



**HAL**  
open science

# General truncated linear statistics for the top eigenvalues of random matrices

Aurélien Grabsch

► **To cite this version:**

Aurélien Grabsch. General truncated linear statistics for the top eigenvalues of random matrices. *Journal of Physics A: Mathematical and Theoretical*, 2022, 55 (12), pp.124001. 10.1088/1751-8121/ac52e3 . hal-03607644

**HAL Id: hal-03607644**

**<https://hal.sorbonne-universite.fr/hal-03607644v1>**

Submitted on 14 Mar 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# General truncated linear statistics for the top eigenvalues of random matrices

**Aurélien Grabsch**

Sorbonne Université, CNRS, Laboratoire de Physique Théorique de la Matière Condensée (LPTMC), 4 Place Jussieu, 75005 Paris, France

**Abstract.** Invariant ensembles, which are characterised by the joint distribution of eigenvalues  $P(\lambda_1, \dots, \lambda_N)$ , play a central role in random matrix theory. We consider the truncated linear statistics  $L_K = \sum_{n=1}^K f(\lambda_n)$  with  $1 \leq K \leq N$ ,  $\lambda_1 > \lambda_2 > \dots > \lambda_N$  and  $f$  a given function. This quantity has been studied recently in the case where the function  $f$  is monotonous. Here, we consider the general case, where this function can be non-monotonous. Motivated by the physics of cold atoms, we study the example  $f(\lambda) = \lambda^2$  in the Gaussian ensembles of random matrix theory. Using the Coulomb gas method, we obtain the distribution of the truncated linear statistics, in the limit  $N \rightarrow \infty$  and  $K \rightarrow \infty$ , with  $\kappa = K/N$  fixed. We show that the distribution presents two essential singularities, which arise from two infinite order phase transitions for the underlying Coulomb gas. We further argue that this mechanism is universal, as it depends neither on the choice of the ensemble, nor on the function  $f$ .

**Contents**

<b>1</b>	<b>Introduction</b>	<b>2</b>
1.1	Main results . . . . .	5
1.2	Outline of the paper . . . . .	7
<b>2</b>	<b>The Coulomb gas method for general truncated linear statistics</b>	<b>8</b>
2.1	Saddle point equations and large deviation function . . . . .	9
2.2	Reformulation and solution of the saddle-point equation . . . . .	10
<b>3</b>	<b>Application to a system of fermions</b>	<b>13</b>
3.1	Optimal density without constraint . . . . .	14
3.2	Phase I: two disjoint supports . . . . .	15
3.3	Phase II: a logarithmic singularity . . . . .	16
3.4	Two infinite order phase transitions . . . . .	18
3.5	Distribution of the potential energy of the $K$ rightmost fermions . . . . .	19
<b>4</b>	<b>Universality</b>	<b>21</b>
<b>5</b>	<b>Conclusion</b>	<b>22</b>
<b>Appendix A A few useful integrals</b>		<b>24</b>
<b>Appendix B Infinite order phase transitions</b>		<b>24</b>

---

**1. Introduction**

Random matrix theory has first been introduced in physics by Wigner and Dyson in the 1950s to model the atomic nucleus [1]. It has now been successfully applied in various domains of physics, such as electronic quantum transport [2–14], quantum information [15–20], statistical physics of fluctuating interfaces [21, 22] or cold atoms [23, 24].

Invariant ensembles play a prominent role in random matrix theory. They correspond to distributions of matrices which are invariant under changes of basis. Consequently, the eigenvalues and eigenvectors are statistically independent, and one can focus on the eigenvalues only. In the most famous invariant ensembles, the joint distribution of the eigenvalues takes the form [25–27]

$$P(\lambda_1, \dots, \lambda_N) \propto \prod_{i < j} |\lambda_i - \lambda_j|^\beta \prod_{n=1}^N e^{-\frac{\beta N}{2} V(\lambda_n)}, \tag{1}$$

Ensemble	Domain	$V(\lambda)$
Gaussian	$\lambda \in \mathbb{R}$	$\lambda^2$
Laguerre	$\lambda \in \mathbb{R}^+$	$\lambda - \nu \ln \lambda$
Jacobi	$\lambda \in [0, 1]$	$-a \ln \lambda - b \ln(1 - \lambda)$
Cauchy	$\lambda \in \mathbb{R}$	$a \ln(1 + \lambda^2)$

**Table 1.** Example of invariant ensembles of random matrices. The joint distribution of eigenvalues is given by Eq. (1).

where  $\beta$  is the Dyson index which classifies real ( $\beta = 1$ ), complex ( $\beta = 2$ ) and quaternionic ( $\beta = 4$ ) matrices, and  $V(\lambda)$  depends on the specific choice of the ensemble (see Table 1 for explicit expressions for the standard ensembles). The prefactor in the exponential has been chosen such that the eigenvalues do not scale with  $N$ , i.e.  $\lambda_n = \mathcal{O}(N^0)$  for  $N \rightarrow \infty$ , and  $\beta$  has been introduced for convenience. Note that joint distributions of eigenvalues of the type (1) can be extended to arbitrary values of  $\beta > 0$  [28], but they only correspond to invariant ensembles for the classical values  $\beta \in \{1, 2, 4\}$ .

Many applications of random matrices, and in particular of invariant ensembles, rely on the study of linear statistics of the eigenvalues, which are quantities of the form

$$L = \sum_{n=1}^N f(\lambda_n), \quad (2)$$

where  $f$  can be any given function (not necessarily linear). Many relevant quantities can indeed be expressed in this form, such as the number of eigenvalues in a given domain [23, 29–33], the conductance and shot noise of a quantum dot [2, 8, 10], or the mutual information in MIMO<sup>1</sup> channels [34, 35]. Several methods have been developed to analyse the statistical properties of linear statistics. The typical fluctuations can be studied using orthogonal polynomials, or Selberg’s integral. In particular, Dyson and Mehta have obtained a general formula for the variance of linear statistics in the Gaussian ensembles [36], and other formulae can be found in the literature [37–41]. More recently, the question of atypical fluctuations, associated to rare events, has been considered [42, 43]. The best suited tool to address this question is the Coulomb gas method [44–46]. The joint distribution of eigenvalues (1) is interpreted as a Gibbs weight for a gas of particles at positions  $\lambda_n$ , which repel logarithmically. The determination of the distribution of the linear statistics reduces to the determination of the configuration of the particles which minimises the energy of the gas under the constraint (2) with fixed  $L$ . An interesting feature of this approach is the possibility of phase transitions in the Coulomb gas (changes of the shape of the density of eigenvalues), which manifest themselves as non-analyticities in the distribution of the linear statistics in the limit  $N \rightarrow \infty$ . For an overview of the different types of transition, see the table at the end of Ref. [47].

---

1. Multiple-input, multiple-output

Recently, several extensions of these linear statistics have been considered, in which the summation in (2) does not run over all the eigenvalues. One of these extensions is based on the work of Bohigas and Pato [48], who considered the effect of randomly removing each eigenvalue with a given probability. This situation is described by the so-called *thinned ensembles*, which have been the focus of several works over the last years [49–51], studying in particular linear statistics (see also [52] for a related problem).

Alternatively, one can choose deterministically a subset of the eigenvalues<sup>2</sup>  $\{\lambda_n\}$ , and compute the associated linear statistics. This situation has been first considered in [47], where only a given number  $K \leq N$  of the largest eigenvalues were selected, leading to consider the *truncated linear statistics*

$$L_K = \sum_{n=1}^K f(\lambda_n), \quad \lambda_1 > \lambda_2 > \dots > \lambda_N. \quad (3)$$

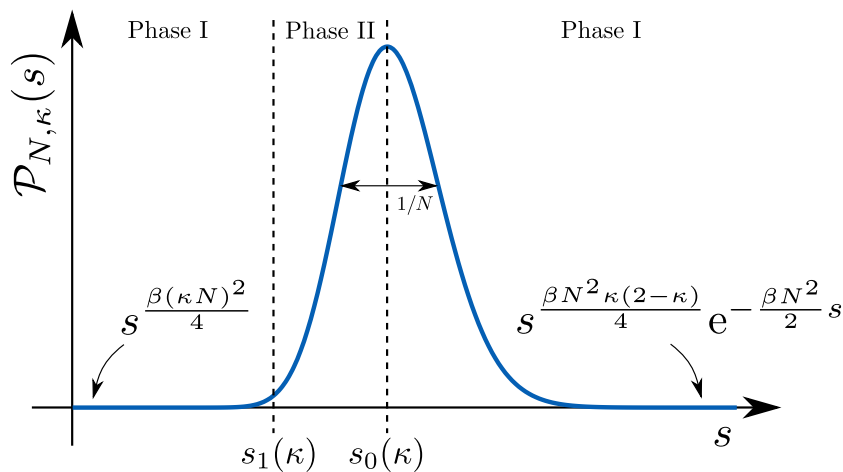
This situation occurs naturally in various contexts, for instance in principal component analysis, where one focuses on a given number of the largest eigenvalues since they contain the most relevant information [31, 53]. The truncated linear statistics (3) interpolates between the usual linear statistics (2) for  $K = N$ , and the largest eigenvalue only for  $K = 1$ , which is also a widely studied quantity [42, 43, 54–59]. The statistical properties of such truncated linear statistics have been studied in Ref. [47] in the bulk regime  $N \rightarrow \infty$  and  $K \rightarrow \infty$  with  $\kappa = K/N$  fixed, using the Coulomb gas method. It was shown that the distribution of  $L_K$  displays a singular behaviour at its typical value, which originates from an infinite order phase transition in the underlying Coulomb gas. This behaviour is universal, meaning that it neither depends on the choice of the ensemble, nor on the choice of the function  $f$ , provided that it is *monotonous*. This problem has also been considered in the edge regime  $N \gg K \gg 1$ , but again for a class of monotonous functions [60]. Finally, truncated linear statistics have been studied for the one dimensional plasma in Ref. [61], which is also a one dimensional gas of particles, but with linear repulsion, and again for a monotonous function  $f$ .

The aim of this paper is to consider the general case where the function  $f$  can be non-monotonous. This extension is crucial to study various important observables, such as the entanglement entropy which corresponds to  $f(\lambda) = -\lambda \ln \lambda$ . Our goal is to determine the distribution of the rescaled truncated linear statistics (3)

$$P_{N,\kappa} \left( s = \frac{L_K}{N} \right) = N! \times \int d\lambda_1 \int^{\lambda_1} d\lambda_2 \dots \int^{\lambda_{N-1}} d\lambda_N P(\lambda_1, \dots, \lambda_N) \delta \left( s - \frac{1}{N} \sum_{n=1}^K f(\lambda_n) \right) \quad (4)$$

---

2. A well-known duality between ensembles of random matrices is obtained by selecting a given subset of eigenvalue. More precisely, the set of every second eigenvalues  $\{\lambda_{2n}\}_{n=1,\dots,N}$  of a matrix of size  $(2N+1) \times (2N+1)$  in the Gaussian Orthogonal Ensemble ( $\beta = 1$ ) is distributed as the eigenvalues of a  $N \times N$  matrix from the Gaussian Symplectic Ensemble ( $\beta = 4$ ). Similar duality relations exist for other ensembles of random matrices [26].



**Figure 1.** Sketch of the distribution of the truncated linear statistics (3), for  $f(\lambda) = \lambda^2$  in the Gaussian ensembles. For  $\beta = 2$ , it is the distribution of the potential energy (43) of the  $K = \kappa N$  rightmost fermions in a harmonic trap. The large deviation function is not analytic, since it presents two essential singularities, located at  $s = s_0(\kappa)$  and  $s = s_1(\kappa)$ . Here, these points are represented for  $\kappa < \frac{1}{2}$ . Otherwise, we have  $s_1(\kappa) > s_0(\kappa)$ , as shown in the phase diagram in Fig. 2.

in the bulk regime  $N \rightarrow \infty$  and  $K \rightarrow \infty$  with  $\kappa = K/N$  fixed. Although we will argue that our results are general, we will mostly focus on a specific example in order to make the analysis more concrete. We will consider the simplest non-monotonous function  $f(\lambda) = \lambda^2$ , and work in the Gaussian ensembles (see Table 1). Besides being the most elementary example, this situation is also motivated by the physics of cold atoms: the eigenvalues  $\{\lambda_n\}$  in the Gaussian ensemble with  $\beta = 2$  can be interpreted as the positions of  $N$  spinless fermions placed in a one-dimensional harmonic trap at zero temperature. The corresponding truncated linear statistics (3) is then the potential energy carried by the  $K$  rightmost fermions.

