Intermediate deviation regime for the full eigenvalue statistics in the complex Ginibre ensemble

Bertrand Lacroix-A-Chez-Toine,[1] Jeyson Andrés Monroy Garzón,[2] Christopher Sebastian Hidalgo Calva,[2]
Isaac Pérez Castillo,[2] Anupam Kundu,[3] Satya N. Majumdar,[1] and Grégory Schehr[1]

[1]*LPTMS, CNRS, Univ. Paris-Sud, Université Paris-Saclay, 91405 Orsay, France*
[2]*Departamento de Física Cuántica y Fotónica, Instituto de Física, Universidad Nacional Autónoma de México,
Cd. de México C.P. 04510, México*
[3]*International Centre for Theoretical Sciences, TIFR, Bangalore 560089, India*

We study the Ginibre ensemble of $N \times N$ complex random matrices and compute exactly, for any finite N, the full distribution as well as all the cumulants of the number N_r of eigenvalues within a disk of radius r centered at the origin. In the limit of large N, when the average density of eigenvalues becomes uniform over the unit disk, we show that for $0 < r < 1$ the fluctuations of N_r around its mean value $\langle N_r \rangle \approx N r^2$ display three different regimes: (i) a typical Gaussian regime where the fluctuations are of order $O(N^{1/4})$, (ii) an intermediate regime where $N_r - \langle N_r \rangle = O(\sqrt{N})$, and (iii) a large deviation regime where $N_r - \langle N_r \rangle = O(N)$. This intermediate behavior (ii) had been overlooked in previous studies and we show here that it ensures a smooth matching between the typical and the large deviation regimes. In addition, we demonstrate that this intermediate regime controls all the (centered) cumulants of N_r, which are all of order $O(\sqrt{N})$. We show that the intermediate deviation function that describes these intermediate fluctuations can be computed explicitly and we demonstrate that it is universal, i.e., it holds for a large class of complex random matrices. Our analytical results are corroborated by precise "importance sampling" Monte Carlo simulations.

I. INTRODUCTION

Since the seminal works of Wishart [1], in statistics, and Wigner [2], in nuclear physics, there has been an ever growing number of topics where random matrix theory has found useful applications [3–5]. This encompasses quantum chaos [6], mesoscopic transport [7], stochastic growth models [8,9], trapped fermions [10–12] or disordered systems [13] in physics, as well as combinatorics or number theory [14] in mathematics, to name just a few. Of crucial importance are the various statistical properties of these random matrices [3,4], which for many of these applications, correspond to deriving the *full counting statistics* (FCS), i.e., the statistics of the number $N_\mathcal{D}$ of eigenvalues of a random matrix contained in a given domain $\mathcal{D}$. For random matrices with real eigenvalues, this domain $\mathcal{D}$ is just an interval on the real line. In contrast, for random matrices with complex eigenvalues, the domain $\mathcal{D}$ would represent a two-dimensional region of the complex plane. Questions related to FCS in random matrix ensembles and also to Coulomb gas models have actually attracted a lot of attention both in physics (see, e.g., [15–18]) and mathematics (see, e.g., [19,20]) during the last few years. In particular, it was realized that these random systems exhibit some "rigidity" or "hyperuniformity" (for reviews, see [21,22]): in these cases, indeed, the variance of $N_\mathcal{D}$ usually grows slower than the volume of the domain $\mathcal{D}$.

The most studied case concerns $N \times N$ random matrices belonging to the Gaussian unitary ensemble (GUE). In this case, the average density of eigenvalues $\rho_N(\lambda)$ approaches the celebrated Wigner's semicircle law $\rho_N(\lambda) \approx \sqrt{2 - \lambda^2}/\pi$ [2], which has a compact support $[-\sqrt{2}, +\sqrt{2}]$. For a domain

located deep inside the bulk of the spectrum, i.e., for a segment $\mathcal{D} = [-L, L]$ around the origin and of size comparable to the interparticle distance, it is well known from Dyson's and Mehta's works [23–26] that the variance of $N_\mathcal{D}$ scales logarithmically with the size of $\mathcal{D}$. Later on, it was shown that the typical fluctuations of $N_\mathcal{D}$, in the bulk, are Gaussian [27–29], while the large deviations in the bulk regime were studied in [30]. It is only recently [10,31] that the fluctuations of $N_\mathcal{D}$ for an interval of size larger than the bulk, i.e., for $L = O(1)$, were investigated, motivated to a large extent by the connection between the GUE and the problem of N noninteracting fermions in a harmonic trap at zero temperature. Interestingly, it was shown that the variance exhibits a nonmonotonic behavior: after a logarithmic growth $\propto \ln(NL)$ in the "extended bulk," i.e., for $0 < L < \sqrt{2}$, it decreases abruptly as L approaches the edge $L \approx \sqrt{2}$, finally decaying exponentially fast (with N) for $L > \sqrt{2}$. Furthermore, the full probability distribution of $N_\mathcal{D}$, $P_\mathcal{D}(K, N) = \text{Prob}(N_\mathcal{D} = K)$, was computed for large N using Coulomb gas techniques [10,31], and it was found that, in the extended bulk regime, it follows a large deviation principle of the form

$$P_\mathcal{D}(K, N) \approx e^{-N^2 \psi_L(K/N)}, \qquad (1)$$

where the function $\psi_L(x)$ was obtained explicitly. In standard large deviation forms, the rate function $\psi_L(x)$ usually has a quadratic behavior around its minimum $x_{\min}$, indicating that the typical fluctuations are described by a Gaussian distribution. In this case, however, it was first shown for the distribution of the index number of the GUE [16,17], that this quadratic form is modulated by a logarithmic singularity,

i.e., $\psi_L(x_{\min} + \delta) \propto \delta^2/|\ln\delta|$ as $\delta \to 0$. A similar behavior was later found for the index distribution of Cauchy random matrices [32]. These results indicate that there are actually two scales associated to the fluctuations of $N_{\mathcal{D}}$: (i) a short scale $N_{\mathcal{D}} - \langle N_{\mathcal{D}} \rangle = O(\sqrt{\ln(NL)})$, which describes the typical fluctuations that are Gaussian and (ii) a larger scale $N_{\mathcal{D}} - \langle N_{\mathcal{D}} \rangle = O(N)$ associated to atypical fluctuations, which are no longer Gaussian but instead described by the large deviation principle given by Eq. (1). The matching between these two regimes is ensured by the logarithmic singularity of the rate function $\psi_L(x)$ near its minimum. A similar scenario was later shown to hold for all the classical rotationally invariant ensembles of random matrix theory (RMT), including the β-Gaussian, the β-Wishart, and the β-Cauchy ensembles [31], all of which have real eigenvalues.

What about the case of random matrices with complex eigenvalues? To answer this question, we focus on the complex Ginibre ensemble [3,4,33], which corresponds to the set of $N \times N$ random matrices $\boldsymbol{M}$ with independent and identically distributed (i.i.d.) complex normal random entries with real and imaginary parts of variance $1/(2N)$. The N eigenvalues are complex and their joint probability distribution function (PDF) is given by

$$P_{\text{joint}}(z_1, \ldots, z_N) = \frac{1}{Z_N} \prod_{i<j} |z_i - z_j|^2 \prod_{k=1}^{N} e^{-NV(z_k)}, \quad (2)$$

with $V(z) = |z|^2$, and Z_N being a normalization factor. This ensemble appears in a variety of contexts, ranging from chaotic dissipative quantum systems [34], the two-dimensional one-component plasma [35,36], the study of normal random matrices (i.e., matrices $\boldsymbol{M}$ that satisfy the commutation relation $[\boldsymbol{M}, \boldsymbol{M}^\dagger] = 0$ [37–40]). More recently, a physical realization of this complex Ginibre ensemble in Eq. (2) was found in Ref. [41]. There, it was shown that the eigenvalues z_i of a complex Ginibre matrix are in one-to-one correspondence with the positions of noninteracting fermions in the ground state of a two-dimensional *rotating* harmonic potential [41].

