Large deviations for the extremal eigenvalues of Ginibre ensembles

Yuanyuan Xu*
AMSS, CAS
yyxu2023@amss.ac.cn

Qiang Zeng†
AMSS, CAS
qzeng@amss.ac.cn

Abstract. We establish large deviation principles for the extremal eigenvalues of the Ginibre ensembles with good rate functions. In contrast to the typical estimates for the extremal eigenvalues, the large deviations for the real Ginibre ensemble come from the eigenvalues lying on the real line. Moreover, we also derive deviation estimates for the second leading term in the asymptotic expansion of the extremal eigenvalues. These polynomially small deviation estimates are universal for any i.i.d. matrices under a mild moment condition.

Keywords: Large deviation, extremal statistics, Ginibre ensemble
MSC number: 60B20, 60G55, 60G70
Date: April 1, 2025

1. Introduction and main results

We consider the Ginibre matrix ensemble, *i.e.*, $n\times n$ matrices $X=\big(x_{ij}\big)_{i,j=1}^n$ with independent, identically distributed (i.i.d.) Gaussian entries, where $\sqrt{n}x_{ij}$ is a standard real or complex Gaussian random variable. We use the parameter $\beta=1,2$ to distinguish the real Ginibre ensemble ($\beta=1$) and the complex Ginibre ensemble ($\beta=2$).

Let $\{\sigma_i\}_{i=1}^n$ denote the eigenvalues of the Ginibre matrix X. The corresponding empirical spectral measure (ESD), $\mu_n:=\frac{1}{n}\sum_{i=1}^n\delta_{\sigma_i}$, converges to the uniform distribution on the unit disk in the complex plane [Gin65, Ede97], which is also known as the celebrated *circular law*. Furthermore, the large deviation principle (LDP) from the circular law was established in [BAZ98] for the real Ginibre ensemble and [PH98] for the complex Ginibre ensemble, with speed n^2 and a good rate function as follows.

Theorem 1.1 ([BAZ98, PH98])**.** *The empirical spectral measures μ_n of the Ginibre ensemble ($\beta=1,2$) satisfy the LDP with speed n^2 and good rate function given by*

$$I(\mu)=\frac{\beta}{2}\Big(\int_{\mathbb{C}}|z|^2\mu(\mathrm{d}^2z)-\int_{\mathbb{C}}\int_{\mathbb{C}}\log|z-w|\mu(\mathrm{d}^2z)\mu(\mathrm{d}^2w)\Big)-\frac{3\beta}{8}. \tag{1.1}$$

Moreover, the uniform distribution σ_D on the unit disk is the unique minimizer of $I(\cdot)$ so that $I(\sigma_D)=0$.

Although the circular law indeed holds for a wide class of matrices with i.i.d. entries [Gir84, Bai97, TV10], the LDP from the circular law is still widely open beyond the Ginibre ensembles. On the other hand, the LDPs for the ESDs and extremal eigenvalues of Hermitian random matrices have been extensively studied since the seminal works [BAG97, BADG01]; see the survey paper [Gui23] for more details.

However, deviation estimates for the extremal eigenvalues of non-Hermitian matrices seem rare (with the exception of a few results discussed below). In this paper, we start the investigation in this direction and focus on the rightmost eigenvalue $\max_{i=1}^{n}\{\operatorname{Re}\sigma_i\}$ and the spectral radius $\rho_n = \max_{i=1}^{n}\{|\sigma_i|\}$. Besides its mathematical interest, the rightmost eigenvalue of a large non-Hermitian matrix is closely related to the stability properties of complex biological systems [May72]. Motivated by this, a series of works [Gem86, BY86, BCCT18, BCGZ22] has proved that both the rightmost eigenvalue and the spectral radius converge to one as the dimension goes to infinity, at a nearly optimal speed $O(n^{-1/2+\epsilon})$ [AEK21]. In particular for Ginibre ensembles, both of them have the following three-term asymptotic expansions [Rid03, RS14, Ben10, CESX22]

$$\max_{i=1}^{n}\operatorname{Re}\sigma_i \overset{\mathrm{d}}{=} 1+\sqrt{\frac{\gamma_n'}{4n}}+\frac{\mathcal{G}_n}{\sqrt{4n\gamma_n'}}, \qquad \gamma_n' := \frac{\log n - 5\log\log n - \log(2\pi^4)}{2}, \tag{1.2}$$

$$\rho_n(X) \overset{\mathrm{d}}{=} 1+\sqrt{\frac{\gamma_n}{4n}}+\frac{\mathcal{G}_n}{\sqrt{4n\gamma_n}}, \qquad \gamma_n := \log n - 2\log\log n - \log 2\pi, \tag{1.3}$$

where $\mathcal{G}_n$ is an asymptotic Gumbel random variable as $n\to\infty$, *i.e.*, $\boldsymbol{P}(\mathcal{G}_n \le t) \to \exp(-\frac{\beta}{2}e^{-t})$. We remark that [CESX22] also proves an effective estimate on the right-tail asymptotics for any $1 \ll t \ll \sqrt{\log n}$; in particular,

$$\lim_{n\to\infty}\boldsymbol{P}\Big(\max_{i=1}^{n}\operatorname{Re}\sigma_i \ge 1+\sqrt{\frac{\gamma_n'}{4n}}+\frac{t}{\sqrt{4n\gamma_n'}}\Big) \sim \frac{\beta}{2}e^{-t}, \qquad t\to\infty. \tag{1.4}$$

It is worth mentioning that the Gumbel fluctuations for general matrices with i.i.d. entries as in (1.2)–(1.3) have been proved very recently in [CEX23]. We also refer to the survey papers [BF22, BF23] on the recent progress on the Ginibre ensembles and the references therein.

It was well known [Gin65] that the eigenvalues $\{\sigma_i\}_{i=1}^{n}$ of the complex Ginibre ensemble X form a determinantal point process, whose density function is given by

$$p_n(z_1,\dots,z_n) = \frac{n^n}{\pi^n n!}e^{-n\sum_i |z_i|^2}\det\Big(\mathcal{K}_n(z_i,z_j)\Big)_{i,j=1}^{n}, \qquad \mathcal{K}_n(z,w) := \sum_{l=0}^{n-1}\frac{(nz\overline{w})^l}{l!}. \tag{1.5}$$

In particular, Kostlan observed [Kos92] that the collection of moduli of the eigenvalues of $\sqrt{2n}X$ has the same distribution as the collection of independent chi-distributed random variables with even degrees, *i.e.*,

$$\{\sqrt{2n}|\sigma_1|, \sqrt{2n}|\sigma_2|, \cdots, \sqrt{2n}|\sigma_n|\} \overset{\mathrm{d}}{=} \{\chi_2, \chi_4, \cdots, \chi_{2n}\}. \tag{1.6}$$

Heuristically, the large deviations of the spectral radius of $\sqrt{2n}X$ are mainly from that of the chi-distributed random variable with the largest degree, *i.e.*,

$$\boldsymbol{P}\Big(\max_{i=1}^{n}|\sigma_i| \ge t\Big) \approx \boldsymbol{P}\Big(\chi_{2n} \ge \sqrt{2n}t\Big) \approx e^{-n\left(t^2-2\log t-1\right)}, \qquad t\ge 1. \tag{1.7}$$

A formal large deviation estimate for the density of the spectral radius of complex Ginibre ensemble has already been obtained [LACTGMS18, CMV16] in the Coulomb gas setting with general potentials, from which an LDP follows immediately.

However, for the real Ginibre ensemble, Kostlan's observation in (1.6) unfortunately fails and the eigenvalues instead form a Pfaffian point process with an explicit correlation kernel given by [BS09, RS14]. Typically there exist $O(\sqrt{n})$ eigenvalues lying on the real line [EKS94], and the asymptotic statistics of the real eigenvalues are very different from those of the complex eigenvalues [BS09]. In particular, the limiting distribution of the largest real eigenvalue has Gaussian right-tail asymptotics [PTZ17, FN07], *i.e.*,

$$\lim_{n\to\infty}\boldsymbol{P}\Big(\max_{\sigma_i\in\mathbb{R}}\sigma_i \ge 1+\frac{t}{\sqrt{n}}\Big) \sim \frac{e^{-t^2}}{4\sqrt{\pi}t}, \qquad t\to\infty, \tag{1.8}$$

which is quite different than that of the complex eigenvalues as shown in (1.4).

The first main result of this paper is the following LDPs for the spectral radius and rightmost eigenvalue of the Ginibre ensembles.