### 1.1. Main results

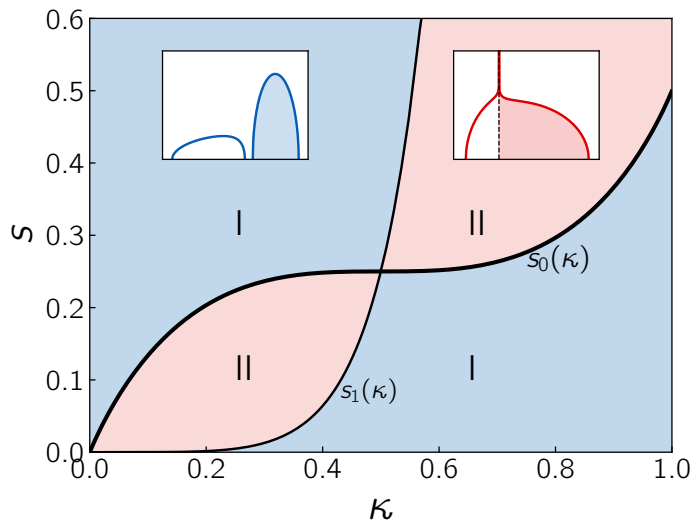
In the limit  $N \rightarrow \infty$ , with  $\kappa = \frac{K}{N}$  fixed, the distribution of the truncated linear statistics (3) with  $f(\lambda) = \lambda^2$  in the Gaussian ensembles ( $V(\lambda) = \lambda^2$ ) takes the form<sup>3</sup>

$$P_{N,\kappa} \left( s = \frac{L_K}{N} \right) \underset{N \rightarrow \infty}{\sim} \exp \left[ -\frac{\beta N^2}{2} \Phi_\kappa(s) \right], \quad (5)$$

where we have introduced the large deviation function  $\Phi_\kappa$ , which has the following

---

3. Throughout the paper, the notation  $P_{N,\kappa}(s) \underset{N \rightarrow \infty}{\sim} \exp[-\frac{\beta N^2}{2} \Phi_\kappa(s)]$  must be understood as  $\lim_{N \rightarrow \infty} \frac{2}{\beta N^2} \ln P_{N,\kappa}(s) = -\Phi_\kappa(s)$ .



**Figure 2.** Phase diagram of the Coulomb gas for the truncated linear statistics (3) with  $f(\lambda) = \lambda^2$  in the Gaussian ensembles ( $V(\lambda) = \lambda^2$ ). The insets represent the shape of the optimal density of eigenvalues in the corresponding phases. In the case of a monotonous function  $f$ , only the line  $s_0(\kappa)$  is present (see Fig. 1.2 in Ref. [47]). The line  $s_0(\kappa)$  is specific to the non-monotonous situation.

behaviours:

$$\Phi_\kappa(s) \simeq \begin{cases} -\frac{\kappa^2}{2} \ln s & \text{for } s \rightarrow 0, \\ \frac{(s - s_0(\kappa))^2}{2F(c_0(\kappa))} & \text{for } s \rightarrow s_0(\kappa), \\ s - \frac{\kappa(2 - \kappa)}{2} \ln s & \text{for } s \rightarrow \infty, \end{cases} \quad (6)$$

where  $s_0(\kappa)$  is the typical value of the truncated linear statistics  $s$ , given by Eq. (47) below, in terms of  $c_0(\kappa)$  (48). The function  $F(c_0)$  controls the variance of  $L_K/N$ , and is given explicitly by (69). The main difference with Ref. [47], in which a monotonous function  $f$  was studied, is that the function  $F$ , and thus the variance of  $L_K$ , displays a surprising non-monotonous behaviour as a function of the fraction  $\kappa$ , as illustrated in Fig. 4 below. A sketch of the distribution (5) of the truncated linear statistics is given in Fig. 1.

This specific form of the distribution arises from a universal mechanism for the underlying Coulomb gas. Indeed, in the limit  $N \rightarrow \infty$ , the distribution of the truncated linear statistics is dominated by one optimal configuration of the eigenvalues (or charges of the Coulomb gas). This optimal configuration is determined by the two parameters: the fraction  $\kappa = K/N$  of eigenvalues under consideration, and  $s$  which controls the constraint in (4). The corresponding phase diagram is shown in Fig. 2.

For a fixed value of  $\kappa$ , the parameter  $s$  drives two consecutive phase transitions for the Coulomb gas, corresponding to a change in the optimal configuration of eigenvalues which dominates (4). The first phase (Phase I) corresponds to an optimal density of

eigenvalues supported on two disjoint supports. For instance, for  $\kappa < \frac{1}{2}$ , it corresponds to the region  $s > s_0(\kappa)$ . As  $s$  decreases, the gap between the two supports reduces, until it completely closes when  $s = s_0(\kappa)$ . For this specific value of  $s$ , the optimal density of eigenvalues is given by Wigner's semicircle law. As  $s$  is further decreased, entering Phase II, a logarithmic divergence emerges at the point where the two interval have merged. Up to now, this scenario is identical to the one described in Ref. [47] in the case of monotonous functions  $f$ .

The specificity of the non-monotonous situation now appears, as  $s$  is further decreased. The location of the logarithmic divergence is displaced, until it reaches the origin (or more generally, the point where  $f'(x) = 0$ ). At this point, the divergence vanishes, but the density still has a logarithmic singularity:  $\rho(x) \underset{x \rightarrow 0}{\simeq} \rho(0) - \alpha x \ln |x|$ , with a constant  $\alpha$ . We denote  $s_1(\kappa)$  the corresponding value of  $s$ . It is given parametrically by Eqs. (61-63). If  $s$  is further decreased, the density splits at the origin into two supports, re-entering Phase I. For  $\kappa > \frac{1}{2}$ , the order of these transitions is inverted.

We further show that the energy of the Coulomb gas, corresponding to the large deviation function  $\Phi_\kappa$ , has an essential singularity at each phase transition,  $s = s_0(\kappa)$  and  $s = s_1(\kappa)$ . We thus speak of *infinite order* phase transitions. The two transition lines intersect at the specific value  $\kappa = \frac{1}{2}$ . At this point, there is no longer a phase transition in the Coulomb gas, as it remains in Phase I both for  $s > s_0(\kappa)$  and  $s < s_0(\kappa)$ .

This scenario is not restricted to the example considered here. We argue in Section 4 (and Appendix B) that this scenario is universal: it holds for any matrix ensemble (choice of  $V$ ) and any linear statistics ( $f$ ), at least in the vicinity of the line  $s_0(\kappa)$  which is the typical value taken by the truncated linear statistics (3). Near this line, the two phases discussed here are always present. These phases are also delimited by a second line  $s_1(\kappa)$  if the function  $f$  admits at least one local extremum in the support of the typical density of eigenvalues. These two lines intersect for some values of  $\kappa$  where  $f'(\langle \lambda_{\kappa N} \rangle) = 0$ , as illustrated in Fig. 5.

Note that, away from the line  $s_0(\kappa)$ , other phases for the Coulomb gas could emerge, depending on the choice of  $f$ .

## 1.2. Outline of the paper

The paper is organised as follows. In Section 2 we introduce the general formalism of the Coulomb gas, applied to the study of truncated linear statistics. In Section 3 we analyse in details an application of this formalism to a specific example of truncated linear statistics, motivated by the study of a system of cold atoms. In Section 4 we argue that the main features observed in this example are actually universal, as they neither depend on the choice of the matrix ensemble, nor on the choice of truncated linear statistics under consideration.



## 2. The Coulomb gas method for general truncated linear statistics

We follow a similar approach to the one used in [47] to study truncated linear statistics with monotonous functions  $f$ . We apply the same method here, based on the Coulomb gas technique, to the general case. The idea of the Coulomb gas method is to rewrite the joint distribution (1) as a Gibbs weight

$$P(\{\lambda_n\}) \propto e^{-\frac{\beta N^2}{2} E_{\text{gas}}(\{\lambda_n\})}, \quad E_{\text{gas}}(\{\lambda_n\}) = \frac{1}{N} \sum_{n=1}^N V(\lambda_n) - \frac{1}{N^2} \sum_{i \neq j} \ln |\lambda_i - \lambda_j|. \quad (7)$$

The energy  $E_{\text{gas}}$  describes a one dimensional gas of particles at positions  $\{\lambda_n\}$ , trapped in a confining potential  $V(\lambda)$  and submitted to repulsive logarithmic interactions between each other. We have placed a factor  $\beta$  in the exponential in (1) so that this energy does not depend on  $\beta$ . In the limit  $N \rightarrow \infty$ , we expect that the typical distribution of eigenvalues finds a balance between the interaction and the confinement energy. This is achieved if  $\lambda_n = \mathcal{O}(N^0)$ , which also implies that  $E_{\text{gas}}(\{\lambda_n\}) = \mathcal{O}(N^0)$ . This is the reason why we placed a factor  $N$  in the definition (1): it ensures that we manipulate quantities which do not scale with  $N$ . We can then introduce the empirical density

$$\rho(x) = \frac{1}{N} \sum_{n=1}^N \delta(x - \lambda_n), \quad (8)$$

which leads us to rewrite the measure (1) as (we neglect the subleading entropic contributions, which are of order  $\mathcal{O}(N^{-1})$  [57, 62])

$$P(\lambda_1, \dots, \lambda_N) d\lambda_1 \cdots d\lambda_N \rightarrow e^{-\frac{\beta N^2}{2} \mathcal{E}[\rho]} \mathcal{D}\rho, \quad (9)$$

where the energy  $\mathcal{E}[\rho]$  is the continuous version of (7),

$$\mathcal{E}[\rho] = - \int \rho(x)\rho(y) \ln |x - y| dx dy + \int \rho(x)V(x)dx. \quad (10)$$

Finally, we rescale the truncated linear statistics (3) as

$$s = \frac{L_K}{N} = \int_c \rho(x)f(x)dx, \quad (11)$$

where  $c = \lambda_K$  is a lower bound ensuring that the summation runs only on the  $K$  largest eigenvalues. It can be determined as

$$\int_c \rho(x)dx = \frac{K}{N} = \kappa. \quad (12)$$

Our aim is to compute the distribution of the rescaled truncated linear statistics  $s$ , which can be expressed in terms of integrals over the density:

$$P_{N,\kappa}(s) = \frac{\int dc \int \mathcal{D}\rho e^{-\frac{\beta N^2}{2} \mathcal{E}[\rho]} \delta\left(\int_c \rho(x)dx - \kappa\right) \delta\left(\int \rho(x)dx - 1\right) \delta\left(\int_c f(x)\rho(x)dx - s\right)}{\int dc \int \mathcal{D}\rho e^{-\frac{\beta N^2}{2} \mathcal{E}[\rho]} \delta\left(\int_c \rho(x)dx - \kappa\right) \delta\left(\int \rho(x)dx - 1\right)} \quad (13)$$

### 2.1. Saddle point equations and large deviation function

When  $N \rightarrow \infty$ , the integrals in (13) are dominated by the minimum of the energy  $\mathcal{E}[\rho]$  under the constraints imposed by the Dirac  $\delta$ -functions. These constraints can be enforced by introducing three Lagrange multipliers  $\mu_0^{(1)}, \mu_0^{(2)}, \mu_1$ . We thus consider

$$\begin{aligned} \mathcal{F}[\rho; \mu_0^{(1)}, \mu_0^{(2)}, \mu_1] = & \mathcal{E}[\rho] + \mu_0^{(1)} \left( \int_c^c \rho(x) dx - (1 - \kappa) \right) \\ & + \mu_0^{(2)} \left( \int_c \rho(x) dx - \kappa \right) + \mu_1 \left( \int_c f(x) \rho(x) dx - s \right) \end{aligned} \quad (14)$$

Let us first focus on the numerator in Eq. (13), and denote  $\rho_\star(x; \kappa, s)$  the density that dominates these integrals. It can be obtained in two steps. First, we find the density  $\tilde{\rho}(x; \mu_0^{(1)}, \mu_0^{(2)}, \mu_1)$  solution of

$$\left. \frac{\delta \mathcal{F}}{\delta \rho(x)} \right|_{\tilde{\rho}} = 0, \quad (15)$$

which yields explicitly the integral equation

$$2 \int \tilde{\rho}(y; \mu_0^{(1)}, \mu_0^{(2)}, \mu_1) \ln |x - y| dy = V(x) + \begin{cases} \mu_0^{(1)} & \text{for } x < c \\ \mu_0^{(2)} + \mu_1 f(x) & \text{for } x > c \end{cases} \quad (16)$$

which can be understood as the energy balance for the particle at point  $x$  between the confinement and the logarithmic repulsion. The Lagrange multipliers  $\mu_0^{(1)}$  and  $\mu_0^{(2)}$  can be interpreted as chemical potentials fixing the fraction of particles respectively below and above  $c$ . The effect of the constraint on  $s$  is to add an additional external potential, proportional to  $\mu_1$ , which acts only on the  $K$  rightmost eigenvalues.