In fact, this correspondence with the noninteracting two dimensional (2D) trapped trapped fermions raises a very interesting and natural question. For trapped fermions, a natural quantity that one can measure is the statistics of the number of fermions in a given domain. This is sometimes referred to as the FCS. Using this mapping to the Ginibre ensemble, the distribution of the number of trapped fermions in domain $\mathcal{D}$ is exactly the same as the distribution of the number of eigenvalues in $\mathcal{D}$ for the Ginibre ensemble. In Ref. [41] only the variance of the number of trapped fermions in a circular domain was computed exactly. In this paper, we show that actually the full distribution of the number of fermions can be computed exactly, giving access to all the moments, thus going beyond the variance. We note that, in general, calculating the FCS of noninteracting trapped fermions is very hard for finite N. For example, even for one-dimensional (1D) noninteracting trapped fermions, the FCS can be computed only in the large N limit [10,31]. However, for the trapped fermions in 2D, using the circular symmetry of the domain $\mathcal{D}$, and the properties of the Ginibre matrices, we are able to compute exactly the distribution of the number of fermions
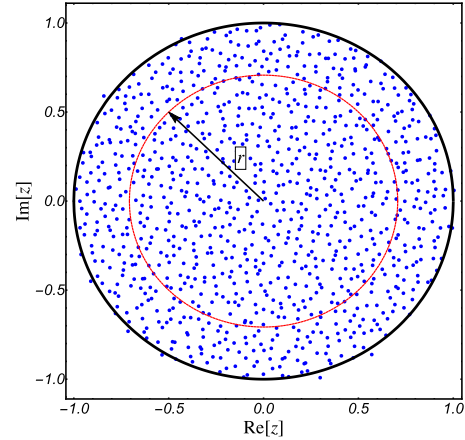

FIG. 1. Snapshot of the position of eigenvalues of the Ginibre ensemble generated by Monte Carlo simulations for $N = 1000$.

for any finite N. The connection to the trapped fermions motivated us to study the large deviation properties of the FCS of the number of eigenvalues in a circular domain $\mathcal{D}$ of the generalized Ginibre ensemble in Eq. (2), where the matrix potential $V(z)$ is isotropic, i.e., $V(z) = v(|z|)$. One of the main results of this paper is to demonstrate the existence of a new intermediate large deviation regime in the FCS of this generalized Ginibre ensemble, which turns out to be universal, i.e., independent of the details of $V(z) = v(|z|)$ as long as it is sufficiently confining.

For the Ginibre ensemble given by Eq. (2) with $V(z) = |z|^2$, in the limit of large N, the limiting density is uniform on the unit disk, i.e., $\rho_N(z) \approx (1/\pi)\Theta(1 - |z|)$, where $\Theta(x)$ is the Heaviside step function. Clearly, the typical distance between two neighboring eigenvalues is $O(1/\sqrt{N})$. In the following, we focus on the FCS inside a disk of radius r centered at the origin, i.e., $\mathcal{D} = \{z \in \mathbb{C}, |z| \leqslant r\}$ with $0 < r < 1$. It turns out that the circular symmetry of the domain $\mathcal{D}$ simplifies considerably the calculations, using our technique, which however is not easily generalizable to noncircular geometry of $\mathcal{D}$. We denote the number of eigenvalues inside this disk by N_r (see Fig. 1). Our main goal is to compute the probability $P_r(K, N) = \text{Prob}(N_r = K)$. Since we naturally expect that $N_r \propto N$, we find convenient to work with the rescaled variable $\kappa = N_r/N$ and introduce its corresponding probability distribution

$$P_r(K, N) = \frac{1}{N} \mathcal{P}_r\left(\kappa = \frac{K}{N}, N\right). \quad (3)$$

In Ref. [42], the function $\mathcal{P}_r(\kappa, N)$ was studied in the limit of large N with $0 < r < 1$ and it was shown to take a scaling form similar to the one given by Eq. (1), that is,

$$\mathcal{P}_r(\kappa, N) \approx e^{-N^2 \Psi_r(\kappa)}, \quad (4)$$

where the rate function $\Psi_r(\kappa)$ was computed explicitly using Coulomb gas techniques [42], yielding

$$\Psi_r(\kappa) = \frac{1}{4}\left| (r^2 - \kappa)(r^2 - 3\kappa) - 2\kappa^2 \ln\left(\frac{\kappa}{r^2}\right) \right|, \quad (5)$$

with $0 \leqslant \kappa \leqslant 1$. Note that a very similar large deviation form (4), with the same rate function (up to trivial rescalings) was

also obtained in the bulk, i.e., for $r = O(1/\sqrt{N})$ in Ref. [43]. Since the density is uniform, one immediately has that $\langle N_r \rangle = N r^2$. This can also be checked from Eq. (5) since $\Psi_r(\kappa)$ has a minimum at $\kappa_{\min} = r^2$. However, contrary to the GUE case, the rate function has a *cubic* behavior near the minimum, i.e.,

$$\Psi_r(\kappa) \approx \frac{|\kappa - r^2|^3}{6r^2}, \quad \kappa \to r^2. \quad (6)$$

This cubic behavior is rather surprising as one would naively expect the typical fluctuations of N_r to be Gaussian, as it was rigorously shown in the bulk when $r = O(1/\sqrt{N})$ [43,44]. Oddly enough, the latter would imply a *quadratic* behavior of the rate function, instead of the cubic one given by Eq. (6). Besides, this cubic behavior would further suggest that the typical fluctuations of N_r are of order $O(N^{1/3})$ [42], which contradicts a recent computation [41] where it was shown that the typical fluctuations of N_r are instead of order $O(N^{1/4})$. In short, the behavior encapsulated in Eq. (6) is thus quite puzzling.

In this paper, we solve this puzzle and show that, at variance with the GUE case, there exist in the present case *three* different scales associated to the fluctuations of $N_r - \langle N_r \rangle$: (i) a shorter scale $O(N^{1/4})$ corresponding to the typical fluctuations which are indeed Gaussian (as in the bulk [43]), (ii) an *intermediate* scale of $O(\sqrt{N})$ associated to "moderate" deviations which are non-Gaussian (see below), and finally (iii) a larger scale $O(N)$ corresponding to large atypical fluctuations, which are described by the large deviation form given by Eqs. (4) and (5). Our findings can be summarized as follows:

$$\mathcal{P}_r(\kappa, N) \approx \begin{cases} e^{-N^{\frac{3}{2}} \frac{\sqrt{\pi}}{2r}(\kappa - r^2)^2}, & |\kappa - r^2| \sim N^{-\frac{3}{4}} \\ e^{-\sqrt{2N} r \Psi_I \left[\sqrt{\frac{N}{2r^2}}(\kappa - r^2) \right]}, & |\kappa - r^2| \sim N^{-\frac{1}{2}} \\ e^{-N^2 \Psi_r(\kappa)}, & |\kappa - r^2| \sim N^0 \end{cases} \quad (7)$$

where we remind that $\kappa = N_r/N$. Note that the fluctuations of N_r is N times larger than the fluctuations of κ. Our main result is thus the existence of an intermediate regime of fluctuations [the second line in (7)], which had been overlooked in previous studies [42,43]. This regime is characterized by the rate function $\Psi_I(x)$ given in Eqs. (29) and (32) below. In the asymptotic regimes of the argument, this rate function behaves as

$$\Psi_I(x) \approx \begin{cases} \sqrt{\frac{\pi}{2}} x^2, & x \to 0 \\ \frac{1}{3} |x|^3, & x \to \pm \infty. \end{cases} \quad (8)$$

Hence, the quadratic form in the first line of (8) matches smoothly with the central Gaussian behavior while the cubic expression in the second line of Eq. (8) matches with the cubic behavior of the large deviation function (6) near its minimum (see Appendix B for details). Moreover, this intermediate deviation regime turns out to be extremely important since, as we later show, it controls all the cumulants of N_r beyond the first one (at the leading order for large N). These read as

$$\langle N_r^p \rangle_c \approx -\sqrt{2N} r \int_{-\infty}^{\infty} dx \, \mathrm{Li}_{1-p} \left(-\frac{\mathrm{erfc}(x)}{\mathrm{erfc}(-x)} \right), \quad (9)$$

with $p > 1$ and where $\mathrm{Li}_n(x) = \sum_{k \geqslant 1} x^k/k^n$ is the polylogarithm function of index n. Using the identity $\mathrm{Li}_{-n}(x) = (-1)^{n+1}\mathrm{Li}_{-n}(1/x)$, for n a positive integer, it is easy to see that the odd order cumulants in Eq. (9) vanish at the leading order $O(\sqrt{N})$. In fact, we argue [see Eqs. (33) and (34)] that, for the Coulomb gas model in Eq. (2) with arbitrary spherically symmetric potential $V(z) = v(|z|)$ [such that $v(r) \gg \ln r^2$ for $r \gg 1$], the cumulants of N_r also behave, for large N, as those given by Eq. (9), up to a v-dependent prefactor [see Eq. (34) below]. Finally, we point out that the very same rate function $\Psi_I(x)$ shows up in a completely different problem, that of the large deviation function of the particle current through the origin in a model of one-dimensional diffusing particles [45].