Theorem 1.2 (LDPs for extremal eigenvalues)**.** *For the real ($\beta=1$) and complex ($\beta=2$) Ginibre ensemble, $\rho_n=\max_{i=1}^n\{|\sigma_i|\}$ satisfies the LDP with speed n and good rate function*

$$I_\beta(t)=\begin{cases}\frac{\beta}{2}\big(t^2-2\log t-1\big), & t\geq 1,\\ +\infty, & t<1.\end{cases}\tag{1.9}$$

The same LDP also holds true for the rightmost eigenvalue $\max_{i=1}^n \operatorname{Re}\sigma_i$.

The above LDP for the spectral radius in the complex case ($\beta=2$) agrees with [LACTGMS18, Eq. (11)]. In the real case ($\beta=1$), the LDP with the rate function I_2 also holds for the largest complex eigenvalue in modulus. However, in contrast to the typical estimates stated in (1.2)-(1.3) and (1.8), the largest real eigenvalue dominates the complex eigenvalues in the large deviation regime, which yields the rate function I_1.

Moreover, we derive the following moderate deviation estimates between the large deviations in Theorem 1.2 and the typical estimates in (1.2)–(1.3) and (1.8) for the Ginibre ensembles.

Theorem 1.3 (Moderate deviations)**.** *Consider the real ($\beta=1$) and complex ($\beta=2$) Ginibre ensemble. For any $\sqrt{\frac{\gamma_n}{4n}}\leq d_n\ll 1$ with γ_n in (1.3), there exist constants $C_1,C_2>0$ such that*

$$\boldsymbol{P}\Big(\max_{i=1}^n|\sigma_i|\geq 1+d_n\Big)\leq \frac{C_1}{\sqrt{n}(d_n)^2}e^{-2n(d_n)^2(1-O(d_n))}+\frac{C_2}{\sqrt{n}d_n}e^{-n(d_n)^2(1-O(d_n))}\mathbb{1}_{\beta=1}.\tag{1.10}$$

For any $\sqrt{\frac{\gamma_n'}{4n}}\leq d_n\ll 1$ with γ_n' in (1.2), there exist constants $C_1',C_2'>0$ such that

$$\boldsymbol{P}\Big(\max_{i=1}^n\operatorname{Re}\sigma_i\geq 1+d_n\Big)\leq \frac{C_1'}{n(d_n)^{5/2}}e^{-2n(d_n)^2(1-O(d_n))}+\frac{C_2'}{\sqrt{n}d_n}e^{-n(d_n)^2(1-O(d_n))}\mathbb{1}_{\beta=1}.\tag{1.11}$$

In particular, in the real case ($\beta=1$), for any $n^{-1/2}\ll d_n\ll 1$, there exists $C>0$ such that

$$\boldsymbol{P}\Big(\max_{\sigma_i\in\mathbb{R}}\sigma_i\geq 1+d_n\Big)\leq\frac{C}{\sqrt{n}d_n}e^{-n(d_n)^2(1-O(d_n))},\tag{1.12}$$

and the same also holds true for $\max_{\sigma_i\in\mathbb{R}}|\sigma_i|$. Finally, for any $\sqrt{\log n/n}\ll d_n\ll 1$, we have

$$\lim_{n\to\infty}\frac{1}{n(d_n)^2}\log\boldsymbol{P}\Big(\max_{i=1}^n|\sigma_i|\geq 1+td_n\Big)=-\beta t^2,\qquad t>0.\tag{1.13}$$

The same statement as in (1.13) also holds true for $\max_{i=1}^n\operatorname{Re}\sigma_i$.

We remark that the deviation estimates in (1.12) for the real eigenvalues ($\beta=1$) are consistent with the Gaussian right-tail asymptotics of the limiting distribution as stated in (1.8).

Finally, we obtain the following deviation estimates for the second leading term in the asymptotic expansions (1.2)–(1.3). Such polynomially small deviation estimates are indeed universal for any i.i.d matrices beyond Ginibre ensembles. More precisely, we consider any matrix X with i.i.d. entries $x_{ab}\overset{\mathrm{d}}{=}n^{-1/2}\chi$, where χ satisfies $\boldsymbol{E}\chi=0$, $\boldsymbol{E}|\chi|^2=1$, and additionally $\boldsymbol{E}\chi^2=0$ in the complex case. Moreover, we also assume the following finite moment condition, *i.e.*, there exists constants $M_k>0$ such that

$$\boldsymbol{E}|\chi|^k\leq M_k,\qquad k\in\mathbb{N}.\tag{1.14}$$

Theorem 1.4 (Small deviations)**.** *Given any real ($\beta=1$) or complex ($\beta=2$) i.i.d. matrix defined as above, for any $s,t>1$, there exists constants $C,C',C_s,C_t>0$ such that*

$$\boldsymbol{P}\Big(\max_{i=1}^n|\sigma_i|\geq 1+s\sqrt{\frac{\gamma_n}{4n}}\Big)\leq C(\log n)^{C_s}\Big(n^{-\frac{s^2-1}{2}}+n^{-\frac{s^2}{4}}\mathbb{1}_{\beta=1}\Big),\tag{1.15}$$

$$\boldsymbol{P}\Big(\max_{i=1}^{n}\mathrm{Re}\,\sigma_i\geq 1+t\sqrt{\frac{\gamma_n'}{4n}}\Big)\leq C'(\log n)^{C_t}\Big(n^{-\frac{t^2-1}{4}}+n^{-\frac{t^2}{8}}\mathbb{1}_{\beta=1}\Big),\tag{1.16}$$

for sufficiently large n, with γ_n, γ_n' given in (1.2)-(1.3). In particular, for any $1\ll s_n\ll\log n$,

$$\boldsymbol{P}\Big(\max_{i=1}^{n}|\sigma_i|\geq 1+\sqrt{\frac{\gamma_n}{4n}}+\frac{s_n}{\sqrt{4n\gamma_n}}\Big)\leq C''e^{-s_n},\tag{1.17}$$

and the same estimate also holds true for $\max_{i=1}^{n}\mathrm{Re}\,\sigma_i$ with γ_n' in (1.2).

Comparing (1.15) with (1.16) at the same level in the complex case (*i.e.*, $\beta=2$ and set $t=\sqrt{2s}$ with $s>1$), the small probability of the rightmost eigenvalue gains an additional $n^{-1/4}$ than that of the spectral radius, mainly due to the volume effect of the complex eigenvalues taken into consideration. However in the real case ($\beta=1$), the difference between them almost vanishes beyond the level $1+\sqrt{\frac{\log n}{2n}}$, since the corresponding small deviations come from the real eigenvalues as estimated in (1.12). Moreover, the estimate in (1.17) also agrees with the Gumbel right-tail asymptotics as stated in (1.4).

Conventions and notations: For n-dependent positive quantities f_n, g_n we use the notation $f_n\ll g_n$ to denote that $\lim_{n\to\infty}(f_n/g_n)=0$. 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 $c,C>0$ independent from n. Throughout the paper $c,C>0$ (resp. $c_t,C_t>0$) denote small and large absolute constants (resp. constants depending only on t), respectively, which may change from line to line. In the following we will use the notations $\boldsymbol{P}^{\mathrm{Gin}(\mathbb{C})}$ (or $\boldsymbol{E}^{\mathrm{Gin}(\mathbb{C})}$) and $\boldsymbol{P}^{\mathrm{Gin}(\mathbb{R})}$ (or $\boldsymbol{E}^{\mathrm{Gin}(\mathbb{R})}$) to denote the probability (or expectation) for the complex and real Ginibre matrix, respectively.

2. Proof of Theorem 1.2

2.1. Proof for the complex Ginibre Ensemble.

To prove Theorem 1.2 in the complex case, it suffices to prove the following proposition.

Proposition 2.1. *For any $t<1$, we have*

$$\limsup_{n\to\infty}\frac{1}{n}\log\boldsymbol{P}^{\mathrm{Gin}(\mathbb{C})}(\rho_n\leq t)=-\infty\tag{2.1}$$

and for any $t\geq 1$,

$$\lim_{n\to\infty}\frac{1}{n}\log\boldsymbol{P}^{\mathrm{Gin}(\mathbb{C})}(\rho_n\geq t)=-(t^2-2\log t-1).\tag{2.2}$$

The above also holds true for $\max_{i=1}^{n}\mathrm{Re}\,\sigma_i$.