Then, we determine the values  $\mu_0^{(1)\star}(\kappa, s)$ ,  $\mu_0^{(2)\star}(\kappa, s)$ ,  $\mu_1^\star(\kappa, s)$  of the Lagrange multipliers in terms of the parameters  $\kappa$  and  $s$  by imposing the constraints:

$$\int_c \tilde{\rho}(x; \mu_0^{(1)\star}, \mu_0^{(2)\star}, \mu_1^\star) dx = \kappa, \quad \int_c^c \tilde{\rho}(x; \mu_0^{(1)\star}, \mu_0^{(2)\star}, \mu_1^\star) dx = 1 - \kappa \quad (17)$$

$$\int_c f(x) \tilde{\rho}(x; \mu_0^{(1)\star}, \mu_0^{(2)\star}, \mu_1^\star) dx = s. \quad (18)$$

Finally, the density which dominates the integrals in the numerator of (13) is given by

$$\rho_\star(x; \kappa, s) = \tilde{\rho}(x; \mu_0^{(1)\star}(\kappa, s), \mu_0^{(2)\star}(\kappa, s), \mu_1^\star(\kappa, s)). \quad (19)$$

For the denominator, we proceed similarly, but without the constraint on  $s$ . The solution can be deduced from the one obtained above by setting  $\mu_1 = 0$ . Explicitly, from the solution (19), it can be obtained by finding the value  $s_0(\kappa)$  which verifies  $\mu_1(\kappa, s_0(\kappa)) = 0$ . We then deduce the density which dominates the denominator of (13) as

$$\rho_0(x) = \rho_\star(x; \kappa, s_0(\kappa)). \quad (20)$$

Having obtained the densities of eigenvalues  $\rho_\star(x; \kappa, s)$  and  $\rho_0(x)$  which dominate respectively the numerator and the denominator of (13), we can evaluate the integrals with a saddle point estimate, which yields

$$P_{N,\kappa}(s) \underset{N \rightarrow \infty}{\sim} \exp \left\{ -\frac{\beta N^2}{2} \Phi_\kappa(s) \right\}, \quad (21)$$

where we have introduced the large deviation function

$$\Phi_\kappa(s) = \mathcal{E}[\rho_\star(x; \kappa, s)] - \mathcal{E}[\rho_0(x)]. \quad (22)$$

This is the difference of energy between the two optimal configurations of eigenvalues dominating the numerator and the denominator of (13), respectively. These energies can be computed from the exact expressions of the densities using Eq. (10), but this is in general a difficult task. However, an important simplification was introduced in Ref. [13], based on the ‘‘thermodynamic’’ identity

$$\frac{d\mathcal{E}[\rho_\star(x; \kappa, s)]}{ds} = -\mu_1^\star(\kappa, s). \quad (23)$$

See Refs. [63,64] for a more detailed discussion of this relation. It can be used to obtain the large deviation function via a simple integration of the Lagrange multiplier  $\mu_1^\star(\kappa, s)$  (which needs to be computed anyway to determine  $\rho_\star(x; \kappa, s)$ ):

$$\Phi_\kappa(s) = \int_s^{s_0(\kappa)} \mu_1^\star(\kappa, t) dt. \quad (24)$$

We will make extensive use of this relation to study of the distribution  $P_{N,\kappa}(s)$ .

To avoid cumbersome notations, the dependence of the density on the parameters  $\kappa$  and  $s$  will be implicit from now on. We will also not distinguish the densities  $\tilde{\rho}(x; \mu_0^{(1)}, \mu_0^{(2)}, \mu_1)$  and  $\rho_\star(x; \kappa, s)$ ; both will be denoted  $\rho_\star(x)$  in the following.

Now that we have laid out the procedure to obtain the distribution  $P_{N,\kappa}(s)$ , the main remaining task is to find the solution of the saddle-point equation (16).

## 2.2. Reformulation and solution of the saddle-point equation

In order to solve the saddle-point equation (16), it is convenient to take its derivative:

$$2 \int \frac{\rho_\star(y)}{x-y} dy = V'(x) + \begin{cases} 0 & \text{for } x < c \\ \mu_1 f'(x) & \text{for } x > c \end{cases} \quad (25)$$

where  $\int$  denotes a Cauchy principal value integral. This equation can be interpreted as the force balance on the eigenvalue located at position  $x$ . The constraint on  $s$  then acts as an additional force  $-\mu_1 f'(x)$  acting on the  $K$  rightmost eigenvalues (the sign comes from the fact that the force is the opposite of the derivative of the potential). It is convenient to split the density  $\rho_\star$  into two densities:  $\rho_R$  describing the  $K$  rightmost eigenvalues under consideration, and  $\rho_L$  for the others (see Figure 3),

$$\rho_R(x) = \frac{1}{N} \sum_{n=1}^K \delta(x - \lambda_n), \quad \rho_L(x) = \frac{1}{N} \sum_{n=K+1}^N \delta(x - \lambda_n). \quad (26)$$

Due to the confining potential, the eigenvalues remain in a bounded region in space, hence the densities  $\rho_L$  and  $\rho_R$  have compact supports. Let us denote  $[a, b]$  the support of  $\rho_L$ , and  $[c, d]$  the support of  $\rho_R$ , where  $c$  is the boundary introduced previously in Eqs. (11,12), as shown in Figure 3. Note that it is possible that the two supports merge, so that  $b = c$ , as shown in Fig. 3 (right). We rewrite Eq. (25) in terms of these densities as

$$2 \int_a^b \frac{\rho_L(y)}{x-y} dy + 2 \int_c^d \frac{\rho_R(y)}{x-y} dy = V'(x) + \mu_1 f'(x) \quad \text{for } x \in [c, d] \quad (27)$$

$$2 \int_a^b \frac{\rho_L(y)}{x-y} dy + 2 \int_c^d \frac{\rho_R(y)}{x-y} dy = V'(x) \quad \text{for } x \in [a, b]. \quad (28)$$

In these two equations, the principal value is only required when  $x$  belongs to the domain of the integral. Such principal value integral equations can be solved using a theorem due to Tricomi [65], which gives an explicit expression for the inversion of Cauchy singular equations of the form

$$\int \frac{\rho(y)}{x-y} dy = g(x), \quad (29)$$

under the assumption that the solution has one single support  $[a, b]$  [65]:

$$\rho(x) = \frac{1}{\pi \sqrt{(x-a)(b-x)}} \left\{ A + \int_a^b \frac{dt \sqrt{(t-a)(b-t)}}{t-x} g(t) \right\}, \quad (30)$$

where  $A = \int_a^b \rho(x) dx$  is a constant. In our case, solving the coupled equations (27,28) requires to perform a double iteration of this theorem, in order to first determine  $\rho_L$  and then  $\rho_R$ , as in Refs. [30,47]. This procedure is rather cumbersome, but it yields explicit expressions for the densities  $\rho_L$  and  $\rho_R$ .

For the case of the Laguerre ensembles of random matrix theory (which corresponds to a specific  $V(x)$  given in Table 1) and for monotonous functions  $f$ , the derivation is performed in the Appendix of Ref. [47]. Here we adapt this procedure for the general situation. The first step is to use Tricomi's theorem to solve (28) for  $\rho_L$ , treating  $\rho_R$  as a known function. Assuming that  $\rho_L$  has a compact support  $[a, b]$ , we directly apply (30), with

$$g(t) = \frac{1}{2} V'(t) - \int_c^d \frac{\rho_R(y)}{t-y} dy \quad (31)$$

and  $A = \int_a^b \rho_L = 1 - \kappa$  from the constraint (17). After permuting the integrals, and using (A.1), we obtain

$$\rho_L(x) = \frac{1}{\pi \sqrt{(x-a)(b-x)}} \left\{ 1 + \frac{1}{2} \int_a^b \frac{dt}{\pi} V'(t) \frac{\sqrt{(t-a)(b-t)}}{t-x} - \int_c^d dy \rho_R(y) \frac{\sqrt{(y-a)(y-b)}}{y-x} \right\}. \quad (32)$$

Using now this expression in the second saddle-point equation (27), combined with the integral (A.2), we obtain an equation for  $\rho_R$  only, valid for  $x \in [c, d]$ :

$$\begin{aligned} \int_c^d dy \frac{\rho_R(y)}{x-y} \sqrt{(y-a)(y-b)} &= \frac{1}{2} \sqrt{(x-a)(x-b)} (V'(x) + \mu_1 f'(x)) \\ &\quad - 1 + \frac{1}{2} \int_a^b \frac{dt V'(t)}{\pi x-t} \sqrt{(t-a)(b-t)}. \end{aligned} \quad (33)$$

We can solve this second equation using again Tricomi's theorem (30), assuming that  $\rho_R$  has a compact support  $[c, d]$ . This yields

$$\begin{aligned} \rho_R(x) &= \frac{1}{\pi \sqrt{(x-a)(x-b)(x-c)(d-x)}} \left\{ -\frac{1}{2} \int_a^b \frac{dt V'(t)}{\pi t-x} \sqrt{(t-a)(b-t)(c-t)(d-t)} \right. \\ &\quad \left. + C + x + \frac{1}{2} \int_c^d \frac{dt V'(t) + \mu_1 f'(t)}{\pi t-x} \sqrt{(t-a)(t-b)(t-c)(d-t)} \right\}, \end{aligned} \quad (34)$$

where  $C$  is a constant that combines the integration constant from Tricomi's theorem and other terms arising from the evaluation of the integrals. The expression for  $\rho_L$  can be obtained by plugging this result into (32),

$$\begin{aligned} \rho_L(x) &= \frac{-1}{\pi \sqrt{(x-a)(b-x)(c-x)(d-x)}} \left\{ -\frac{1}{2} \int_a^b \frac{dt V'(t)}{\pi t-x} \sqrt{(t-a)(b-t)(c-t)(d-t)} \right. \\ &\quad \left. + C + x + \frac{1}{2} \int_c^d \frac{dt V'(t) + \mu_1 f'(t)}{\pi t-x} \sqrt{(t-a)(t-b)(t-c)(d-t)} \right\}. \end{aligned} \quad (35)$$

The constant  $C$  can then be determined in the following way. Since  $c$  corresponds to the value  $\lambda_K$  of the  $K^{\text{th}}$  eigenvalue, it can freely fluctuate. Therefore, we do not expect that the density of eigenvalues diverges as  $\rho_x(x) \sim (x-c)^{-1/2}$  for  $x \rightarrow c$ , as this type of behaviour typically occurs near a hard edge, which is a hard constraint on the eigenvalues (such as  $\lambda_n > 0$ ). Therefore, the bracket in (34) must vanish for  $x = c$ . This determines the value of the constant  $C$ , which we can now use to simplify the expressions (34,35). We can actually express the total density  $\rho_\star = \rho_L \cup \rho_R$  in a compact form:

$$\begin{aligned} \rho_\star(x) &= \frac{1}{2\pi} \sqrt{\frac{c-x}{(x-a)(b-x)(d-x)}} \left\{ 2 + \int_a^b \frac{dt V'(t)}{\pi t-x} \sqrt{\frac{(t-a)(b-t)(d-t)}{c-t}} \right. \\ &\quad \left. + \int_c^d \frac{dt V'(t) + \mu_1 f'(t)}{\pi t-x} \sqrt{\frac{(t-a)(t-b)(d-t)}{t-c}} \right\}, \end{aligned} \quad (36)$$

where the principal value must be applied only when  $x$  is in the domain of integration. This gives the general solution of the saddle-point equation (25), in any invariant ensemble (1) and for any truncated linear statistics  $f$ , under the assumptions that both  $\rho_L$  and  $\rho_R$  have a compact support. This will be the case for the example discussed below, but some situations might lead to more complex solutions, which would require to iterate Tricomi's theorem again for each additional compact support. The constants  $a$ ,  $b$  and  $d$  in (36) will be determined by the boundary conditions (such as vanishing of

the density at the edge), while  $c$  and  $\mu_1$  will be fixed by the constraints (17,18) which become

$$\int_c^d \rho_R(x) dx = \kappa, \quad \int_c^d \rho_R(x) f(x) dx = s. \quad (37)$$

Note that we have already used that  $\rho_L$  normalises to  $1 - \kappa$  in the derivation above, so only the condition on  $\rho_R$  remains.

