

Universality of local spectral statistics of products of random matrices

Gernot Akemann ^{*}*Faculty of Physics, Bielefeld University, Postfach 100131, D-33501 Bielefeld, Germany*Zdzislaw Burda [†]*Faculty of Physics and Applied Computer Science, AGH University of Science and Technology, al. Mickiewicza 30, PL-30059 Krakow, Poland*Mario Kieburg [‡]*School of Mathematics and Statistics, University of Melbourne, 813 Swanston Street, Parkville, Melbourne VIC 3010, Australia*

(Received 7 September 2020; accepted 8 November 2020; published 30 November 2020)

We derive exact analytical expressions for correlation functions of singular values of the product of M Ginibre matrices of size N in the double scaling limit $M, N \rightarrow \infty$. The singular value statistics is described by a determinantal point process with a kernel that interpolates between Gaussian unitary ensemble statistic and Dirac-delta (picket-fence) statistic. In the thermodynamic limit $N \rightarrow \infty$, the interpolation parameter is given by the limiting quotient $a = N/M$. One of our goals is to find an explicit form of the kernel at the hard edge, in the bulk, and at the soft edge for any a . We find that in addition to the standard scaling regimes, there is a transitional regime which interpolates between the hard edge and the bulk. We conjecture that these results are universal, and that they apply to a broad class of products of random matrices from the Gaussian basin of attraction, including correlated matrices. We corroborate this conjecture by numerical simulations. Additionally, we show that the local spectral statistics of the considered random matrix products is identical with the local statistics of Dyson Brownian motion with the initial condition given by equidistant positions, with the crucial difference that this equivalence holds only locally. Finally, we have identified a mesoscopic spectral scale at the soft edge which is crucial for the unfolding of the spectrum.

DOI: [10.1103/PhysRevE.102.052134](https://doi.org/10.1103/PhysRevE.102.052134)

I. INTRODUCTION

Statistical properties of random matrix ensembles have been intensively studied over years and a deep understanding of the underlying principles has been achieved. Much less is known on the principles governing random matrix dynamics. There are two notable exceptions. In the early days of random matrix theory, Dyson studied Brownian motion in matrix spaces [1]; especially, the evolution of the eigenvalues of Hermitian matrices were considered starting from specific or randomized initial conditions. The evolution is given by adding Hermitian matrices made of independently identically normal distributed matrix entries to this initial matrix. The kernel of the corresponding determinantal point process, describing the evolution of the eigenvalues in the bulk, with initial condition of a fixed spectrum, was derived in Ref. [2]. Recently, also the first step towards the formulation of a non-Hermitian version of Dyson's Brownian motion has been taken [3].

The second example is the Dorokhov-Mello-Pereyra-Kumar (DMPK) equation [4,5] for the joint-probability density function of transmission eigenvalues in a quantum

wire. This equation has the form of a Fokker-Planck equation describing Brownian motion of eigenvalues propagating in a narrow wire as a function of its length, which plays the role of time. The equation can be derived by constructing the transfer matrix for the whole wire as a product of independent transfer matrices for thin wire slices assuming isotropic propagation through each thin slice. The assumption of isotropy means that the flux in each ingoing channel is in average uniformly distributed among all outgoing channels. One obtains basically the same equation also for a multiplicative stochastic model in the context of May-Wigner stability [6]. A common feature of Dyson's Brownian motion and of quantum transport in a quantum wire is that, mathematically, both are formulated as evolution equations for eigenvalues in one dimension, which is the time or the length of the system, respectively.

In this work, we study yet another example of this type. It is an evolution of singular values of the product of random matrices. If one interprets these individual matrices as incremental transfer matrices for a time step Δt of a system with N degrees of freedom, the product of M matrices can be comprehended as a transfer matrix at time $M \Delta t$. We derive an exact analytical expression for the kernels of the determinantal point processes describing the local statistics of singular values at the hard edge, in the bulk and at the soft edge for the product of Ginibre matrices in the double scaling limit $N, M \rightarrow \infty$. Previous results were obtained either at fixed N, M when $N \rightarrow \infty$ (e.g., see [7–14]), leading to the Meijer

^{*}akemann@physik.uni-bielefeld.de[†]zdzislaw.burda@agh.edu.pl[‡]m.kieburg@unimelb.edu.au

G kernel at the hard edge [13] or the sine and Airy kernel in the bulk and at the soft edge [14], respectively. Or, the limit with N fixed with $M \rightarrow \infty$ was considered (e.g., see [10,15–21]), leading to picket fence statistics [15]. A review on more recent developments is given in Ref. [22]. Here, we will consider the double scaling limit M and $N \rightarrow \infty$, simultaneously. The results in this limit were announced in our paper [23] and, in a parallel development, partly derived in the mathematical work [24]. In this work, we will give a detailed derivation of our results in Ref. [23], where we cover the entire spectrum, including the vicinity of the hard edge and the bulk close to the soft edge, that were not contained in Ref. [24]. In particular, we extend previous results for standard random matrix statistics from fixed M to $N \gg M$, as well as previous results for picket fence statistics at fixed N to $M \gg N$, in the respective double scaling limits. Furthermore, we present deeper insights into these statistics, including the issue of universality, unfolding, and what they mean. For a first work on the complex eigenvalues statistic, we refer to [25].

For instance, we argue that the results hold for a broader class of multiplicative stochastic processes. This is corroborated by Monte Carlo simulations, we have carried out, of several ensembles that include non-Gaussian ensembles as well as a certain degree of statical dependence between the matrices that are multiplied. Indeed, in a recent work [26], it was shown that also a product of complex Jacobi matrices (truncated unitary matrices) leads to the same picture. What seems to lie behind this universality, and came even more as surprise for us, is that the local kernels are those of the additive stochastic processes such as Dyson’s Brownian motion [2,27]. Those results of the additive processes describe the microscopic statistics of eigenvalues from an initial condition given by a nondegenerate deterministic matrix. The initial condition of the Dyson Brownian motion model is the one of the picket fence statistics (equidistant eigenvalues); for example, the eigenvalue level density is

$$\rho_{\text{pf}}(y) = \frac{1}{N} \sum_{j=1}^N \delta(y - j). \tag{1}$$

Interestingly, both cases, the multiplicative as well as the additive one, yield the same limiting microscopic eigenvalue statistics in the limit $M, N \rightarrow \infty$. This holds not only for the bulk, which has been computed for Dyson’s Brownian motion in Ref. [2], but extends to the soft edge, too. The soft edge for the Dyson Brownian motion has not been analyzed before; we will fill this gap in this work. The correlations depend only on the limiting value of the parameter

$$a = \lim_{N \rightarrow \infty} \frac{N}{M(N)}, \tag{2}$$

where $a = 0$ corresponds to the picket fence statistics [cf. Eq. (1)] and $a = \infty$ corresponds to the Gaussian unitary ensemble (GUE) local spectral statistics. The interesting and critical scaling is when the number M of matrices multiplied is proportional to the matrix dimension N . Hence, the additive and multiplicative processes have the same limiting local statistics and thus they belong to the same universality class, given by the interpolation of the picket fence

statistics and the GUE statistics. While the bulk statistics of this interpolating kernel has been derived by Johansson in Ref. [2], as mentioned above, the soft edge statistics has not been done, yet. We will give a brief derivation of this result and show that also at the soft edge the agreement of the kernels between the multiplicative and additive process holds.

Another insight we have already argued before in Ref. [23] and understand now very well is that actually not the ratio N/M is crucial but the ratio of the average width of the distributions of individual eigenvalues about the point where one zooms in and the local mean level spacing. With the case of products of independent complex Ginibre matrices one can quantify this by the broadened picket fence spectrum, i.e., $0 < a \ll 1$, where the level density is not any more a sum of Dirac delta functions (1), but a sum of log-normal distributions [15]

$$\rho_Y(y) = \frac{1}{N} \sum_{j=1}^N \frac{1}{\sqrt{2\pi\sigma_j^2}y} \exp\left[-\frac{[\ln(y) - \bar{\lambda}_j]^2}{2\pi\sigma_j^2}\right], \tag{3}$$

with mean and standard deviation

$$\bar{\lambda}_j = \frac{\psi(j)}{2}, \quad \sigma_j = \sqrt{\frac{\psi'(j)}{4M}}. \tag{4}$$

The digamma function $\psi(z) = \partial_z \ln \Gamma(z)$, with $\Gamma(z)$ being the gamma function, plays a crucial role for the Lyapunov exponents of products of Ginibre matrices. This may change for other matrix ensembles. Yet, it has been recently observed [28] that the asymptotic behavior of the width $\sigma_j \approx 1/\sqrt{4Mj}$ [compare (6) below] seems to be universal, as for products of real and complex Wigner matrices the zeros of the characteristic polynomial and the positions for large Lyapunov exponents match those of the corresponding Ginibre matrices.

A consequence of (3) is that the general width-to-spacing (WSR) ratio of two consecutive eigenvalues at the mean positions $\bar{\lambda}_j$ and $\bar{\lambda}_{j+1}$ is [15,23]

$$\text{WSR}_j = \frac{1}{2} \frac{\sigma_{j+1} + \sigma_j}{\bar{\lambda}_{j+1} - \bar{\lambda}_j}, \tag{5}$$

and Eq. (2) has, then, to be replaced by $a = \text{WSR}_j^2$ as the overlap of the distributions of the individual eigenvalues varies in the position $\bar{\lambda}_j$. Hence, the hard edge about the origin will always exhibit picket fence statistics while for large $j \gg 1$ the transition parameter simplifies to $a \approx j/M$, which follows from the asymptotic expansions [29, 6.3.18 and 6.4.11]

$$\begin{aligned} \psi(z) &= \ln(z) - \frac{1}{2z} + O\left(\frac{1}{z^2}\right) \quad \text{and} \\ \psi^{(l)}(z) &= (-1)^{l-1} \left[\frac{(l-1)!}{z^l} + \frac{l!}{z^{l+1}} \right. \\ &\quad \left. + O\left(\frac{1}{z^{l+2}}\right) \right] \quad \text{for } |z| \gg 1. \end{aligned} \tag{6}$$

We will see in the derivations in the ensuing sections that indeed j/M is highly important in the bulk and it only happens at the soft edge that M/N takes the role of the transition parameter.

As a final insight, we have found that close to the soft edge, but still in the bulk, a mesoscopic scale of spectral statistics arises. The microscopic statistics will be not affected and continues to agree with the bulk statistics, especially the interpolating kernel between picket fence and GUE statistics still applies. However, here the unfolding of the spectrum deviates from the bulk unfolding. We have already mentioned this observation in Ref. [23] but at that time did not find an analytical way to derive the proper unfolding. In this work, we have filled in this gap. This insight is valuable and important because it gives the proper unfolding and allows for the identification of the universality of results. It has also relevance for unraveling the conundrum that the macroscopic level density seems to follow always the same law, and never shows a square root behavior at the soft edge, albeit for $N \gg M$ it is known that locally one finds the Airy kernel, whose asymptotic form into the bulk describes a square root. This narrow region at the soft edge comprises a tail made of a certain number of eigenvalue contributions. In this work we estimate the fraction of eigenvalues which contribute to be of order N/M . When $M \gg N$, this tail is not present. The eigenvalues in this narrow tail have been formerly not considered and we have now derived their corresponding mesoscopic level density.

This article is structured as follows. In Sec. II we briefly review the determinantal point process of the product of M complex Ginibre matrices. Especially, we give two particular representations of the kernel that will be the starting point of our analysis. These representations are derived in Appendix A. Before we go over to studying the local spectral statistics in the bulk (Sec. IV), at the hard edge (Sec. V), and at the soft edge (Sec. VI), we first derive the proper unfolding for the various double scaling limits in Sec. III. In this section we also unveil that there is a mesoscopic spectral regime close to the soft edge. So, in addition to the discussion presented in Ref. [23], we are now able to unfold the spectrum at the soft edge analytically. Our claim that these local spectral statistics are universal is corroborated by the Monte Carlo simulation of several matrix ensembles including non-Gaussian as well as correlated matrices. These simulations are explained and discussed in Sec. VIII. Prior to that we dedicate one section to the discussion of a puzzling duality between local statistics for matrix products and Dyson’s Brownian motion in Sec. VII. Therein, we also derive the result of the local soft edge kernel for Dyson’s Brownian motion with the picket fence spectrum as its initial condition. In Sec. IX, we summarize our findings and give an outlook on open problems. Further technical details are collected in Appendices B–D.

II. PRELIMINARIES

Consider the discrete-time evolution of an open physical system with N degrees of freedom. The state of the system at time t is described by an N -dimensional state vector $|v\rangle_t$ that evolves according to a recursive equation $|v\rangle_t = X_t |v\rangle_{t-1}$, with a transfer matrix X_t . The map between an initial state $|v\rangle_0$ and the state $|v\rangle_M = X |v\rangle_0$ after M steps is given by the evolution operator

$$X = X_M \dots X_1. \tag{7}$$

Let us assume that the transfer matrices can be modeled by random matrices. Since the system is open, the evolution is nonunitary so that, e.g., the norm of a state is not conserved.

We are interested in the singular value statistics of X or, equivalently, in the eigenvalue statistics of the associated Hermitian operator

$$Y = X^\dagger X, \tag{8}$$

which controls the growth of the norm $\langle v|v\rangle_M = \langle v|Y|v\rangle_0$. The eigenvalue statistics of the product matrix Y is in one-to-one correspondence with the statistics of the Lyapunov matrix

$$L = \frac{1}{2M} \ln[(X_M \dots X_1)^\dagger X_M \dots X_1] = \frac{1}{2M} \ln[Y]. \tag{9}$$

We concentrate on the thermodynamic limit $N \rightarrow \infty$, but at the same time assume that the number of matrices (time steps) in the product is an increasing function of the matrix size $M = M(N)$.

The microscopic spectral statistics of the Hermitian operator Y is expected to be universal for a large class of transfer matrices, including the case of independent matrices X_j with independent normal random variables as entries, as we shall argue later. Consequently, it is useful to consider an ensemble from this class which is analytically tractable. To be more precise, we assume that the transfer matrices X_t , $t = 1, \dots, M$, are identically distributed independent complex Ginibre matrices [30] with independent and identically distributed Gaussian elements, i.e.,

$$P(X_j) = \frac{\exp[-\text{tr} X_j X_j^\dagger]}{\pi^{N^2}}, \quad X_j \in \mathbb{C}^{N \times N}. \tag{10}$$

As shown in Refs. [7,8], this ensemble is completely solvable, in the sense that all eigenvalue correlation functions of Y of any order $k = 1, \dots, N$ can be given in closed, explicit expressions for any M and N . They form a so-called determinantal point process [31], with the joint probability density of eigenvalues $y_1, \dots, y_n > 0$ of Y given by

$$P_{Y,N}(y_1, \dots, y_N) = \frac{1}{N!} \det [K_Y(y_i, y_j)]_{i,j=1,\dots,N}. \tag{11}$$

Its kernel K_Y is given by (A1) [7,8], and in Appendix A it is shown to be equivalent to the forms (14) and (15) given below. The k -point correlation functions of such a point process take an elegant determinantal form [31], as well,

$$\begin{aligned} R_{Y,k}(y_1, \dots, y_k) & \equiv \frac{N!}{(N-k)!} \int_0^\infty dy_{k+1} \dots \int_0^\infty dy_N P_{Y,N}(y_1, \dots, y_N) \\ & = \det [K_Y(y_i, y_j)]_{i,j=1,\dots,k}. \end{aligned} \tag{12}$$

As an example, the normalized level density is given by $\rho_Y(y) = R_{Y,1}(y)/N = K_Y(y, y)/N$. The first subscript in the above quantities indicates that in this case the correlation functions (kernel) are meant for the eigenvalues of the matrix Y . We use this convention throughout our work to distinguish between quantities for matrices Y , L and others that will be discussed. We would like to underline that the correlation functions (12) are invariant with respect to an equivalence

transformation of the kernel

$$K_Y(x, y) \rightarrow \frac{g(x)}{g(y)} K_Y(x, y), \tag{13}$$

where $g(x)$ is a nonsingular function. We shall use this invariance several times to simplify the form of the kernels.