This paper is organized as follows: In Sec. II we introduce the complex Ginibre ensemble and derive an exact expression for the cumulant generating function of N_r, valid for any finite N. In Sec. III we analyze this exact result in the asymptotic limit of large N and describe in detail the three different regimes of typical, intermediate, and atypical fluctuations. In Sec. IV we discuss the occurrence of the intermediate deviation function $\Psi_I(x)$ in diffusive systems, and, in Sec. V we compare our analytical results to (reweighted) Monte Carlo (MC) simulations, before summarizing our findings and pointing out some remarks in Sec. VI. Some details of the derivations have been relegated to the Appendices.

II. EXACT RESULTS FOR FINITE N

We start with the joint PDF in Eq. (2) with a generic spherically symmetric potential $V(z) = v(|z|)$. Let $r_i = |z_i|$ be the modulus of the eigenvalue z_i. The joint distribution of the r_i's can be computed explicitly from Eq. (2) by integrating over the phases $\theta_i = \arg z_i$ of the eigenvalues. This yields [46,47] (see also Appendix A)

$$P_{\mathrm{rad}}(r_1, \ldots, r_N) = \int_0^{2\pi} \cdots \int_0^{2\pi} \prod_{i=1}^{N} r_i \, d\theta_i \, P_{\mathrm{joint}}(z_1, \ldots, z_N)$$

$$= \frac{1}{N!} \sum_{\sigma \in \mathcal{S}_N} \prod_{k=1}^{N} \frac{r_k^{2\sigma(k)-1}}{h_{\sigma(k)}} e^{-Nv(r_k)}, \quad (10)$$

where $\mathcal{S}_N$ denotes the permutation group of N elements while the coefficients h_k's are given by

$$h_k = \int_0^{\infty} r^{2k-1} e^{-Nv(r)} dr. \quad (11)$$

The expression of $P_{\mathrm{rad}}(r_1, \ldots, r_N)$ is manifestly invariant under any permutation of the r_i's (note that it can also be written as a permanent).

We are interested to compute the number of eigenvalues inside a disk of radius r and centered at the origin, i.e., the number of eigenvalues N_r inside this disk

$$N_r = \sum_{i=1}^{N} \Theta(r - r_i). \quad (12)$$

Starting from the expression (10), one can easily compute the moments of N_r. For instance, using the invariance property of the joint distribution in Eq. (10) under permutation of r_i's, one

can easily show that the mean is given by

$$\langle N_r \rangle = \sum_{k=1}^{N} L_k(r) , \ L_k(r) = \frac{1}{h_k} \int_0^r du \, u^{2k-1} e^{-Nv(u)}, \quad (13)$$

where the notation $\langle (\dots) \rangle$ corresponds to averaging with respect to the joint PDF given by Eq. (10). To derive both the full distribution of N_r and a compact expression for higher order cumulants, it is more convenient to introduce the (centered) *cumulant generating function* (CGF) $\chi_r(\mu, N)$ as

$$\chi_r(\mu, N) = \ln \langle e^{-\mu(N_r - \langle N_r \rangle)} \rangle. \quad (14)$$

Since N_r, as defined in Eq. (12), is a symmetric function of the moduli r_i's, one can easily obtain the following expression for the CGF:

$$\chi_r(\mu, N) = \sum_{k=1}^{N} \{ \ln[M_k(r) + e^{-\mu} L_k(r)] + \mu L_k(r) \}, \quad (15)$$

where $M_k(r) = 1 - L_k(r)$. Furthermore, from $\chi_r(\mu, N)$, one can formally derive an explicit expression for the probability $P_r(K, N) = \text{Prob}(N_r = K)$ of having exactly K eigenvalues inside the disk of radius r centered at the origin, with $0 \leqslant K \leqslant N$. Indeed, noting that $\langle e^{-\mu N_r} \rangle$ can also be written as the generating function of the variable $z = e^{-\mu}$ of the probabilities $P_r(K, N)$ with respect to K, one has that

$$\langle e^{-\mu N_r} \rangle = \sum_{K=0}^{N} e^{-\mu K} P_r(K, N). \quad (16)$$

From Eq. (15), we see that $\langle e^{-\mu N_r} \rangle$ is a simple polynomial of degree N in the variable $z = e^{-\mu}$, that is, $\langle e^{-\mu N_r} \rangle = \prod_{k=1}^{N} [M_k(r) + e^{-\mu} L_k(r)]$. Therefore, we obtain the explicit expression

$$P_r(K, N) = \left[\prod_{k=1}^{N} M_k \right] e_K \left(\frac{L_1}{M_1}, \frac{L_2}{M_2}, \dots, \frac{L_N}{M_N} \right), \quad (17)$$

where we used the shorthand notations $L_k \equiv L_k(r)$, $M_k \equiv M_k(r)$, and where

$$e_K(x_1, \dots, x_N) = \sum_{1 \leqslant \lambda_1 < \dots < \lambda_K \leqslant N} \prod_{k=1}^{K} x_{\lambda_k} \quad (18)$$

is the elementary symmetric polynomial of N variables and degree K. In addition, by expanding Eq. (15) in powers of μ, it is possible to compute all the cumulants $\langle N_r^p \rangle_c$ of order $p > 1$, yielding [41]

$$\langle N_r^p \rangle_c = -\sum_{k=1}^{N} \text{Li}_{1-p} \left(-\frac{L_k(r)}{M_k(r)} \right), \quad (19)$$

where $\text{Li}_n(x)$ is the polylogarithm function.

The formulas (17) and (19) are exact for any N and for any spherically symmetric potential $v(r)$, the dependence on the latter being contained in the expressions $L_k(r)$ and $M_k(r) = 1 - L_k(r)$ given by Eq. (13). In particular, for the complex Ginibre ensemble, which corresponds to $v(r) = r^2$, we have that

$$M_k(r) = \frac{\Gamma(k, Nr^2)}{\Gamma(k)}, \quad L_k(r) = 1 - \frac{\Gamma(k, Nr^2)}{\Gamma(k)}, \quad (20)$$

where $\Gamma(k, z) = \int_z^{\infty} t^{k-1} e^{-t} dt$ is the upper incomplete gamma function. Note that for $K = N$ and $K = 0$, the formula (17) becomes fairly simple. In fact, it is easy to see that $P_r(N, N) = \text{Prob}[r_{\max} \leqslant r]$ and $P_r(0, N) = \text{Prob}[r_{\min} \geqslant r]$, with $r_{\max} = \max_{1 \leqslant i \leqslant N} r_i$ and $r_{\min} = \min_{1 \leqslant i \leqslant N} r_i$, that is, they correspond to the cumulative distributions of $r_{\max}$ and $r_{\min}$, respectively. Therefore, from Eq. (17), one has that

$$P_r(N, N) = \text{Prob}[r_{\max} \leqslant r] = \prod_{k=1}^{N} L_k(r), \quad (21)$$

$$P_r(0, N) = \text{Prob}[r_{\min} \geqslant r] = \prod_{k=1}^{N} M_k(r). \quad (22)$$

For other values of K, the expression of $P_r(K, N)$ is a bit more cumbersome but it can be evaluated numerically rather easily. We note that in the limit $N \to \infty$ with $r = O(1/\sqrt{N})$, our formulas (17) and (18) yield back the result obtained in Ref. [48].

To analyze the large N behavior of $P_r(K, N)$, it turns out that this expression (17) is actually of little use. Instead, it is much more convenient to study the large N behavior of the CGF $\chi_r(\mu, N)$ in Eq. (15). This is the purpose of the next section.

III. LIMIT OF LARGE N

From now on, unless stated otherwise, we focus on the case of Ginibre matrices, corresponding to $v(r) = r^2$ in Eq. (10). The starting point of our analysis is the exact expression for the CGF given by Eq. (15) where $L_k(r)$ and $M_k(r)$ are given in Eq. (20). We consider separately the three regimes corresponding to (i) typical, (ii) intermediate, and (iii) large fluctuations of N_r away from its mean value $\langle N_r \rangle$.