The density  $\rho_0(x)$ , which dominates the denominator of (13), is obtained as the solution of equation (25) with  $\mu_1 = 0$ . This equation is already in the form (29), and under the assumption that the solution has a compact support  $[a, b]$ , its solution is given by (30) with  $g(t) = V'(t)/2$ .

We will see below that the general solution (36) gives rise to two different types of solutions (one with  $b < c$  and the other for  $b = c$ ), which we will interpret as different phases for the Coulomb gas. Instead of discussing the meaning of these phases and their implication for the distribution of the truncated linear statistics (3) on these general expressions, we will consider a concrete example. We will discuss the generality of the results obtained on this example in Section 4.

### 3. Application to a system of fermions

In order to illustrate our analysis of truncated linear statistics, we will study in details a specific example which arises from the physics of cold atoms. Consider a system of  $N$  spinless fermions in one dimension, confined by a potential  $\mathcal{V}(x)$ , described by the Hamiltonian

$$\mathcal{H} = \sum_{i=1}^N \left( -\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x_n^2} + \mathcal{V}(x_n) \right). \quad (38)$$

The ground state of this system can be expressed in terms of the one-particle eigenfunctions  $\psi_k$  as a Slater determinant

$$\Psi_0(x_1, \dots, x_N) = \frac{1}{N!} \det[\psi_i(x_j)]_{1 \leq i, j \leq N}. \quad (39)$$

This allows to establish a connection, for specific choices of confining potential  $\mathcal{V}(x)$ , between the positions of the trapped fermions and the eigenvalues of random matrices. For instance, for a harmonic trap  $\mathcal{V}(x) = \frac{1}{2}m\omega^2 x^2$ , the joint distribution of the positions of the fermions is given by

$$|\Psi_0(x_1, \dots, x_N)|^2 \propto \prod_{i < j} (x_i - x_j)^2 \prod_{n=1}^N e^{-m\omega x_n^2/\hbar}. \quad (40)$$

Introducing

$$\lambda_n = x_n \sqrt{\frac{m\omega}{N\hbar}}, \quad (41)$$

the joint distribution of the positions (40) reduces to the joint distribution of eigenvalues (1) for the Gaussian Unitary Ensemble, corresponding to

$$V(\lambda) = \lambda^2 \quad \text{and} \quad \beta = 2. \quad (42)$$

This relation has been used to study various observables, such as the number of particles in a given interval [23]. For a review, see [24]. For higher dimensional systems or systems at finite temperature, the connection with random matrices is lost. One can nevertheless use determinantal point processes to study systems of noninteracting fermions in these cases [24, 66–69]. Here, we will focus on the zero temperature case, in one dimension, where the system is in its ground state. We can therefore treat the positions of the fermions as eigenvalues of random matrices from the Gaussian Unitary Ensemble.

As an example of observable, we consider the potential energy carried by the  $K$  rightmost fermions:

$$E_P(K) = \sum_{n=1}^K \frac{1}{2} m \omega^2 x_n^2 = \frac{N^2 \hbar \omega}{2} s, \quad s = \frac{1}{N} \sum_{n=1}^K \lambda_n^2. \quad (43)$$

For  $K = N$ , the distribution of this observable can be studied by standard techniques, and one can show that it follows a Gamma distribution [70]

$$P_{N,\kappa=1}(s) = \frac{N^{N^2}}{\Gamma\left(\frac{N^2}{2}\right)} s^{\frac{N^2}{2}-1} e^{-N^2 s}. \quad (44)$$

For  $K < N$ , the observable (43) is a truncated linear statistics (3) with  $f(\lambda) = \lambda^2$ . We now focus on the study of the distribution of this observable in the regime  $K \rightarrow \infty$ ,  $N \rightarrow \infty$  with  $\kappa = K/N$  fixed. Although this observable has a physical meaning only when  $\beta = 2$ , we will obtain its distribution for any  $\beta$  since the derivation does not depend on this parameter.

### 3.1. Optimal density without constraint

The first step is to obtain the optimal density of eigenvalues  $\rho_0(x)$  in the absence of constraint. It is the density that dominates the denominator of (13). It can be obtained by using the same procedure as the one used in Ref. [47] for a different example. This density verifies the saddle point equation (25) with  $\mu_1 = 0$ :

$$\int \frac{\rho_0(y)}{x-y} dy = x, \quad (45)$$

which can be solved using Tricomi's theorem (29,30). The density is the celebrated semicircle distribution [25–27]

$$\rho_0(x) = \frac{1}{\pi} \sqrt{2-x^2}. \quad (46)$$

This density, obtained from  $\mu_1 = 0$ , corresponds to the maximum of the probability  $P_{N,\kappa}$ , and therefore to the most probable value of  $s$ , given by

$$s_0(\kappa) = \int_{c_0}^{\sqrt{2}} \rho_0(x) f(x) dx = \frac{1}{2\pi} \arccos \frac{c_0}{\sqrt{2}} + \frac{c_0(1-c_0^2)}{4\pi} \sqrt{2-c_0^2}, \quad (47)$$

where  $c_0$  is fixed by the fraction  $\kappa$  of eigenvalues we consider,

$$\kappa = \int_{c_0}^{\sqrt{2}} \rho_0(x) dx = \frac{1}{\pi} \arccos \frac{c_0}{\sqrt{2}} - \frac{c_0}{2\pi} \sqrt{2 - c_0^2}. \quad (48)$$

These two equations give a parametric representation of the line  $s_0(\kappa)$  in the  $(\kappa, s)$  plane. It is the thick solid line represented in Figure 2, which has the following limiting behaviours

$$s_0(\kappa) \simeq \begin{cases} 2\kappa & \text{for } \kappa \rightarrow 0, \\ \frac{1}{2} - 2(1 - \kappa) & \text{for } \kappa \rightarrow 1. \end{cases} \quad (49)$$

We will see in the following that  $s_0(\kappa)$  defines a phase transition line.

### 3.2. Phase I: two disjoint supports

We now turn to the general situation  $\mu_1 \neq 0$ , for which the solution of the saddle point equation is given by (36). We first consider the situation where the density  $\rho_*$  has two disjoint supports, i.e.  $b < c$ . In this case, the general equations of Section 2 give for the optimal density,

$$\begin{aligned} \rho_*(x) = \text{sign}(x - b) & \frac{\sqrt{(x - a)(b - x)(c - x)(d - x)}}{\pi} \\ & \times \int_{[a,b] \cup [c,d]} \frac{dt \text{sign}(t - b)}{\pi} \frac{t(1 + \mu_1 \Theta(t - c))}{t - x} \frac{1}{\sqrt{(t - a)(t - b)(t - c)(d - t)}}, \end{aligned} \quad (50)$$

where  $\Theta$  is the Heaviside step function. We also have the conditions coming from the vanishing of the density at  $x = a$ ,  $x = d$  and  $x = b$ ,

$$1 + \int_{[a,b] \cup [c,d]} \frac{dt}{\pi} t(1 + \mu_1 \Theta(t - c)) \sqrt{\frac{(d - t)(t - b)}{(t - a)(t - c)}} = 0, \quad (51)$$

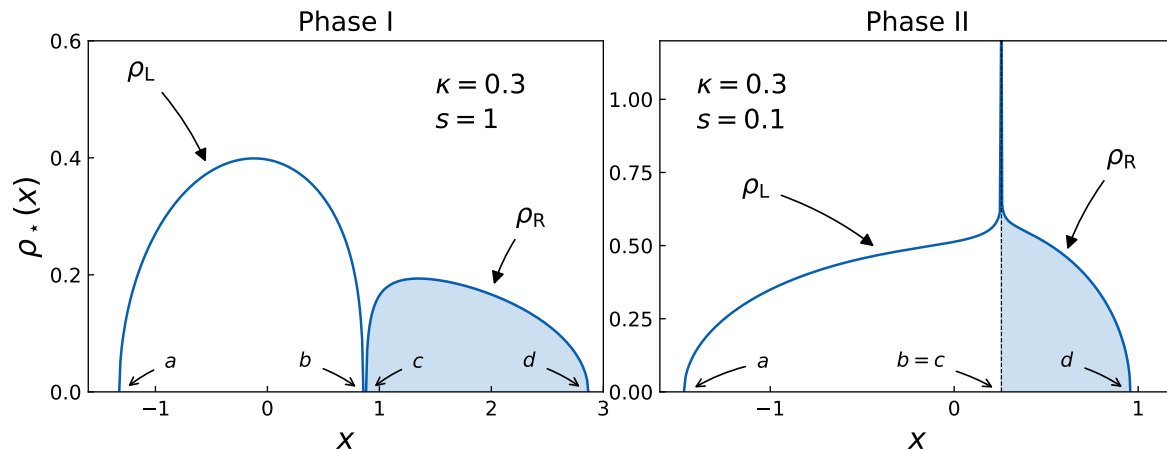
$$\int_{[a,b] \cup [c,d]} \frac{dt}{\pi} t(1 + \mu_1 \Theta(t - c)) \sqrt{\frac{t - b}{(t - a)(t - c)(d - t)}} = 0, \quad (52)$$

$$\int_{[a,b] \cup [c,d]} \frac{dt}{\pi} \text{sign}(t - c) \frac{t(1 + \mu_1 \Theta(t - c))}{\sqrt{(t - a)(t - b)(t - c)(d - t)}} = 0, \quad (53)$$

and the constraints (37). These expressions can be written explicitly in terms of elliptic integrals, but they are more compact in the integral form given above. The density (50) is plotted in Fig. 3 (left).

This phase exists as long as the two supports remain disjoint, that is  $b < c$ . We can actually obtain a necessary condition for this phase to exist using a physical argument. We have indeed seen that the saddle point equation (25) can be understood as a force balance. The additional force in (25) acting on the  $K$  rightmost eigenvalues is  $-\mu_1 f'(x) = -\mu_1 x$ , where  $x$  is the location of the eigenvalue subjected to this force.





**Figure 3.** Optimal density of eigenvalues  $\rho_* = \rho_L \cup \rho_R$ , solution of the saddle point equation (25), or equivalently Eqs. (27,28), for  $\kappa = 0.3$  and different values of  $s$ . Left: density for  $s = 1$ , corresponding to Phase I. Right: density for  $s = 0.1$ , corresponding to Phase II. In both cases, the fraction  $\kappa$  of the largest eigenvalues are described by the density  $\rho_R$ , supported on  $[c, d]$ . The other eigenvalues are described by the density  $\rho_L$ , supported on  $[a, b]$ .

Near the boundary  $x = c$ , this force is thus  $-\mu_1 c$ . For the solution to have two supports, this force needs to be positive. This gives the condition

$$\mu_1 c < 0 \quad (54)$$

for the existence of Phase I.

The limit of existence for this phase is  $b = c$ , when the two supports merge into a single one. This will give rise to another phase of the Coulomb gas, which we now study.

### 3.3. Phase II: a logarithmic singularity

The second phase of interest here consists of a density with single support  $[a, d]$ . It is deduced from the general expression (36) by setting  $b = c$ . The optimal density of eigenvalues take a simpler form in this case:

$$\begin{aligned} \rho_*(x) = & \frac{\sqrt{(d-x)(x-a)}}{\pi} \left( 1 + \frac{2\mu_1}{\pi} \arcsin \sqrt{\frac{d-c}{d-a}} \right) \\ & + \frac{\mu_1 x}{\pi^2} \ln \left| \frac{\sqrt{(x-a)(d-c)} + \sqrt{(c-a)(d-x)}}{\sqrt{(x-a)(d-c)} - \sqrt{(c-a)(d-x)}} \right|, \end{aligned} \quad (55)$$

along with the conditions coming from the vanishing of the density at  $x = a$  and  $x = d$ ,

$$\begin{aligned} 1 + (1 + \mu_1) \frac{(d-a)(3a+d)}{8} \\ + \frac{\mu_1}{4\pi} \left( (d-2c-3a)\sqrt{(c-a)(d-c)} - (d-a)(3a+d) \arcsin \sqrt{\frac{c-a}{d-a}} \right) = 0, \end{aligned} \quad (56)$$

$$\frac{a+d}{2} + \frac{\mu_1}{\pi} \left( \sqrt{(c-a)(d-c)} + (a+d) \arcsin \sqrt{\frac{d-c}{d-a}} \right) = 0, \quad (57)$$

and the constraints (37). This gives four equations, to determine the four parameters  $a$ ,  $c$ ,  $d$  and  $\mu_1$ . The density (55) is plotted in Fig. 3 (right). An important feature of the density (55) is that it exhibits a logarithmic divergence at  $x = c$ :

$$\rho_\star(x) \underset{x \rightarrow c}{\simeq} -\frac{\mu_1 c}{\pi^2} \ln |x - c| \quad \text{for } c \neq 0. \quad (58)$$

This behaviour had already been found in [47] in the case of a monotonous linear statistics. Additionally, when  $c = 0$ , the density no longer diverges, but presents a different logarithmic singularity:

$$\rho_\star(x) \underset{x \rightarrow c}{\simeq} \frac{\sqrt{-ad}}{\pi} - \frac{\mu_1 x}{\pi^2} \ln |x| \quad \text{for } c = 0. \quad (59)$$

This type of singularity has, to the best of our knowledge, never been found previously in the density of eigenvalues of random matrices. It arises here because of the non-monotonicity of the function  $f$  in the truncated linear statistics (3).