The kernel K_Y can be expressed in an explicit way in terms of Meijer G functions (see [7,8] and Appendix A for details). Here, we use two equivalent and closely related representations which are particularly well suited for the various double scaling limits in $M, N \rightarrow \infty$ to be taken:

$$K_Y(x, y) = \frac{1}{y} \sum_{j=0}^{N-1} \int_{\gamma_t} \frac{dt}{2\pi i} \frac{\sin[\pi(j-t)]}{\pi(j-t)} e^{i\pi \operatorname{sign}[\operatorname{Im}(t)]t} \times \exp[-\mathcal{S}(j; x) + \mathcal{S}(t; y)] \tag{14}$$

and

$$K_Y(x, y) = \frac{1}{y} \int_{\gamma_t} \frac{dt}{2\pi i} \int_{\gamma_s} \frac{ds}{2\pi i} \frac{1}{s-t} \frac{\sin(\pi t)}{\sin(\pi s)} \frac{e^{i\pi \operatorname{sign}[\operatorname{Im}(t)]t}}{e^{i\pi \operatorname{sign}[\operatorname{Im}(s)]s}} \times \exp[-\mathcal{S}(s; x) + \mathcal{S}(t; y)], \tag{15}$$

where

$$\mathcal{S}(z; \alpha) = -i\pi \operatorname{sign}[\operatorname{Im}(z)]z - \ln[\alpha]z + (M+1) \times \ln[\Gamma(1+z)] + \ln[\Gamma(N-z)], \tag{16}$$

with $\alpha = x, y$ and $z = j, t, s$. The term $\mathcal{S}(z; \alpha)$ is called action in the remainder of the paper. The sign function is given by $\operatorname{sign}(\chi) = \chi/|\chi|$ for $\chi \in \mathbb{R} \setminus \{0\}$ and vanishes for $\chi = 0$. The contour γ_t is an integration parallel to the imaginary axis along $c + i\mathbb{R}$, with $-1 < c < 0$ chosen such that it does not cross the closed contour γ_s in Eq. (15). The contour γ_s encircles the closed interval $[0, N-1]$ counterclockwise. These formulas are derived in Appendix A [cf. [24] for (15)].

We would like to highlight that the prefactors of the exponentials in Eqs. (14) and (15) do not grow or shrink exponentially. They have only simple poles and zeros. Hence, they do not contribute in the saddle point equation when making an asymptotic expansion. Moreover, we would like to emphasize that, although the splitting into the exponents and prefactors is nonanalytic, the integrand as a whole is a meromorphic function [see also (A7) and (A6), respectively].

Our goal is to analyze local (microscopic) properties of the kernel (14) in the double scaling limit $M, N \rightarrow \infty$, depending on how the limit is taken in terms of $M = M(N)$, and where in the spectrum we zoom in. Before we discuss the local level statistics, let us derive in detail the relevant results on the macroscopic level density for the product of M Ginibre matrices in Sec. III. This preparation is necessary in order to take the local limits, where we have to unfold with respect to the macroscopic or mesoscopic level density.

III. MACROSCOPIC AND MESOSCOPIC LEVEL DENSITY AND UNFOLDING

In this section, we will use a saddle point analysis to determine the macroscopic level density. It is a key ingredient for the following discussion of the local statistics in the bulk and at the edges. In the discussion we will distinguish two cases, which differ in how the saddle point scales with M , in

Secs. III A and III B. We will also define what we mean by the mesoscopic density and explain when it occurs in Sec. III C. In particular, it will be used to unfold the spectrum at the soft edge.

For the macroscopic level density we start from (15), with $x = y > 0$, and perform a saddle point analysis of the action \mathcal{S} in Eq. (16). We look for the points z_s that satisfy

$$\partial_{z_s} \mathcal{S}(z_s; y) = -\ln(y) + (M+1)\psi(1+z_s) - \psi(N-z_s) - i\pi \operatorname{sign}[\operatorname{Im}(z_s)] = 0. \tag{17}$$

A solution in the upper half-plane has a complex conjugate partner z_s^* in the lower half-plane. As discussed in Appendix B, the imaginary part $\operatorname{Im}[z_s]$ of the saddle point solution is of the order

$$\operatorname{Im}[z_s] = O\left(\frac{\operatorname{Re}[z_s]}{M}\right). \tag{18}$$

The real part $\operatorname{Re}[z_s]$ lies in the interval $\operatorname{Re}[z_s] \in]-1, N[$. The lower end of the interval corresponds to the hard edge and the upper one to the soft edge of the spectrum. We see that z_s is governed by its real part. This will be used in the following two subsections in a case by case discussion, where $\operatorname{Re}[z_s]$ is taken either much larger than M , in Sec. III A, or much smaller than or at most of the same order as M , in Sec III B. Let us underline at this point that we have not assumed how M and N are related in the limit $M, N \rightarrow \infty$.

The hard edge is represented by $\operatorname{Re}[z_s] = 0$, while the soft edge by $\operatorname{Re}[z_s] = N-1$. Hence, for the above estimates we require that $\operatorname{Re}[z_s] \in]0, N-1[$. While for the contour γ_s , this is not difficult to satisfy, the contour γ_t has to be deformed accordingly. For the analysis in the bulk we want to stay away from the edges of the spectrum. We will thus assume for the limiting level density that when $N \gg 1$ it holds

$$\operatorname{Re}[z_s] \gg 1 \quad \text{and} \quad N-1-\operatorname{Re}[z_s] \gg 1. \tag{19}$$

A. Case: $\operatorname{Re}[z_s] \gg M$

In this section we derive the limiting macroscopic level density when $\operatorname{Re}[z_s] \gg M$, and identify the proper unfolding in two parts of the bulk of the spectrum.

We shift the contour γ_t parallel to the real axis, in particular we only set the parameter $c = \operatorname{Re}[z_s]$, such that it runs through both saddle points z_s and z_s^* . As above and without loss of generality, we assume that $\operatorname{Im}(z_s) \geq 0$. Since also the closed contour γ_s has to run through these two saddle points, both contours have to cross each other. Originally, in the derivation of Eq. (15) in Appendix A, the contours were chosen not to cross, in order not to pick up the pole at $1/(s-t)$. To compensate this newly created residuum we have to subtract it whenever the integration path of t lies inside γ_s , so that the kernel takes the form

$$K_Y(y, y) = \frac{1}{y} \left[\int_{z_s^*}^{z_s} \frac{dt}{2\pi i} + \int_{\operatorname{Re}[z_s]-i\infty}^{\operatorname{Re}[z_s]+i\infty} \frac{dt}{2\pi i} \times \int_{\gamma_s} \frac{ds}{2\pi i} \frac{1}{s-t} \frac{\sin(\pi t)}{\sin(\pi s)} \frac{e^{i\pi \operatorname{sign}[\operatorname{Im}(t)]t}}{e^{i\pi \operatorname{sign}[\operatorname{Im}(s)]s}} \times \exp[-\mathcal{S}(s; y) + \mathcal{S}(t; y)] \right]. \tag{20}$$

The first integral can be readily carried out, to give $\text{Im}[z_s]/\pi$. This will turn out to be the dominant contribution, as we will argue in the following.

For the second integral in Eq. (20), we expand around the saddle points $s = s_0 + \delta s$ and $t = t_0 + i\delta t$ with any combination of $s_0, t_0 = z_s, z_s^*$. This leads to a sum of four contributions to the integral. Notice that due to the form of γ_t the perturbation $i\delta t$ always runs parallel to the imaginary axis in the same direction as the axis. In contrast, due to the form of γ_s , the perturbation δs is real and runs antiparallel to the real axis for z_s , and parallel for z_s^* , leading to a relative minus sign. In this expansion around the saddle point the combination of actions will be replaced by

$$\begin{aligned}
 & -\mathcal{S}(s; y) + \mathcal{S}(t; y) \\
 & \approx -\mathcal{S}(s_0; y) + \mathcal{S}(t_0; y) \\
 & - \frac{(M+1)\psi'(1+s_0) + \psi'(N-s_0)}{2} \delta s^2 \\
 & - \frac{(M+1)\psi'(1+t_0) + \psi'(N-t_0)}{2} \delta t^2. \tag{21}
 \end{aligned}$$

When $s_0 = t_0$, that is both are either z_s or z_s^* , the leading contribution cancels. In the other case, the leading part of the exponent becomes $-\mathcal{S}(z_s; y) + \mathcal{S}(z_s^*; y) = 2i \text{Im} \mathcal{S}(z_s; y)$, or its complex conjugate.

Before we write the integrals, let us consider the prefactors. For $s_0 = t_0 = z_s$, we have

$$\frac{1}{s-t} = \frac{1}{\delta s - i\delta t} \tag{22}$$

and

$$\frac{\sin[\pi(z_s + i\delta t)] \exp\{i\pi \text{sign}[\text{Im}(z_s + i\delta t)]\text{Im}(z_s + i\delta t)\}}{\sin[\pi(z_s + \delta s)] \exp\{i\pi \text{sign}[\text{Im}(z_s + \delta s)](z_s + \delta s)\}} \approx 1. \tag{23}$$

As follows from (18), for $\text{Re}[z_s] \gg M$ both real and imaginary parts of z_s are large and dominate over the perturbations δs and $i\delta t$. Equations (22) and (23) obviously also hold for $z_s \rightarrow z_s^*$.

In the mixed case $s_0 = t_0^* = z_s$, we obtain

$$\frac{1}{s-t} = \frac{1}{z_s + \delta s - z_s^* - i\delta t} \approx \frac{1}{2i \text{Im}[z_s]} \tag{24}$$

or its complex conjugate for $s_0^* = t_0 = z_s$. For the sine functions we have

$$\begin{aligned}
 & \sin[\pi(z_s + i\delta t)] e^{i\pi \text{sign}[\text{Im}(z_s + i\delta t)](z_s + i\delta t)} \\
 & = \frac{1}{2i} (e^{i\pi(z_s + i\delta t)} - e^{-i\pi(z_s + i\delta t)}) e^{i\pi(z_s + i\delta t)} \approx -\frac{1}{2i}, \tag{25}
 \end{aligned}$$

as the second term dominates, recalling that $\text{Im}[z_s] > 0$ is large. For the same factor with $z_s \rightarrow z_s^*$ we arrive at

$$\begin{aligned}
 & \sin[\pi(z_s^* + i\delta t)] e^{i\pi \text{sign}[\text{Im}(z_s^* + i\delta t)](z_s^* + i\delta t)} \\
 & = \frac{1}{2i} (e^{i\pi(z_s^* + i\delta t)} - e^{-i\pi(z_s^* + i\delta t)}) e^{-i\pi(z_s^* + i\delta t)} \approx +\frac{1}{2i}, \tag{26}
 \end{aligned}$$

as now the first term dominates. We thus have

$$\frac{\sin[\pi(z_s + i\delta t)] \exp\{i\pi \text{sign}[\text{Im}(z_s + i\delta t)](z_s + i\delta t)\}}{\sin[\pi(z_s^* + \delta s)] \exp\{i\pi \text{sign}[\text{Im}(z_s^* + \delta s)](z_s^* + \delta s)\}} \approx -1, \tag{27}$$

as well as for its complex conjugate. Therefore, we can now write for the kernel

$$\begin{aligned}
 K_Y(y, y) & \stackrel{M, N \gg 1}{\approx} \frac{1}{y} \left\{ \frac{\text{Im}[z_s]}{\pi} - \int_{-\infty}^{\infty} \frac{d\delta t}{2\pi} \int_{-\infty}^{\infty} \frac{d\delta s}{\pi} \frac{1}{\delta s - i\delta t} \text{Im} \left[e^{-\frac{1}{2}[(M+1)\psi'(1+z_s) + \psi'(N-z_s)](\delta s^2 + \delta t^2)} \right. \right. \\
 & + \frac{\int_{-\infty}^{\infty} \frac{d\delta t}{2\pi} \int_{-\infty}^{\infty} \frac{d\delta s}{2\pi i} \frac{e^{2i \text{Im}[\mathcal{S}(z_s; y)] - \frac{1}{2}[(M+1)\psi'(1+z_s) + \psi'(N-z_s)]\delta s^2 - \frac{1}{2}[(M+1)\psi'(1+z_s^*) + \psi'(N-z_s^*)]\delta t^2}}{2i \text{Im}[z_s]} \\
 & \left. \left. - \int_{-\infty}^{\infty} \frac{d\delta t}{2\pi} \int_{-\infty}^{\infty} \frac{d\delta s}{2\pi i} \frac{e^{-2i \text{Im}[\mathcal{S}(z_s; y)] - \frac{1}{2}[(M+1)\psi'(1+z_s^*) + \psi'(N-z_s^*)]\delta s^2 - \frac{1}{2}[(M+1)\psi'(1+z_s) + \psi'(N-z_s)]\delta t^2}}{-2i \text{Im}[z_s]} \right\}. \tag{28}
 \end{aligned}$$

The second integral in the first line vanishes, as one can see in polar coordinates $\delta s - i\delta t = r e^{i\varphi}$, where the integral over the angle becomes $\int_0^{2\pi} d\varphi e^{-i\varphi} = 0$. In the exponents in the second and third lines we may expand the digamma function and its derivatives via Eq. (6). Taking into account the dominance of the real part over the imaginary one (18), $\text{Re}[z_s] \gg \text{Im}[z_s] = O(\text{Re}[z_s]/M) \gg 1$, we have

$$\begin{aligned}
 (M+1)\psi'(1+z_s) + \psi'(N-z_s) & \approx \frac{M}{\text{Re}[z_s]} + \frac{1}{N - \text{Re}[z_s]} \\
 & \approx \frac{M}{\text{Re}[z_s]}, \tag{29}
 \end{aligned}$$

and analogously for z_s^* . Both terms on the right-hand side are small and positive, due to the conditions $\text{Re}[z_s] \gg M$ and $N - \text{Re}[z_s] \gg 1$. After performing the Gaussian integrals we thus obtain for the limiting kernel

$$K_Y(y, y) \stackrel{M, N \gg 1}{\approx} \frac{1}{y} \left[\frac{\text{Im}[z_s]}{\pi} - \frac{\text{Re}[z_s]}{2\pi M \text{Im}[z_s]} \cos\{2 \text{Im}[\mathcal{S}(z_s; y)]\} \right]. \tag{30}$$

Because $\text{Re}[z_s]/(2\pi M \text{Im}[z_s])$ is of order one, the second term in Eq. (30) is subleading as compared to the first one which is growing with $\text{Im}[z_s] \gg 1$. We eventually arrive at the following asymptotic expression for the macroscopic level

density:

$$\rho_Y(y) = \frac{1}{N} K_Y(y, y) \stackrel{M, N \gg 1}{\approx} \frac{\text{Im}[z_s(y)]}{\pi N y}. \quad (31)$$

Hence, the relation between the saddle point z_s and the macroscopic level density is very simple, and a remarkable relation to the corresponding resolvent is pointed out in Appendix C. Let us remark that, initially, we have not imposed any relation between N and M , apart from both $M, N \rightarrow \infty$. However, the condition $\text{Re}[z_s] \gg M$ together with the requirement to stay away from the hard edge (19), $N - 1 - \text{Re}[z_s] \gg 1 \Leftrightarrow N \gg \text{Re}[z_s]$, implies that this case corresponds to $N \gg M$.