A. Regime of typical fluctuations

In the limit of large N, the density of eigenvalue for the complex Ginibre ensemble is uniform over the unit disk $\rho(r) \approx \frac{1}{\pi}$ and, therefore,

$$\langle N_r \rangle \approx N r^2, \quad (23)$$

which can also be obtained from the exact result given in Eq. (13). To compute the variance, one can extract the coefficient of μ^2 in the small μ expansion of the CGF $\chi_\mu(r)$ and one obtains the exact formula

$$\text{Var}(N_r) = \langle N_r^2 \rangle - \langle N_r \rangle^2 = \sum_{k=1}^{N} L_k(r)[1 - L_k(r)]. \quad (24)$$

In the large N limit, one can show [41] that it behaves as

$$\text{Var}(N_r) \approx \sqrt{\frac{N}{\pi}} r. \quad (25)$$

The fact that the variance grows like $\propto r$, i.e., much slower than the area of the disk ($\propto r^2$) demonstrates the hyperuniformity of the fluctuations in this system [21,22]. Besides, a more detailed analysis of the CGF allows to show that all the cumulants, other than the first one, grow at most like $\sqrt{N}$ [41]. This indicates that for large N the rescaled variable $(N_r - \langle N_r \rangle)/\sqrt{\text{Var}(N_r)}$ converges to a centered Gaussian

random variable with unit variance (since its cumulants of order greater than 2 vanish when $N \to \infty$). Therefore, using Eqs. (23) and (25), we obtain the result given in the first line of Eq. (7). Below, in Sec. V, we compare analytical expressions of the mean (13) and variance (24) with Monte Carlo simulations. Even though we derived this limiting normal distribution only for the Ginibre case, i.e., for the potential $v(r) = r^2$, we show later that this Gaussian form is actually universal, i.e., holds for all $v(r)$ such that $v(r) \gg \ln r^2$ for $r \gg 1$.

B. Regime of intermediate fluctuations

The intermediate regime corresponds to taking the limit of large N but keeping μ fixed in the expression for the CGF $\chi_r(\mu, N)$ given in (15). In this limit, the discrete sum over k is dominated by the region $k \sim N r^2$ and of width $|k - N r^2| = O(\sqrt{N})$. In this region, $L_k(r)$ can be approximated by the form [40]

$$L_k(r) \approx \frac{1}{2}\mathrm{erfc}\left(\frac{k - Nr^2}{\sqrt{2N\,r^2}}\right). \quad (26)$$

By substituting the asymptotic form (26) into Eq. (15) and replacing the discrete sum over k by an integral over $x = (k - Nr^2)/\sqrt{2Nr^2}$, one obtains, for $N \to \infty$ while keeping μ fixed, that

$$\chi_r(\mu, N) \approx \sqrt{2N}r\,\chi(\mu), \quad (27)$$

where the scaling function $\chi(\mu)$ is given by

$$\chi(\mu) = \int_{-\infty}^{\infty} dx \left\{ \ln\left[1 + \frac{e^{-\mu} - 1}{2}\mathrm{erfc}(x) \right] + \frac{\mu}{2}\mathrm{erfc}(x) \right\}. \quad (28)$$

It can be easily verified that the integral over x is well defined. Furthermore, by writing the integral over x as $\int_{-\infty}^{\infty} dx = \int_{-\infty}^{0} dx + \int_{0}^{\infty} dx$, performing the change of variable $x \to -x$ in the first integral and using the identity $\mathrm{erfc}(x) + \mathrm{erfc}(-x) = 2$, the scaling function $\chi(\mu)$ can be conveniently written as

$$\chi(\mu) = \int_{0}^{\infty} ds \ln\left[1 + \sinh^2\left(\frac{\mu}{2}\right)\mathrm{erfc}(s)\mathrm{erfc}(-s) \right]. \quad (29)$$

Under this form (29), we see that $\chi(\mu)$ is an even function of μ, which implies that the odd cumulants vanish at leading order in this limit. By expanding $\chi(\mu)$ in Eq. (29) in powers of μ, one can show rather straightforwardly [41] that the cumulants of even order are given by the formula (9). For later use, it is also useful to extract the asymptotic behaviors of $\chi(\mu)$. These read as (see Appendix B)

$$\chi(\mu) \approx \begin{cases} \dfrac{\mu^2}{2\sqrt{2\pi}}, & \mu \to 0 \\[3mm] \dfrac{2}{3}|\mu|^{3/2}, & \mu \to \pm\infty. \end{cases} \quad (30)$$

To obtain the PDF of N_r, we need to compute the inverse Laplace transform

$$\mathcal{P}_r(\kappa, N) \approx \int_{\mathcal{C}} \frac{d\mu}{2i\pi} e^{N\mu(\kappa - r^2)} e^{\sqrt{2N}r\chi(\mu)}, \quad (31)$$

where $\mathcal{C}$ is the Bromwich contour in the μ-complex plane. In the regime $(\kappa - r^2) = O(N^{-1/2})$, which corresponds precisely to the intermediate regime, the argument of both exponentials in (31) are of the same order $O(\sqrt{N})$ and the Bromwich integral can be evaluated using a saddle-point approximation. This yields the scaling form in the second line of Eq. (7) where the rate function $\Psi_I(x)$ characterizing the fluctuations in the intermediate regime is given by the Legendre transform

$$\Psi_I(x) = -\min_{\mu \in \mathbb{R}} \{x\mu + \chi(\mu)\}, \ x = \sqrt{\frac{N}{2r^2}}(\kappa - r^2). \quad (32)$$

Since the expression (29) of $\chi(\mu)$ is rather involved, it seems fairly cumbersome to obtain a more explicit expression of the rate function $\Psi_I(x)$. Luckily, $\Psi_I(x)$ can easily be evaluated numerically from the expression (32) (see Sec. V below). In addition, using the asymptotic behaviors of $\chi(\mu)$ in Eq. (30), one can show that the intermediate deviation rate function behaves asymptotically as in Eq. (8).

While the previous analysis concerns the Ginibre ensemble with $V(z) = |z|^2$, one can, in fact, show that a similar result can be obtained for a spherically symmetric potential $V(z) = v(|z|)$ so that $v(r) \gg \ln r^2$ as $r \to \infty$. Indeed, for such a potential, the large N behavior of $L_k(r)$ can be obtained by a saddle-point method [40]

$$L_k(r) \approx \frac{1}{2}\mathrm{erfc}[\sqrt{2\pi N\rho(r)}(u_k - r)], \quad (33)$$

where u_k is the solution of $u_k v'(u_k) = 2k/N$ and $\rho(r) = \frac{1}{4\pi r}\partial_r(rv'(r))$ is the average density of the Coulomb gas. For the quadratic potential $v(r) = r^2$ one has $u_k = \sqrt{k/N}$ together with $\rho(r) = 1/\pi$ and, for r close to $u_k = \sqrt{k/N}$ the formula (33) yields back the one given by (26). Inserting the asymptotic form (33) in Eq. (15) and performing the same manipulations as done for the Ginibre case, one finds that the CGF reads as, for large N,

$$\chi_r(\mu, N) = \ln\left\langle e^{-\mu(N_r - \langle N_r \rangle)} \right\rangle \approx \sqrt{2\pi N\rho(r)}\,r\,\chi(\mu), \quad (34)$$

with $\chi(\mu)$ being the same function $\chi(\mu)$ as in Eq. (29). This shows that, in the intermediate regime, the CGF is, for large N, universal up to a multiplicative prefactor $\propto\sqrt{\rho(r)}$ which contains the whole v dependence. An immediate consequence of this result (34) is that the cumulants of N_r are also universal (up to the same multiplicative prefactor), i.e., they are given by Eq. (9) after substituting $\sqrt{2N}$ by $\sqrt{2\pi N\rho(r)}$.

From this universal behavior in the intermediate fluctuation regime, one can also access the typical fluctuation regime by setting the argument x of the intermediate deviation function $\Psi_I(x)$ small. It follows from the first line of Eq. (8) that, for small x, $\Psi_I(x) \sim \sqrt{\pi/2}\,x^2$ to leading order, indicating that the typical fluctuations are described by a universal normal distribution, irrespective of the details of the potential $v(r)$, as long as $v(r) \gg \ln r^2$ for $r \gg 1$. This then demonstrates the validity of the universal central limit theorem in the typical fluctuation regime.