Recall from [Meh04] that the eigenvalues of the complex Ginibre ensemble in (1.5) form a determinantal point process with k-point correlation function given by

$$p_n^{(k)}(z_1,\ldots,z_k)=\det\Big(K_n(z_i,z_j)\Big)_{i,j=1}^{k},\qquad K_n(z,w):=\frac{n}{\pi}e^{-n(|z|^2+|w|^2)/2}\mathfrak{e}_{n-1}(nz\overline{w}),\tag{2.3}$$

where the polynomial $\mathfrak{e}_{n-1}$ is defined by

$$\mathfrak{e}_{n-1}(z):=\sum_{k=0}^{n-1}\frac{z^k}{k!}=e^z\frac{\Gamma(n,z)}{\Gamma(n)},\qquad\Gamma(n,z):=\int_z^{\infty}t^{n-1}e^{-t}\mathrm{d}t,\tag{2.4}$$

with the integration contour going from $z\in\mathbb{C}$ to the real infinity. These polynomials have the following precise asymptotic estimates.

Lemma 2.2 (Proposition 3 in [BG07])**.** *Uniformly in $t>0$, we have*

$$e^{-nt}\mathfrak{e}_n(nt) = \mathbb{1}_{0\le t<1} + \frac{1}{\sqrt{2}}\frac{\mu(t)t}{t-1}\mathrm{erfc}\big(\sqrt{n}\mu(t)\big)\left(1+O\Big(\frac{1}{\sqrt{n}}\Big)\right), \tag{2.5}$$

with $\mu(t)=|t-1-\log t|^{1/2}$ and

$$\mathrm{erfc}(t) := \frac{2}{\sqrt{\pi}}\int_t^\infty e^{-s^2}\mathrm{d}s = \frac{e^{-t^2}}{\sqrt{\pi}t}\left(1-\frac{1}{2t^2}+O(|t|^{-4})\right). \tag{2.6}$$

Proof of Proposition 2.1. The first estimate (2.1) follows directly from the LDP for the empirical spectral measures in Theorem 1.1. Indeed, we may take a bounded Lipschitz function $f:\mathbb{C}\to\mathbb{R}$ supported in $\{z\in\mathbb{C}:\mathrm{Re}\, z\ge t\}$ such that $f(z)=1$ for $\mathrm{Re}\, z>\frac{t+1}{2}$ where $0<t<1$. By the LDP for the empirical spectral measures μ_n,

$$\boldsymbol{P}(\rho_n\le t)\le \boldsymbol{P}\Big(\max_{i=1}^{n}\mathrm{Re}\,\sigma_i\le t\Big)\le \boldsymbol{P}\Big(\Big|\int f\mathrm{d}\mu_n-\frac{1}{\pi}\int f(z)\mathrm{d}^2z\Big|>\delta\Big)\le e^{-cn^2}$$

for some $\delta>0$, where $c>0$ depends on f and δ.

We next prove the second estimate (2.2) for $t\ge 1$. For $t=1$, it follows directly from (1.3) that $\boldsymbol{P}(\rho_n\ge 1)\ge c$ for some constant $c>0$. We hence focus on proving (2.2) for $t>1$. To obtain the upper bound, we write

$$\boldsymbol{P}(\rho_n\ge t) =\boldsymbol{P}\big(\#\{1\le i\le n:|\sigma_i|\ge t\}\ge 1\big)\le \boldsymbol{E}\big[\#\{1\le i\le n:|\sigma_i|\ge t\}\big]. \tag{2.7}$$

Note that the one point correlation function for the eigenvalue process is given by $K_n(z,z)$ in (2.3). Hence for $t>1$, we have

$$\begin{aligned}\boldsymbol{E}\big[\#\{1\le i\le n:|\sigma_i|\ge t\}\big] &= \int_{|z|\ge t}K_n(z,z)\mathrm{d}^2z=\int_{|z|\ge t}\frac{n}{\pi}e^{-n|z|^2}\mathfrak{e}_{n-1}(n|z|^2)\mathrm{d}^2z\\ &=\int_{|z|\ge t}\frac{n}{\pi\sqrt{2\pi n}}\frac{1}{|z|^2-1}e^{-n\left(|z|^2-2\log|z|-1\right)}\mathrm{d}^2z\Big(1+O(n^{-1/2})\Big)\\ &=\sqrt{\frac{2}{\pi}}\int_t^\infty\frac{r}{r^2-1}e^{-n(r^2-2\log r-1)+\frac12\log n}\mathrm{d}r\Big(1+O(n^{-1/2})\Big)\\ &\le C_te^{-n\left(t^2-2\log t-1\right)-\frac12\log n},\end{aligned} \tag{2.8}$$

where we also used that from Lemma 2.2

$$\mathfrak{e}_{n-1}(nz)=\frac{(ez)^n}{\sqrt{2\pi n}(z-1)}\Big(1+O\Big(\frac{1}{\sqrt{n}}\Big)\Big),$$

and the following estimate from integration by parts

$$\frac{t^2e^{-n(t^2-2\log t)}}{2n(t^2-1)^2}\Big[1-\frac{t^2+1}{n(t^2-1)^2}\Big]\le\int_t^\infty\frac{r}{r^2-1}e^{-n(r^2-2\log r)}\mathrm{d}r\le\frac{t^2e^{-n(t^2-2\log t)}}{2n(t^2-1)^2}, \tag{2.9}$$

for $t>1$. One could obtain a similar lower bound from (2.9) as in (2.8), *i.e.,*

$$\boldsymbol{E}\big[\#\{1\le i\le n:|\sigma_i|\ge t\}\big]\ge C_t'e^{-n\left(t^2-2\log t-1\right)-\frac12\log n}. \tag{2.10}$$

Combining (2.8) with (2.7), we obtain the desired upper bound

$$\boldsymbol{P}(\rho_n\ge t)\le C_te^{-n\left(t^2-2\log t-1\right)-\frac12\log n}. \tag{2.11}$$

To obtain the matching lower bound, we have

$$\begin{aligned}\boldsymbol{P}(\rho_n\ge t)=\boldsymbol{P}\big(\#\{1\le i\le n:|\sigma_i|\ge t\}\ge 1\big)&\ge\frac{1}{n}\boldsymbol{E}\big[\#\{1\le i\le n:|\sigma_i|\ge t\}\big]\\ &\ge C_t'e^{-n\left(t^2-2\log t-1\right)-\frac32\log n},\end{aligned} \tag{2.12}$$

where we also used (2.10). This together with (2.11) proves (2.2).

To end the proof, we consider the rightmost eigenvalue. Similar to (2.7)–(2.8), we have

$$\begin{aligned}\boldsymbol{P}\big(\max_{i=1}^{n}\operatorname{Re}\sigma_i \geq t\big) =&\boldsymbol{E}\big[\#\{1\leq i\leq n: \operatorname{Re}\sigma_i\geq t\}\big] = \int_{\mathbb{R}}\int_{x\geq t} K_n(x+\mathrm{i}y, x+\mathrm{i}y)\mathrm{d}x\mathrm{d}y\\ &\leq \int_{|z|\geq t} K_n(z,z)\mathrm{d}^2 z \leq C_t e^{-n\left(t^2-2\log t-1\right)-\frac{1}{2}\log n},\end{aligned} \tag{2.13}$$

for $t>1$. One could obtain a similar lower bound as in (2.12), *i.e.*,

$$\begin{aligned}\boldsymbol{P}\big(\max_{i=1}^{n}\operatorname{Re}\sigma_i \geq t\big) \geq& \frac{1}{n}\boldsymbol{E}\big[\#\{1\leq i\leq n: \operatorname{Re}\sigma_i\geq t\}\big]\\ &= \frac{1}{n}\int_{\mathbb{R}}\int_{x\geq t}\frac{n}{\pi\sqrt{2\pi n}}\frac{1}{x^2+y^2-1}e^{-n\left(x^2+y^2-\log(x^2+y^2)-1\right)}\mathrm{d}x\mathrm{d}y\Big(1+O(n^{-1/2})\Big)\\ &\geq C\frac{1}{\sqrt{n}}\int_{x\geq t}\left(\int_{\mathbb{R}}\frac{1}{x^2+y^2-1}e^{-ny^2}\mathrm{d}y\right)e^{-n\left(x^2-\log(x^2)-1\right)}\mathrm{d}x\\ &\geq C\frac{1}{n}\int_t^{\infty}\frac{1}{x^2-1}e^{-n\left(x^2-\log(x^2)-1\right)}\mathrm{d}x\\ &\geq C_t' e^{-n\left(t^2-2\log t-1\right)-2\log n},\end{aligned} \tag{2.14}$$

where we used that $\log(x^2+y^2)\geq\log(x^2)$, and for sufficiently large n,

$$\int_0^{\infty}\frac{1}{x^2+y^2-1}e^{-ny^2}\mathrm{d}y = \frac{2n}{\sqrt{x^2-1}}\int_0^{\infty}\arctan\Big(\frac{y}{\sqrt{x^2-1}}\Big)ye^{-ny^2}\mathrm{d}y \geq \frac{C}{\sqrt{n}(x^2-1)},$$

