&:=4(a+1)\left(a^2\tau+a\xi\tau+2a\tau+\xi+1\right),\\p_{2,0,1}&:=2\big(a^3\tau^2+2a^2\xi\tau+4a^2\tau+2a\xi^2+2a\xi\tau\\&\quad+4a\xi+3a\tau+2a+\xi^3+4\xi^2+5\xi+2\big)\\p_{1,0,1}&:=2\big(a^3\tau^2+2a^2\xi\tau+4a^2\tau+a\xi^2\tau+3a\xi\tau\\&\quad+2a\xi+3a\tau+2a+\xi^2+3\xi+2\big),\\p_{2,2,1}&:=4(a+1)(a+\xi+2),\\p_{2,0,2}&:=a^2\tau+2a\xi+4a+\xi^2+4\xi+4.\end{aligned}$$

Acknowledgement

We would like to thank Folkmar Bornemann for valuable comments on a preliminary version of the present paper. We are grateful to the anonymous referees for their careful reading of our manuscript. Their suggestions to include the numerical experiments of Figs. 5 and 6 and add more details to the supersymmetric derivation significantly improved our paper.

References

[1]J. Alt, L. Erdős, and T. Krüger, *Spectral radius of random matrices with independent entries*, Prob. Math. Physics **2**, 221–280 (2019), arXiv:1907.13631.

[2]D. Armentano, C. Beltrán, P. Bürgisser, F. Cucker, and M. Shub, *A stable, polynomial-time algorithm for the eigenpair problem*, J. Eur. Math. Soc. (JEMS) **20**, 1375–1437 (2018), MR3801817.

[3]J.-M. Azaïs and M. Wschebor, *Upper and lower bounds for the tails of the distribution of the condition number of a Gaussian matrix*, SIAM J. Matrix Anal. Appl. **26**, 426–440 (2004/05), MR2124157.

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

[5]Z. D. Bai and Y. Q. Yin, *Limiting behavior of the norm of products of random matrices and two problems of Geman-Hwang*, Probab. Theory Related Fields **73**, 555–569 (1986), MR863545.

[6]J. Banks, J. Garza-Vargas, A. Kulkarni, and N. Srivastava, *Pseudospectral shattering, the sign function, and diagonalization in nearly matrix multiplication time*, preprint (2019), arXiv:1912.08805.

[7]J. Banks, A. Kulkarni, S. Mukherjee, and N. Srivastava, *Gaussian regularization of the pseudospectrum and Davies' conjecture*, preprint (2019), arXiv:1906.11819.

[8]J. Banks, J. G. Vargas, A. Kulkarni, and N. Srivastava, *Overlaps, eigenvalue gaps, and pseudospectrum under real Ginibre and absolutely continuous perturbations*, preprint (2020), arXiv:2005.08930.

[9]F. A. Berezin, *Introduction to superanalysis*, Vol. 9, Mathematical Physics and Applied Mathematics, Edited and with a foreword by A. A. Kirillov, With an appendix by V. I. Ogievetsky, Translated from the Russian by J. Niederle and R. Kotecký, Translation edited by Dimitri Leĭtes (D. Reidel Publishing Co., Dordrecht, 1987), pp. xii+424, MR914369.

[10]C. Bordenave, P. Caputo, D. Chafaï, and K. Tikhomirov, *On the spectral radius of a random matrix: an upper bound without fourth moment*, Ann. Probab. **46**, 2268–2286 (2018), MR3813992.

[11]C. Bordenave, D. Chafaï, and D. García-Zelada, *Convergence of the spectral radius of a random matrix through its characteristic polynomial*, preprint (2020), arXiv:2012.05602.

[12]A. Borodin and C. D. Sinclair, *The Ginibre ensemble of real random matrices and its scaling limits*, Comm. Math. Phys. **291**, 177–224 (2009), MR2530159.

[13]P. Bourgade and G. Dubach, *The distribution of overlaps between eigenvectors of Ginibre matrices*, Probab. Theory Related Fields **177**, 397–464 (2020), MR4095019.

[14]J. T. Chalker and B. Mehlig, *Eigenvector statistics in non-Hermitian random matrix ensembles*, Phys. Rev. Lett. **81**, 3367–3370 (1998).

[15]Z. Chen and J. J. Dongarra, *Condition numbers of Gaussian random matrices*, SIAM J. Matrix Anal. Appl. **27**, 603–620 (2005), MR2208323.

[16]G. Cipolloni, L. Erdős, and D. Schröder, *Density of small singular values of the shifted real ginibre ensemble*, preprint (2021), arXiv:2105.13720.