C. Regime of atypical fluctuations

The regime of atypical fluctuations of N_r, for $N_r - \langle N_r \rangle = O(N)$, was studied in Ref. [42] using a Coulomb gas method,

leading to the result in Eqs. (4) and (5). Here, we show that this large deviation regime can also be obtained directly from the exact expression for the CGF in Eq. (15) in the limit of large N with $\mu = O(N)$. To do so, we first set $\mu = \lambda N$, with $\lambda = O(1)$ for large N. Moreover, one can show that, in this limit, the sum over k in Eq. (15) is dominated by $k = O(N)$ so that expressions for $L_k(r)$ and $M_k(r)$ have the following uniform asymptotic expansions [49]:

$$M_{xN}(r) = e^{-N(r^2 - x - x \ln \frac{r^2}{x}) + o(N)}, \qquad x \leqslant r^2 \quad (35)$$

$$M_{xN}(r) = 1 - e^{-N(r^2 - x - x \ln \frac{r^2}{x}) + o(N)}, \quad x \geqslant r^2. \quad (36)$$

To simplify the notations, we set

$$\eta(r, x) = r^2 - x - x \ln \frac{r^2}{x}, \quad (37)$$

which is a positive function. Inserting these asymptotic forms into Eq. (15) and replacing the discrete sum over k by an integral over $x = k/N$ we obtain the scaling form

$$\chi_r(\mu, N) \approx N^2 \, \Xi\left(\frac{\mu}{N}, r\right), \quad (38)$$

where the function $\Xi(\lambda, r)$ reads as

$$\Xi(\lambda, r) = \int_0^{r^2} dx \left[\frac{1}{N} \ln \left(e^{-N\eta(r,x)} + e^{-N\lambda} \right) + \lambda r^2 \right]$$
$$+ \int_{r^2}^1 dx \left[\frac{1}{N} \ln \left(1 + e^{-N[\lambda + \eta(r,x)]} \right) \right]. \quad (39)$$

Since $\eta(r, x) > 0$, the second integral vanishes in the large N limit for $\lambda > 0$. The contribution of the first integral can be expressed in terms of the solution $g(\lambda, r)$ in x and in the interval $0 \leqslant x \leqslant r^2$ of the equation

$$\eta(r, x) = r^2 - x - x \ln \frac{r^2}{x} = |\lambda|. \quad (40)$$

In the first integral of Eq. (39), the term in $\eta(r, x)$ will remain exponentially smaller than the term in λ for $0 < x < g(\lambda, r)$ while it is exponentially larger for $g(\lambda, r) < x < r^2$. This yields

$$\Xi(\lambda, r) = \int_{g(\lambda, r)}^{r^2} dx [\lambda - \eta(r, x)], \quad \lambda \geqslant 0. \quad (41)$$

The PDF can be obtained from this scaling form of the cumulant generating function by Laplace inversion

$$\mathcal{P}_r(\kappa, N) \approx \int_\mathcal{C} \frac{d\lambda}{2i\pi} e^{N^2[\lambda(\kappa - r^2) + \Xi(\lambda, r)]}, \quad (42)$$

where $\mathcal{C}$ is the Bromwich contour. Evaluating this integral by a saddle-point approximation finally yields the scaling form in the third line of Eq. (7), where the rate function $\Psi_r(\kappa)$ is obtained for atypical fluctuations to the left of the typical regime $0 \leqslant \kappa \leqslant r^2$ as

$$\Psi_r(\kappa) = -\min_\lambda \{\lambda(\kappa - r^2) + \Xi(\lambda, r)\}. \quad (43)$$

In particular, the saddle point λ^* for $0 \leqslant \kappa \leqslant r^2$ is such that $g(\lambda^*, r) = \kappa$. Inserting this result in Eq. (43), we obtain

$$\Psi_r(\kappa) = \int_\kappa^{r^2} dx \left(r^2 - x - x \ln \frac{r^2}{x} \right)$$
$$= \frac{1}{4} \left[r^4 - 4r^2\kappa + 3\kappa^2 + 2\kappa^2 \ln \left(\frac{r^2}{\kappa} \right) \right]. \quad (44)$$

To obtain the behavior of $\Psi_r(\kappa)$ for $r^2 \leqslant \kappa \leqslant 1$, one needs to evaluate the saddle point for $\lambda^* \leqslant 0$ and it therefore requires to obtain $\Xi(\lambda, r)$ for $\lambda \leqslant 0$. Following a similar analysis to the case $\lambda \geqslant 0$, we obtain that

$$\Xi(\lambda, r) = -\int_{r^2}^{h(r, \lambda)} dx [\lambda + \eta(r, x)], \quad \lambda \leqslant 0 \quad (45)$$

where $h(r, \lambda)$ is the solution in x and for $r^2 \leqslant x \leqslant 1$ of Eq. (40). The saddle point λ^* in this case is such that $h(r, \lambda^*) = \kappa$ which leads to the expression

$$\Psi_r(\kappa) = \frac{1}{4} \left[-r^4 + 4r^2\kappa - 3\kappa^2 - 2\kappa^2 \ln \left(\frac{r^2}{\kappa} \right) \right], \quad (46)$$

valid for $r^2 \leqslant \kappa \leqslant 1$. Finally, for all values of $0 \leqslant \kappa \leqslant 1$, the rate function $\Psi_r(\kappa)$ can be expressed as in Eq. (5), thus recovering the result of [42]. Note that in the limit $\kappa = 1$, the rate function $\Psi_r(\kappa = 1)$ is the large deviation function of $r_{\max}$ already obtained in [36] with a similar technique, viz.,

$$\Psi_r(\kappa = 1) = -\frac{1}{N^2} \ln \text{Prob}[r_{\max} \leqslant r] \quad (47)$$

$$= -\frac{1}{4}(r^4 - 4r^2 + 3) - \ln(r). \quad (48)$$

On the other hand, for $\kappa = 0$ we obtain the large deviation function of $r_{\min}$:

$$\Psi_r(\kappa = 0) = -\frac{1}{N^2} \ln \text{Prob}[r_{\min} \geqslant r] = \frac{r^4}{4}. \quad (49)$$

IV. CONNECTION WITH CURRENT FLUCTUATIONS IN A DIFFUSIVE SYSTEM

Interestingly, it turns out that the scaling function describing the cumulant generating function in Eq. (29) also appears in a seemingly unrelated problem of diffusing particles on a line, studied in [45], that we briefly describe here. Consider a set of independent Brownian motions on a line, starting at $t = 0$ from initial positions x_0 which are drawn from the following steplike density:

$$\rho^{\text{BM}}(x_0) = \rho_a \Theta(-x_0) + \rho_b \Theta(x_0). \quad (50)$$

The system then evolves diffusively for later times $t > 0$, with a diffusion constant D which we set to one, for simplicity. The probability to find a Brownian particle at a position x at time t, having started at position x_0 at time $t = 0$, is given by the standard propagator

$$P_t(x|x_0) = \frac{e^{-\frac{(x - x_0)^2}{4t}}}{\sqrt{4\pi t}}. \quad (51)$$

Notice that this problem has two sources of randomness: one due to the natural stochastic nature of the Brownian paths,
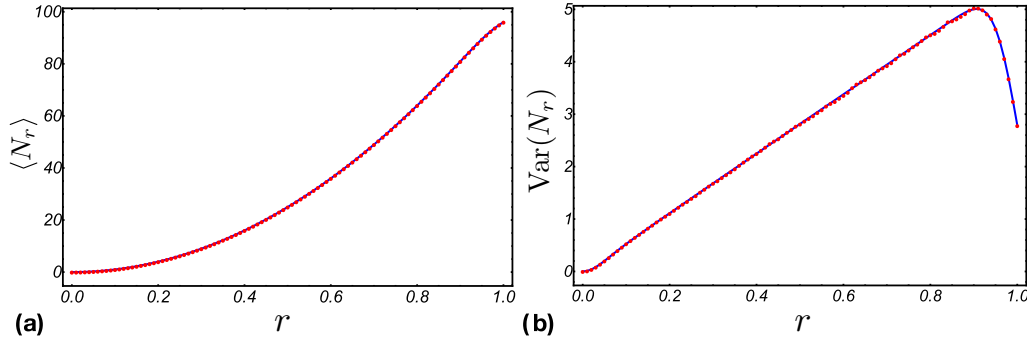
FIG. 2. Plot comparing the exact expressions for the first cumulants $\langle N_r \rangle$ (left panel) and $\mathrm{Var}(N_r)$ (right panel) of N_r as a function of r (solid blue lines) with Monte Carlo simulations. The latter were performed for $N = 100$ and averaging over 10^5 samples.