together with the following estimates from integration by parts

$$\frac{te^{-n(t^2-2\log t)}}{2n(t^2-1)^2}\Big[1-\frac{3t^2+1}{2n(t^2-1)^2}\Big] \leq \int_t^{\infty}\frac{1}{r^2-1}e^{-n(r^2-2\log r)}\mathrm{d}r \leq \frac{te^{-n(t^2-2\log t)}}{2n(t^2-1)^2}, \tag{2.15}$$

for $t>1$. We hence finish the proof of Proposition 2.1. □

2.2. **Proof for the real Ginibre Ensemble.** Given a real Ginibre matrix of dimension n with $n=L+2M$, where L is the number of real eigenvalues, and M is the number of conjugated pairs of complex eigenvalues. The explicit formula for the joint density distribution of L real eigenvalues and M complex eigenvalues in the upper half plane was first introduced in [LS91, Ede97]. As the analog of (2.3), these eigenvalues indeed form a Pfaffian point process with the so-called (l,m)-correlation functions [BS09, Theorem 8] (or see [RS14, Propositions 2.1–2.2]). In particular, the one point correlation functions for the complex and real eigenvalue process are given by [Ede97, EKS94], respectively, *i.e.*,

$$S_n^{\mathbb{C},\mathbb{C}}(z,z) := \frac{\sqrt{2}n^{3/2}|\operatorname{Im} z|}{\sqrt{\pi}}e^{2n(\operatorname{Im} z)^2}\operatorname{erfc}(\sqrt{2n}|\operatorname{Im} z|)e^{-n|z|^2}\mathfrak{e}_{n-2}(n|z|^2), \tag{2.16}$$

$$S_n^{\mathbb{R},\mathbb{R}}(x,x) := \sqrt{\frac{n}{2\pi}}e^{-nx^2}\mathfrak{e}_{n-2}(nx^2) + \frac{n^{\frac{n}{2}}|x|^{n-1}e^{-\frac{n}{2}x^2}}{2^{\frac{n}{2}}\Gamma\big(\frac{n}{2}\big)\Gamma\big(\frac{n-1}{2}\big)}\int_0^{\frac{nx^2}{2}}u^{\frac{n-3}{2}}e^{-u}\mathrm{d}u. \tag{2.17}$$

We remark that these are also part of the Pfaffian kernels introduced in [BS09].

Proposition 2.3. *For any $t>1$, we have*

$$\lim_{n\to\infty}\frac{1}{n}\log\boldsymbol{P}^{\mathrm{Gin}(\mathbb{R})}\Big(\max_{\sigma_i\in\mathbb{C}\setminus\mathbb{R}}|\sigma_i|\geq t\Big) = -\big(t^2-2\log t-1\big), \tag{2.18}$$

$$\lim_{n\to\infty}\frac{1}{n}\log\boldsymbol{P}^{\mathrm{Gin}(\mathbb{R})}\Big(\max_{\sigma_i\in\mathbb{R}}\sigma_i\geq t\Big) = -\frac{t^2-2\log t-1}{2}. \tag{2.19}$$

The same statement as in (2.18) also holds true for $\max_{\sigma_i\in\mathbb{C}\setminus\mathbb{R}}\operatorname{Re}\sigma_i$, and (2.19) also holds true for $\max_{\sigma_i\in\mathbb{R}}|\sigma_i|$.

Proof of Proposition 2.3. The proof is similar to that of Proposition 2.1. We first consider the complex eigenvalues in (2.18). From Lemma 2.2 note that

$$\begin{aligned} S_n^{\mathbb{C},\mathbb{C}}(z,z) &= \sqrt{2n\pi}e^{2n(\operatorname{Im} z)^2}|\operatorname{Im} z|\operatorname{erfc}(\sqrt{2n}|\operatorname{Im} z|)K_{n-1}(z,z)e^{1-|z|^2}\Big(1+O(n^{-1/2})\Big) \\ &= e^{1-|z|^2}K_{n-1}(z,z)\left(1+O\left(\min\left\{1,\frac{1}{n(\operatorname{Im} z)^2}\right\}+\frac{1}{\sqrt{n}}\right)\right), \end{aligned} \tag{2.20}$$

with K_n given by (2.3) in the complex case, where we also used the asymptotic estimate $\operatorname{erfc}(x)=e^{-x^2}/(\sqrt{\pi}x)(1+O(x^{-2}))$ and the bound $\operatorname{erfc}(x)\le e^{-x^2}/(\sqrt{\pi}x)$. This, together with the arguments as in (2.7)–(2.14), implies (2.18) and its analog for $\max_{\sigma_i\in\mathbb{C}\setminus\mathbb{R}}\operatorname{Re}\sigma_i$.

We next focus on the real eigenvalues in (2.19). Since the one point correlation function for the real eigenvalue process in (2.17) is symmetric, the statement for $\max_{\sigma_i\in\mathbb{R}}|\sigma_i|$ follows immediately from the one-sided estimate in (2.19). Note that the first term on the right side of (2.17) is given by

$$\sqrt{\frac{n}{2\pi}}e^{-nx^2}\mathfrak{e}_{n-2}(nx^2)=\sqrt{\frac{\pi}{2n}}K_{n-1}(x,x)e^{1-x^2}\Big(1+O(n^{-1/2})\Big). \tag{2.21}$$

Using that $\Gamma(z)\Gamma\big(z+1/2\big)=2^{1-2z}\sqrt{\pi}\Gamma(2z)$, we write the second term on the right side of (2.17)

$$\frac{n^{\frac{n}{2}}x^{n-1}e^{-\frac{n}{2}x^2}}{2^{\frac{n}{2}}\Gamma\big(\frac{n}{2}\big)\Gamma\big(\frac{n-1}{2}\big)}\int_0^{\frac{nx^2}{2}}u^{\frac{n-3}{2}}e^{-u}\mathrm{d}u=\frac{2^{\frac{n}{2}-2}n^{\frac{n}{2}}x^{n-1}e^{-\frac{n}{2}x^2}}{\sqrt{\pi}(n-2)!}\Big(\int_0^\infty-\int_{\frac{n}{2}x^2}^\infty\Big)u^{\frac{n-3}{2}}e^{-u}\mathrm{d}u. \tag{2.22}$$

Note that the function $u^{\frac{n-3}{2}}e^{-\frac{n-3}{n}u}$ is decreasing in the regime $u\ge nx^2/2$ with $x\ge 1$. Thus the last integral over $[nx^2/2,\infty)$ in (2.22) is bounded from above by

$$\int_{\frac{n}{2}x^2}^\infty u^{\frac{n-3}{2}}e^{-u}\mathrm{d}u\le\Big(\frac{nx^2}{2}\Big)^{\frac{n-3}{2}}e^{-\frac{(n-3)x^2}{2}}\int_{\frac{n}{2}x^2}^\infty e^{-\frac{3}{n}u}\mathrm{d}u\le Cn\Big(\frac{nx^2}{2}\Big)^{\frac{n-3}{2}}e^{-\frac{(n-3)x^2}{2}},$$

which is much smaller than the whole integral over $[0,\infty)$. Thus by Stirling's formula, we have

$$\begin{aligned} \frac{n^{\frac{n}{2}}x^{n-1}e^{-\frac{n}{2}x^2}}{2^{\frac{n}{2}}\Gamma\big(\frac{n}{2}\big)\Gamma\big(\frac{n-1}{2}\big)}\int_0^{\frac{nx^2}{2}}u^{\frac{n-3}{2}}e^{-u}\mathrm{d}u &= \frac{2^{\frac{n}{2}-2}n^{\frac{n}{2}}x^{n-1}e^{-\frac{n}{2}x^2}}{\sqrt{\pi}(n-2)!}\Gamma\big(\frac{n-1}{2}\big)(1+o(1)) \\ &\sim\frac{1}{x}e^{-\frac{n}{2}\big(x^2-2\log x-1\big)+\frac{1}{2}\log n}. \end{aligned} \tag{2.23}$$

Combining (2.21), (2.23) with (2.17), we have

$$S_n^{\mathbb{R},\mathbb{R}}(x,x)\sim\sqrt{\frac{\pi}{2n}}K_{n-1}(x,x)e^{1-x^2}+\frac{1}{x}e^{-\frac{n}{2}\big(x^2-2\log x-1\big)+\frac{1}{2}\log n}. \tag{2.24}$$