In order to better understand the result of the saddle point analysis of the kernel (31), let us investigate the implications of this scaling on the relation between the argument $y \in \mathbb{R}_+$ and the saddle point z_s , as it follows directly from (17). Recalling our choice $\text{Im}[z_s] \geq 0$, we can exploit the asymptotic expansion (6) of the digamma function to obtain

$$\begin{aligned} 0 &\approx -\ln(y) + (M + 1)\ln[1 + z_s] - \ln[N - z_s] - i\pi \\ \Rightarrow y &\approx \frac{z_s^{M+1}}{N - z_s} \gg \frac{M^{M+1}}{N - M}. \end{aligned} \quad (32)$$

The last inequality follows from $\text{Re}[z_s] \gg M$. We also see that the limit $y \rightarrow \infty$ coincides with $\text{Re}[z_s]/N \rightarrow 1$ for the saddle point equation (17).

B. Case: $\text{Re}[z_s] \leq O(M)$

In this section, we derive the limiting macroscopic level density and find the proper unfolding in the case $\text{Re}[z_s] \leq O(M)$, meaning $\text{Re}[z_s]$ is maximally of order M . Here, we have to be more careful since we need to separate the microscopic from the mesoscopic and macroscopic scale. The microscopic behavior is encoded in the poles of the term $1/\sin(\pi s)$ in the integration variable s where the integrand (15) gets strongly peaked. They become visible in the integral because the imaginary part of the saddle point $\text{Im}[z_s] = O(\text{Re}[z_s]/M)$ is of order one or smaller (18). For this reason, we start from the series representation of the kernel (14), in which these poles are explicitly evaluated. The deformation of the t integral through the saddle point does not pose a problem, as the preexponential factor no longer has a pole. The index j represents now the position in the spectrum and replaces the real part $\text{Re}[z_s]$ of the saddle point of the s integration in Eq. (15) considered so far.

We have to be aware that in this representation the action $S(j; y)$ cannot be minimized at the saddle point z_s since j is an integer, $j = 0, \dots, N - 1$. Therefore, we consider only the real part of the saddle point equation (17) which is

$$\ln(y) = (M + 1)\psi[1 + z_0] - \psi[N - z_0] \quad (33)$$

with $z_0 \in] - 1, N[$. Those integers j that are closest to z_0 will contribute the most, that is why we take z_0 to be real.

The uniqueness of the solution for z_0 is slightly simpler than for z_s . For the known monotonicity property of the digamma function of real argument (cf. [32] Sec. 5.3(i) and see (B5) for $t = 0$), it is clear that the right-hand side of (33) is strictly increasing on $] - 1, N[$ and goes to $-\infty$ for $z_0 \rightarrow -1$ and to $+\infty$ for $z_0 \rightarrow N$. Therefore, there is a unique solution

which can be given in an integral form over the Heaviside step function

$$z_0 = \int_{-1}^N \Theta(\ln(y) - (M + 1)\psi[1 + z] + \psi[N - z]) dz - 1. \quad (34)$$

Since Eq. (33) gives a one-to-one relation between y and z_0 and, hence, the summation index j , we have already found the proper unfolding in N and M , given we can show that the contributions of the index j are tightly concentrated around z_0 .

Before we come to this, let us discuss the scaling bound of z_0 . It follows from Eq. (33) by expanding the digamma function, assuming that also z_0 is large, and exponentiating the equation

$$y \approx \frac{z_0^{M+1}}{N - z_0} \leq O\left(\frac{M^{M+1}}{N - M}\right), \quad (35)$$

given that not only $\text{Re}[z_s] \leq O(M)$ but also $z_0 \leq O(M)$ which is consistent with (32) which was the scaling bound in the opposite regime. We will make use of this insight in the ensuing discussion.

To decide which summands in Eq. (14) eventually contribute, we need to study the difference of the exponents

$$\begin{aligned} \Delta &= S(j; y) - S(z_0; y) \\ &= (M + 1) \left[\ln\left(\frac{\Gamma[1 + j]}{\Gamma[1 + z_0]}\right) - \psi[1 + z_0](j - z_0) \right] \\ &\quad + \ln\left(\frac{\Gamma[N - j]}{\Gamma[N - z_0]}\right) + \psi[N - z_0](j - z_0) \geq 0 \end{aligned} \quad (36)$$

for $z_0 \in] - 1, N[$ and $j = 0, 1, \dots, N - 1$. Here, we have inserted already (33). The condition to stay in the bulk of the spectrum is

$$z_0 \gg 1 \quad \text{and} \quad N - z_0 \gg 1, \quad (37)$$

as $z_0 = -1$ corresponds to the hard edge at $y = 0$ and $z_0 = N - 1$ to the soft edge at $y = \infty$. Note that the maximal value of $j = N - 1$ leads to a natural upper bound of the spectrum which is the soft edge. This means we also consider $j \gg 1$ and $N - 1 - j \gg 1$.

We can rewrite (36) as

$$\begin{aligned} \Delta &= (\delta j)^2 \int_0^1 d\lambda (1 - \lambda) [(M + 1)\psi'(1 + z_0 + \delta j\lambda) \\ &\quad + \psi'(N - z_0 - \delta j\lambda)], \end{aligned} \quad (38)$$

where $\delta j = j - z_0$. This equation can be readily derived via integration by parts. In the leading order when $|\delta j| \ll z_0$, we get $\Delta \approx (\delta j)^2 [(M + 1)\psi'(1 + z_0) + \psi'(N - z_0)]/2 \approx (\delta j)^2 M/2z_0$, for z_0 far from the hard and soft edges. We have neglected higher order terms $(\delta j)^k$, for $k = 3, 4, \dots$. Using this approximation we can find out how many summands about z_0 contribute, namely,

$$\Delta \approx \frac{1}{2} (\delta j)^2 \frac{M}{z_0} \leq O(1) \Leftrightarrow |\delta j| \leq O\left(\sqrt{\frac{z_0}{M}}\right) < O(z_0). \quad (39)$$

Due to strict convexity of Δ we can disregard $\delta j \geq O(z_0)$ as they are strongly suppressed in the exponent (14).

We can insert the scaling of $\delta j = j - z_0 = O(\sqrt{z_0/M})$ and of $t - z_0 = i\delta t = O(\sqrt{z_0/M}) \geq O(z_0/M) = \text{Im}[z_s]$ into the action in Eq. (14), and perform a Taylor expansion

$$-\mathcal{S}(j; y) + \mathcal{S}(t; y) \stackrel{M, N \gg 1}{\approx} -\frac{M}{2z_0}(\delta j^2 + \delta t^2) - i\pi \text{sign}(\delta t), \tag{40}$$

where we have exploited (37). For the kernel (14) we obtain

$$K_Y(y, y) \stackrel{M, N \gg 1}{\approx} \frac{1}{y} \sum_{j=0}^{N-1} \int_{-\infty}^{\infty} \frac{d\delta t}{2\pi} \frac{\sin[\pi(j - z_0 - i\delta t)]}{\pi(j - z_0 - i\delta t)} \times \exp\left[-\frac{M}{2z_0}[(j - z_0)^2 + \delta t^2]\right]. \tag{41}$$

This integral can be evaluated by first rephrasing

$$\frac{\sin[\pi(j - z_0 - i\delta t)]}{\pi(j - z_0 - i\delta t)} = \int_{-1}^1 \frac{dr}{2} e^{i\pi(j - z_0 - i\delta t)r}, \tag{42}$$

then, integrating over δt and afterward over r leading to

$$K_Y(y, y) \stackrel{M, N \gg 1}{\approx} \frac{1}{2\pi y} \sum_{j=0}^{N-1} \text{Re} \left\{ \text{erfi} \left[\pi \sqrt{\frac{z_0}{2M}} + i \sqrt{\frac{M}{2z_0}}(j - z_0) \right] \right\}. \tag{43}$$

The function $\text{erfi}(x) = -i \text{erf}(ix)$ is the imaginary error function.

So far we have not separated the microscopic scale from any scale that is larger than the local mean level spacing. However, we established the relation of j with z_0 , and of z_0 with y . In this way, we know that when staying away from the two edges at $z_0 = j = 0$ and $z_0 = j = N - 1$ there are infinitely many eigenvalues on both sides. Therefore, the kernel becomes discretely translation invariant under the shift $z_0 \rightarrow z_0 + 1$; in particular, we have

$$K_Y(y(z_0), y(z_0)) \approx K_Y(y(z_0 + 1), y(z_0 + 1)), \tag{44}$$

which is valid for (37). Here,

$$y(z_0) = \exp[(M + 1)\psi(1 + z_0) - \psi(N - z_0)]$$

is defined by the saddle point equation (33). Indeed, for $z_0 \ll M$ the imaginary error functions in Eq. (43) become a sum of Dirac delta functions $\delta(z_0 - j)$ while for $z_0 \propto M$ the change of the summands in a shift $z_0 \rightarrow z_0 + 1$ will be of order $1/M$.

From this approximate translational invariance, we can read off that on any scale larger than the mean level spacing, which is one for the variable z_0 , the distribution in the variable z_0 is uniformly distributed on $[0, N - 1]$. We can also rephrase this statement and say that the limiting distribution in the original eigenvalue y of the product matrix Y is

$$\rho_{\text{bulk}}(y) = \frac{1}{N} \frac{dz_0}{dy} = \frac{1}{Ny} \frac{dz_0}{d \ln(y)} = \frac{1}{Ny} \frac{1}{(M + 1)\psi'(1 + z_0) + \psi'(N - z_0)} \approx \frac{z_0}{N(M + 1)y} = \frac{\int_{-1/N}^1 \Theta[y - \exp[(M + 1)\psi(1 + N\hat{y}) - \psi(N - N\hat{y})]] d\hat{y} - 1/N}{(M + 1)y} \tag{45}$$

when $\text{Re}[z_s] \approx z_0 \leq O(M)$. The unfolded variable is given by $\hat{y} = z_0/N$ and therefore for finite N and M we have

$$y = \exp[(M + 1)\psi(1 + N\hat{y}) - \psi(N - N\hat{y})]. \tag{46}$$

As a final remark, we would like to emphasize that there are various spectral scales encoded in Eq. (46). For example, when $N\hat{y} \propto M$, the last equation can be approximated by

$$y = \frac{\hat{y}}{1 - \hat{y}} (N\hat{y})^M \exp\left[\frac{1}{2} \frac{M}{N\hat{y}}\right] \tag{47}$$

which unfolds the spectrum in the bulk. Depending on how $N\hat{y}$ scales with respect to M , one has to go to higher orders in the asymptotic series (6). The scaling discussed in this section corresponds to a fraction $O(M/N)$ of eigenvalues, which lie close to the hard edge.

Let us conclude this section with a remark, that the expansion about the point z_0 is fully justified as it covers the contributions from the saddle point z_s , as discussed in Appendix D.

C. Case: $\text{Re}[z_s] \gg M$ revisited: Unfolding

In Sec. III A, we took a double scaling limit where both $N, M \rightarrow \infty$, without specifying the rate $M = M(N)$. By assuming that $\text{Re}[z_s] \gg M$ and considering a part of the spectrum away from the soft and hard edges, it emerged that

$N \gg M$. It is useful to look at the regime $N \gg M$ from a slightly different perspective by exploiting a known result for the limiting level density $\rho^{(M)}(\zeta)$ of the matrix Y/N^M for $N \rightarrow \infty$ [33] at fixed M . This approach was also taken in Ref. [14,24]. For the bulk, as we shall see, we recover the results that we have discussed in Sec. III A. However, we additionally find an unfolding map close to the soft edge that is very useful for the analysis of the soft edge statistics, that we carry out in Sec. VI.

The eigenvalue density $\rho^{(M)}(\zeta)$ is given by the following parametrization [33]:

$$\rho^{(M)}(\zeta) = \frac{1}{\pi \zeta} \frac{\sin[\phi/(M + 1)] \sin[\phi]}{\sin[M\phi/(M + 1)]}, \quad \text{with} \\ \zeta = \frac{\sin^{M+1}[\phi]}{\sin[\phi/(M + 1)] \sin^M[M\phi/(M + 1)]} \quad \text{and} \\ \phi \in]0, \pi[. \tag{48}$$

It has a support $\zeta \in [0, (M + 1)^{M+1}/M^M]$. Note that we have rescaled the parameter ϕ by $1/(M + 1)$ compared to [33]. The hard edge located at $\zeta = 0$ corresponds to $\phi = \pi$, whereas the soft edge at $\zeta = (M + 1)^{M+1}/M^M$ corresponds to $\phi = 0$. It is convenient to consider two different limits of the level density (48), that we call deep bulk (db) and soft bulk (sb). The corresponding unfolded variables will be specified for

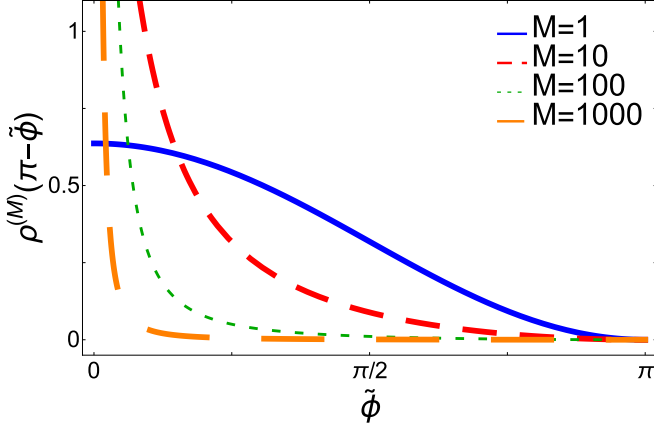


FIG. 1. The macroscopic level density (49) obtained in the limit $N \rightarrow \infty$ for various values of a fixed number of matrices M multiplied. For increasing M the spectrum concentrates at $\tilde{\phi} = \pi - \phi = 0$ which constitutes the deep bulk of the spectrum. Yet, the tail supported inside the interval $]0, \pi[$ still comprises roughly N/M eigenvalues which is finite when $N = O(M)$ and even creates a mesoscopic, soft bulk of eigenvalues when $N \gg M$. We have reflected the angle so that the hard edge lies at the origin which agrees with the picture of the hard edge of the product matrix Y .

these two limits in the next two subsections. This can be made transparent when we give (48) in terms of the angle ϕ , instead of the eigenvalue ζ of the matrix Y/N^M . In particular we consider

$$\begin{aligned} \rho^{(M)}(\phi) &= \frac{1}{\pi} \frac{\sin[\phi/(M+1)] \sin[\phi]}{\sin[M\phi/(M+1)]} \\ &\times \left[\frac{1}{(M+1) \tan[\phi/(M+1)]} \right. \\ &\left. + \frac{M^2}{(M+1) \tan[M\phi/(M+1)]} - \frac{M+1}{\tan[\phi]} \right] \quad \text{with} \\ \phi &\in]0, \pi[, \end{aligned} \quad (49)$$

where we have multiplied the Jacobian from the transformation $\zeta \rightarrow \phi$. In Fig. 1, we reflected this density so that the hard edge is again at the origin on the left of the plot.