and a second one due to the initial conditions. Let us then focus on the current Q_t through the origin at $x = 0$, which corresponds to the number of particles that have crossed the origin, up to time t. If we consider the initial conditions as quenched disorder, as in the theory of disordered systems, then the natural definition of the cumulant generating function of the current ought to be [in the same spirit as Eq. (34)] [45]

$$\chi^{\mathrm{q}}(\mu, t) = \overline{\ln \langle e^{-\mu Q_t} \rangle_{\mathrm{th}}}, \tag{52}$$

where the superscript "q" refers to a quenched average, and $\langle \dots \rangle_{\mathrm{th}}$ and $\overline{(\dots)}$ correspond to the average over the different Brownian trajectories for fixed initial conditions, and the average over initial conditions, respectively. It turns out that the quenched disorder is dominated by *typical* initial conditions, and the authors in Ref. [45] were able to derive an exact expression for $\chi^{\mathrm{q}}(\mu, t)$ for arbitrary ρ_a and ρ_b [see Eq. (50)]:

$$\frac{\chi^{\mathrm{q}}(\mu, t)}{2\sqrt{t}} = \rho_a \int_{-\infty}^0 \ln\left[\frac{1}{2}\mathrm{erfc}(s) + \frac{e^{-\mu}}{2}\mathrm{erfc}(-s) \right] ds$$
$$+ \rho_b \int_0^\infty \ln\left[\frac{1}{2}\mathrm{erfc}(-s) + \frac{e^{\mu}}{2}\mathrm{erfc}(s) \right] ds. \tag{53}$$

Setting $\rho_a = \rho_b = \rho$, changing variable $s \to -s$ in the first integral, and gathering all the terms in the logarithm we see that $\overline{\ln \langle e^{-\mu Q_t} \rangle_{\mathrm{th}}} = 2\sqrt{t}\rho \, \chi(\mu)$. Thus, in this mapping, the time t plays the role of N, while $\rho = r$ [50]. Since the fluctuations of the integrated current in single file systems are closely related to the distribution of the displacement of a tagged particle [51], the CGF $\chi(\mu)$ also appears in the computation of the distribution of the displacement X_t made of a tagged particle after time t [52,53]. This formal similarity between the cumulant generating function of diffusive current in the one-dimensional model and that for the number of particles in a circular disk in the two-dimensional Ginibre ensemble is rather intriguing. It would be interesting to explore if there is a deeper one-to-one correspondence between these two problems.

V. COMPARISON WITH MONTE CARLO SIMULATIONS

We first compare our theoretical predictions, exact for any N, for the mean value and variance of N_r, given by Eqs. (13) and (24), respectively, with Monte Carlo simulations. These observables are dominated by typical fluctuations of N_r and they can be precisely computed by using a standard

Metropolis algorithm to simulate the joint PDF of the Ginibre ensemble (2) with $V(z) = |z|^2$. The comparison is shown in Fig. 2, for a system of size $N = 100$, and, as it can be appreciated, the agreement is very good.

To go beyond the first two moments and access the intermediate and large deviation regimes, we need to use "importance sampling" techniques. Here, we follow the procedure introduced in [54].

To generate atypical values of N_r, the idea is to use a "reweighting" scheme such that "atypical" turns "typical." This is achieved by introducing the following auxiliary joint PDF for the Ginibre ensemble

$$P_{\mathrm{joint}}^{(\beta)}(z_1, \dots, z_N) = \frac{e^{-\beta \sum_{i=1}^N \Theta(r - r_i)}}{Z_\beta}$$
$$\times \prod_{i<j} |z_i - z_j|^2 \prod_{k=1}^N e^{-NV(z_k)}, \tag{54}$$

where the parameter β plays a role similar to the inverse temperature, but in this case it can take both positive and negative values. In Eq. (54), Z_β is a normalization constant, i.e., the partition function of the biased Coulomb gas (which, as we will see, does not need to be computed numerically). With respect to this auxiliary joint PDF one can also derive the corresponding PDF of N_r, which we denote here $\mathcal{P}_r^{(\beta)}(\kappa, N)$. Clearly, the original distribution $\mathcal{P}_r(\kappa, N)$ and the new one are related by the formula

$$\mathcal{P}_r(\kappa, N) = e^{\beta N \kappa} Z_\beta \mathcal{P}_r^{(\beta)}(\kappa, N). \tag{55}$$

The role of β is thus to force κ to fluctuate around a typical value with respect to the auxiliary distribution $P_{\mathrm{joint}}^{(\beta)}(z_1, \dots, z_N)$, now atypical with respect to $P_{\mathrm{joint}}(z_1, \dots, z_N)$. We then proceed as follows: one performs a Monte Carlo simulation with respect to $P_{\mathrm{joint}}^{(\beta)}(z_1, \dots, z_N)$ and constructs the histogram of $\mathcal{P}_r^{(\beta)}(\kappa, N)$. Using this distribution for some $\beta \neq 0$ in Eq. (55), one gets the distribution (up to an unknown normalization constant Z_β) in the unbiased problem (i.e., for $\beta = 0$) at a region of κ which is atypical in the original problem. Thus, by changing β, one can access different atypical regions of κ values and gluing the histograms (after proper reweighting [54]) together, one obtains the distribution of $\mathcal{P}_r(\kappa, N)$. When varying β it is important that the sequence
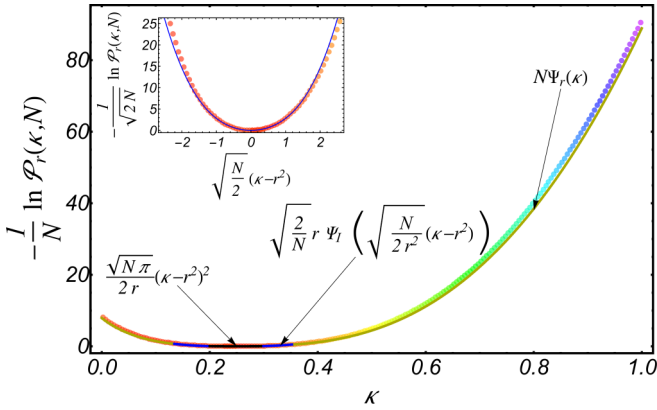
FIG. 3. Plot of $-\ln \mathcal{P}_r(\kappa, N)/N$ against κ corresponding MC simulations (circular color markers) and the exact analytical expression of the typical regime (solid black curve), intermediate regime (solid blue curve), and atypical regime (solid olive green curve). For the MC simulations we have used $N = 500$ and values of β ranging from -333 up to 117 in steps of 4.5. For each temperature, Metropolis algorithm is used for 5×10^4 MC steps to achieve equilibration, after which we generate samples of N_r for another 5×10^4 MC steps. The inset figure shows a more detailed plot corresponding to the intermediate regime.

of histograms do overlap, so that they can be glued together, making the whole distribution of κ properly normalized.

We have applied this reweighting Monte Carlo method to estimate $\mathcal{P}_r(\kappa, N)$ for a system of $N = 500$ eigenvalues and for $r = \frac{1}{2}$. The results are plotted in Fig. 3, where we show the theoretical result of $-\ln \mathcal{P}_r(\kappa, N)/N$ of the three different regimes (solid black line for the typical regime, solid blue curve for the intermediate regime, and solid olive green curve for the atypical regime) together with the Monte Carlo estimates (solid circles in gradient colors representing the varying values of the inverse temperature β). Moreover, the inset of Fig. 3 shows a zoom on the intermediate regime. As it can be appreciated, there is an excellent agreement between theory and Monte Carlo simulations.