Therefore, we obtain an upper bound similarly to (2.8), *i.e.*,

$$\begin{aligned} \boldsymbol{E}\big[\#\{\sigma_i\in\mathbb{R}:\sigma_i\ge t\}\big] &= \int_{x>t}S_n^{\mathbb{R},\mathbb{R}}(x,x)\mathrm{d}x \\ &\le C\int_t^\infty\frac{1}{x^2-1}e^{-n(x^2-2\log x-1)}\mathrm{d}x+C\sqrt{n}\int_t^\infty\frac{1}{x}e^{-\frac{n}{2}\big(x^2-2\log x-1\big)}\mathrm{d}x \\ &\le C_te^{-\frac{n}{2}\big(t^2-2\log t-1\big)-\frac{1}{2}\log n}, \end{aligned} \tag{2.25}$$

where we also used (2.15) together with the following estimate

$$\frac{e^{-\frac{n}{2}(t^2-2\log t)}}{n(t^2-1)}\Big(1-\frac{2t^2}{n(t^2-1)^2}\Big)\le\int_t^\infty\frac{1}{x}e^{-\frac{n}{2}\big(x^2-2\log x\big)}\mathrm{d}x\le\frac{e^{-\frac{n}{2}(t^2-2\log t)}}{n(t^2-1)}. \tag{2.26}$$

From here we also find a lower bound in the form of (2.25). Using that

$$\frac{1}{n}\boldsymbol{E}\big[\#\{\sigma_i\in\mathbb{R}:\sigma_i\ge t\}\big]\le\boldsymbol{P}\Big(\max_{\sigma_i\in\mathbb{R}}\sigma_i\ge t\Big)\le\boldsymbol{E}\big[\#\{\sigma_i\in\mathbb{R}:\sigma_i\ge t\}\big],$$

we obtain the desired upper and lower bound to prove (2.19) and hence finish the proof of Proposition 2.3. □

Proof of Theorem 1.2 for $\beta = 1$. As in the complex case, from the LDP for the empirical spectral measures in the real case we know

$$\limsup_{n\to\infty} \boldsymbol{P}(\rho_n \le t) \le \limsup_{n\to\infty} \boldsymbol{P}\Big(\max_{i=1}^{n} \operatorname{Re}\sigma_i \le t\Big) = -\infty, \qquad t < 1.$$

The asserted LDP for ρ_n follows from Proposition 2.3 since

$$\boldsymbol{P}\Big(\max_{\sigma_i\in\mathbb{R}} |\sigma_i| \ge t\Big) \le \boldsymbol{P}(\rho_n \ge t) \le \boldsymbol{P}\Big(\max_{\sigma_i\in\mathbb{R}} |\sigma_i| \ge t\Big) + \boldsymbol{P}\Big(\max_{\sigma_i\in\mathbb{C}\setminus\mathbb{R}} |\sigma_i| \ge t\Big).$$

For the rightmost eigenvalue, we observe that

$$\boldsymbol{P}\Big(\max_{\sigma_i\in\mathbb{R}} \sigma_i \ge t\Big) \le \boldsymbol{P}\Big(\max_{i=1}^{n} \operatorname{Re}\sigma_i \ge t\Big) \le \boldsymbol{P}(\rho_n \ge t).$$

The proof is completed using Proposition 2.3 again. □

3. Proofs of Theorems 1.3 and 1.4

3.1. Proof of Theorem 1.3.

We first consider the complex Ginibre ensemble. Similar to the arguments used in (2.7)–(2.11), for any $\sqrt{\frac{\gamma_n}{4n}} \le d_n \ll 1$, we have

$$\begin{aligned}
\boldsymbol{P}^{\mathrm{Gin}(\mathbb{C})}\Big(\max_{i=1}^{n} |\sigma_i| \ge 1 + d_n\Big) &\le \boldsymbol{E}^{\mathrm{Gin}(\mathbb{C})}\big[\#\{1 \le i \le n : |\sigma_i| \ge 1 + d_n\}\big] \\
&= \int_{|z|\ge 1+d_n} K_n(z,z)\mathrm{d}^2 z \\
&= \int_{|z|\ge 1+d_n} \frac{n}{\pi\sqrt{2\pi n}} \frac{1}{|z|^2 - 1} e^{-n\left(|z|^2 - 2\log|z| - 1\right)} \mathrm{d}^2 z \Big(1 + O(n^{-1/2})\Big) \\
&\lesssim \frac{1}{\sqrt{n}(d_n)^2} e^{-n[(d_n)^2 + 2d_n - 2\log(1+d_n)]}
\end{aligned} \tag{3.1}$$

where we used the upper estimate of (2.9) in the last step. Using a simple Taylor expansion $\log(1+x) = x - \frac{x^2}{2} + O(x^3)$, we obtain that

$$\boldsymbol{P}^{\mathrm{Gin}(\mathbb{C})}\Big(\max_{i=1}^{n} |\sigma_i| \ge 1 + d_n\Big) \lesssim \frac{1}{\sqrt{n}(d_n)^2} e^{-2n(d_n)^2(1 - O(d_n))}. \tag{3.2}$$

This proves (1.10) for the complex case ($\beta = 2$).

The rightmost eigenvalue (1.11) can be estimated similarly. For any $\sqrt{\frac{\gamma_n'}{4n}} \le d_n \ll 1$, we have

$$\begin{aligned}
\boldsymbol{P}^{\mathrm{Gin}(\mathbb{C})}\big(\max_{i=1}^{n} \operatorname{Re}\sigma_i \ge 1 + d_n\big) \le& \int_{\mathbb{R}} \int_{x\ge 1+d_n} K_n(x+\mathrm{i}y, x+\mathrm{i}y)\mathrm{d}x\mathrm{d}y \\
=& \int_{\mathbb{R}} \int_{x\ge 1+d_n} \frac{n}{\pi\sqrt{2\pi n}} \frac{1}{x^2+y^2-1} e^{-n\left(x^2+y^2-\log(x^2+y^2)-1\right)} \mathrm{d}x\mathrm{d}y \Big(1 + O(n^{-1/2})\Big) \\
\lesssim& \sqrt{n} \int_{|y|\le 10\sqrt{d_n}} \int_{x\ge 1+d_n}^{10} \frac{1}{x^2-1} e^{-n\left(x^2+y^2-\log(x^2+y^2)-1\right)} \mathrm{d}x\mathrm{d}y \\
&+ O\Big(\frac{1}{\sqrt{n}(d_n)^2} e^{-2n(10d_n)^2(1-O(d_n))}\Big),
\end{aligned} \tag{3.3}$$

where the integral with $y \ge 10\sqrt{d_n}$ or $x \ge 10$ is bounded as in (3.1)–(3.2) using that $|z|^2 = x^2 + y^2 \ge 1 + 100d_n$. By a simple Taylor expansion for $y \le 10\sqrt{d_n}$ and $1 + d_n \le x \le 10$, we have

$$x^2 + y^2 - \log(x^2+y^2) - 1 = x^2 + y^2 - 2\log x - \log\left(1 + \frac{y^2}{x^2}\right) - 1$$

$$\geq x^2 - 2\log x - 1 + Cy^2(x-1). \tag{3.4}$$

Therefore we obtain from (3.3) that

$$\begin{aligned} \boldsymbol{P}^{\mathrm{Gin}(\mathbb{C})}\big(\max_{i=1}^{n} \operatorname{Re}\sigma_i \geq 1+d_n\big) &\lesssim \sqrt{n}\int_{x\geq 1+d_n} \frac{1}{x^2-1} e^{-n\left(x^2-2\log x-1\right)}\mathrm{d}x \left(\int_{|y|\leq 10\sqrt{d_n}} e^{-Cny^2 d_n}\mathrm{d}y\right) \\ &\quad + O\Big(\frac{1}{\sqrt{n}(d_n)^2} e^{-2n(10d_n)^2(1-O(d_n))}\Big) \\ &\lesssim \frac{1}{n(d_n)^{5/2}} e^{-2n(d_n)^2(1-O(d_n))}, \end{aligned} \tag{3.5}$$

where we also used the upper estimate in (2.15). This proves (1.11) in the complex case.

We next consider the real case ($\beta = 1$). Recall the one-point correlation function for the complex eigenvalue process in (2.20). The same estimates as in (3.2) and (3.5) also hold true for the complex eigenvalues of the real Ginibre ensemble. We hence focus on the real eigenvalues. Similarly to (2.25), we obtain that

$$\begin{aligned} \boldsymbol{P}^{\mathrm{Gin}(\mathbb{R})}\big(\max_{\sigma_i\in\mathbb{R}} \sigma_i \geq 1+d_n\big) &\leq \boldsymbol{E}\big[\#\{\sigma_i\in\mathbb{R} : \sigma_i \geq 1+d_n\}\big] \\ &\lesssim \frac{1}{nd_n^2} e^{-n[(d_n)^2+2d_n-2\log(1+d_n)]} + \frac{1}{\sqrt{n}d_n} e^{-\frac{n}{2}[(d_n)^2+2d_n-2\log(1+d_n)]} \\ &\leq \frac{1}{nd_n^2} e^{-2n(d_n)^2(1-O(d_n))} + \frac{1}{\sqrt{n}d_n} e^{-n(d_n)^2(1-O(d_n))}, \end{aligned} \tag{3.6}$$

where we also used the estimates in (2.15) and (2.26). Since the second term on the right side of (3.6) is dominant, this proves (1.12) and hence (1.11) in the real case. One could obtain a similar upper bound for $\max_{\sigma_i\in\mathbb{R}}|\sigma_i|$ as in (1.12) and hence prove (1.10) in the real case.