A necessary condition for this phase to exist is that the density (55) should remain positive for all  $x \in [a, d]$ . In particular, from the behaviour (58), this imposes that

$$\mu_1 c > 0. \quad (60)$$

This is the complementary condition of the one obtained for Phase I, see Eq. (54). Since the expression (55) for Phase II can be obtained by taking the limit  $b \rightarrow c$  in the general expressions of the density in Phase I (36), these two phases should share a common boundary in the  $(\kappa, s)$  plane, which is thus given by  $\mu_1 c = 0$ . This gives two possibilities:

- $\mu_1 = 0$ , which corresponds to the line  $s = s_0(\kappa)$  in Fig. 2. This line was already present in the case of monotonous functions studied in Ref. [47];
- $c = 0$ , which gives a new line  $s = s_1(\kappa)$  in the phase diagram (Fig. 2). As we will discuss below, this line actually corresponds to the condition  $f'(c) = 0$ . The existence of this second line is thus specific to the study of truncated linear statistics with a non monotonous function  $f$ .

Combining Eqs. (56,57) with the constraints (37), we can write this line  $c = 0$  in the parametric form

$$\kappa = \frac{2}{\pi} \arccos \left( \sqrt{\frac{-a}{d-a}} \right), \quad (61)$$

$$s_1(\kappa) = (3d^2 + 2ad + 3a^2 - 8) \frac{(-ad)^{3/2}}{16(a+d)\pi}, \quad (62)$$

where  $a$  and  $d$  are related by

$$\arccos \left( \sqrt{\frac{-a}{d-a}} \right) + (2+ad) \frac{\sqrt{-ad}}{2(a+d)} = 0. \quad (63)$$

It has the following asymptotic behaviour

$$s_1(\kappa) \simeq \begin{cases} \frac{2\pi^4 \kappa^5}{45} & \text{for } \kappa \rightarrow 0, \\ \frac{3}{2\pi^2(1-\kappa)^3} & \text{for } \kappa \rightarrow 1. \end{cases} \quad (64)$$

Note that we can recover the condition (60) by reversing the physical argument given in Section 3.2: the eigenvalues near  $x = c$ , for  $x > c$ , feel the force  $-\mu_1 c$ . If this force is positive, it pushes the eigenvalues to the right causing the opening of a gap. In order to reverse the situation and get an accumulation of eigenvalues near  $x = c$ , as it is the case here, this force must be negative. This condition yields (60).

### 3.4. Two infinite order phase transitions

We have seen that the two lines  $s_0(\kappa)$  and  $s_1(\kappa)$  delimit regions in the  $(\kappa, s)$  plane in which the optimal density of eigenvalues  $\rho_*(x; \kappa, s)$  takes different forms. We can thus interpret these lines as phase transitions for the Coulomb gas. We now turn to the analysis of the order of these transitions.

*Line  $s_0(\kappa)$*  — We first consider the line on which  $\mu_1 = 0$ , corresponding to the most probable value taken by the truncated linear statistics (3). On this line, the typical density of eigenvalues is given by Wigner’s semicircle law. One can show that all the derivatives of the energy of the Coulomb gas  $\mathcal{E}[\rho_*(x; \kappa, s)]$ , and thus of the large deviation function  $\Phi_\kappa(s)$  are continuous on this line (see Appendix B). However, this function is not analytic: it possesses an essential singularity in Phase I (corresponding to  $s = s_0^+$  for  $\kappa < \frac{1}{2}$  and  $s = s_0^-$  for  $\kappa > \frac{1}{2}$ ):

$$\Phi_\kappa(s_0(\kappa) + \varepsilon) - \Phi_\kappa(s_0(\kappa) - \varepsilon) = \mathcal{O}(\varepsilon e^{\gamma_0(\kappa)/\varepsilon}) \quad \text{for } \kappa \neq \frac{1}{2}, \quad (65)$$

where  $\gamma_0(\kappa)$  is a constant. Therefore, in the standard terminology of statistical physics, it corresponds to an *infinite order* phase transition. This exact same transition has been observed in [47] for truncated linear statistics associated with a monotonous function  $f$ . Here, we obtain exactly the same behaviour for all values of  $\kappa \neq \frac{1}{2}$ . Indeed, for these values of  $\kappa$ , the typical position of the  $K^{\text{th}}$  largest eigenvalue (the last to contribute to  $s$ ) is  $\lambda_K \neq 0$ , away from the point where  $f$  has a minimum. Therefore, small fluctuations of  $\lambda_K$  do not probe the non-monotonicity of  $f$ , and the behaviour of  $\Phi_\kappa$  near  $s_0(\kappa)$  is identical to the one observed in the monotonous case, which has been shown to be universal [47].

*Line  $s_1(\kappa)$*  — The second line, corresponding to  $f'(c) = 0$  (that is,  $c = 0$  here, corresponding to  $\kappa = \frac{1}{2}$ ) is specific to the study of the case of non-monotonous functions  $f$ . It is the main novelty that arises in this case.

On this line, the density presents a logarithmic singularity, as shown in Eq. (59). It is a new specific feature to the case of truncated linear statistics with a non-monotonous function  $f$ .

One can show (see [Appendix B](#)) that the large deviation function exhibits another essential singularity on this line (again located in Phase I):

$$\Phi_\kappa(s_1(\kappa) + \varepsilon) - \Phi_\kappa(s_1(\kappa) - \varepsilon) = \mathcal{O}(\varepsilon e^{\gamma_1(\kappa)/\varepsilon}) \quad \text{for } \kappa \neq \frac{1}{2}, \quad (66)$$

where  $\gamma_1(\kappa)$  is a constant. This shows that  $s_1(\kappa)$  also corresponds to an *infinite order* phase transition.

*Intersection of the two lines for  $\kappa = \frac{1}{2}$*  — The two phase transition lines intersect for  $\kappa = \frac{1}{2}$ . Indeed, for this specific value of  $\kappa$ , in the absence of constraint ( $\mu_1 = 0$ ), the  $K^{\text{th}}$  largest eigenvalue is typically located at  $\langle \lambda_K \rangle = c = 0$ . At this point, the essential singularities vanish, as well as the phase transition. Indeed, only Phase I exists for this specific value of  $\kappa$ .

### 3.5. Distribution of the potential energy of the $K$ rightmost fermions

Using the results above on the optimal density  $\rho_\star(x)$ , we can study the distribution of the truncated linear statistics under consideration: the potential energy of the  $K$  rightmost fermions in a harmonic trap.

*First cumulants* — The value  $s_0(\kappa)$  corresponds to the most probable value taken by the truncated linear statistics, or equivalently by the potential energy of the  $K$  rightmost fermions. It implies that

$$\langle E_P(K) \rangle_{N, K \rightarrow \infty} \simeq \frac{N^2 \hbar \omega}{2} s_0(\kappa) \simeq \frac{\hbar \omega}{2} \begin{cases} 2KN & \text{for } K \rightarrow 0, \\ 2KN - \frac{3}{2}N^2 & \text{for } K \rightarrow N. \end{cases} \quad (67)$$

We can study the fluctuations around this value by expanding Eqs. [\(37,56,57\)](#) in the limit  $\mu_1 \rightarrow 0$ . We obtain

$$s = s_0(\kappa) + F(c_0)\mu_1 + \mathcal{O}(\mu_1^2), \quad (68)$$

with

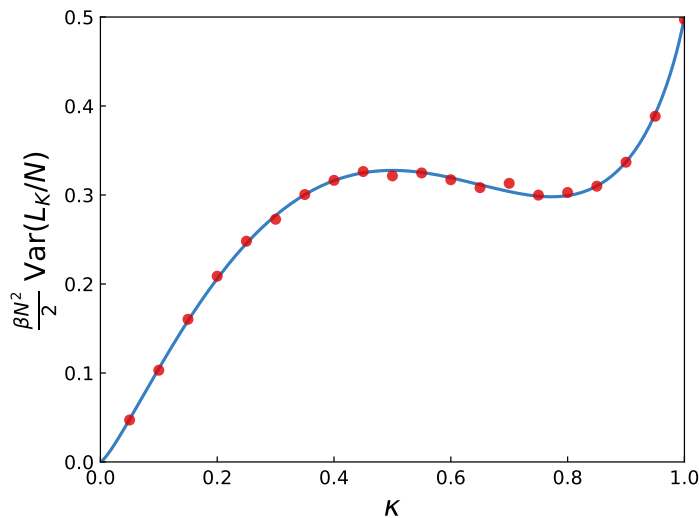
$$F(c_0) = \frac{1}{8\pi^2} \left( 3c_0^4 - 14c_0^2 - 4\sqrt{2 - c_0^2} (c_0^2 - 1) c_0 \arccos\left(\frac{c_0}{\sqrt{2}}\right) + 4 \arccos\left(\frac{c_0}{\sqrt{2}}\right)^2 + 16 \right) \quad (69)$$

and  $c_0$  is related to  $\kappa$  via [\(48\)](#). Inverting the series [\(68\)](#), we deduce the expression of the large deviation function near  $s_0(\kappa)$  via direct integration over  $s$  thanks to the thermodynamic identity [\(24\)](#):

$$\Phi_\kappa(s) \underset{s \rightarrow s_0(\kappa)}{\simeq} \frac{(s - s_0(\kappa))^2}{2F(c_0)} + \mathcal{O}((s - s_0(\kappa))^3). \quad (70)$$

From this result, we straightforwardly deduce

$$\text{Var}\left(\frac{L_K}{N}\right) = \frac{2}{\beta N^2} F(c_0) \simeq \frac{2}{\beta N^2} \begin{cases} \left(\frac{18}{\pi}\right)^{2/3} \kappa^{4/3} & \text{for } \kappa \rightarrow 0, \\ \frac{1}{2} - 4(1 - \kappa) & \text{for } \kappa \rightarrow 1, \end{cases} \quad (71)$$



**Figure 4.** Variance of the rescaled truncated linear statistics (43) as a function of  $\kappa = K/N$ , obtained from the parametric form (48,69,71) (solid line). The points are obtained from the numerical computation of the variance of (3) by averaging over 20000 realisations of  $200 \times 200$  matrices in the Gaussian Unitary Ensemble ( $\beta = 2$ ).

with  $\beta = 2$  here to describe the system of fermions. This variance is represented as a function of  $\kappa$  in Fig. 4. It displays a non-monotonic behaviour, which has never been observed in previous studies on truncated linear statistics [47, 52, 61]. This new feature has been confirmed by numerical simulations (see Fig. 4).

*Behaviour for  $s \rightarrow 0$*  — Looking at the phase diagram in Fig. 2, we see that the limit  $s \rightarrow 0$  is reached in Phase I. We thus start from the expressions obtained in this case. From the expression of the truncated linear statistics (43), we see that this limit corresponds to  $\lambda_n \rightarrow 0$  for all  $n \leq K$ . Therefore, for the optimal density  $\rho_*$  it corresponds to  $c \rightarrow 0$  and  $d \rightarrow 0$ . In order to get the leading behaviour of the large deviation function, we only need to expand Eq. (53), which imposes  $c = -d$ , and then expand the constraints (37), which give

$$\kappa \simeq \mu_1 \frac{d^2}{2}, \quad s \simeq \mu_1 \frac{d^4}{8}. \quad (72)$$

We straightforwardly deduce that

$$\mu_1 \underset{s \rightarrow 0}{\simeq} \frac{\kappa^2}{2s} + \mathcal{O}(1), \quad (73)$$

which yields the behaviour

$$\Phi_\kappa(s) \underset{s \rightarrow 0}{\simeq} -\frac{\kappa^2}{2} \ln s + \mathcal{O}(1), \quad (74)$$

by direct integration, thanks to the thermodynamic identity (23).