VI. CONCLUSIONS AND OUTLOOK

In this paper we have derived exact expressions, valid for any N, of the distribution and all the cumulants of the number N_r of eigenvalues of a $N \times N$ complex Ginibre matrices inside a disk of radius r. In the limit of large N, and for $0 < r < 1$, we showed that there are actually three different regimes of fluctuations: in addition to the typical (Gaussian) and the large deviation regimes, we have unveiled the existence of an intermediate (non-Gaussian) regime that smoothly interpolates between the two. Our main results are thus twofold. First of all, we solved a puzzling problem of matching raised by previous results obtained in Refs. [42,43]. Second, our findings are at variance with the behavior observed in invariant models, as for instance the β-Gaussian, β-Wishart, or β-Cauchy ensembles, in which there is no intermediate regime and, instead, the typical (Gaussian) and the large deviation regimes match smoothly [10,31]. We hope that our exact results for the FCS in the complex Ginibre ensemble will be interesting for the rather wide community working on the FCS

in matrix models and related Coulomb gas systems, and more generally on random or disordered hyperuniform systems [21,22]. Besides, the results found in this work could be experimentally relevant, as it has recently been shown that the positions of the eigenvalues of the complex Ginibre ensemble (2) are in one-to-one correspondence with the positions of fermions in a 2D-rotating harmonic trap [41], a system which, in principle, can be realized experimentally (see, e.g., [55]).

The intermediate regime we have uncovered in our analysis is reminiscent of the one observed for the eigenvalue with largest modulus $r_{\max}$ in the complex Ginibre ensemble [36,40], where the existence of an intermediate regime was also found. In fact, a similar intermediate regime for $r_{\max}$ was also found for 2D-trapped fermions, which constitutes another instance of 2D-determinantal point process with spherical symmetry [56,57]. This is in contrast with matrices with real eigenvalues, such as the classical Gaussian ensembles, where the large deviations of the top eigenvalue typically match smoothly with its typical behavior [58].

One may wonder whether such an intermediate regime for the number of eigenvalues (or particles) in a given domain also exists for other matrix models or Coulomb gases. In particular, a problem of matching between the typical and the large deviation regimes was recently found in [59] for the positions of the bulk particles in the 1D Coulomb gas in an external convex potential. As it was pointed out in [59], this mismatch naturally indicates the presence of an intermediate regime, but the authors in this work were unable to mathematically describe such regime. Oddly enough, it seems that in this particular model, there is no intermediate phase for the fluctuations of the position $r_{\max}$ of the rightmost particle [60]. Thus, it would be worth to revisit this problem and try to characterize its intermediate regime completely for bulk particles. Finally, it is natural to ask whether an intermediate regime for N_r may also be found in the real or symplectic Ginibre ensembles. In particular, for the symplectic Ginibre ensemble, the techniques used in this paper can be extended in a straightforward manner using an identity similar to Eq. (A2) (see, e.g., [3,61]).

APPENDIX A: FINITE N RESULTS FOR THE CUMULANT GENERATING FUNCTION

We first provide a short derivation of the formula given in Eq. (10). For this purpose, we use the Vandermonde identity

to obtain

$$\prod_{i<j} |z_i - z_j|^2 = \det_{1 \leqslant i,j \leqslant N} z_i^{j-1} \det_{1 \leqslant k,l \leqslant N} \bar{z}_l^{k-1}$$

$$= \sum_{\sigma_1, \sigma_2 \in S_N} \text{sign}(\sigma_1) \, \text{sign}(\sigma_2) \prod_{l=1}^N z_l^{\sigma_1(l)-1} \bar{z}_l^{\sigma_2(l)-1}, \tag{A1}$$

where we used the Leibniz formula $\det_{1 \leqslant i,j \leqslant N} a_{ij} = \sum_{\sigma \in S_N} \text{sign}(\sigma) \prod_{l=1}^N a_{l,\sigma(l)}$. Injecting this formula (A1) in Eq. (2) and integrating over the phases $\theta_i = \arg z_i$, we obtain

$$P_{\text{rad}}(r_1, \ldots, r_N) = \int_0^{2\pi} \cdots \int_0^{2\pi} \prod_{i=1}^N r_i \, d\theta_i \, P_{\text{joint}}(z_1, \ldots, z_N)$$

$$= \frac{1}{N!} \sum_{\sigma \in S_N} \prod_{k=1}^N \frac{r_k^{2\sigma(k)-1}}{h_{\sigma(k)}} e^{-Nv(r_k)}, \tag{A2}$$

where we used the identity $\int_0^{2\pi} e^{i(m-n)\theta} d\theta = 2\pi \delta_{m,n}$ and where h_k is given in Eq. (11). This yields Eq. (10) given in the text.

We now compute the moment generating function which reads as, by definition,

$$\langle e^{-\mu N_r} \rangle = \int_0^\infty \cdots \int_0^\infty \prod_{i=1}^N dr_i P_{\text{rad}}(r_1, \ldots, r_N) e^{-\mu \sum_{i=1}^N \theta(r-r_i)}. \tag{A3}$$

Inserting the expression for the joint PDF of the radii given in Eq. (A2) we obtain

$$\langle e^{-\mu N_r} \rangle = \frac{1}{N!} \sum_{\sigma \in S_N} \prod_{k=1}^N \int_0^\infty du \frac{u^{2\sigma(k)-1}}{h_{\sigma(k)}} e^{-Nv(u)-\mu\theta(r-u)}$$

$$= \prod_{k=1}^N \left[e^{-\mu} L_k(r) + M_k(r) \right], \tag{A4}$$

where $L_k(r)$ is given in Eq. (13) and $M_k(r) = 1 - L_k(r)$. Taking the logarithm, and using the expression for $\langle N_r \rangle$ in (13), we obtain the result given in Eq. (15).

APPENDIX B: MATCHING BETWEEN THE THREE DIFFERENT REGIMES

Let us finally check that, as stated in the text, the rate function characterizing the intermediate phase interpolates between the typical and the atypical regimes. One naturally expects to recover the typical regime for small values of μ, while the large argument regime is determined by the large $|\mu|$ behavior of the function $\chi(\mu)$. Now, in the first case, expanding $\chi(\mu)$ for small μ we obtain

$$\chi(\mu) = \frac{\mu^2}{4} \sqrt{\frac{2}{\pi}} + O(\mu^4) = \frac{\mu^2}{2\sqrt{2\pi}} + O(\mu^4), \tag{B1}$$

where we have used that

$$\int_0^\infty du \, \text{erfc}(u) \text{erfc}(-u) = \sqrt{\frac{2}{\pi}}. \tag{B2}$$

This implies that the rate function $\Psi_I(x)$ controlling the intermediate phase becomes

$$\Psi_I(x) = -\min_\mu \{\mu x + \chi(\mu)\} \approx -\min_\mu \left\{ \mu x + \frac{\mu^2}{2\sqrt{2\pi}} \right\}$$

$$\approx \sqrt{\frac{\pi}{2}} x^2, \quad \text{as } x \to 0 \tag{B3}$$

where we have used that the minimum is achieved at $\mu^* = -x\sqrt{2\pi}$. Thus, we have that $\mathcal{P}_r(\kappa, N) \approx e^{-\sqrt{2N} r \sqrt{\frac{\pi}{2}} x^2}$ and recalling that $x = \sqrt{N/(2r^2)}(\kappa - r^2)$ we finally obtain

$$\mathcal{P}_r(\kappa, N) \approx e^{-N^{\frac{3}{2}} \frac{\sqrt{\pi}}{2r} (\kappa - r^2)^2}, \tag{B4}$$

which corresponds to the typical regime of fluctuations, as given in the first line of Eq. (7).

On the other hand, for large values of $|\mu|$, we first realize that the integrand appearing in $\chi(\mu)$ in Eq. (29) behaves differently whenever $s < \sqrt{|\mu|}$ or $s > \sqrt{|\mu|}$. It turns out that only the first region $s < \sqrt{|\mu|}$ contributes, at leading order for large $|\mu|$, yielding

$$\chi(\mu) \approx \int_0^{\sqrt{|\mu|}} ds \ln \left[1 + \frac{2}{\sqrt{\pi} s} e^{|\mu| - s^2} \right]$$

$$\approx \int_0^{\sqrt{|\mu|}} du \ln \left[\frac{2}{\sqrt{\pi} u} e^{|\mu| - u^2} \right] \approx \frac{2}{3} |\mu|^{3/2}. \tag{B5}$$

Thus, the expression for $\Psi_I(x)$ is obtained by solving the following minimization problem:

$$\Psi_I(x) = -\min_\mu \{\mu x + \chi(\mu)\} \approx -\min_\mu \left\{ \mu x + \frac{2}{3} |\mu|^{3/2} \right\}$$

$$\approx \frac{|x|^3}{3}, \quad \text{as } x \to \pm\infty \tag{B6}$$

where in the last step we have used that the value of μ which minimizes the above expression is $\mu^* = -\text{sign}(x) x^2$ (which is positive for negative x and negative for positive x). Finally, in terms of the original variables $x = \sqrt{N/(2r^2)}(\kappa - r^2)$, the tail of the intermediate regime reads as

$$\mathcal{P}_r(\kappa, N) \approx \exp\left(-\sqrt{2N} r \frac{|x|^3}{3} \right)$$

$$\approx \exp\left(-N^2 \frac{|\kappa - r^2|^3}{6r^2} \right), \tag{B7}$$

which exactly coincides with the small argument behavior of the large deviation regime [i.e., the third line in (7)] since the large deviation function exhibits exactly the same cubic behavior (6). We thus conclude that the intermediate regime matches perfectly the typical and atypical regimes.