Finally, we prove the last statement (1.13). Using similar arguments as in (2.12) and (2.14), we obtain the following (non-optimal) lower bound (*c.f.*, (3.1), (3.5) and (3.6), respectively)

$$\begin{aligned} \boldsymbol{P}^{\mathrm{Gin}(\mathbb{C})}\Big(\max_{i=1}^{n} |\sigma_i| \geq 1+d_n\Big) &\geq \frac{1}{n}\boldsymbol{E}^{\mathrm{Gin}(\mathbb{C})}\big[\#\{1\leq i\leq n : |\sigma_i| \geq 1+d_n\}\big] \\ &\gtrsim \frac{1}{n^{3/2}(d_n)^2} e^{-2n(d_n)^2(1-O(d_n))}, \end{aligned} \tag{3.7}$$

$$\boldsymbol{P}^{\mathrm{Gin}(\mathbb{C})}\big(\max_{i=1}^{n} \operatorname{Re}\sigma_i \geq 1+d_n\big) \gtrsim \frac{1}{n^2(d_n)^{5/2}} e^{-2n(d_n)^2(1-O(d_n))}, \tag{3.8}$$

$$\boldsymbol{P}^{\mathrm{Gin}(\mathbb{R})}\big(\max_{\sigma_i\in\mathbb{R}} \sigma_i \geq 1+d_n\big) \gtrsim \frac{1}{n^{3/2}d_n} e^{-n(d_n)^2(1-O(d_n))}. \tag{3.9}$$

Hence for any $\sqrt{\log n/n} \ll d_n \ll 1$, using that $n(d_n)^2 \gg \log n$, the last statement (1.13) follows directly from combining the upper estimates (3.2), (3.5)–(3.6) with the lower estimates (3.7)–(3.9). We hence finish the proof of Theorem 1.3.

3.2. **Proof of Theorem 1.4.** The estimates in (1.15)–(1.17) for the Ginibre ensembles follow directly from (1.10)–(1.11), respectively. In particular, along the proof we indeed obtain the following Ginibre estimates, *i.e.*, for $s, t > 1$,

$$\boldsymbol{E}^{\mathrm{Gin}}\Big[\#\Big\{|\sigma_i| \geq 1+s\sqrt{\frac{\gamma_n}{4n}}\Big\}\Big] \lesssim \frac{1}{s^2}\Big(\frac{\log n}{\sqrt{n}}\Big)^{s^2-1} + \frac{\big(\log n\big)^{\frac{1}{2}(s^2-1)}}{sn^{\frac{s^2}{4}}}\mathbb{1}_{\beta=1}, \tag{3.10}$$

$$\boldsymbol{E}^{\mathrm{Gin}}\Big[\#\Big\{\operatorname{Re}\sigma_i \geq 1+t\sqrt{\frac{\gamma_n'}{4n}}\Big\}\Big] \lesssim \frac{1}{t^{\frac{5}{2}}}\Big(\frac{(\log n)^{\frac{5}{4}}}{n^{\frac{1}{4}}}\Big)^{t^2-1} + \frac{(\log n)^{\frac{5}{8}t^2-\frac{1}{2}}}{tn^{\frac{t^2}{8}}}\mathbb{1}_{\beta=1}, \tag{3.11}$$

and for any $1 \ll s_n \ll \log n$,

$$\boldsymbol{E}^{\mathrm{Gin}}\Big[\#\big\{|\sigma_i| \geq 1+\sqrt{\frac{\gamma_n}{4n}}+\frac{s_n}{4n\gamma_n}\big\}\Big] \lesssim e^{-s_n}, \tag{3.12}$$

together with the same estimate for $\operatorname{Re}\sigma_i$ with γ_n' in (1.2), where $\boldsymbol{E}^{\mathrm{Gin}}$ denotes the expectation for the Ginibre ensembles. In the following lemma, we extend the above Ginibre results to any i.i.d. matrices, from which Theorem 1.4 follows.

Lemma 3.1. *Under the same conditions as in Theorem 1.4, then for any $s,t>1$, there exist $C_s, C_t > 0$ such that*

$$\boldsymbol{E}\Big[\#\big\{|\sigma_i| \geq 1+s\sqrt{\frac{\gamma_n}{4n}}\big\}\Big] \lesssim (\log n)^{C_s}\Big(n^{-\frac{s^2-1}{2}}+n^{-\frac{s^2}{4}}\mathbb{1}_{\beta=1}\Big), \tag{3.13}$$

$$\boldsymbol{E}\Big[\#\big\{\operatorname{Re}\sigma_i \geq 1+t\sqrt{\frac{\gamma_n'}{4n}}\big\}\Big] \lesssim (\log n)^{C_t}\Big(n^{-\frac{t^2-1}{4}}+n^{-\frac{t^2}{8}}\mathbb{1}_{\beta=1}\Big). \tag{3.14}$$

Moreover, the Ginibre estimates as in (3.12) also hold true for i.i.d. matrices.

To simplify the arguments, we will only prove Lemma 3.1 in the complex case ($\beta=2$). The proof for the real case ($\beta=1$) is similar, so we omit it (see more details in [CESX23, Section 2.4]). We will follow the same strategy as in the proof of [CEX23, Theorem 4.1] with minor modifications. Actually, the proof is much easier since we only focus on the expected number of eigenvalues in a given regime without considering its fluctuations.

In the proof we will also use the concept of “with very high probability” for an n-dependent event, meaning that for any fixed $D>0$ the probability of the event is bigger than $1-n^{-D}$ if $n \geq n_0(D)$. Moreover, we will use the following standard definition of *stochastic domination*; its standard arithmetic properties can be found in Proposition 6.5 in [EY17].

Definition 3.2. *Let $\mathcal{X} \equiv \mathcal{X}^{(N)}$ and $\mathcal{Y} \equiv \mathcal{Y}^{(N)}$ be two sequences of nonnegative random variables. We say $\mathcal{Y}$ stochastically dominates $\mathcal{X}$ if, for all (small) $\epsilon>0$ and (large) $D>0$,*

$$\boldsymbol{P}\big(\mathcal{X}^{(N)} > N^{\epsilon}\mathcal{Y}^{(N)}\big) \leq N^{-D}, \tag{3.15}$$

for sufficiently large $N \geq N_0(\epsilon, D)$, and we write $\mathcal{X} \prec \mathcal{Y}$ or $\mathcal{X} = O_\prec(\mathcal{Y})$.

Proof of Lemma 3.1 in the complex case. From [AEK21, Theorem 2.1], we know that, for any small $\tau>0$ and large $D>0$,

$$\boldsymbol{P}\Big(\max_{i=1}^{n}|\sigma_i| \geq 1+n^{-\frac{1}{2}+\tau}\Big) \leq n^{-D}. \tag{3.16}$$

Fix any $s \geq 1$ and define an annulus $\Omega_s := \{z \in \mathbb{C} : 1+s\sqrt{\frac{\gamma_n}{4n}} \leq 1+n^{-1/2+\tau}\}$. Then we have

$$\boldsymbol{E}\Big[\#\big\{|\sigma_i| \geq 1+s\sqrt{\frac{\gamma_n}{4n}}\big\}\Big] = \boldsymbol{E}\Big[\#\big\{\sigma_i \in \Omega_s\big\}\Big] + O(n^{-D}). \tag{3.17}$$

To count the number of eigenvalues in Ω_s, we choose two test functions $f_s^{\pm} \in \mathbb{C}_c^2(\mathbb{C})$ satisfying

$$\mathbb{1}_{\Omega_s^-} \leq f_s^- \leq \mathbb{1}_{\Omega_s} \leq f_s^+ \leq \mathbb{1}_{\Omega_s^+}, \tag{3.18}$$

which are supported on $\Omega_s^{\pm} := \{z \in \mathbb{C} : 1+s\sqrt{\frac{\gamma_n}{4n}} \mp \frac{1}{\sqrt{n}} \leq |z| \leq 1+n^{-\frac{1}{2}+\tau}\}$ respectively, and

$$\max_{\alpha_1,\alpha_2\in\mathbb{N}}\Big\{n^{-\frac{\alpha_1+\alpha_2}{2}}\max_{z\in\mathbb{C}}\Big|\partial_z^{\alpha_1}\partial_{\bar{z}}^{\alpha_2}f_s^{\pm}(z)\Big|\Big\} = O(1).$$