*Behaviour for  $s \rightarrow +\infty$*  — From the phase diagram in Fig. 2, the limit  $s \rightarrow \infty$  is also reached in Phase I. But this time, it corresponds to making the eigenvalues  $\lambda_n$  for  $n \leq K$  large. For the density  $\rho_*$ , it corresponds to  $c, d \rightarrow \infty$ . Expanding Eqs. (51,52,53), as well as the constraints (37), we obtain:

$$1 + \frac{b-a}{8}(3a+b)\sqrt{\frac{c}{d}} + (1+\mu_1)\frac{d-c}{8}(3c+d) = 0, \quad (75)$$

$$\frac{b-a}{8}\frac{3a+b}{\sqrt{cd}} + (1+\mu_1)\frac{c+d}{2} = 0, \quad (76)$$

$$b = -a, \quad (77)$$

$$\kappa = (1+\mu_1)\frac{(d-c)^2}{8}, \quad (78)$$

$$s = (1+\mu_1)\frac{(d-c)^2}{128}(5c^2 + 6cd + 5d^2). \quad (79)$$

Combining these expressions, we deduce that

$$\mu_1 = -1 + \frac{\kappa(2-\kappa)}{2s} + \mathcal{O}(s^{-2}), \quad (80)$$

which yields the behaviour

$$\Phi_\kappa(s) \underset{s \rightarrow +\infty}{\simeq} s - \frac{\kappa(2-\kappa)}{2} \ln s + \mathcal{O}(1), \quad (81)$$

for the large deviation function, after using again the thermodynamic identity (23).

A sketch of the distribution of the truncated linear statistics (43) is shown in Fig. 1, with all the behaviours identified in this Section. Unlike the case of the monotonous linear statistics studied in Ref. [47], both asymptotic behaviours  $s \rightarrow 0$  and  $s \rightarrow \infty$  depend explicitly on the fraction  $\kappa$  of eigenvalues under consideration. Indeed, in Ref. [47], these two limits corresponded to two different phases of the Coulomb gas (equivalent to Phase I and Phase II here). Here, they are both reached in Phase I, in which the  $K$  eigenvalues of interest are separated from the others and yield the dominating contribution to the energy of the Coulomb gas (and thus to the large deviation function  $\Phi_\kappa$ ).

Finally, for  $\kappa = 1$ , we recover the known results for the full linear statistics [70].

#### 4. Universality

In Section 3 we have applied the Coulomb gas formalism to the study of an example of truncated linear statistics with a non monotonous function  $f$ , motivated by the study of a gas of cold fermions. We now argue that several features of the Coulomb gas, and thus of the large deviation function, are actually universal in the sense that they do not depend on the specific choice of the linear statistics (i.e. the function  $f$ ) or on the matrix ensemble (i.e. the potential  $V$ ). Let us again denote  $\rho_0$  the typical density of eigenvalues in the ensemble under consideration, and  $c_0(\kappa) = \langle \lambda_{\kappa N} \rangle$  the position of the

smallest eigenvalue to contribute to the truncated linear statistics (3). This position can be determined from  $\rho_0$  via the relation

$$\kappa = \int_{c_0(\kappa)} \rho_0(x) dx. \tag{82}$$

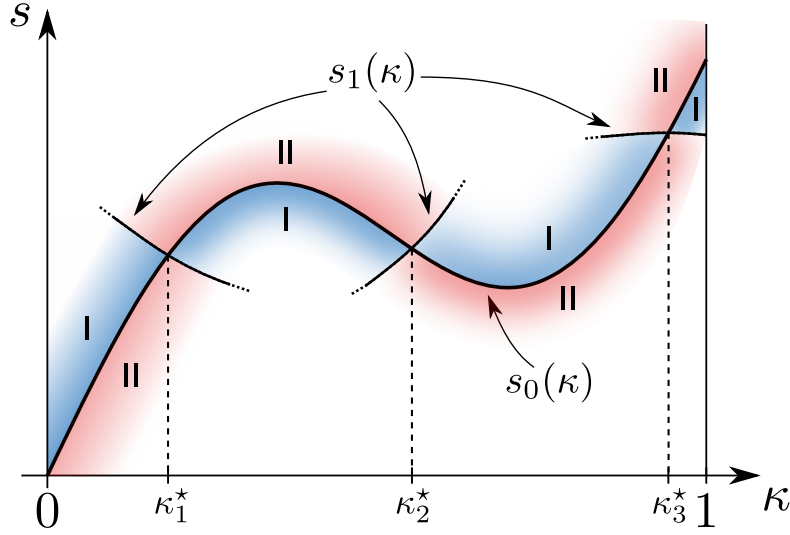
The following features are expected to occur in the study of any truncated linear statistics.

- The line  $s_0(\kappa)$ , corresponding to  $\mu_1 = 0$ , gives the typical value of the truncated linear statistics restricted to the fraction  $\kappa$  of the largest eigenvalues. It is shown in [Appendix B](#) that this line also corresponds to a transition line between two phases: a phase in which the density is supported on two disjoint supports (Phase I), and one in which the density exhibits a logarithmic divergence (Phase II). Moreover, it corresponds to an infinite order phase transition for all values of  $\kappa$  such that  $f'(c_0(\kappa)) \neq 0$ . This extends the results of Ref. [47] to the case of a non-monotonous function  $f$ .
- The line  $s_1(\kappa)$  is also present for any non-monotonous function  $f$ , at least in the vicinity of the typical line  $s_0(\kappa)$ . Indeed, away from this line, the two phases studied in this article could stop to exist and let other configurations of the Coulomb gas emerge (such as densities with more than two supports). Nevertheless, for values of  $s$  close to  $s_0(\kappa)$ , the line  $s_1(\kappa)$  is expected to exist, and is also an infinite order phase transition for the Coulomb gas, as is shown in [Appendix B](#).
- The two lines  $s_0(\kappa)$  and  $s_1(\kappa)$  intersect for values of  $\kappa$  which verify  $f'(c_0(\kappa)) = 0$ . There are as many intersections as local extrema of the function  $f$  in the support of  $\rho_0$ . At these points, there is no longer a phase transition in the Coulomb gas, as it remains in the same phase both for  $s < s_0(\kappa)$  and  $s > s_0(\kappa)$ .
- The positions of Phase I and Phase II with respect to the line  $s_0(\kappa)$  are exchanged after a crossing with the line  $s_1(\kappa)$ .

We show in Fig. 5 a sketch of the corresponding phase diagram of the Coulomb gas, in the vicinity of the line  $s_0(\kappa)$ , in the case where the lines  $s_0(\kappa)$  and  $s_1(\kappa)$  intersect for 3 different values of  $\kappa$ .

## 5. Conclusion

We have studied the distribution of the truncated linear statistics (3) with  $f(\lambda) = \lambda^2$  in the Gaussian ensembles (Eq. (1) with  $V(\lambda) = \lambda^2$ ) in the large  $N$  limit, with  $\kappa = \frac{K}{N}$  fixed, using the Coulomb gas method. We have shown that, for  $\kappa \neq \frac{1}{2}$ , the large deviation function admits two essential singularities, at  $s = s_0(\kappa)$  and  $s = s_1(\kappa)$ , which result from two infinite order phase transitions for the Coulomb gas. For the specific value  $\kappa = \frac{1}{2}$ , these phase transitions merge and disappear. We have argued that this picture is universal, and holds for any choice of linear statistics  $f$  or invariant matrix ensemble (1), at least in the vicinity of the typical value  $s_0(\kappa)$  of the linear statistics.



**Figure 5.** Sketch of the phase diagram for a truncated linear statistics for which  $f$  has 3 local extrema on the support of the typical density of eigenvalues  $\rho_0$ . The line  $s_0(\kappa)$  corresponds to the typical value taken by the truncated linear statistics  $s$ , i.e.  $\mu_1 = 0$ . The second line  $s_1(\kappa)$  is present only for non-monotonous functions  $f$ . It corresponds to  $f'(c) = 0$ . Note that Phase I and Phase II have been placed arbitrarily, and their precise positions depend on the specific choice of the function  $f$ .

As it was already noticed in Ref. [47], the limits  $\kappa \rightarrow 0$  and  $\kappa \rightarrow 1$  are singular. Indeed, looking for instance at  $\kappa \rightarrow 1$  on the phase diagram in Fig. 2, we would expect a phase transition at  $s = s_0(1) = \frac{1}{2}$ , and thus an essential singularity in the distribution of  $s$  for  $\kappa = 1$ . However, that is not the case as this distribution is analytic, as easily seen on Eq. (44). This is due to the fact that the two limits  $\kappa \rightarrow 1$  and  $N \rightarrow \infty$  do not commute. This motivates a more detailed study of this joint limit, in order to understand how the truncated linear statistics reduces to the full linear statistics when  $\kappa \rightarrow 1$ .

The singularities observed on the large deviation function are expected to be regularised by a smooth function when looking on a scale  $s - s_0(\kappa) = \mathcal{O}(N^{-\eta})$  with an exponent  $\eta > 0$  (and similarly for  $s_1(\kappa)$ ). The most famous example is the case of the largest eigenvalue: its large deviation function displays a singularity which originates from a third order phase transition for the Coulomb gas when looking at fluctuations of order  $\lambda_{\max} - \langle \lambda_{\max} \rangle = \mathcal{O}(1)$  (for the distribution (1) for which  $\lambda_n = \mathcal{O}(N^0)$ ). This singularity is regularised when looking on a scale  $\Delta\lambda = \lambda_{\max} - \langle \lambda_{\max} \rangle = \mathcal{O}(N^{-2/3})$ , where  $\Delta\lambda$  obeys the Tracy-Widom distribution [59]. We expect an equivalent function here that regularises the distribution of the truncated linear statistics at  $s_0(\kappa)$  and  $s_1(\kappa)$ , when looking on a smaller scale. It would be of particular interest to study the behaviour of this function in vicinity of the point(s) where the two phase transition lines  $s_0(\kappa)$  and  $s_1(\kappa)$  intersect, as there is no singular behaviour at this point. It could thus help to understand how these essential singularities emerge in the distribution of



truncated linear statistics.

### Appendix A. A few useful integrals

The following integrals are used in the manuscript.

$$\int_a^b \frac{dt}{\pi} \frac{\sqrt{(t-a)(b-t)}}{t-x} \frac{1}{t-y} = -1 + \frac{\sqrt{(y-a)(y-b)}}{|y-x|}, \quad x \in [a, b], y \notin [a, b], \quad (\text{A.1})$$

$$\int_a^b \frac{dt}{\pi} \frac{1}{(x-t)\sqrt{(t-a)(b-t)}} = \begin{cases} -\frac{1}{\sqrt{(a-x)(b-x)}}, & x < a \\ 0, & a < x < b \\ \frac{1}{\sqrt{(x-a)(x-b)}}, & x > b \end{cases} \quad (\text{A.2})$$

### Appendix B. Infinite order phase transitions

We interpret the two lines  $s_0(\kappa)$  and  $s_1(\kappa)$  as phase transitions lines for the Coulomb gas. We now show that these transitions are of infinite order, for any invariant matrix ensemble (choice of  $V$ ) and any linear statistics ( $f$ ). A similar analysis was performed in Ref. [47] in the Laguerre ensembles (see Table 1), and for a monotonous function  $f$ . We extend this discussion to the general case, and in particular to the study of the line  $s_1(\kappa)$  which emerges due to the non-monotony of  $f$ .

Let us start from the general expression of the density (36). It depends on five parameters:  $a, b, c, d$  and  $\mu_1$ . These parameters are determined by the two constraints (37), along with conditions on the boundaries  $a, b$  and  $d$  of the support (the condition at  $x = c$  has already been used in the derivation of the expression (36) for the density). The determination of the value of  $b$  depends on the phase we consider:

- In Phase I in which the supports are disjoint ( $b < c$ ), the value of  $b$  is determined by imposing that the density does not diverge as  $(b-x)^{-1/2}$  for  $x \rightarrow b^-$  (as it was done for  $c$ ), since this kind of behaviour only appears near a hard edge.

This gives

$$2 = \int_a^b \frac{dt}{\pi} V'(t) \sqrt{\frac{(t-a)(d-t)}{(b-t)(c-t)}} - \int_c^d \frac{dt}{\pi} (V'(t) + \mu_1 f'(t)) \sqrt{\frac{(t-a)(d-t)}{(t-b)(t-c)}}. \quad (\text{B.1})$$

- In Phase II,  $b$  is directly obtained as  $b = c$ .