[1] J. Wishart, Biometrika **20A**, 32 (1928).

[2] E. P. Wigner, Ann. Math. **67**, 325 (1958).

[3] M. L. Mehta, *Random Matrices* (Elsevier, Amsterdam, 2004), Vol. 142.

[4] P. J. Forrester, *Log-Gases and Random Matrices (LMS-34)* (Princeton University Press, Princeton, NJ, 2010).

[5] G. Akemann, J. Baik, and P. Di Francesco, *The Oxford Handbook of Random Matrix Theory* (Oxford University Press, Oxford, 2011).

[6] P. Bourgade and J. P. Keating, *Chaos* (Springer, Berlin, 2013), pp. 125–168.

[7] C. W. Beenakker, Rev. Mod. Phys. **69**, 731 (1997).

[8] S. N. Majumdar, in *Complex Systems: Lecture Notes of the Les Houches Summer School: July 2007*, edited by J.-P. Bouchaud, M. Mézard, and J. Dalibard (Elsevier, Amsterdam, 2007), Vol. 85.

[9] T. Kriecherbauer and J. Krug, J. Phys. A: Math. Gen. **43**, 403001 (2010).

[10] R. Marino, S. N. Majumdar, G. Schehr, and P. Vivo, Phys. Rev. Lett. **112**, 254101 (2014).

[11] I. P. Castillo, Phys. Rev. E **90**, 040102(R) (2014).

[12] D. S. Dean, P. Le Doussal, S. N. Majumdar, and G. Schehr, J. Phys. A: Math. Theor. **52**, 144006 (2019).

[13] G. Biroli, J.-P. Bouchaud, and M. Potters, J. Stat. Mech.: Theory Exp. (2007) P07019.

[14] J. Keating, in *Stochastic Processes and Random Matrices: Lecture Notes of the Les Houches Summer School: July 2015*, edited by G. Schehr, A. Altland, Y. V. Fyodorov, N. O'Connell, and L. F. C. Cugliandolo (Oxford University Press, Oxford, 2017), Vol. 104.

[15] S. Torquato and F. H. Stillinger, Phys. Rev. E **68**, 041113 (2003).

[16] S. N. Majumdar, C. Nadal, A. Scardicchio, and P. Vivo, Phys. Rev. Lett. **103**, 220603 (2009).

[17] S. N. Majumdar, C. Nadal, A. Scardicchio, and P. Vivo, Phys. Rev. E **83**, 041105 (2011).

[18] D. Hexner and D. Levine, Phys. Rev. Lett. **114**, 110602 (2015).

[19] S. Chatterjee, Probab. Theory Relat. Fields, 1 (2019).

[20] T. Leblé and S. Serfaty, Geometric Funct. Anal. **28**, 443 (2018).

[21] S. Ghosh and J. L. Lebowitz, Indian J. Pure Appl. Math. **48**, 609 (2017).

[22] S. Torquato, Phys. Rep. **745**, 1 (2018).

[23] F. J. Dyson, J. Math. Phys. **3**, 140 (1962).

[24] F. J. Dyson, J. Math. Phys. **3**, 157 (1962).

[25] F. J. Dyson, J. Math. Phys. **3**, 166 (1962).

[26] F. J. Dyson and M. L. Mehta, J. Math. Phys. **4**, 701 (1963).

[27] O. Costin and J. L. Lebowitz, Phys. Rev. Lett. **75**, 69 (1995).

[28] M. M. Fogler and B. I. Shklovskii, Phys. Rev. Lett. **74**, 3312 (1995).

[29] A. B. Soshnikov, J. Stat. Phys. **100**, 491 (2000).

[30] D. Holcomb and B. Valkó, *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques* (Institut Henri Poincaré, Paris, 2017), Vol. 53, pp. 1181–1195.

[31] R. Marino, S. N. Majumdar, G. Schehr, and P. Vivo, Phys. Rev. E **94**, 032115 (2016).

[32] R. Marino, S. N. Majumdar, G. Schehr, and P. Vivo, J. Phys. A: Math. Theor. **47**, 055001 (2014).

[33] J. Ginibre, J. Math. Phys. **6**, 440 (1965).

[34] F. Haake, *Quantum Signatures of Chaos* (Springer, Berlin, 2013), Vol. 54.

[35] P. Forrester, Phys. Rep. **301**, 235 (1998).

[36] F. D. Cunden, F. Mezzadri, and P. Vivo, J. Stat. Phys. **164**, 1062 (2016).

[37] L.-L. Chau and Y. Yu, Phys. Lett. A **167**, 452 (1992).

[38] L.-L. Chau and O. Zaboronsky, Commun. Math. Phys. **196**, 203 (1998).

[39] Y. Ameur, H. Hedenmalm, and N. Makarov, Duke Math. J. **159**, 31 (2011).

[40] B. Lacroix-A-Chez-Toine, A. Grabsch, S. N. Majumdar, and G. Schehr, J. Stat. Mech.: Theory Exp. (2018) 013203.

[41] B. Lacroix-A-Chez-Toine, S. N. Majumdar, and G. Schehr, Phys. Rev. A **99**, 021602(R) (2019).

[42] R. Allez, J. Touboul, and G. Wainrib, J. Phys. A: Math. Theor. **47**, 042001 (2014).

[43] T. Shirai, J. Stat. Phys. **123**, 615 (2006).

[44] A. Soshnikov, Russ. Math. Surveys **55**, 923 (2000).

[45] B. Derrida and A. Gerschenfeld, J. Stat. Phys. **137**, 978 (2009).

[46] E. Kostlan, Linear Algebra Applicat. **162–164**, 385 (1992).

[47] D. Chafaï and S. Péché, J. Stat. Phys. **156**, 368 (2014).

[48] P. Forrester, Phys. Lett. A **169**, 21 (1992).

[49] N. Temme, Math. Comput. **29**, 1109 (1975).

[50] There is a misprint in Eqs. (52) and (54) in Ref. [45], where t should be replaced by $\sqrt{t}$.

[51] T. Sadhu and B. Derrida, J. Stat. Mech.: Theory Exp. (2015) P09008.

[52] P. L. Krapivsky, K. Mallick, and T. Sadhu, J. Stat. Phys. **160**, 885 (2015).

[53] J. Cividini and A. Kundu, J. Stat. Mech.: Theory Exp. (2017) 083203.

[54] A. K. Hartmann and M. Mézard, Phys. Rev. E **97**, 032128 (2018).

[55] V. Schweikhard, I. Coddington, P. Engels, V. P. Mogendorff, and E. A. Cornell, Phys. Rev. Lett. **92**, 040404 (2004).

[56] D. S. Dean, P. Le Doussal, S. N. Majumdar, and G. Schehr, J. Stat. Mech.: Theory Exp. (2017) 063301.

[57] B. Lacroix-A-Chez-Toine, P. Le Doussal, S. N. Majumdar, and G. Schehr, Europhys. Lett. **120**, 10006 (2017).

[58] S. N. Majumdar and G. Schehr, J. Stat. Mech.: Theory Exp. (2014) P01012.

[59] R. D. H. Rojas, C. S. H. Calva, and I. P. Castillo, Phys. Rev. E **98**, 020104(R) (2018).

[60] A. Dhar, A. Kundu, S. N. Majumdar, S. Sabhapandit, and G. Schehr, Phys. Rev. Lett. **119**, 060601 (2017).

[61] B. Rider, Probab. Theory Relat. Fields **130**, 337 (2004).