1. Macroscopic level density and unfolding in the deep bulk

As it can be seen from Fig. 1, the macroscopic density concentrates at $\phi \approx \pi$ for large M . In order to take the large M limit and derive a properly scaled level density, one needs to magnify the scale close to π . We do this by introducing a new variable θ as follows: $\phi = \pi - \theta/M$ with $\theta \in [0, \pi M]$. A straightforward but lengthy calculation using a Taylor expansion of (49) in $1/M$ leads to

$$\frac{1}{M} \rho^{(M)}\left(\pi - \frac{\theta}{M}\right) \stackrel{M \gg 1}{\approx} \frac{\pi}{(\theta + \pi)^2} = \rho_{\text{db}}(\theta). \quad (50)$$

This scaling limit is called deep bulk limit (db) because almost all eigenvalues for $N \gg M \gg 1$ are described by the density ρ_{db} (50). Obviously, the part of the spectrum captured by this level density does not include the soft edge. As we shall see, it does not actually include the hard edge, either.

The corresponding unfolding in N and M , starting from our original eigenvalues of the matrix Y and expanding the variable ζ in Eq. (48) to higher order in $1/M$, leads to

$$\begin{aligned} y &= N^M \zeta \\ &= N^M \frac{\sin^{M+1}[\pi - \theta/M]}{\sin[(\pi - \theta/M)/(M+1)] \sin^M[(M\pi - \theta)/(M+1)]} \\ &\stackrel{M \gg 1}{\approx} \frac{e\theta}{\pi} \left(N \frac{\theta}{\pi + \theta} \right)^M. \end{aligned} \quad (51)$$

Thus, unfolding is given by the change of variables $\hat{y} = 1 - \pi/(\theta + \pi) = \theta/(\theta + \pi) \in]0, M/(M+1)[$. In terms of y it reads as

$$y = e \frac{\hat{y}}{1 - \hat{y}} (N\hat{y})^M \gg \frac{M^{M+1}}{N - M} \Rightarrow \hat{y} \gg M/N. \quad (52)$$

The scaling bound of y in Eq. (32) implies a scaling bound for the unfolded spectral variable \hat{y} , given in the last inequality. A similar scaling as in Eq. (52) has been found in Ref. [23], apart from the prefactor $e\hat{y}/(1 - \hat{y})$. Note also that the unfolding (52) differs by an exponential factor from the unfolding (47) as the scale of $\text{Re}[z_s]$ is different. The origin of this breakdown lies in the difference of z_0 and z_s which are not close any more when $\text{Re}[z_s] \gg M$ such that the full unfolding formula (46) breaks down. However, the factor becomes irrelevant for the macroscopic spectral statistics in the limit $M \rightarrow \infty$. Such prefactors do not have any impact on the resulting unfolding of the kernel in the double limit $M, N \rightarrow \infty$ because, in order to take this limit, one first has to take the M th root and then rescale it by $1/N$, i.e., $y^{1/M}/N = [e\hat{y}/(1 - \hat{y})]^{1/M} \hat{y}$. So we see that the root of the prefactor converges to 1 as long as we stay away from the hard edge at $\hat{y} = 0$ and from the soft edge $\hat{y} = 1$. Nevertheless, the fact that something different happens at the edges is a hint that a new mesoscopic scale may appear.

Let us quantify this by computing the normalization of the macroscopic density (50):

$$\int_0^{\pi M} d\theta \rho_{\text{db}}(\theta) = -\frac{\pi}{\theta + \pi} \Big|_0^{\pi M} = 1 - \frac{1}{M+1}. \quad (53)$$

Hence, we are missing a fraction of the order of $1/(M+1)$ eigenvalues, which clearly vanishes only when $M \rightarrow \infty$. The missing eigenvalues are located close to the soft edge because θ cannot become arbitrarily large, as $\phi = \pi - \theta/M \in [0, \pi]$. This is the reason why we have to separately study a mesoscopic scaling regime close to the soft edge. We will do this in the next subsection.

There are also missing eigenvalues close to the hard edge at $\hat{y} = \theta = 0$. Indeed, the derivation above breaks down when the condition $\text{Re}[z_s] \gg M$ is violated. To see this we translate the scaling bound (52) to θ ,

$$\theta \gg \frac{\pi M}{N - M}. \quad (54)$$

Therefore, θ cannot become too small. The fraction we are missing is given naively by M/N , as follows from

$$\begin{aligned} \int_{\pi M/(N-M)}^{\pi M} d\theta \rho_{\text{db}}(\theta) &= -\frac{\pi}{\theta + \pi} \Big|_{\pi M/(N-M)}^{\pi M} \\ &= \frac{N - M}{N} - \frac{1}{M + 1} \\ &= 1 - \frac{M}{N} - \frac{1}{M + 1}, \end{aligned} \quad (55)$$

where we have corrected the integral (53) by modifying its lower bound. This also nicely highlights that for $M \gg N$ the entire spectrum has to be dealt with in a different way than it is done in this section.

2. Mesoscopic level density and unfolding close to the soft edge

The aim of this section is to work out an unfolding map close to the soft edge. To that end we take a pointwise limit of $\rho^{(M)}(\phi)$ (49) with ϕ fixed and $M \rightarrow \infty$. This leads us to define a level density ρ_{sb} for the soft bulk (sb). By expanding (49) in $1/M$, we obtain

$$\begin{aligned} \rho^{(M)}(\phi) &\stackrel{M \gg 1}{\approx} \frac{1}{M\pi} \left[1 - \frac{2\phi}{\tan[\phi]} + \frac{\phi^2}{\sin^2[\phi]} \right] + O\left(\frac{1}{M^2}\right) \\ &= \rho_{sb}(\phi) + O\left(\frac{1}{M^2}\right). \end{aligned} \tag{56}$$

First of all, the density ρ_{sb} vanishes like $1/M$. Second, it is not integrable at the hard edge $\phi = \pi$ because it diverges there like $1/(\pi - \phi)^2$, from the last term. However, the hard edge is not of our interest in this section.

The corresponding parametrization in N and M of the original eigenvalues y of the matrix Y follows from expanding ζ in Eq. (48) in powers of $1/M$:

$$y = N^M \zeta \stackrel{M \gg 1}{\approx} (M + 1)N^M \frac{\sin(\phi)}{\phi} \exp\left[\frac{\phi}{\tan(\phi)}\right]. \tag{57}$$

The corresponding level density reads as

$$\rho_Y(y) = \rho_{sb}(\phi) \left| \frac{\partial \phi}{\partial y} \right| = \frac{\phi(y)}{\pi(M + 1)y}, \tag{58}$$

which follows easily by first computing $\partial y / \partial \phi$ from (57).

We can also go back to the deep bulk by letting $\phi \rightarrow \pi$. When we scale this limit like $\phi = \pi - (\pi + \theta)/M$, due to the divergence of $\rho_{sb}(\phi) \sim \pi/[M(\pi - \phi)^2]$, we can recover the density in the deep bulk (50). The shift of θ by π is a relict hinting to the order of the limits. Since we can go back to the deep bulk regime in this way, we can exclude that we have missed any other intermediate mesoscopic scaling regime.

The unfolding in the mesoscopic regime is given by

$$\hat{y} = \int_0^{\phi(y)} d\phi' \rho_{sb}(\phi') = \frac{1}{(M + 1)\pi} \left(\phi(y) - \frac{\phi(y)^2}{\tan(\phi(y))} \right). \tag{59}$$

For the dependence on the original eigenvalue y we need to invert the relation (57) for $\phi = \phi(y)$. Formally one can write the solution as

$$\phi(y) = \int_0^\pi d\phi \Theta \left\{ (M + 1)N^M \frac{\sin(\phi)}{\phi} \exp\left[\frac{\phi}{\tan(\phi)}\right] - y \right\} \tag{60}$$

by noticing that the function on the right-hand side of (57) is strictly decreasing on the interval $[0, \pi]$. We shall use this unfolding in Secs. VI and VIII while discussing the soft edge local statistics.

Finally, when combining (57) and (58) one finds that the density has a square root behavior at the upper edge $y_+ = N^M(M + 1)^{M+1}/M^M \approx e(M + 1)N^M$ of the support

$$\rho_Y(y) \stackrel{y_+ - y \ll y_+}{\approx} \frac{\sqrt{2}}{\pi(M + 1)y_+} \sqrt{1 - \frac{y}{y_+}}. \tag{61}$$

This square root behavior describes a substantial portion of the spectrum when $M \leq O(N)$, namely, about N/M eigenvalues. For instance for $N \gg M$, this number is infinitely large, indicating the onset of bulk statistics but with respect to the unfolding pointed out in this section which is evidently different from the one in Sec. III C 1.

IV. BULK STATISTICS

The local statistics are always defined by choosing a base point $y_0 \in \mathbb{R}_+$ and zooming into the vicinity of this point. Zooming in means here unfolding and this implies that the macroscopic level density becomes flat. Therefore, we first compute the saddle point z_s (with $\text{Im}[z_s] \geq 0$) with respect to the base point y_0 with the aid of Eq. (17).

A. Sine kernel

As we have seen, the corresponding level density is for $\text{Re}[z_s] \gg M$ given by (31). This already tells us what the correct unfolded variables are for the two spectral variables in the kernel (15),

$$x=y_0 \exp\left[\frac{\pi}{\text{Im}[z_s]}\hat{x}\right] \text{ and } y=y_0 \exp\left[\frac{\pi}{\text{Im}[z_s]}\hat{y}\right] \text{ with } \hat{x}, \hat{y}=O(1) \tag{62}$$

because their corresponding density is flat, i.e.,

$$N\rho(y)dy = \frac{\text{Im}[z_s]}{\pi y_0 \exp[\pi\hat{y}/\text{Im}[z_s]]} d\left\{ y_0 \exp\left[\frac{\pi}{\text{Im}[z_s]}\hat{y}\right] \right\} \approx d\hat{y}. \tag{63}$$

The dependence of $z_s(y)$ on \hat{y} is vanishing due to $\text{Im}[z_s] = O(\text{Re}[z_s]/M) \gg 1$. Thus, we can replace $z_s(y)$ by $z_s(y_0)$ which is the reason why we drop its argument.

The microscopic bulk statistics can be readily obtained by plugging the change of variables (62) into the kernel (15):

$$K_Y(x, y) \frac{dy}{d\hat{y}} = \frac{\pi}{\text{Im}[z_s]} \int_{\gamma'} \frac{dt}{2\pi i} \int_{\gamma_s} \frac{ds}{2\pi i} \frac{1}{s - t} \frac{\sin(\pi t) e^{i\pi \text{sign}[\text{Im}(t)]t}}{\sin(\pi s) e^{i\pi \text{sign}[\text{Im}(s)]s}} \exp\left[\frac{\pi\hat{x}}{\text{Im}[z_s]}s - \frac{\pi\hat{y}}{\text{Im}[z_s]}t \right] \exp[-\mathcal{S}(s; y_0) + \mathcal{S}(t; y_0)]. \tag{64}$$

The factor $dy/d\hat{y}$ is the resulting Jacobian of the change of variable. As for the level density we shift the contour γ_t to $\text{Re}[z_s] + i\mathbb{R}$ while the contour γ_s runs through z_s and z_s^* . Hence, we obtain a residuum which leads to

$$K_Y(x, y) \frac{dy}{d\hat{y}} = \frac{\pi}{\text{Im}[z_s]} \int_{z_s^*}^{z_s} \frac{dt}{2\pi i} \exp\left[\frac{\pi(\hat{x} - \hat{y})}{\text{Im}[z_s]} t\right] + \frac{\pi}{\text{Im}[z_s]} \int_{\text{Re}[z_s]-i\infty}^{\text{Re}[z_s]+i\infty} \frac{dt}{2\pi i} \int_{\gamma_s} \frac{ds}{2\pi i} \frac{1}{s-t} \frac{\sin(\pi t) e^{i\pi \text{sign}[\text{Im}(t)]t}}{\sin(\pi s) e^{i\pi \text{sign}[\text{Im}(s)]s}} \exp\left[\frac{\pi\hat{x}}{\text{Im}[z_s]} s - \frac{\pi\hat{y}}{\text{Im}[z_s]} t\right] e^{-\mathcal{S}(s;y_0) + \mathcal{S}(t;y_0)}. \quad (65)$$

The first term is the one we are looking for since it evaluates to

$$\frac{\pi}{\text{Im}[z_s]} \int_{z_s^*}^{z_s} \frac{dt}{2\pi i} \exp\left[\frac{\pi(\hat{x} - \hat{y})}{\text{Im}[z_s]} t\right] = \exp\left[\pi \frac{\text{Re}[z_s]}{\text{Im}[z_s]} (\hat{x} - \hat{y})\right] \frac{\sin[\pi(\hat{x} - \hat{y})]}{\pi(\hat{x} - \hat{y})}. \quad (66)$$

We exploit the invariance (13) to get rid of the exponential factor in the expression on the right-hand side by multiplying it by $\exp[-\pi \text{Re}[z_s](\hat{x} - \hat{y})/\text{Im}[z_s]]$.

The second term multiplied by this new factor vanishes then in the large N, M limit. This can be seen by expanding the integration variables $s = s_0 + \delta s$ and $t = t_0 + i\delta t$ about the two saddle points $s_0, t_0 = z_s, z_s^*$. In particular, we employ (21) for the spectral variable y_0 . The integrals are indeed bounded because for $s_0 = t_0 = z_s$ (and similarly for $s_0 = t_0 = z_s^*$, with $i \rightarrow -i$), and we have

$$\exp\left[\frac{\pi\hat{x}}{\text{Im}[z_s]} s - \frac{\pi\hat{y}}{\text{Im}[z_s]} t - \pi \frac{\text{Re}[z_s]}{\text{Im}[z_s]} (\hat{x} - \hat{y})\right] = \exp\left[i(\hat{x} - \hat{y}) + \frac{\pi\hat{x}}{\text{Im}[z_s]} \delta s - i \frac{\pi\hat{y}}{\text{Im}[z_s]} \delta t\right] \xrightarrow{M, N \rightarrow \infty} e^{i(\hat{x} - \hat{y})}. \quad (67)$$

Note that $\delta s, \delta t = O(\sqrt{\text{Im}[z_s]})$ because of the scaling of the quadratic terms (21). For $s_0 = t_0^* = z_s$ (and similarly for $s_0 = t_0^* = z_s^*$, with $i \rightarrow -i$), we have

$$\exp\left[\frac{\pi\hat{x}}{\text{Im}[z_s]} s - \frac{\pi\hat{y}}{\text{Im}[z_s]} t - \pi \frac{\text{Re}[z_s]}{\text{Im}[z_s]} (\hat{x} - \hat{y})\right] = \exp\left[i(\hat{x} + \hat{y}) + \frac{\pi\hat{x}}{\text{Im}[z_s]} \delta s - i \frac{\pi\hat{y}}{\text{Im}[z_s]} \delta t\right] \xrightarrow{M, N \rightarrow \infty} e^{i(\hat{x} + \hat{y})}. \quad (68)$$

Since these double contour integrals are multiplied by $1/\text{Im}[z_s] = O(M/\text{Re}[z_s]) = o(1)$ the terms vanish so that we find the celebrated sine kernel [34]

$$K_{\text{sin}}(\hat{x}, \hat{y}) = \lim_{M, N \rightarrow \infty} \exp\left[-\pi \frac{\text{Re}[z_s]}{\text{Im}[z_s]} (\hat{x} - \hat{y})\right] K_Y(x, y) \frac{dy}{d\hat{y}} = \frac{\sin[\pi(\hat{x} - \hat{y})]}{\pi(\hat{x} - \hat{y})} \quad (69)$$

for all cases when the base point y_0 satisfies the scaling bounds $y_0 \gg M^{M+1}/(N - M)$ and $1 - y_0/[e(M + 1)N^M] \gg (M/N)^{2/3}$. This extends the result of [14] obtained when M is fixed. The first bound shows the transition to the other double scaling limits. The latter bound at the soft edge follows from the fact that it always vanishes like a square root, particularly the mesoscopic level density has at the soft edge the form (61).