There remains to determine the outer edges  $a$  and  $d$ . There are different possibilities, depending on the random matrix ensemble under consideration (and hence on the choice of  $V$ ). In the main text, since we worked in the Gaussian ensembles, we imposed that the density vanishes at  $x = a$  and  $x = d$ . In general, these conditions take the form

$$2 + \int_{[a,b] \cup [c,d]} \frac{dt}{\pi} (V'(t) + \mu_1 f'(t)) \Theta(t-c) \sqrt{\frac{(b-t)(d-t)}{(t-a)(c-t)}} = 0, \quad (\text{B.2})$$

$$2 - \int_{[a,b] \cup [c,d]} \frac{dt}{\pi} (V'(t) + \mu_1 f'(t) \Theta(t-c)) \sqrt{\frac{(b-t)(t-a)}{(c-t)(d-t)}} = 0. \quad (\text{B.3})$$

Note that other types of boundary conditions can apply. For instance, in the Jacobi ensembles in which the eigenvalues are restricted to  $[0, 1]$  (see Table 1), we could have instead the conditions  $a = 0$  and  $d = 1$ . In the following, we will consider only the case where the density vanishes at  $x = a$  and  $x = d$ , and thus we impose (B.2, B.3), since the following procedure can be straightforwardly adapted to the other types of boundary conditions.

In order to analyse the order of the two transitions occurring at  $s = s_0(\kappa)$  and  $s = s_1(\kappa)$ , we study the behaviour of

$$\Phi_\kappa(s_0(\kappa) + \varepsilon) - \Phi_\kappa(s_0(\kappa) - \varepsilon) \quad \text{and} \quad \Phi_\kappa(s_1(\kappa) + \varepsilon) - \Phi_\kappa(s_1(\kappa) - \varepsilon) \quad (\text{B.4})$$

for  $\varepsilon \rightarrow 0$ . The two lines  $s_0(\kappa)$  and  $s_1(\kappa)$  corresponding to transitions between Phase I and Phase II, the limit  $\varepsilon \rightarrow 0$  in Phase I corresponds to the limit  $b \rightarrow c$  in both cases. Expanding the conditions on  $a$ ,  $b$  and  $d$  (B.1, B.2, B.3) in this limit, we obtain for Phase I:

$$\mu_1 f'(c) = \frac{\alpha}{\ln(c-b)} \quad \Rightarrow \quad c - b = e^{\alpha/(\mu_1 f'(c))}, \quad (\text{B.5})$$

$$2 + \int_a^d \frac{dt}{\pi} (V'(t) + \mu_1 f'(t) \Theta(t-c)) \sqrt{\frac{d-t}{t-a}} = \mathcal{O}\left(\frac{e^{\alpha/(\mu_1 f'(c))}}{\mu_1 f'(c)}\right), \quad (\text{B.6})$$

$$2 - \int_a^d \frac{dt}{\pi} (V'(t) + \mu_1 f'(t) \Theta(t-c)) \sqrt{\frac{t-a}{d-t}} = \mathcal{O}\left(\frac{e^{\alpha/(\mu_1 f'(c))}}{\mu_1 f'(c)}\right), \quad (\text{B.7})$$

where we have denoted

$$\alpha = \frac{\pi}{\sqrt{(c-a)(d-c)}} \left( 2 + \int_a^d \frac{dt}{\pi} \frac{V'(t)}{t-c} \sqrt{(t-a)(d-t)} \right). \quad (\text{B.8})$$

Similarly, using the expression of the density (36), the constraints (37) take the form:

$$\begin{aligned} \kappa = \int_c^d \frac{dx}{2\pi} \frac{1}{\sqrt{(x-a)(d-x)}} \left\{ 2 + \int_a^d \frac{dt}{\pi} \frac{V'(t)}{t-x} \sqrt{(t-a)(d-t)} \right. \\ \left. + \mu_1 \int_c^d \frac{dt}{\pi} \frac{f'(t)}{t-x} \sqrt{(t-a)(d-t)} \right\} + \mathcal{O}\left(\frac{e^{\alpha/(\mu_1 f'(c))}}{\mu_1 f'(c)}\right), \quad (\text{B.9}) \end{aligned}$$

$$\begin{aligned} s = \int_c^d \frac{dx}{2\pi} \frac{f(x)}{\sqrt{(x-a)(d-x)}} \left\{ 2 + \int_a^d \frac{dt}{\pi} \frac{V'(t)}{t-x} \sqrt{(t-a)(d-t)} \right. \\ \left. + \mu_1 \int_c^d \frac{dt}{\pi} \frac{f'(t)}{t-x} \sqrt{(t-a)(d-t)} \right\} + \mathcal{O}\left(\frac{e^{\alpha/(\mu_1 f'(c))}}{\mu_1 f'(c)}\right). \quad (\text{B.10}) \end{aligned}$$

In Phase II, we obtain exactly the same expressions but without the terms in  $\mathcal{O}(e^{\alpha/(\mu_1 f'(c))}/(\mu_1 f'(c)))$ . From these expressions, we now consider each transition line independently.

*Line  $s_0(\kappa)$ :*  $\mu_1 = 0$  — Let us first study the transition occurring on the line  $s_0(\kappa)$ . It corresponds to  $\mu_1 = 0$  and hence the most probable value of the linear statistics. It is given in a parametric form by

$$s_0(\kappa) = \int_{c_0}^{d_0} f(x)\rho_0(x)dx, \quad \kappa = \int_{c_0}^{d_0} \rho_0(x)dx \quad (\text{B.11})$$

in terms of the position  $c_0 = \langle \lambda_K \rangle$  of the last eigenvalue to contribute to  $s$ , where we have denoted

$$\rho_0(x) = \frac{1}{\pi \sqrt{(x-a)(d-x)}} \left( 1 + \int_a^d \frac{dt}{2\pi} \frac{V'(t)}{t-x} \sqrt{(t-a)(d-t)} \right) \quad (\text{B.12})$$

the typical density of eigenvalues in the absence of constraint.

For values of  $c_0$  such that  $f'(c_0) \neq 0$ , Eqs. (B.6,B.7,B.9,B.10) are identical to the ones studied in Ref. [47] in the case of a monotonous linear statistics. The fact that we recover these exact same equations can be understood as follows: the small fluctuations of  $\lambda_K$  around the typical value  $c_0$  are not sufficient to probe the non-monotony of  $f$  if  $f'(c_0) \neq 0$ . Therefore, for these values, the line  $s_0(\kappa)$  corresponds to an infinite order phase transition, as it was shown in Ref. [47]. The idea to prove this result is to expand Eqs. (B.6,B.7,B.9,B.10) for  $s \rightarrow s_0(\kappa)$ , i.e. for  $\mu_1 \rightarrow 0$ , and combine them in order to get

$$s - s_0(\kappa) = F_0(\kappa)\mu_1 + \mathcal{O}(\mu_1^2) + \mathcal{O}\left(\frac{e^{\alpha/(\mu_1 f'(c_0))}}{\mu_1}\right) \quad (\text{Phase I}), \quad (\text{B.13})$$

$$s - s_0(\kappa) = F_0(\kappa)\mu_1 + \mathcal{O}(\mu_1^2) \quad (\text{Phase II}). \quad (\text{B.14})$$

where  $F_0(\kappa)$  is a constant fully determined by this expansion. For the case considered in the main text ( $f(x) = V(x) = x^2$ ),  $F_0$  is given parametrically by (48,69). For the expansion in Phase I, we have kept the subleading term  $\mathcal{O}\left(\frac{e^{\alpha/(\mu_1 f'(c_0))}}{\mu_1}\right)$ , as it is the only one which differs from the expansion in Phase II, since all the terms in  $\mathcal{O}(\mu_1^n)$  are identical in both expressions. Inverting these expansions, we obtain  $\mu_1$  as a function of  $s$ , which can be integrated to yield  $\Phi_\kappa(s)$  by using the thermodynamic identity (24). This gives

$$\frac{d}{d\varepsilon} [\Phi_\kappa(s_0(\kappa) + \varepsilon) - \Phi_\kappa(s_0(\kappa) - \varepsilon)] = \mathcal{O}(e^{\gamma_0(\kappa)/\varepsilon}/\varepsilon), \quad (\text{B.15})$$

which after integration becomes

$$\Phi_\kappa(s_0(\kappa) + \varepsilon) - \Phi_\kappa(s_0(\kappa) - \varepsilon) = \mathcal{O}(\varepsilon e^{\gamma_0(\kappa)/\varepsilon}), \quad (\text{B.16})$$

where we have denoted

$$\gamma_0(\kappa) = \frac{\alpha}{f'(c_0)} F_0(\kappa), \quad (\text{B.17})$$

with  $c_0$  determined from  $\kappa$  via (B.11). This proves that all the derivatives of  $\Phi_\kappa$  are continuous on the line  $s_0(\kappa)$ , but the function is not analytic at this point, due to the essential singularity present in Phase I. However, for specific values of  $\kappa$  such that  $f'(c_0) = 0$  this singularity vanishes and there is no transition. Indeed, the optimal density of eigenvalues is given by the one of Phase I in both cases  $s > s_0(\kappa)$  and  $s < s_0(\kappa)$  (see for instance Fig. 2 for  $\kappa = \frac{1}{2}$ ).

*Line  $s_1(\kappa)$ :  $f'(c) = 0$*  — We now turn to the case of the second transition line, which is specific to the study of non-monotonous truncated linear statistics. The behaviour of the large deviation function for  $s$  near  $s_1(\kappa)$  can be obtained in Phase I from (B.6,B.7,B.9,B.10) by expanding these equations around  $c = \check{c}$  such that  $f'(\check{c}) = 0$ . For  $c = \check{c}$ , we have  $\mu_1 = \check{\mu}_1$ . For  $\check{\mu}_1 \neq 0$ , we can combine these expansions in order to write (B.10) as

$$s - s_1(\kappa) = F_1(\kappa)(\mu_1 - \check{\mu}_1) + \mathcal{O}((\mu_1 - \check{\mu}_1)^2) + \mathcal{O}\left(\frac{e^{\tilde{\gamma}_1(\kappa)/(\mu_1 - \check{\mu}_1)}}{\mu_1 - \check{\mu}_1}\right) \quad (\text{Phase I}), \quad (\text{B.18})$$

where  $F_1(\kappa)$  and  $\tilde{\gamma}_1(\kappa)$  are two constants. We have kept the subleading last term, as for Phase II, we have the same expansion, but without this last term:

$$s - s_1(\kappa) = F_1(\kappa)(\mu_1 - \check{\mu}_1) + \mathcal{O}((\mu_1 - \check{\mu}_1)^2) \quad (\text{Phase II}). \quad (\text{B.19})$$

In these two expansions all the regular powers  $\mathcal{O}((\mu_1 - \check{\mu}_1)^n)$  are identical, the only difference is the essential singularity present in Phase I only. The large deviations function can be deduced from these expansions via the thermodynamic identity (23), which yields

$$\frac{d}{d\varepsilon} [\Phi_\kappa(s_1(\kappa) + \varepsilon) - \Phi_\kappa(s_1(\kappa) - \varepsilon)] = \mathcal{O}(e^{\gamma_1(\kappa)/\varepsilon}/\varepsilon), \quad (\text{B.20})$$

which after integration gives

$$\Phi_\kappa(s_1(\kappa) + \varepsilon) - \Phi_\kappa(s_1(\kappa) - \varepsilon) = \mathcal{O}(\varepsilon e^{\gamma_1(\kappa)/\varepsilon}), \quad (\text{B.21})$$

where we have denoted  $\gamma_1(\kappa) = F_1(\kappa)\tilde{\gamma}_1(\kappa)$ . This proves that the line  $s_1(\kappa)$  is also associated to an infinite order phase transition.

## References

- [1] E. P. Wigner, On the statistical distribution of the widths and spacings of nuclear resonance levels, *Math. Proc. Cambridge Philos. Soc.* **47**(4), 790–798 (1951).
- [2] C. W. J. Beenakker, Random-matrix theory of quantum transport, *Rev. Mod. Phys.* **69**, 731–808 (1997).
- [3] T. Guhr, A. Müller-Groeling, and H. A. Weidenmüller, Random-matrix theories in quantum physics: common concepts, *Phys. Rep.* **299**(4/6), 189–425 (1998).
- [4] I. L. Aleiner, P. W. Brouwer, and L. I. Glazman, Quantum effects in Coulomb blockade, *Phys. Rep.* **358**(5-6), 309–440 (2002).
- [5] P. A. Mello and N. Kumar, *Quantum transport in mesoscopic systems – Complexity and statistical fluctuations*, Oxford University Press, 2004.
- [6] P. W. Brouwer, Generalized circular ensemble of scattering matrices for a chaotic cavity with nonideal leads, *Phys. Rev. B* **51**, 16878–16884 (1995).
- [7] P. A. Mello and H. U. Baranger, Interference phenomena in electronic transport through chaotic cavities: An information-theoretic approach, *Waves Random Media* **9**, 105–162 (1999).
- [8] H.-J. Sommers, W. Wicczorek, and D. V. Savin, Statistics of conductance and shot noise power for chaotic cavities, *Acta Phys. Pol. A* **112**, 691 (2007).
- [9] P. Vivo, S. N. Majumdar, and O. Bohigas, Distributions of Conductance and Shot Noise and Associated Phase Transitions, *Phys. Rev. Lett.* **101**, 216809 (2008).
- [10] B. A. Khoruzhenko, D. V. Savin, and H.-J. Sommers, Systematic approach to statistics of conductance and shot-noise in chaotic cavities, *Phys. Rev. B* **80**, 125301 (2009).