It is straightforward to check that the inequalities in (3.18) imply that

$$\boldsymbol{E}\Big[\#\big\{\sigma_i \in \Omega_s^-\big\}\Big] \leq \boldsymbol{E}\Big[\sum_{i=1}^{n} f_s^-(\sigma_i)\Big] \leq \boldsymbol{E}\Big[\#\big\{\sigma_i \in \Omega_s\big\}\Big]$$

$$\leq \boldsymbol{E}\Big[\sum_{i=1}^n f_s^+(\sigma_i)\Big] \leq \boldsymbol{E}\Big[\#\big\{\sigma_i \in \Omega_s^+\big\}\Big]. \tag{3.19}$$

It then suffices to study the expected linear statistics $\boldsymbol{E}\Big[\sum_{i=1}^n f_s^\pm(\sigma_i)\Big]$. Note that (3.19) also holds true when X is a Ginibre matrix. From the Ginibre estimates in (3.10)–(3.11), we know $\boldsymbol{E}^{\mathrm{Gin}}\big[\#\{\sigma_i \in \Omega_s^\pm\}\big]$ have the same upper bound as $\boldsymbol{E}^{\mathrm{Gin}}\big[\#\{\sigma_i \in \Omega_s\}\big]$, *i.e.*,

$$\boldsymbol{E}^{\mathrm{Gin}(\mathbb{C})}\Big[\#\big\{\sigma_i \in \Omega_s^\pm\big\}\Big] \lesssim (\log n)^{C_s} n^{-\frac{s^2-1}{2}}, \tag{3.20}$$

which, together with (3.19), implies that

$$\boldsymbol{E}^{\mathrm{Gin}(\mathbb{C})}\Big[\sum_{i=1}^n f_s^\pm(\sigma_i)\Big] \lesssim (\log n)^{C_s} n^{-\frac{s^2-1}{2}}. \tag{3.21}$$

The following proof uses the same strategy as in [CEX23, Sections 6–7]. For the reader's convenience, we recall the definitions and notations used there. By Girko's formula [Gir84], for any test function $f \in C_c^2(\mathbb{C})$, the linear statistics $\mathcal{L}_f := \sum_{i=1}^n f(\sigma_i)$ can be written as

$$\mathcal{L}_f = -\frac{1}{4\pi}\int_{\mathbb{C}} \Delta_z f(z) \int_0^T \operatorname{Im}\operatorname{Tr} G^z(\mathrm{i}\eta)\mathrm{d}\eta\mathrm{d}^2 z + O_\prec(n^{-100}), \qquad T := n^{100}, \tag{3.22}$$

where G^z is the *resolvent* or *Green function* of the Hermitised matrix H^z:

$$G^z(w) := (H^z - w)^{-1}, \qquad H^z := \begin{pmatrix} 0 & X-z \\ X^* - \overline{z} & 0 \end{pmatrix}, \qquad w \in \mathbb{C}\setminus\mathbb{R}. \tag{3.23}$$

Choose $\eta_c := n^{-L}$ for a sufficiently large $L > 100$ depending on (1.14). Then we define the following regularized quantity to approximate $\mathcal{L}_f$ in the sense of expectation, *i.e.*,

$$\widehat{\mathcal{L}}_f := -\frac{1}{4\pi}\int \Delta_z f(z)\Big(\int_{\eta_c}^T \operatorname{Im}\operatorname{Tr}\big(G^z(\mathrm{i}\eta) - M^z(\mathrm{i}\eta)\big)\mathrm{d}\eta\Big) q_z \mathrm{d}^2 z, \tag{3.24}$$

where the matrix M^z is a deterministic approximation of G^z (see more details in [CEX23, Section 3.1]), and q_z is a regularized truncating function so that the smallest singular value of $X - z$ does not fall too much below its typical size $n^{-3/4}$, *i.e.*, for a fixed small $\epsilon > 0$ and $0 < \zeta < \epsilon/100$,

$$q_z := q\left(\int_{-E_0}^{E_0} \operatorname{Im}\operatorname{Tr} G^z(y + \mathrm{i}\eta_0)\mathrm{d}y\right), \qquad E_0 := n^{-3/4-\epsilon}, \qquad \eta_0 := n^{-3\zeta}E_0, \tag{3.25}$$

where $q : \mathbb{R}_+ \to [0,1]$ is a smooth and non-increasing cut-off function such that

$$q(x) = 1, \quad \text{if} \quad 0 \leq x \leq 1/9; \qquad q(x) = 0, \quad \text{if} \quad x \geq 2/9. \tag{3.26}$$

Repeating the same arguments as in Step 1–6 from [CEX23, Section 6], we obtain the analog of [CEX23, Proposition 6.1], *i.e.*, for any test function $f = f_s^+$ or $f = f_s^-$ chosen in (3.18),

$$\boldsymbol{E}\big|\mathcal{L}_f - \widehat{\mathcal{L}}_f\big| = O\Big(n^{-c\epsilon} n^{-\frac{s^2-1}{2}}\Big), \tag{3.27}$$

for some small constant $c > 0$. Compared to [CEX23, Proposition 6.1], we gain an additional smallness $n^{-\frac{s^2-1}{2}}$ ($s \geq 1$) using that $|z| > 1 + s\sqrt{\gamma_n/(4n)} + O(n^{-1/2})$ from (3.18) and the precise tail asymptotics of the smallest singular value λ_1^z of $X - z$ from [CEX23, Proposition 3.9]:

$$\boldsymbol{P}\big(\lambda_1^z \leq E_0\big) \lesssim n^{3/2} E_0^2 e^{-n(|z|^2-1)^2/2} + n^{-D}, \tag{3.28}$$

with $E_0 = n^{-3/4-\epsilon}$, for any $n^{-1/2} \ll |z| - 1 \leq n^{-1/2+\tau}$ and any large $D > 0$.

In the following, we aim to compare the expected quantity $\boldsymbol{E}[\widehat{\mathcal{L}}_f]$ with the corresponding Ginibre expectation $\boldsymbol{E}^{\mathrm{Gin}(\mathbb{C})}[\widehat{\mathcal{L}}_f]$, via a continuous interpolating flow. More precisely, we consider the Ornstein–Uhlenbeck matrix flow

$$\mathrm{d}H_t^z = -\frac{1}{2}(H_t^z + Z)\mathrm{d}t + \frac{1}{\sqrt{n}}\mathrm{d}\mathcal{B}_t, \quad Z := \begin{pmatrix} 0 & zI \\ \bar{z}I & 0 \end{pmatrix}, \quad \mathcal{B}_t := \begin{pmatrix} 0 & B_t \\ B_t^* & 0 \end{pmatrix}, \tag{3.29}$$

with initial condition $H_{t=0}^z = H^z$ in (3.23) with the i.i.d. matrix X, where B_t is an $n \times n$ matrix with i.i.d. standard complex valued Brownian motion entries. The matrix flow H_t^z interpolates between the initial matrix H^z and the same matrix with X replaced with an independent complex Ginibre ensemble. The proof is the same as in [CEX23, Section 7] with a minor modification on the regime of z, so we omit the details. Repeating the same arguments as in [CEX23, Section 7], we obtain an analog of [CEX23, Proposition 6.2], *i.e.,* for any $f = f_s^{\pm}$

$$\Big|\boldsymbol{E}\big[\widehat{\mathcal{L}}_f\big] - \boldsymbol{E}^{\mathrm{Gin}(\mathbb{C})}\big[\widehat{\mathcal{L}}_f\big]\Big| = O\Big(n^{-1/4+C\epsilon} n^{-\frac{s^2-1}{2}}\Big), \tag{3.30}$$

for some large constant $C > 0$. Again, compared to [CEX23, Proposition 6.2], we achieve an additional smallness $n^{-\frac{s^2-1}{2}}$ $(s \geq 1)$ using (3.28) for $|z| > 1 + s\sqrt{\gamma_n/(4n)} + O(n^{-1/2})$. Combining (3.30) with (3.27), for a sufficiently small $\epsilon > 0$, we obtain that

$$\Big|\boldsymbol{E}\big[\mathcal{L}_f\big] - \boldsymbol{E}^{\mathrm{Gin}(\mathbb{C})}\big[\mathcal{L}_f\big]\Big| = O\Big(n^{-c\epsilon} n^{-\frac{s^2-1}{2}}\Big), \qquad f = f_s^{\pm}. \tag{3.31}$$