Satisfying both bounds implies the double scaling $N \gg M \gg 1$. Therefore, the sine kernel cannot always be found.

B. Picket fence statistics in the bulk

For this regime, we require the stricter condition $\text{Re}[z_s(y_0)] = o(M)$, rather than $\text{Re}[z_s(y_0)] \leq O(M)$ which was the scaling in Sec. III B. We choose an unfolding of the form

$$x = y_0 \exp[(M + 1)\psi'(1 + z_0)\hat{x}] \quad \text{and} \quad y = y_0 \exp[(M + 1)\psi'(1 + z_0)\hat{y}] \quad \text{with} \quad \hat{x}, \hat{y} = O(1) \quad (70)$$

with z_0 given by (33) with $y \rightarrow y_0$. This change indeed flattens the bulk density (45) in the present scaling limit.

As we have seen in Appendix D, the saddle point z_s is infinitesimally close to the real point z_0 . The contributing summands are those with an index j that satisfies the scaling $|j - z_0| = O(\sqrt{\text{Re}[z_s]/M})$. Because of $\text{Re}[z_s] \approx z_0 \ll M$, the kernel does not vanish only when $z_0 = 0, 1, \dots, N - 1$. Indeed, the Gaussian approximation (40) of the action for the sum representation of the kernel (14) simplifies to a sum of Dirac delta functions. One can show this by the following computation. The kernel in the new variables is

$$K_Y(x, y) \frac{dy}{d\hat{y}} = (M + 1)\psi'(1 + z_0) \sum_{j=0}^{N-1} \int_{\gamma_t} \frac{dt}{2\pi i} \frac{\sin[\pi(j - t)]}{\pi(j - t)} e^{i\pi \text{sign}[\text{Im}(t)]t} \exp[(M + 1)\psi'(1 + z_0)(\hat{x}j - \hat{y}t)] \times \exp[-\mathcal{S}(j; y_0) + \mathcal{S}(t; y_0)] \approx (M + 1)\psi'(1 + z_0) \sum_{j=0}^{N-1} \int_{-\infty}^{\infty} \frac{d\delta t}{2\pi} \frac{\sin[\pi(j - z_0 - i\delta t)]}{\pi(j - z_0 - i\delta t)} \exp\{M\psi'(1 + z_0)[\hat{x}j - \hat{y}(z_0 + i\delta t)]\} \times \exp\left[-\frac{M\psi'(1 + z_0)}{2} [(j - z_0)^2 + \delta t^2]\right]. \quad (71)$$

First, we shift the integration variable $\delta t \rightarrow \delta t - i\hat{y}$:

$$K_Y(x, y) \frac{dy}{d\hat{y}} \approx (M + 1)\psi'(1 + z_0) \exp \left[\frac{M\psi'(1 + z_0)}{2} ([\hat{x} + z_0]^2 - [\hat{y} + z_0]^2) \right] \times \sum_{j=0}^{N-1} \int_{-\infty}^{\infty} \frac{d\delta t}{2\pi} \frac{\sin[\pi(j - z_0 - \hat{y} - i\delta t)]}{\pi(j - z_0 - \hat{y} - i\delta t)} \exp \left[-\frac{M\psi'(1 + z_0)}{2} ([j - z_0 - \hat{x}]^2 + \delta t^2) \right]. \tag{72}$$

In the next step, we can replace the two Gaussians by two Dirac delta functions where we can evaluate the integral over δt . This leads to the result

$$K_Y(x, y) \frac{dy}{d\hat{y}} \approx \exp \left[\frac{M\psi'(1 + z_0)}{2} ([\hat{x} + z_0]^2 - [\hat{y} + z_0]^2) \right] \sum_{j=0}^{N-1} \frac{\sin[\pi(j - z_0 - \hat{y})]}{\pi(j - z_0 - \hat{y})} \delta(j - z_0 - \hat{x}). \tag{73}$$

The factor in front of the sum can be skipped on virtue of the invariance (13) of the kernel. Additionally, we consider the bulk which means that $z_0 - 1, N - 1 - z_0 \gg 1$ so that we can extend the sum into both directions to $\pm\infty$ in the limit $N, M \rightarrow \infty$ when splitting $j - z_0 = \hat{j} + \nu$ with $\hat{j} \in \mathbb{Z}$ and $\nu \in] -0.5, 0.5]$. The parameter $|\nu|$ is thus the distance of z_0 to its closest integer. Hence, we eventually arrive at

$$K_{\text{pf}}(\hat{x}, \hat{y}) = \lim_{M, N \rightarrow \infty} \exp \left[-\frac{M\psi'(1 + z_0)}{2} ([\hat{x} + z_0]^2 - [\hat{y} + z_0]^2) \right] K_Y(x, y) \frac{dy}{d\hat{y}} = \sum_{\hat{j}=-\infty}^{\infty} \frac{\sin[\pi(\hat{j} + \nu - \hat{y})]}{\pi(\hat{j} + \nu - \hat{y})} \delta(\hat{j} + \nu - \hat{x}). \tag{74}$$

This is the kernel of an equidistant spectrum which looks like a picket fence in both directions, thus, the name picket fence statistics (pf). The sine function in the summand is essential since it guarantees that no two eigenvalues lie at the same position. For instance, for the microscopic one- and two-point functions we get

$$R_1(\hat{y}) = \sum_{\hat{j}=-\infty}^{\infty} \frac{\sin[\pi(\hat{j} + \nu - \hat{y})]}{\pi(\hat{j} + \nu - \hat{y})} \delta(\hat{j} + \nu - \hat{y}) = \sum_{\hat{j}=-\infty}^{\infty} \delta(\hat{j} + \nu - \hat{y}) \tag{75}$$

and

$$R_2(\hat{x}, \hat{y}) = R_1(\hat{x})R_1(\hat{y}) - \sum_{\hat{j}, \hat{l}=-\infty}^{\infty} \frac{\sin[\pi(\hat{j} + \nu - \hat{y})]}{\pi(\hat{j} + \nu - \hat{y})} \frac{\sin[\pi(\hat{l} + \nu - \hat{x})]}{\pi(\hat{l} + \nu - \hat{x})} \delta(\hat{j} + \nu - \hat{x})\delta(\hat{l} + \nu - \hat{y}) = R_1(\hat{x})R_1(\hat{y}) - \sum_{\hat{j}, \hat{l}=-\infty}^{\infty} \delta_{\hat{j}\hat{l}} \delta(\hat{j} + \nu - \hat{x})\delta(\hat{l} + \nu - \hat{y}) = \sum_{\hat{j} \neq \hat{l}} \delta(\hat{j} + \nu - \hat{x})\delta(\hat{l} + \nu - \hat{y}), \tag{76}$$

respectively.

Let us underline that the kernel (74) always holds when the real part of the saddle point solution satisfies $1 \ll \text{Re}[z_s] = z_0 \ll M$ and $N - \text{Re}[z_s] \gg 1$. This translates into the original eigenvalue y_0 of the product matrix Y into the form $\ln(Ny_0) \gg M$ and $(M + 1)\psi[N] - \psi[1] - \ln(y_0) \gg \max\{M/N, 1\}$ meaning that one stays away from the hard and soft edge, respectively, as well as $y_0 \ll (M + 1)^{M+1}/[N - M]$ which represents the regime where the distributions of the individual eigenvalues do not overlap anymore. The condition $(M + 1)\psi[N] - \psi[1] - \ln(y_0) \gg \max\{M/N, 1\}$ follows from Eq. (33) at the maximal value $z_0 = N - 1$ and taking the difference of this equation for $z_0 - \delta j$ with δj of order one. Namely, for $\delta j = O(1)$ one still sees the upper boundary of the spectrum.

From these scaling boundaries we see that there is always a part of the spectrum of Y whose spectral statistics follows the kernel (74) regardless of the relation between M and N in the double scaling limit. For $M \gg N$ the entire bulk is governed by the picket fence local spectral statistics. This extends the results of [15] where this was shown for fixed N only.

C. Critical regime in the bulk

The two bulk statistics above have a transition regime which is given by the fact that $\text{Re}[z_s] \approx z_0 = O(M)$, which is included in the scaling $\text{Re}[z_s] \leq O(M)$ in Sec. III B. In this regime, we unfold the variables like in Eq. (70) since the level density is the

same [see (45)]. Replacing $\psi'(1 + z_0)$ by its asymptotic behavior $1/z_0 + 1/(2z_0^2) + \dots$, Eq. (6), we have

$$x = y_0 \exp\left[\frac{M+1}{z_0}\widehat{x}\right] \quad \text{and} \quad y = y_0 \exp\left[\frac{M+1}{z_0}\widehat{y}\right] \quad \text{with} \quad \widehat{x}, \widehat{y} = O(1). \tag{77}$$

The term M/z_0 is of order one for $z_0 = O(M)$. All other terms can be neglected. The next-to-leading one is of order $M/z_0^2 \sim O(1/M)$ so it vanishes in the limit $M \rightarrow \infty$.

We can already start from the intermediate result (72), which was derived for $\text{Re}[z_s] = o(M)$ but still holds here as it only takes into account the computations in Sec. III B which hold for $\text{Re}[z_s] \leq O(M)$. It reads now

$$K_Y(x, y) \frac{dy}{d\widehat{y}} \approx \frac{M+1}{z_0} \exp\left[\frac{M}{2z_0}([\widehat{x} + z_0]^2 - [\widehat{y} + z_0]^2)\right] \times \sum_{j=0}^{N-1} \int_{-\infty}^{\infty} \frac{d\delta t}{2\pi} \frac{\sin[\pi(j - z_0 - \widehat{y} - i\delta t)]}{\pi(j - z_0 - \widehat{y} - i\delta t)} \exp\left[-\frac{M}{2z_0}([j - z_0 - \widehat{x}]^2 + \delta t^2)\right]. \tag{78}$$

This time we cannot replace the two Gaussians by Dirac delta functions since their variance is of order one. To carry out the δt integral, nevertheless, we exploit the same trick as in Eq. (42) and arrive at the result

$$K_{\text{cb}}(\widehat{x}, \widehat{y}) = \lim_{\substack{M, N \rightarrow \infty \\ z_0/M \rightarrow a}} \exp\left[-\frac{M}{z_0}(v + z_0)(\widehat{x} - \widehat{y})\right] K_Y(x, y) \frac{dy}{d\widehat{y}} = \frac{1}{2\pi a} \sum_{\widehat{j}=-\infty}^{\infty} \text{Re} \left\{ \text{erfi} \left[\pi \sqrt{\frac{a}{2}} + i \sqrt{\frac{1}{2a}}(\widehat{j} + v - \widehat{y}) \right] \right\} \exp\left[\frac{1}{a}(\widehat{x} - \widehat{y})\widehat{j}\right]. \tag{79}$$

The subscript cb stands for critical bulk. We used the invariance (13) to introduce the prefactor in the last expression. As in the picket fence case the index j is closely bound to z_0 , $j - z_0 = \widehat{j} + v$ must be of order one. This in combination with the fact that $1 + z_0, N - 1 - z_0 \gg 1$ has allowed us to extend the sum to \mathbb{Z} .

The result (79) holds for those y_0 which are of the order $(M + 1)^{M+1}/[N - M]$ which is the scale where the single eigenvalues start to feel their neighboring eigenvalues. Additionally, we need to stay away from the soft edge so that it is also $(M + 1)\psi[N] - \psi[1] - \ln(y_0) \gg \max\{M/N, 1\}$. This condition is, however, of relevance only when $N = O(M)$. Only then this critical regime reaches the soft edge.

The result (79) was already presented in Ref. [23]. In Sec. VII, we give a detailed derivation showing that the same kernel is obtained from Dyson’s Brownian motion. There is also an alternative version of this kernel shown in Ref. [24] which is expressed in terms of the Jacobi theta function

$$\vartheta(z; \tau) = \sum_{j=-\infty}^{\infty} \exp[\pi i j^2 \tau + 2\pi i j z]. \tag{80}$$

For this purpose, we do not carry out the r integral resulting from the trick (42) so that the kernel has the form

$$K_{\text{cb}}(\widehat{x}, \widehat{y}) = \sqrt{\frac{1}{8\pi a}} \sum_{\widehat{j}=-\infty}^{\infty} \int_{-1}^1 dr \exp\left[-\frac{1}{2a}\widehat{j}^2 + i\left(\pi r - i\frac{1}{a}(-v + \widehat{x})\right)\widehat{j} + \frac{a}{2}\left(\pi r - i\frac{1}{a}(-v + \widehat{y})\right)^2\right]. \tag{81}$$

In this expression one can identify the series with the Jacobi theta function (80) which yields [24]

$$K_{\text{cb}}(\widehat{x}, \widehat{y}) = \sqrt{\frac{1}{8\pi a}} \int_{-1}^1 dr \vartheta\left(\frac{r}{2} - i\frac{1}{2\pi a}(-v + \widehat{x}); i\frac{1}{2\pi a}\right) \exp\left[\frac{a}{2}\left(\pi r - i\frac{1}{a}(-v + \widehat{y})\right)^2\right]. \tag{82}$$

Let us finally mention that the kernel (79) exhibits again a discrete translation symmetry $(\widehat{x}, \widehat{y}) \rightarrow (\widehat{x} + 1, \widehat{y} + 1)$ as does the picket fence statistics. It reflects the fact that we have indeed properly unfolded the spectrum since the averaged mean level distance is one. When taking the ratio $a = z_0/M \rightarrow 0$ we indeed regain the picket fence kernel (74). This can be easiest seen in Eq. (81) where the Gaussian in \widehat{j} can be replaced by a Dirac delta function and the remaining integral in r is a simple exponential function leading to the sinus cardinalis in Eq. (74).

The sine-kernel result (69) can be found when sending $a = z_0/M \rightarrow \infty$. Then, the sum (81) in \widehat{j} becomes quasicontinuous and, hence, a simple Gaussian. Carrying out this Gaussian integral in \widehat{j} leads again to an exponential integrand in r giving the sinus cardinalis in Eq. (69).