[17]G. Cipolloni, L. Erdős, and D. Schröder, *Edge universality for non-Hermitian random matrices*, Probab. Theory Related Fields **179**, 1–28 (2021), MR4221653.

[18]G. Cipolloni, L. Erdős, and D. Schröder, *Fluctuation around the circular law for random matrices with real entries*, Electron. J. Probab. **26**, Paper No. 24, 61 (2021), MR4235475.

[19]G. Cipolloni, L. Erdős, and D. Schröder, *Optimal lower bound on the least singular value of the shifted Ginibre ensemble*, Prob. Math. Physics **1**, 101–146 (2020), arXiv:1908.01653.

[20]P. A. Deift, G. Menon, S. Olver, and T. Trogdon, *Universality in numerical computations with random data*, Proc. Natl. Acad. Sci. USA **111**, 14973–14978 (2014), MR3276499.

[21]J. W. Demmel, *The probability that a numerical analysis problem is difficult*, Math. Comp. **50**, 449–480 (1988), MR929546.

[22]A. Edelman, *Eigenvalues and condition numbers of random matrices*, SIAM J. Matrix Anal. Appl. **9**, 543–560 (1988), MR964668.

[23]A. Edelman and B. D. Sutton, *Tails of condition number distributions*, SIAM J. Matrix Anal. Appl. **27**, 547–560 (2005), MR2179688.

[24]Y. V. Fyodorov, *On statistics of bi-orthogonal eigenvectors in real and complex Ginibre ensembles: combining partial Schur decomposition with supersymmetry*, Comm. Math. Phys. **363**, 579–603 (2018), MR3851824.

[25]S. Ge, *The eigenvalue spacing of IID random matrices and related least singular value results*, PhD thesis (UCLA) (2017).

[26]S. Geman, *The spectral radius of large random matrices*, Ann. Probab. **14**, 1318–1328 (1986), MR866352.

[27]V. L. Girko, *The circular law*, Teor. Veroyatnost. i Primenen. **29**, 669–679 (1984), MR773436.

[28]H. H. Goldstine and J. von Neumann, *Numerical inverting of matrices of high order. II*, Proc. Amer. Math. Soc. **2**, 188–202 (1951), MR41539.

[29] A. Greenbaum, *Behavior of slightly perturbed Lanczos and conjugate-gradient recurrences*, Linear Algebra Appl. **113**, 7–63 (1989), MR978581.

[30] A. Greenbaum and Z. Strakoš, *Predicting the behavior of finite precision Lanczos and conjugate gradient computations*, SIAM J. Matrix Anal. Appl. **13**, 121–137 (1992), MR1146656.

[31] W. Hackbusch, *Iterative solution of large sparse systems of equations*, Second, Vol. 95, Applied Mathematical Sciences (Springer, [Cham], 2016), pp. xxiii+509, MR3495481.

[32] M. R. Hestenes and E. Stiefel, *Methods of conjugate gradients for solving linear systems*, J. Research Nat. Bur. Standards **49**, 409–436 (1953) (1952), MR0060307.

[33] V. Jain, A. Sah, and M. Sawhney, *On the real Davies' conjecture*, preprint (2020), arXiv:2005.08908.

[34] P. Littelmann, H.-J. Sommers, and M. R. Zirnbauer, *Superbosonization of invariant random matrix ensembles*, Comm. Math. Phys. **283**, 343–395 (2008), MR2430637.

[35] K. Luh and S. O'Rourke, *Eigenvectors and controllability of non-Hermitian random matrices and directed graphs*, Electron. J. Probab. **26**, Paper No. 16, 43 (2021), MR4235467.

[36] B. Mehlig and J. T. Chalker, *Statistical properties of eigenvectors in non-Hermitian Gaussian random matrix ensembles*, J. Math. Phys. **41**, 3233–3256 (2000), MR1755501.

[37] A. Sankar, D. A. Spielman, and S.-H. Teng, *Smoothed analysis of the condition numbers and growth factors of matrices*, SIAM J. Matrix Anal. Appl. **28**, 446–476 (2006), MR2255338.

[38] D. A. Spielman and S.-H. Teng, *Smoothed analysis of algorithms: why the simplex algorithm usually takes polynomial time*, J. ACM **51**, 385–463 (2004), MR2145860.

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

[40] L. N. Trefethen and M. Embree, *Spectra and pseudospectra* (Princeton University Press, Princeton, NJ, 2005), pp. xviii+606, MR2155029.