- [11] P. Vivo, S. N. Majumdar, and O. Bohigas, Probability distributions of linear statistics in chaotic cavities and associated phase transitions, *Phys. Rev. B* **81**, 104202 (2010).
- [12] P. Vivo and E. Vivo, Transmission eigenvalue densities and moments in chaotic cavities from random matrix theory, *J. Phys. A: Math. Theor.* **41**(12), 122004 (2008).
- [13] A. Grabsch and C. Texier, Capacitance and charge relaxation resistance of chaotic cavities—Joint distribution of two linear statistics in the Laguerre ensemble of random matrices, *Europhys. Lett.* **109**(5), 50004 (2015).
- [14] F. D. Cunden, P. Facchi, and P. Vivo, Joint statistics of quantum transport in chaotic cavities, *Europhys. Lett.* **110**(5), 50002 (2015).
- [15] D. N. Page, Average entropy of a subsystem, *Phys. Rev. Lett.* **71**, 1291–1294 (1993).
- [16] P. Facchi, U. Marzolino, G. Parisi, S. Pascazio, and A. Scardicchio, Phase Transitions of Bipartite Entanglement, *Phys. Rev. Lett.* **101**, 050502 (2008).
- [17] A. De Pasquale, P. Facchi, G. Parisi, S. Pascazio, and A. Scardicchio, Phase transitions and metastability in the distribution of the bipartite entanglement of a large quantum system, *Phys. Rev. A* **81**, 052324 (2010).
- [18] C. Nadal, S. N. Majumdar, and M. Vergassola, Phase Transitions in the Distribution of Bipartite Entanglement of a Random Pure State, *Phys. Rev. Lett.* **104**, 110501 (2010).
- [19] C. Nadal, S. N. Majumdar, and M. Vergassola, Statistical Distribution of Quantum Entanglement for a Random Bipartite State, *J. Stat. Phys.* **142**(2), 403–438 (2011).
- [20] P. Facchi, G. Florio, G. Parisi, S. Pascazio, and K. Yuasa, Entropy-driven phase transitions of entanglement, *Phys. Rev. A* **87**, 052324 (2013).
- [21] C. Nadal and S. N. Majumdar, Nonintersecting Brownian interfaces and Wishart random matrices, *Phys. Rev. E* **79**, 061117 (2009).
- [22] C. Nadal, *Matrices aléatoires et leurs applications à la physique statistique et physique quantique*, PhD thesis, Université Paris-Sud, 2011.
- [23] R. Marino, S. N. Majumdar, G. Schehr, and P. Vivo, Number statistics for  $\beta$ -ensembles of random matrices: applications to trapped fermions at zero temperature, *Phys. Rev. E* **94**, 032115 (2016).
- [24] D. S. Dean, P. L. Doussal, S. N. Majumdar, and G. Schehr, Noninteracting fermions in a trap and random matrix theory, *J. Phys. A: Math. Theor.* **52**(14), 144006 (2019).
- [25] M. L. Mehta, *Random matrices*, Elsevier, Academic, New York, third edition, 2004.
- [26] P. J. Forrester, *Log-gases and random matrices*, Princeton University Press, 2010.
- [27] G. Akemann, J. Baik, and P. Di Francesco, *The Oxford handbook of random matrix theory*, Oxford University Press, 2011.
- [28] I. Dumitriu and A. Edelman, Matrix models for beta ensembles, *J. Math. Phys.* **43**(11), 5830–5847 (2002).
- [29] S. N. Majumdar, C. Nadal, A. Scardicchio, and P. Vivo, Index distribution of Gaussian random matrices, *Phys. Rev. Lett.* **103**, 220603 (2009).
- [30] S. N. Majumdar, C. Nadal, A. Scardicchio, and P. Vivo, How many eigenvalues of a Gaussian random matrix are positive?, *Phys. Rev. E* **83**, 041105 (2011).
- [31] S. N. Majumdar and P. Vivo, Number of Relevant Directions in Principal Component Analysis and Wishart Random Matrices, *Phys. Rev. Lett.* **108**, 200601 (2012).
- [32] R. Marino, S. N. Majumdar, G. Schehr, and P. Vivo, Index distribution of Cauchy random matrices, *J. Phys. A: Math. Theor.* **47**(5), 055001 (2014).
- [33] R. Marino, S. N. Majumdar, G. Schehr, and P. Vivo, Phase Transitions and Edge Scaling of Number Variance in Gaussian Random Matrices, *Phys. Rev. Lett.* **112**, 254101 (2014).
- [34] P. Kazakopoulos, P. Mertikopoulos, A. L. Moustakas, and G. Caire, Living at the Edge: A Large Deviations Approach to the Outage MIMO Capacity, *IEEE Trans. Info. Theo.* **57**(4), 1984–2007 (2011).
- [35] A. Karadimitrakakis, A. L. Moustakas, and P. Vivo, Outage Capacity for the Optical MIMO Channel, *IEEE Trans. Info. Theo.* **60**(7), 4370–4382 (2014).
- [36] F. J. Dyson and M. L. Mehta, Statistical Theory of the Energy Levels of Complex Systems. IV,

- J. Math. Phys.* **4**(5), 701–712 (1963).
- [37] C. W. J. Beenakker, Universality in the random-matrix theory of quantum transport, *Phys. Rev. Lett.* **70**, 1155–1158 (1993).
- [38] C. W. J. Beenakker, Random-matrix theory of mesoscopic fluctuations in conductors and superconductors, *Phys. Rev. B* **47**, 15763–15775 (1993).
- [39] C. Beenakker, Universality of Brézin and Zee’s spectral correlator, *Nucl. Phys. B* **422**(3), 515 – 520 (1994).
- [40] E. L. Basor and C. A. Tracy, Variance calculations and the Bessel kernel, *J. Stat. Phys.* **73**(1), 415–421 (1993).
- [41] B. Jancovici and P. J. Forrester, Derivation of an asymptotic expression in Beenakker’s general fluctuation formula for random-matrix correlations near an edge, *Phys. Rev. B* **50**, 14599–14600 (1994).
- [42] D. S. Dean and S. N. Majumdar, Large Deviations of Extreme Eigenvalues of Random Matrices, *Phys. Rev. Lett.* **97**, 160201 (2006).
- [43] P. Vivo, S. N. Majumdar, and O. Bohigas, Large deviations of the maximum eigenvalue in Wishart random matrices, *J. Phys. A: Math. Theor.* **40**(16), 4317 (2007).
- [44] F. J. Dyson, Statistical Theory of the Energy Levels of Complex Systems. I, *J. Math. Phys.* **3**(1), 140–156 (1962),  
 F. J. Dyson, Statistical Theory of the Energy Levels of Complex Systems. II, *J. Math. Phys.* **3**(1), 157–165 (1962),  
 F. J. Dyson, Statistical Theory of the Energy Levels of Complex Systems. III, *J. Math. Phys.* **3**(1), 166–175 (1962).
- [45] G. Ben Arous and A. Guionnet, Large deviations for Wigner’s law and Voiculescu’s non-commutative entropy, *Prob. Theo. Relat. Fields* **108**(4), 517–542 (1997).
- [46] G. Ben Arous and O. Zeitouni, Large deviations from the circular law, *ESAIM: Prob. Stat.* **2**, 123–134 (1998).
- [47] A. Grabsch, S. N. Majumdar, and C. Texier, Truncated Linear Statistics Associated with the Top Eigenvalues of Random Matrices, *J. Stat. Phys.* **167**(2), 234–259 (2017), updated version [arXiv:1609.08296](https://arxiv.org/abs/1609.08296).
- [48] O. Bohigas and M. P. Pato, Randomly incomplete spectra and intermediate statistics, *Phys. Rev. E* **74**, 036212 (2006).
- [49] T. Berggren and M. Duits, Mesoscopic Fluctuations for the Thinned Circular Unitary Ensemble, *Math. Phys. Anal. Geom.* **20**(3), 19 (2017).
- [50] C. Charlier and T. Claeys, Thinning and conditioning of the circular unitary ensemble, *Random Matrices: Theory Appl.* **06**(02), 1750007 (2017).
- [51] G. Lambert, Incomplete determinantal processes: from random matrix to Poisson statistics, [preprint arXiv:1612.00806](https://arxiv.org/abs/1612.00806) (2016).
- [52] A. Grabsch, S. N. Majumdar, and C. Texier, Truncated Linear Statistics Associated with the Eigenvalues of Random Matrices II. Partial Sums over Proper Time Delays for Chaotic Quantum Dots, *J. Stat. Phys.* **167**(6), 1452–1488 (2017).
- [53] L. I. Smith, A tutorial on principal components analysis, 2002.
- [54] C. A. Tracy and H. Widom, Level-spacing distributions and the Airy kernel, *Commun. Math. Phys.* **159**(1), 151–174 (1994).
- [55] C. A. Tracy and H. Widom, On orthogonal and symplectic matrix ensembles, *Commun. Math. Phys.* **177**(3), 727–754 (1996).
- [56] K. Johansson, From Gumbel to Tracy-Widom, *Probab. Theory Rel.* **138**(1), 75–112 (2007).
- [57] D. S. Dean and S. N. Majumdar, Extreme value statistics of eigenvalues of Gaussian random matrices, *Phys. Rev. E* **77**, 041108 (2008).
- [58] G. Borot, B. Eynard, S. N. Majumdar, and C. Nadal, Large deviations of the maximal eigenvalue of random matrices, *J. Stat. Mech: Theory Exp.* **2011**(11), P11024 (2011).
- [59] S. N. Majumdar and G. Schehr, Top eigenvalue of a random matrix: large deviations and third

- order phase transition, *J. Stat. Mech.* **2014**(1), P01012 (2014).
- [60] A. Krajenbrink and P. Le Doussal, Linear statistics and pushed Coulomb gas at the edge of  $\beta$ -random matrices: Four paths to large deviations, *Europhys. Lett.* **125**(2), 20009 (2019).
- [61] A. Flack, S. N. Majumdar, and G. Schehr, Truncated linear statistics in the one dimensional one-component plasma, *J. Phys. A: Math. Theor.* **54**(43), 435002 (2021).
- [62] F. J. Dyson, The Threefold Way. Algebraic Structure of Symmetry Groups and Ensembles in Quantum Mechanics, *J. Math. Phys.* **3**, 1199–1215 (1962).
- [63] F. D. Cunden, P. Facchi, and P. Vivo, A shortcut through the Coulomb gas method for spectral linear statistics on random matrices, *J. Phys. A Math. Theor.* **49**(13), 135202 (2016).
- [64] A. Grabsch and C. Texier, Distribution of spectral linear statistics on random matrices beyond the large deviation function – Wigner time delay in multichannel disordered wires, *J. Phys. A: Math. Theor.* **49**, 465002 (2016).
- [65] F. G. Tricomi, *Integral equations*, Interscience, London, 1957.
- [66] D. S. Dean, P. L. Doussal, S. N. Majumdar, and G. Schehr, Universal ground-state properties of free fermions in a  $d$ -dimensional trap, *Europhys. Lett.* **112**(6), 60001 (2015).
- [67] D. S. Dean, P. Le Doussal, S. N. Majumdar, and G. Schehr, Finite-Temperature Free Fermions and the Kardar-Parisi-Zhang Equation at Finite Time, *Phys. Rev. Lett.* **114**, 110402 (2015).
- [68] D. S. Dean, P. Le Doussal, S. N. Majumdar, and G. Schehr, Noninteracting fermions at finite temperature in a  $d$ -dimensional trap: Universal correlations, *Phys. Rev. A* **94**, 063622 (2016).
- [69] A. Grabsch, S. N. Majumdar, G. Schehr, and C. Texier, Fluctuations of observables for free fermions in a harmonic trap at finite temperature, *SciPost Phys.* **4**, 014 (2018).
- [70] J. Grella, S. N. Majumdar, and G. Schehr, Kinetic Energy of a Trapped Fermi Gas at Finite Temperature, *Phys. Rev. Lett.* **119**, 130601 (2017).