V. HARD EDGE

For the hard edge, the digamma function $\psi[1 + z_0]$ as well as its derivatives cannot be approximated by the asymptotic (6) while the digamma function $\psi[N - z_0]$ is essentially $\ln(N)$ and its derivatives are subleading. Therefore, the reduced saddle

point equation (33) simplifies to

$$\psi[1 + z_0] = \frac{\ln(Ny)}{M + 1}. \tag{83}$$

Hence, the spectral variable scales like $y = \exp[(M + 1)\psi(1 + z_0)]/N$. When comparing this result with those in Ref. [23] one can easily notice a difference by a factor $\exp[\psi(1 + z_0)]$. As long as $\ln(z_0) \ll M$, this term always vanishes in the unfolding because its M th root goes to one. Yet, there will be eventually corrections when $\ln(z_0) \geq O(M)$. This latter regime is not the case for the hard edge scaling. As we will see, the variable z_0 has to be still very close to an integer, so that the smooth part of the level density for y follows:

$$\rho_{\text{hard}}(y) = \frac{1}{N(M + 1)y \psi'(1 + \int_0^\infty ds \Theta\{y - \exp[(M + 1)\psi(1 + s)]\}/N)}. \tag{84}$$

In principal, one can also consider this density as the mesoscopic part of the density (45) close to the hard edge. Therefore, the unfolding is given by

$$x = \exp[(M + 1)\psi[1 + \hat{x}]]/N \text{ and } y = \exp[(M + 1)\psi[1 + \hat{y}]]/N \text{ with } \hat{x}, \hat{y} = O(1). \tag{85}$$

Essentially, we can make use of the analysis of Secs. III B and IV B since the present discussion does not differ much from it. The only but essential difference is that \hat{x} and \hat{y} are already the solutions of (33) so that we expand the summation index j about \hat{x} and the integration variable t about \hat{y} . Here, we begin with the sum representation of the kernel (14). The actions can be approximated then by

$$-\mathcal{S}(j; x) + \mathcal{S}(t; y) \stackrel{M, N \gg 1}{\approx} -\mathcal{S}(\hat{x}; x) + \mathcal{S}(\hat{y}; y) - \frac{(M + 1)\psi'(1 + \hat{x})}{2}(j - \hat{x})^2 - \frac{(M + 1)\psi'(1 + \hat{y})}{2}\delta t^2 - i\pi \text{ sign}(\delta t) \tag{86}$$

with $t = \hat{y} + i\delta t$. Plugging this into the kernel (14), we obtain

$$K_Y(x, y) \frac{dy}{d\hat{y}} = (M + 1)\psi'(1 + \hat{y}) \exp[-\mathcal{S}(\hat{x}; x) + \mathcal{S}(\hat{y}; y)] \times \sum_{j=0}^{N-1} \int_{-\infty}^{\infty} \frac{dt}{2\pi} \frac{\sin[\pi(j - \hat{y} - i\delta t)]}{\pi(j - \hat{y} - i\delta t)} \exp\left[-\frac{(M + 1)\psi'(1 + \hat{x})}{2}(j - \hat{x})^2 - \frac{(M + 1)\psi'(1 + \hat{y})}{2}\delta t^2\right]. \tag{87}$$

The Gaussian can be anew replaced by Dirac delta functions since their variance shrinks like $1/M$. Moreover, we can take the limit of the upper boundary of j to infinity because it has to be close to \hat{x} which is of order one. Multiplying the kernel with the factor $\sqrt{\psi'(1 + \hat{x})/\psi'(1 + \hat{y})} \exp[\mathcal{S}(\hat{x}; x) - \mathcal{S}(\hat{y}; y)]$ [see (13)] leads us to the final result of this section

$$K_{\text{pt}}^+(\hat{x}, \hat{y}) = \lim_{M, N \rightarrow \infty} \sqrt{\frac{\psi'(1 + \hat{x})}{\psi'(1 + \hat{y})}} \exp[\mathcal{S}(\hat{x}; x) - \mathcal{S}(\hat{y}; y)] K_Y(x, y) \frac{dy}{d\hat{y}} = \sum_{j=0}^{\infty} \frac{\sin[\pi(j - \hat{y})]}{\pi(j - \hat{y})} \delta(j - \hat{x}). \tag{88}$$

This kernel agrees with (74), apart from the shift ν , and, indeed, it is the spectral statistics of a picket fence spectrum with a lower bound, hence the superscript $+$. The sine function again ensures that no two eigenvalues sit at the same position.

We would like to emphasize that this part of the spectrum always shows up regardless how M and N are sent to infinity as only the relation between $\text{Re}[z_s]$ and M has been important for the derivation.

VI. SOFT EDGE

We start again from (14) and exploit the knowledge that the action is convex on the interval $] - 1, N[$. At the soft edge the spectral variable y grows so strongly in the large N, M limit (see Sec. IV), that z_0 , which is defined by (33), is close to the value $N - 1$, in particular $N - 1 - z_0 = O(1)$. The question is again how many summands can contribute. To solve this problem, we consider the scaling bound on $\Delta \leq O(1)$ [Eq. (39)] defined in Eq. (38). It has been obtained from the difference of the action at j and z_0 which has to be of order one or less to be contributing, as $\Delta \gg 1$ will be exponentially suppressed in the sum. In the present case we have to replace z_0 by $N - 1 - \hat{z}_0$ with $\hat{z}_0 = O(1)$ such that we consider

$$|\delta j| \leq O\left(\frac{1}{\sqrt{\int_0^1 d\lambda (1 - \lambda)[(M + 1)\psi'(N - \hat{z}_0 + \delta j\lambda) + \psi'(\hat{z}_0 + 1 - \delta j\lambda)]}}\right) \tag{89}$$

with $\delta j = j - N + 1 + \hat{z}_0$. When assuming that $|\delta j| < O(N)$ and plugging this into (89), we get the new tighter bound

$$|\delta j| \leq O\left(\min\left\{\sqrt{\frac{N}{M}}, \sqrt{|\delta j|}\right\}\right). \tag{90}$$

We can combine this inequality with the knowledge that the smallest nonvanishing $|\delta j|$ is equal to 1. This leads to $|\delta j| \leq O(1)$ which is immediate for $N \leq M$. For $N \gg M$, we can plug $|\delta j| \leq O(\sqrt{N/M})$ back into (90) which shows $|\delta j| \leq O(\sqrt{|\delta j|})$ and, thus, $|\delta j| \leq O(1)$.

For the integration variable $t = z_0 + i\delta t$, we are looking for the maximum t_{\max} of $\text{Re}[S(z_0 + i\delta t, y)]$ with respect to δt . Its derivative yields the imaginary part of the saddle point equation (17) with the fixed real part $\text{Re}(z_s) \rightarrow z_0$. The upper bound (B10) for the positions of the maxima still holds so that we have still the scaling bound $t_{\max} \leq O(z_0/M) = O(N/M)$. The lower bound in Eq. (B10) is not useful anymore because z_0 can be very close to $N - 1$.

Let us point out that $\text{Im}[S'(z_0 + i\delta t, y)]$ is strictly decreasing for $\delta t \in \mathbb{R}_-$ and for $\delta t \in \mathbb{R}_+$, separately. Thus, $\text{Re}[S(z_0 + i\delta t, y)]$ is concave on both of the two half-axes \mathbb{R}_- and \mathbb{R}_+ so that the maxima are unique on each of the two parts. What we have to check is the width of the maximum which is given by the second derivative of $\text{Re}[S(z_0 + i\delta t, y)]$ with respect to δt at t_{\max} :

$$-\text{Re}[S''(z_0 + it_{\max}, y)] = -(M + 1)\text{Re}[\psi'(1 + z_0 + it_{\max})] + \text{Re}[\psi'(N - z_0 - it_{\max})] = O\left(\max\left\{\frac{M}{N}, 1\right\}\right), \quad (91)$$

where we have used $t_{\max} \leq O(N/M)$ and the asymptotic expansion (6). Hence, when expanding about $\delta t = 0$ it is safe to say that we choose δt of order $O(\max\{N/M, 1\})$. It is very important that we take N/M and not the width of the maximum $\sqrt{N/M}$ since for the case $N \gg M$ the maximum might be further away from the real axis than the width of the maximum is covering.

In summary, we choose $j = N - 1 - \hat{j}$ with \hat{j} being of order one and also shift the integration variable as follows: $t = N - 1 - \hat{j} + i\hat{t}$. Then, the kernel becomes

$$\begin{aligned} \left(\frac{y}{x}\right)^{N-1} K_Y(x, y) &= \frac{1}{y} \sum_{\hat{j}=0}^{N-1} \int_{-\infty}^{\infty} \frac{d\hat{t}}{2\pi} \frac{\sinh[\pi\hat{t}]}{\pi\hat{t}} x^{-\hat{j}} \frac{\Gamma[N - \hat{j} + i\hat{t}]}{(N - 1 - \hat{j})!} \frac{\Gamma[1 + \hat{j} - i\hat{t}]}{\hat{j}!} \\ &= \frac{1}{y} \sum_{\hat{j}=0}^{N-1} \int_{-\infty}^{\infty} \frac{d\hat{t}}{2\pi i\hat{t}y^{i\hat{t}}} \left\{ \frac{\Gamma[N + i\hat{t}]}{(N - 1)!} \exp\left[-\sum_{l=1}^{\hat{j}} \ln\left(1 + i\frac{\hat{t}}{N - l}\right)\right] \right\}^{M+1} \frac{(-y/x)^{\hat{j}}}{\hat{j}!\Gamma[i\hat{t} - \hat{j}]}. \end{aligned} \quad (92)$$

In the second line we have exploited Euler's reflection formula (A8) for $\Gamma[1 + \hat{j} - i\hat{t}]$. We will remove a factor $(y/x)^{N-1}$ from this intermediate result by the invariance (13) out of convenience as we will see below.

In the next step, we can approximate the exponential

$$\exp\left[-(M + 1) \sum_{l=1}^{\hat{j}} \ln\left(1 + i\frac{\hat{t}}{N - l}\right)\right] \stackrel{M, N}{\approx} \exp\left[-i(M + 1) \frac{\hat{t}}{N} \hat{j}\right] \quad (93)$$

because \hat{t} is of order N/M or smaller so that all higher order terms in the expansion above vanish in the limit regardless how M and N are related. Afterward, the sum over \hat{j} can be extended to a series over \mathbb{N}_0 since all terms which are not of order one will be suppressed in the large N, M limit as we have noticed in the previous discussion. The benefit of this extension is the binomial series which can be carried out now, i.e.,

$$\sum_{\hat{j}=0}^{N-1} \frac{(-e^{-i(M+1)\hat{t}/N} y/x)^{\hat{j}}}{\hat{j}!\Gamma[i\hat{t} - \hat{j}]} \approx \sum_{\hat{j}=0}^{\infty} \frac{(-e^{-i(M+1)\hat{t}/N} y/x)^{\hat{j}}}{\hat{j}!\Gamma[i\hat{t} - \hat{j}]} = \frac{1}{\Gamma[i\hat{t}]} \left(1 - \frac{y}{x} e^{-i(M+1)\hat{t}/N}\right)^{i\hat{t}-1}. \quad (94)$$

We eventually arrive at the following intermediate result:

$$\left(\frac{y}{x}\right)^{N-1} K_Y(x, y) \stackrel{M, N \gg 1}{\approx} \frac{1}{y} \int_{-\infty}^{\infty} \frac{d\hat{t}}{2\pi} \frac{y^{-i\hat{t}}}{\Gamma[1 + i\hat{t}]} \left(\frac{\Gamma[N + i\hat{t}]}{(N - 1)!}\right)^{M+1} \left(1 - \frac{y}{x} e^{-i(M+1)\hat{t}/N}\right)^{i\hat{t}-1}. \quad (95)$$

This representation is ideal to perform the three different double scaling limits in the ensuing subsections.

A. Airy kernel: $N \gg M$

As we have seen for the double scaling $N \gg M$, we may have to go in the integration over \hat{t} up to the scale N/M . This is certainly a rough estimate, but it at least covers the contributing parts of the integrand. Indeed, we will see that the scale where \hat{t} is contributing is actually smaller.

In Sec. III C 2 we have seen that the soft edge lies at $y_0 = y_+ = e(M + 1)N^M$. Moreover, it vanishes like the square root (61), so that the local scale is additionally multiplied by $(M/N)^{2/3}$. Therefore, we consider the change of variables

$$x = e(M + 1)N^M e^{(M/N)^{2/3}\chi} \quad \text{and} \quad y = e(M + 1)N^M e^{(M/N)^{2/3}\zeta} \quad \text{with} \quad \chi, \eta = O(1). \quad (96)$$

The coordinates χ and ζ are not yet properly unfolded, and they will need to be corrected for the square root behavior. The kernel will, however, exhibit the well-known form of the Airy kernel [35]. To see this we expand the logarithm of the integrand

of (95) about $\hat{t} = 0$, apart from $\Gamma[1 + \hat{t}]$ where the argument \hat{t} is large compared to 1, so that we use Stirlings formula for this term. In particular, we employ the expansions

$$\begin{aligned}
 -\ln(\Gamma[1 + \hat{t}]) &= -\hat{t} \ln(\hat{t}) + \hat{t} - \frac{1}{2} \ln(\hat{t}) - \frac{\ln(2\pi)}{2} + O\left(\frac{1}{|\hat{t}|}\right), \\
 (M + 1) \ln\left(\frac{\Gamma[N + \hat{t}]}{(N - 1)!}\right) &= (M + 1) \left(i \ln(N) \hat{t} - \frac{1}{2N} \hat{t}^2 \right) + O\left(\frac{M|\hat{t}|}{N} + \frac{M|\hat{t}|^2}{N^2} + \frac{M|\hat{t}|^3}{N^2}\right),
 \end{aligned} \tag{97}$$

and

$$\begin{aligned}
 &(\hat{t} - 1) \ln(1 - e^{(M/N)^{2/3}(\zeta - \chi) - i(M+1)\hat{t}/N}) \\
 &= (\hat{t} - 1) \left[\left(\frac{M}{N}\right)^{2/3} (\zeta - \chi) + \ln(1 - e^{-i(M+1)\hat{t}/N}) + \ln\left(1 + \frac{e^{-(M/N)^{2/3}(\zeta - \chi)} - 1}{1 - e^{-i(M+1)\hat{t}/N}}\right) \right] \\
 &= (\hat{t} - 1) \ln\left(i \frac{M + 1}{N} \hat{t}\right) + \frac{1}{2} \frac{M}{N} \hat{t}^2 - \frac{i}{24} \left(\frac{M}{N}\right)^2 \hat{t}^3 + i \frac{1}{2} \left(\frac{M}{N}\right)^{2/3} (\zeta - \chi) \hat{t} + \frac{i}{2} \left(\frac{N}{M}\right)^{2/3} \frac{(\zeta - \chi)^2}{\hat{t}} \\
 &\quad - \left(\frac{N}{M}\right)^{1/3} (\zeta - \chi) + O\left[\frac{M}{N} |\hat{t}| + \left(\frac{N}{M}\right)^{1/3} \frac{1}{|\hat{t}|} + \left(\frac{M}{N}\right)^{5/3} |\hat{t}|^2 + \left(\frac{M}{N}\right)^3 |\hat{t}|^4 + \frac{N}{M} \frac{1}{|\hat{t}|^2}\right].
 \end{aligned} \tag{98}$$

Combining these three terms with $y^{-\hat{t}}$, the kernel takes the form

$$\begin{aligned}
 \frac{y^{N-1} \exp[(N/M)^{1/3} \zeta]}{x^{N-1} \exp[(N/M)^{1/3} \chi]} K_Y(x, y) \frac{dy}{d\zeta} &= \left(\frac{N}{M}\right)^{1/3} \int_{-\infty}^{\infty} \frac{d\hat{t}}{(2\pi i \hat{t})^{3/2}} \exp\left[-\frac{i}{24} \left(\frac{M}{N}\right)^2 \hat{t}^3 - \frac{i}{2} \left(\frac{M}{N}\right)^{2/3} (\zeta + \chi) \hat{t} + \frac{i}{2} \left(\frac{N}{M}\right)^{2/3} \frac{(\zeta - \chi)^2}{\hat{t}}\right] \\
 &\quad \times \exp\left\{O\left[\frac{M}{N} |\hat{t}| + \left(\frac{N}{M}\right)^{1/3} \frac{1}{|\hat{t}|} + \left(\frac{M}{N}\right)^{5/3} |\hat{t}|^2 + \left(\frac{M}{N}\right)^3 |\hat{t}|^4 + \frac{|\hat{t}|^2}{N}\right]\right\}.
 \end{aligned} \tag{99}$$

When \hat{t} scales like $(N/M)^{2/3}$ the first part of the exponential is of order one while the corrections are all smaller than one and, hence, vanish in the large N, M limit. This is the aforementioned scaling that is smaller than the crude approximation by N/M . Therefore, we choose $\hat{t} = (N/M)^{2/3} \delta t$ with $\delta t = O(1)$.