This, together with the Ginibre estimate in (3.21), yields that

$$\boldsymbol{E}\Big[\sum_{i=1}^{n} f_s^{\pm}(\sigma_i)\Big] \lesssim (\log n)^{C_s} n^{-\frac{s^2-1}{2}}. \tag{3.32}$$

Combining this with the inequalities in (3.19) and (3.17), we have proved (3.13) for $\beta = 2$. Similarly, one could extend the Ginibre estimates as in (3.12) to i.i.d. matrices as well. More precisely, we choose slightly different domains than those in (3.18), *i.e.,* $\Omega_s^{\pm} = \{z \in \mathbb{C} : 1 + \sqrt{\frac{\gamma_n}{4n}} + \frac{s_n}{\sqrt{4n\gamma_n}} \mp \frac{1}{\sqrt{n}} \leq |z| \leq 1 + n^{-\frac{1}{2}+\tau}\}$. Repeating the above arguments using (3.28) for $|z| > 1 + \sqrt{\gamma_n/(4n)} + s_n/\sqrt{4n\gamma_n}$ with $1 \ll s_n \ll \log n$, we then obtain

$$\Big|\boldsymbol{E}\big[\mathcal{L}_f\big] - \boldsymbol{E}^{\mathrm{Gin}(\mathbb{C})}\big[\mathcal{L}_f\big]\Big| = O\Big(n^{-c\epsilon} e^{-s_n}\Big), \qquad f = f_s^{\pm}. \tag{3.33}$$

Together with (3.17) and (3.19), we hence have extended the Ginibre estimate in (3.12) to i.i.d. matrices.

Finally for the rightmost eigenvalue, one could prove a similar statement as in (3.31) with a modified error bound $O\Big(n^{-c\epsilon} n^{-(s^2-1)/4}\Big)$ using that $|z| > 1 + s\sqrt{\gamma_n'/(4n)} + O(n^{-1/2})$ with γ_n' in (1.2) (see modification details in [CEX23]), from which the second statement (3.14) follows. This completes the proof of Lemma 3.1 in the complex case. ☐

References

[AEK21] J. Alt, L. Erdős, and T. Krüger, *Spectral radius of random matrices with independent entries*, Probab. Math. Phys. **2** (2021), no. 2, 221–280. MR4408013

[BADG01] G. Ben Arous, A. Dembo, and A. Guionnet, *Aging of spherical spin glasses*, Probab. Theory Related Fields **120** (2001), no. 1, 1–67. MR1856194

[BAG97] G. Ben Arous and A. Guionnet, *Large deviations for Wigner's law and Voiculescu's non-commutative entropy*, Probab. Theory Related Fields **108** (1997), no. 4, 517–542. MR1465640

[Bai97] Z. D. Bai, *Circular law*, Ann. Probab. **25** (1997), no. 1, 494–529. MR1428519

[BAZ98] G. Ben Arous and O. Zeitouni, *Large deviations from the circular law*, ESAIM Probab. Statist. **2** (1998), 123–134. MR1660943

[BCCT18] 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** (2018), no. 4, 2268–2286. MR3813992

[BCGZ22] C. Bordenave, D. Chafaï, and D. García-Zelada, *Convergence of the spectral radius of a random matrix through its characteristic polynomial*, Probab. Theory Related Fields **182** (2022), no. 3-4, 1163–1181. MR4408512

[Ben10] M. Bender, *Edge scaling limits for a family of non-Hermitian random matrix ensembles*, Probab. Theory Related Fields **147** (2010), no. 1-2, 241–271. MR2594353

[BF22] S.-S. Byun and P. J Forrester, *Progress on the study of the ginibre ensembles i: Ginue*, arXiv preprint arXiv:2211.16223 (2022).

[BF23] ———, *Progress on the study of the ginibre ensembles ii: Ginoe and ginse*, arXiv preprint arXiv:2301.05022 (2023).

[BG07] R. Boyer and W. M. Y. Goh, *On the zero attractor of the Euler polynomials*, Adv. in Appl. Math. **38** (2007), no. 1, 97–132. MR2288197

[BS09] A. Borodin and C. D. Sinclair, *The Ginibre ensemble of real random matrices and its scaling limits*, Comm. Math. Phys. **291** (2009), no. 1, 177–224. MR2530159

[BY86] 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** (1986), no. 4, 555–569. MR863545

[CESX22] G. Cipolloni, L. Erdős, D. Schröder, and Y. Xu, *Directional extremal statistics for Ginibre eigenvalues*, J. Math. Phys. **63** (2022), no. 10, Paper No. 103303, 11. MR4496015

[CESX23] ———, *On the rightmost eigenvalue of non-Hermitian random matrices*, Ann. Probab. **51** (2023), no. 6, 2192–2242. MR4666294

[CEX23] G. Cipolloni, L. Erdős, and Y. Xu, *Universality of extremal eigenvalues of large random matrices*, arXiv preprint arXiv:2312.08325 (2023).

[CMV16] F. D. Cunden, F. Mezzadri, and P. Vivo, *Large deviations of radial statistics in the two-dimensional one-component plasma*, J. Stat. Phys. **164** (2016), no. 5, 1062–1081. MR3534484

[Ede97] A. Edelman, *The probability that a random real Gaussian matrix has k real eigenvalues, related distributions, and the circular law*, J. Multivariate Anal. **60** (1997), no. 2, 203–232. MR1437734

[EKS94] A. Edelman, E. Kostlan, and M. Shub, *How many eigenvalues of a random matrix are real?*, J. Amer. Math. Soc. **7** (1994), no. 1, 247–267. MR1231689

[EY17] L. Erdős and H.-T. Yau, *A dynamical approach to random matrix theory*, Courant Lecture Notes in Mathematics, vol. 28, Courant Institute of Mathematical Sciences, New York; American Mathematical Society, Providence, RI, 2017. MR3699468

[FN07] P. J Forrester and T. Nagao, *Eigenvalue statistics of the real ginibre ensemble*, Physical review letters **99** (2007), no. 5, 050603.

[Gem86] S. Geman, *The spectral radius of large random matrices*, Ann. Probab. **14** (1986), no. 4, 1318–1328. MR866352

[Gin65] J. Ginibre, *Statistical ensembles of complex, quaternion, and real matrices*, J. Mathematical Phys. **6** (1965), 440–449. MR173726

[Gir84] V. L. Girko, *The circular law*, Teor. Veroyatnost. i Primenen. **29** (1984), no. 4, 669–679. MR773436

[Gui23] A. Guionnet, *Rare events in random matrix theory*, ICM—International Congress of Mathematicians. Vol. 2. Plenary lectures, 2023, pp. 1008–1052. MR4680275

[Kos92] E. Kostlan, *On the spectra of Gaussian matrices*, 1992, pp. 385–388. Directions in matrix theory (Auburn, AL, 1990). MR1148410

[LACTGMS18] B. Lacroix-A-Chez-Toine, A. Grabsch, S. N. Majumdar, and G. Schehr, *Extremes of 2d Coulomb gas: universal intermediate deviation regime*, J. Stat. Mech. Theory Exp. **1** (2018), 013203, 39. MR3761607

[LS91] N. Lehmann and H.-J. Sommers, *Eigenvalue statistics of random real matrices*, Physical review letters **67** (1991), no. 8, 941.

[May72] R. M May, *Will a large complex system be stable?*, Nature **238** (1972), no. 5364, 413–414.

[Meh04] M. L. Mehta, *Random matrices*, Third, Pure and Applied Mathematics (Amsterdam), vol. 142, Elsevier/Academic Press, Amsterdam, 2004. MR2129906

[PH98] D. Petz and F. Hiai, *Logarithmic energy as an entropy functional*, Advances in differential equations and mathematical physics (Atlanta, GA, 1997), 1998, pp. 205–221. MR1606719

[PTZ17] M. Poplavskyi, R. Tribe, and O. Zaboronski, *On the distribution of the largest real eigenvalue for the real Ginibre ensemble*, Ann. Appl. Probab. **27** (2017), no. 3, 1395–1413. MR3678474

[Rid03] B. Rider, *A limit theorem at the edge of a non-Hermitian random matrix ensemble*, 2003, pp. 3401–3409. Random matrix theory. MR1986426

[RS14] B. Rider and C. D. Sinclair, *Extremal laws for the real Ginibre ensemble*, Ann. Appl. Probab. **24** (2014), no. 4, 1621–1651. MR3211006

[TV10] T. Tao and V. Vu, *Random matrices: universality of ESDs and the circular law*, Ann. Probab. **38** (2010), no. 5, 2023–2065. With an appendix by Manjunath Krishnapur. MR2722794