Finally, we arrive at the limiting soft edge kernel for $N \gg M$ which is the Airy kernel

$$\begin{aligned}
 K_{\text{Airy}}(\chi, \zeta) &= \lim_{N, M \rightarrow \infty} \frac{y^{N-1} \exp[(N/M)^{1/3} \zeta]}{x^{N-1} \exp[(N/M)^{1/3} \chi]} K_Y(x, y) \frac{dy}{d\zeta} \\
 &= \int_{-\infty}^{\infty} \frac{d\delta t}{[2\pi(c + i\delta t)]^{3/2}} \exp\left[\frac{1}{24}(c + i\delta t)^3 - \frac{1}{2}(\zeta + \chi)(c + i\delta t) - \frac{1}{2} \frac{(\zeta - \chi)^2}{c + i\delta t}\right],
 \end{aligned} \tag{100}$$

where $c > 0$ is a positive shift to guarantee the convergence. The integration can be made absolutely integrable by tilting the two half-axes.

The integral above is none of the standard representations of the Airy kernel. That can be obtained by introducing a Gaussian integral for the last term in the exponential. After additionally rescaling $c + i\delta t \rightarrow 2(c + i\delta t)$ we have

$$\begin{aligned}
 K_{\text{Airy}}(\chi, \zeta) &= \int_{-\infty}^{\infty} \frac{d\delta t}{2\pi(c + i\delta t)} \int_{-\infty}^{\infty} \frac{d\delta s}{2\pi} \exp\left[\frac{(c + i\delta t)^3}{6} - \zeta(c + i\delta t) + (c + i\delta t)(c - i\delta s)^2 - (\chi - \zeta)(c - i\delta s)\right] \\
 &= \int_{-\infty}^{\infty} \frac{d\delta t}{2\pi(c + i\delta t)} \int_{-\infty}^{\infty} \frac{d\delta s}{2\pi} \exp\left[\frac{(c + i\delta t)^3}{6} + \frac{(c - i\delta s)^3}{6} - \zeta(c + i\delta t) - \chi(c - i\delta s)\right] \\
 &= 2^{1/3} \int_0^{\infty} dr \text{Ai}[2^{1/3}(\zeta + r)] \text{Ai}[2^{1/3}(\chi + r)].
 \end{aligned} \tag{101}$$

In the second step we have shifted, first, $c + i\delta t \rightarrow (2c + i\delta t - i\delta s)$, and in the final one we have introduced the integral $1/(2c + i\delta t - i\delta s) = \int_0^{\infty} dr \exp[-(2c + i\delta t - i\delta s)r]$. Moreover, we have rescaled $(c + i\delta t) \rightarrow 2^{1/3}(c + i\delta t)$ and $(c - i\delta s) \rightarrow 2^{1/3}(c - i\delta s)$ and used the integral representation [32] of the Airy function Ai . The final line of (101) is one of the common representations of the Airy kernel.

As already mentioned, the coordinates (96) are not properly unfolded. When going into the bulk, the level spacing will shrink, due to the square root increase. This is quite unfortunate while connecting these statistics with the uniform picket fence spectrum, which we will derive in the next subsection for the opposite scaling relation $M \gg N$. To amend this problem, we change variables,

$$x = e(M + 1)N^M \left\{ 1 + \text{sign}(\hat{x}) \left[\frac{3\pi(M + 1)}{\sqrt{8N}} |\hat{x}| \right]^{2/3} \right\} \quad \text{and} \quad y = e(M + 1)N^M \left\{ 1 + \text{sign}(\hat{y}) \left[\frac{3\pi(M + 1)}{\sqrt{8N}} |\hat{y}| \right]^{2/3} \right\} \tag{102}$$

with $\widehat{x}, \widehat{y} = O(1)$ instead of (96), that follows from the mesoscopic level density (61) very close to the soft edge. The unfolding outside the support of the mesoscopic level density is somewhat artificial, but in this way one can still catch the tail of the largest eigenvalue. There is certainly one disadvantage, that we get a coordinate singularity at the origin, which is chosen to be the edge of the mesoscopic level density.

The kernel in the new coordinates can be readily obtained by noticing that

$$\zeta = \left(\frac{N}{M}\right)^{2/3} \ln \left\{ 1 + \text{sign}(\widehat{y}) \left[\frac{3\pi(M+1)}{\sqrt{8N}} |\widehat{y}| \right]^{2/3} \right\} \approx \text{sign}(\widehat{y}) \frac{(3\pi|\widehat{y}|)^{2/3}}{2} \tag{103}$$

and similarly for the relation between \widehat{x} and χ . Hence, we have after unfolding

$$\begin{aligned} \widehat{K}_{\text{Airy}}(\widehat{x}, \widehat{y}) &= K_{\text{Airy}}(\chi, \zeta) \frac{d\zeta}{d\widehat{y}} \\ &= \frac{2^{2/3}}{3|\widehat{y}|^{1/3}} \frac{\text{Ai}(\text{sign}(\widehat{x})(3\pi|\widehat{x}|/2)^{2/3})\text{Ai}'(\text{sign}(\widehat{y})(3\pi|\widehat{y}|/2)^{2/3}) - \text{Ai}'(\text{sign}(\widehat{x})(3\pi|\widehat{x}|/2)^{2/3})\text{Ai}(\text{sign}(\widehat{y})(3\pi|\widehat{y}|/2)^{2/3})}{\text{sign}(\widehat{x})|\widehat{x}|^{2/3} - \text{sign}(\widehat{y})|\widehat{y}|^{2/3}}. \end{aligned} \tag{104}$$

The limit $\widehat{x} \rightarrow \widehat{y}$ yields the unfolded microscopic level density which is equal to

$$\widehat{\rho}_{\text{Airy}}(\widehat{y}) = \left(\frac{2\pi^2}{3|\widehat{y}|}\right)^{1/3} \left\{ \left[\text{Ai}'\left(\text{sign}(\widehat{y})\left(\frac{3\pi}{2}|\widehat{y}|\right)^{2/3}\right) \right]^2 - \text{sign}(\widehat{y})\left(\frac{3\pi}{2}|\widehat{y}|\right)^{2/3} \left[\text{Ai}\left(\text{sign}(\widehat{y})\left(\frac{3\pi}{2}|\widehat{y}|\right)^{2/3}\right) \right]^2 \right\}. \tag{105}$$

As can be easily checked, the local mean level spacing equals one even for the largest eigenvalues (mapped to the origin here).

B. Picket fence at the soft edge: $N \ll M$

For the other extreme case $M \gg N$, the maxima of the action are very close to the origin, where the soft edge is located. The properly unfolded scaling variables are

$$x = \exp[(M+1)\psi(N) + M\psi'(N)\widehat{x}] \quad \text{and} \quad y = \exp[(M+1)\psi(N) + M\psi'(N)\widehat{y}] \quad \text{with} \quad \widehat{x}, \widehat{y} = O(1), \tag{106}$$

in particular it is $z_0 \approx N - 1$. This unfolding is essentially the same as in Eq. (70) up to the scaling of z_0 . This is the reason why we can still use the intermediate result (73) because the arguments still hold. The only difference is that $j = N - 1 - \widehat{j}$ has to be in the vicinity of $N - 1$, i.e., $\widehat{j} = O(1)$. This means that we can extend the sum over j from $-\infty$ to 0. This leads to the result

$$\begin{aligned} K_{\text{pf}}^-(\widehat{x}, \widehat{y}) &= \lim_{M, N \rightarrow \infty} \exp \left[-\frac{M\psi'(N)}{2} ([\widehat{x} + N - 1]^2 - [\widehat{y} + N - 1]^2) \right] K_Y(x, y) \frac{dy}{d\widehat{y}} \\ &= \sum_{\widehat{j}=-\infty}^0 \frac{\sin[\pi(-\widehat{j} + \widehat{y})]}{\pi(-\widehat{j} + \widehat{y})} \delta(-\widehat{j} + \widehat{x}) \\ &= K_{\text{pf}}^+(-\widehat{x}, -\widehat{y}). \end{aligned} \tag{107}$$

We see that up to a reflection it agrees with the one at the hard edge (88). It is not very surprising since we are at the upper edge of the picket fence spectrum. Therefore, it is also properly unfolded, implying the mean level spacing is one.

C. Critical regime at the soft edge: $N \propto M$

In the critical regime when $M = O(N)$, the integration variable in \widehat{t} in Eq. (95) is of order one. Hence, we only need to expand the term $\ln(\Gamma[N + i\widehat{t}])$ in \widehat{t} and choose the variables

$$x = N^{M+1}e^\chi \quad \text{and} \quad y = N^{M+1}e^\zeta \quad \text{with} \quad \chi, \zeta = O(1). \tag{108}$$

We would like to mention that the new variables χ and ζ are again not properly unfolded, yet, since they show a transition between the uniformly distributed picket fence spectrum at $M \gg N$ and the square root behavior for $N \gg M$ (see Sec. IV). The kernel in this critical regime at the soft edge (cs) is then

$$K_{\text{cs}}(\chi, \zeta) = \lim_{\substack{M, N \rightarrow \infty \\ N/M \rightarrow a}} \left(\frac{y}{x}\right)^{N-1} K_Y(x, y) \frac{dy}{d\zeta} = \int_{-\infty}^{\infty} \frac{d\widehat{t}}{2\pi} \frac{(1 - \exp[\zeta - \chi - i\widehat{t}/a])^{i\widehat{t}-1}}{\Gamma[1 + i\widehat{t}]} \exp \left[-\frac{1}{2a}\widehat{t}^2 - i\left(\zeta + \frac{1}{2a}\right)\widehat{t} \right]. \tag{109}$$

We have only approximated $(\Gamma[N + i\widehat{t}]/\Gamma[N])^{M+1} \approx N^{i(M+1)\widehat{t}} \exp[-(M+1)\widehat{t}^2/(2N) - iM\widehat{t}/(2N)]$ and $a = N/M$. Apart from some rescaling this is the result in Ref. [23].

Another representation of this kernel has been derived in Ref. [24] and has the form of a double contour integral

$$K_{cs}(\chi, \zeta) = \oint_{\mathcal{C}} \frac{ds}{2\pi i} \int_{-\infty}^{\infty} \frac{dt}{2\pi} \frac{1}{1+it-s} \frac{\Gamma[s]}{\Gamma[1+it]} \frac{\exp[1/(2a)(1+it)^2 - (\zeta + 1/(2a))(1+it)]}{\exp[1/(2a)s^2 - (\chi + 1/(2a))s]}, \quad (110)$$

where the contour \mathcal{C} encircles the poles at $s = 0, -1, -2, \dots$ of $\Gamma[s]$. One can readily show that both expressions agree by taking the residues at $s = 0, -1, -2, \dots$, yielding

$$K_{cs}(\chi, \zeta) = \sum_{\hat{j}=0}^{\infty} \int_{-\infty}^{\infty} \frac{dt}{2\pi} \frac{1}{1+it+\hat{j}} \frac{(-1)^{\hat{j}}}{\hat{j}!\Gamma[1+it]} \frac{\exp\{1/(2a)(1+it)^2 - [\zeta + 1/(2a)](1+it)\}}{\exp\{1/(2a)\hat{j}^2 + [\chi + 1/(2a)]\hat{j}\}}. \quad (111)$$

Next, we shift $it = \hat{t} - \hat{j} - 1$ and perform the resulting binomial series with the help of (94). Then, we arrive at (109).

Let us see how we get the other scaling limits from Secs. VI A and VI B. When we rescale $\chi = M\hat{x}/N$ and $\zeta = M\hat{y}/N$ and take the limit $N/M = a \rightarrow 0$, we can regain the result (107). To this aim, we expand the bracket in Eq. (109) via reading (94) backward and then shift $\hat{t} \rightarrow \hat{t} - i(\hat{j} + \hat{y} + 1/2)$. This leads to four Gaussian terms:

$$\begin{aligned} \frac{M}{N} K_{cs}\left(\frac{M}{N}\hat{x}, \frac{M}{N}\hat{y}\right) &= \frac{M}{N} \sum_{\hat{j}=0}^{\infty} \int_{-\infty}^{\infty} \frac{d\hat{t}}{2\pi} (-1)^{\hat{j}} \frac{\exp\left[-\frac{\hat{t}^2}{2} - \frac{M}{N}\hat{j}(\hat{x} - \hat{y}) - \frac{M}{2N}(\hat{j} + \hat{y} + \frac{1}{2})^2\right]}{(\hat{j} + \hat{y} + 1/2 + \hat{t})\hat{j}!\Gamma[\hat{y} + 1/2 + \hat{t}]} \\ &\approx \frac{M}{N} \sum_{\hat{j}=0}^{\infty} \int_{-\infty}^{\infty} \frac{d\hat{t}}{2\pi} (-1)^{\hat{j}} \frac{\exp\left[-\frac{M}{2N}\hat{t}^2 - \frac{M}{2N}(\hat{j} + \hat{x} + \frac{1}{2})^2 + \frac{M}{2N}(\hat{x} + \frac{1}{2})^2 - \frac{M}{2N}(\hat{y} + \frac{1}{2})^2\right]}{(\hat{j} + \hat{y} + 1/2 + \hat{t})\hat{j}!\Gamma[\hat{y} + 1/2 + \hat{t}]} \end{aligned} \quad (112)$$

The Gaussians for \hat{t} and \hat{j} can be replaced by Dirac delta functions as their variance shrinks with $N/M \rightarrow 0$:

$$\frac{M}{N} K_{cs}\left(\frac{M}{N}\hat{x}, \frac{M}{N}\hat{y}\right) \approx \exp\left[\frac{M}{2N}\left(\hat{x} + \frac{1}{2}\right)^2 - \frac{M}{2N}\left(\hat{y} + \frac{1}{2}\right)^2\right] \sum_{\hat{j}=0}^{\infty} (-1)^{\hat{j}} \frac{\delta(\hat{j} + \hat{x} + 1/2)}{(\hat{j} + \hat{y} + 1/2)\hat{j}!\Gamma[\hat{y} + 1/2]}. \quad (113)$$

For the gamma function $\Gamma[\hat{y} + 1/2]$ we apply Euler's reflection formula (A8) and combine it with the sign $(-1)^{\hat{j}}$. Additionally, we can replace $\hat{j}!$ by $\Gamma[1/2 - \hat{x}]$, which leads us to the final result

$$\lim_{M/N \rightarrow \infty} \frac{M}{N} \frac{\Gamma[1/2 - \hat{x}] \exp\left[-\frac{M}{2N}\left(\hat{x} + \frac{1}{2}\right)^2\right]}{\Gamma[1/2 - \hat{y}] \exp\left[-\frac{M}{2N}\left(\hat{y} + \frac{1}{2}\right)^2\right]} K_{cs}\left(\frac{M}{N}\hat{x}, \frac{M}{N}\hat{y}\right) = K_{\text{pf}}\left(\hat{x} + \frac{1}{2}, \hat{y} + \frac{1}{2}\right). \quad (114)$$

The shift by $\frac{1}{2}$ results from the fact that the square root behavior in the other double scaling limits pushes the largest eigenvalue slightly into the bulk, away from the edge of the mesoscopic support of the level density.

In the opposite limit $N/M = a \rightarrow \infty$, we choose

$$\chi = 1 - \ln\left(\frac{N}{M}\right) + \left(\frac{M}{N}\right)^{2/3} \chi' \quad \text{and} \quad \zeta = 1 - \ln\left(\frac{N}{M}\right) + \left(\frac{M}{N}\right)^{2/3} \zeta'. \quad (115)$$

We need to expand only the ratio in the integral (109) by using the expansions in Eq. (97) and (98). This directly leads to the limit

$$\lim_{N/M \rightarrow \infty} \left(\frac{M}{N}\right)^{2/3} \frac{\exp\left[-\left(\frac{N}{M}\right)^{1/3} \zeta'\right]}{\exp\left[-\left(\frac{N}{M}\right)^{1/3} \chi'\right]} K_{cs}\left(1 - \ln\left(\frac{N}{M}\right) + \left(\frac{M}{N}\right)^{2/3} \chi', 1 - \ln\left(\frac{N}{M}\right) + \left(\frac{M}{N}\right)^{2/3} \zeta'\right) = K_{\text{Airy}}(\chi', \zeta'). \quad (116)$$

As both limits need a rescaling, we immediately notice that the choice (108) has not been the proper unfolding, otherwise, the mean level spacing would have been the same. As for the Airy kernel in Sec. VI A, we employ the unfolding on the mesoscopic scale, Eq. (59). Despite the fact that it has been derived for $N \gg M$ it still works out for $N = O(M)$ because the number of eigenvalues that sit in the tail which is suppressed by $1/M$ is of order $N/M = O(1)$. In Fig. 2, we show that the scaling

$$\phi(\hat{y}) = \int_{-\pi}^{\pi} \Theta\left[\hat{y} - \frac{N}{M\pi}\left(\phi' - \frac{\phi'^2}{\tan(\phi')}\right)\right] d\phi' - \pi \quad \text{with} \quad \zeta = \ln\left(\frac{M}{N}\right) + 1 + \text{sign}(\hat{y})\left[\ln\left(\frac{\sin[\phi(\hat{y})]}{\phi(\hat{y})}\right) + \frac{\phi(\hat{y})}{\tan[\phi(\hat{y})]} - 1\right] \quad (117)$$

yields indeed a properly unfolded microscopic level density

$$\hat{\rho}_{cs}(\hat{y}) = K_{cs}(\zeta, \zeta) \left| \frac{d\zeta}{d\hat{y}} \right| = \frac{M\pi}{N|\phi(\zeta(\hat{y}))|} K_{cs}(\zeta(\hat{y}), \zeta(\hat{y})). \quad (118)$$

In this figure, we notice that the true soft edge is slightly shifted inside by $\frac{1}{2}$, when going from the Airy statistics to the picket fence statistics. The deeper reason for this is that the width of the distributions of the individual eigenvalues becomes increasingly narrow, and the mean of the largest eigenvalue was always inside the bulk of the spectrum. This is a rather fascinating effect,

which is shared with a nonzero vacuum energy of the harmonic oscillator. The question is whether there is some deeper physical meaning behind this observation.

VII. DUALITY BETWEEN LEVEL STATISTICS FOR PRODUCTS AND BROWNIAN MOTION

The evolution of eigenvalues of a Gaussian matrix whose elements perform independent Brownian motions is known as Dysonian Brownian motion. The problem of Dysonian Brownian motion with the initial condition given by equidistant eigenvalues was solved in Ref. [2]. The statistics of eigenvalues at time T is identical as for the matrix

$$H = H_0 + \sqrt{T}H_1, \tag{119}$$

where $H_0 = \text{diag}(0, 1, 2, \dots, N - 1)$ and H_1 is an $N \times N$ matrix from the Gaussian unitary ensemble, with the second moment $\langle \text{tr} H_1^2 \rangle = N^2$. The joint probability density of the eigenvalues $E = \text{diag}(E_1, \dots, E_N)$ of H is given by

$$P_H(E) = \frac{\exp\left[-\sum_{j=0}^{N-1} j^2/(2T)\right]}{N! \prod_{j=0}^{N-1} (\sqrt{2\pi T} j!)} \Delta_N(E) \Delta_N(e^{E/T}) \exp\left[-\frac{\text{tr} E^2}{2T}\right], \tag{120}$$

where $\Delta_N(E) = \det[E_a^{b-1}]_{a,b=1,\dots,N}$ is the Vandermonde determinant. The probability density (120) can be expressed as a determinant of the kernel [see Eq. (3.14) in Ref. [2]]

$$\begin{aligned} K_H^{(N)}(E_1, E_2) &= \sum_{j=0}^{N-1} e^{-(j-E_1)^2/(2T)} \int_{-\infty}^{\infty} \frac{dw}{2\pi T} e^{i(w-E_2)^2/(2T)} \prod_{l \neq j} \frac{iw-l}{j-l} \\ &= \sum_{j=0}^{N-1} \int_{-\infty}^{\infty} \frac{dw}{2\pi T} \frac{\sin[\pi(j-iw)]}{\pi(j-iw)} e^{i(w-E_2)^2/(2T) - (j-E_1)^2/(2T)} \frac{\Gamma[N-iw]\Gamma[iw+1]}{\Gamma[N-j]\Gamma[j+1]}. \end{aligned} \tag{121}$$

We would like to mention that in Ref. [2] the initial conditions were symmetrically chosen about the origin while in our case the origin is at the position of the lowest eigenvalue. Apart from this trivial shift, the kernel (121) is the one in Ref. [2].

One can study the limiting forms of the kernel in the local scale at the edges and in the bulk. In particular, it was shown in Ref. [2] that the kernel takes the following form in the bulk:

$$\tilde{K}_{\text{cb}}(\hat{x}, \hat{y}) = \frac{1}{\pi} \sum_{m=-\infty}^{\infty} \text{Re} \left[\frac{\exp\{-2\pi^2 am(m-1) + i\pi[(v-\hat{y}) + (2m-1)(v-\hat{x})]\}}{2\pi ma + i(\hat{x}-\hat{y})} \right]. \tag{122}$$

The simplest way to derive this result from the kernel (121) is to zoom in at the center of the spectrum where it is locally flat. One can do this by setting $N = 2n + 1$ and choosing the base point close to n , i.e., $E_1 = n + \nu + \hat{x}$, $E_2 = n + \nu + \hat{y}$, and $\nu \in]-0.5, 0.5[$. Employing the expansion $w = n + \delta w$ and $j = n + m$ for $\delta w, m = O(1)$ and the identity (42), one arrives after some manipulations at (122) as was already done in Ref. [2].

It has been surprising for us to discover that the kernel for Dyson’s Brownian motion (122) is equivalent to the kernel for the product of Ginibre matrices (81) that we discussed in Sec. IV C. The equivalence can be derived by applying the Poisson summation formula to (79). Writing

$$\begin{aligned} F(h) &= \int_{-\infty}^{\infty} ds \text{erfi} \left[\pi \sqrt{\frac{a}{2}} + i \sqrt{\frac{1}{2a}}(s + \nu - \hat{y}) \right] \exp \left[\frac{1}{a}(\hat{x} - \hat{y})s - ihs \right] \\ &= \frac{2}{h + i(1/a)(\hat{x} - \hat{y})} \exp \left[-\frac{a}{2} \left(h + i \frac{1}{a}(\hat{x} - \hat{y}) \right)^2 + a \left(h + i \frac{1}{a}(\hat{x} - \hat{y}) \right) \left(\pi + i \frac{1}{a}(\nu - \hat{y}) \right) \right], \end{aligned} \tag{123}$$

we have

$$\begin{aligned} K_{\text{cb}}(\hat{x}, \hat{y}) &= \frac{1}{2\pi a} \sum_{m=-\infty}^{\infty} \text{Re} [F(2\pi m)] \\ &= \frac{1}{\pi a} \sum_{m=-\infty}^{\infty} \text{Re} \left[\frac{\exp \left\{ -\frac{a}{2} \left[\pi(2m-1) + i \frac{1}{a}(\hat{x} - \nu) \right]^2 + \frac{a}{2} \left[\pi + i \frac{1}{a}(\nu - \hat{y}) \right]^2 \right\}}{2\pi m + i(1/a)(\hat{x} - \hat{y})} \right] \\ &= \frac{\exp \left[\frac{1}{2a}(\hat{x} - \nu)^2 - \frac{1}{2a}(\nu - \hat{y})^2 \right]}{\pi} \sum_{m=-\infty}^{\infty} \text{Re} \left[\frac{\exp \{-2\pi^2 am(m-1) + i\pi[(\nu - \hat{y}) + (2m-1)(\nu - \hat{x})]\}}{2\pi ma + i(\hat{x} - \hat{y})} \right] \\ &= \exp \left[\frac{1}{2a}(\hat{x} - \nu)^2 - \frac{1}{2a}(\nu - \hat{y})^2 \right] \tilde{K}_{\text{cb}}(\hat{x}, \hat{y}). \end{aligned} \tag{124}$$

We see that up to an irrelevant factor $\exp[(\hat{x} - \nu)^2/(2a) - (\nu - \hat{y})^2/(2a)]$ see (13), the two kernels are identical. The identification of time T in the Dysonian Brownian motion with the parameter M/z_0 is rather straightforward, namely, $T = a = z_0/M$.

The number of matrices M in the product is proportional to time, if one interprets the product as a transfer matrix, but we also see that it is inversely proportional to time when one maps the kernel to that of Brownian motion. It is a sort of duality. For increasing M the picket fence statistics crystallizes [23]. The duality manifests also as a map between real space modes in one picture and Fourier modes in the other one which is provided by the Poisson summation formula.

It is worth mentioning that a relation between the Dysonian random walk and a multiplicative stochastic process can be found also in Ref. [6] where the joint probability density for singular values of a product matrix is identical to (120).

As we have demonstrated above, the local statistics of the singular values of the random matrix product in the bulk is described by the same kernel for Dysonian Brownian motion with the initial condition given by a picket fence. We can extend that statement even to the soft edge. For this purpose, we study the behavior of the kernel (121) for energy levels near the upper edge of the spectrum

$$E_1 = T[\chi + \ln(N)] \quad \text{and} \quad E_2 = T[\zeta + \ln(N)] \quad \text{with} \quad \chi, \zeta = O(1). \tag{125}$$

We choose the summation index $j = N - 1 - \hat{j}$ and the integration variable is $iw = N - 1 - \hat{j} + i\hat{t}$, with $\hat{j}, \hat{t} = O(1)$, in Eq. (121). The ratio of the gamma functions can be approximated like $\Gamma[N - \hat{j} + i\hat{t}]/\Gamma[N - \hat{j}] \approx N^{i\hat{t}}$ which cancels with the logarithmic shift in the spectral variables [cf. Eq. (125)]. The other gamma function $\Gamma[1 + \hat{j} - i\hat{t}]$ can be rewritten as $(-1)^{\hat{j}}\pi/(\Gamma[i\hat{t} - \hat{j}]\sin[i\pi\hat{t}])$ via Euler's reflection formula (A8) which also cancels the resulting sine function. The binomial series (94) leads to

$$\lim_{N \rightarrow \infty} T \frac{\exp\{[\chi + \ln(N) - (N - 1)/T]^2\}}{\exp\{[\zeta + \ln(N) - (N - 1)/T]^2\}} K_H^{(N)}(T[\chi + \ln(N)], T[\zeta + \ln(N)]) = K_{cs}\left(\chi - \frac{1}{2T}, \zeta - \frac{1}{2T}\right) \tag{126}$$

showing that the Dyson Brownian motion with $T = a = N/M$ gives the same limiting result at the soft edge as the multiplicative model discussed here. The critical behavior at the soft edge shown above was not discussed in Ref. [2].

VIII. NUMERICAL SIMULATIONS AND UNIVERSALITY

It is tempting to conjecture that the local statistics is universal, which means that it holds not only for the product of independent Gaussian matrices, but for a wider class of multiplicative stochastic processes in matrix space. In order to support this conjecture we have performed Monte Carlo simulations of six different kinds of products of random matrices. They are listed below.

(1) The product of independent and identically distributed complex Ginibre matrices [see (10)] has been the main object of interest for our analytical study. Thus, it is most natural to consider it as our first ensemble to simulate, meaning we compute the squared singular values of the product matrix

$X^{(M)}$ that is recursively defined by

$$X^{(j)} = X_j X^{(j-1)} \quad \text{with} \quad X^{(1)} = X_1 \quad \text{for} \quad j = 2, \dots, M, \tag{127}$$

and each X_j is drawn from the Gaussian distribution (10).

(2) As a second product matrix, we generate $X^{(M)}$ as in Eq. (127), except that each X_j is itself a product

$$X_j = \tilde{X}_j D_j, \tag{128}$$

where \tilde{X}_j are independent Ginibre matrices drawn from (10) and D_j are independent diagonal matrices whose diagonal matrix entries are independently and uniformly drawn from the interval [0.5,1.5]. Hence, $X^{(M)}$ is an alternating product of complex Ginibre matrices and real diagonal matrices, that would alone yield Poisson spectral statistics.

(3) The next matrix product $X^{(M)}$, that as a first example involves correlations, is constructed via the recursive relation

$$X^{(j)} = (X_j + X_{j-1})X^{(j-1)} \quad \text{with} \\ X^{(1)} = X_1 \quad \text{for} \quad j = 2, \dots, M, \tag{129}$$

with independent X_j drawn from (10). We call it a short ranged memory model since the consecutive matrices in the product $X_j + X_{j-1}$ and $X_{j+1} + X_j$ are correlated through X_j , which contributes to both of them.

(4) A longer ranged memory model is the one with

$$X^{(j)} = \left(\sum_{l=1}^j 2^{l-j} X_l\right) X^{(j-1)} \quad \text{with} \\ X^{(1)} = X_1 \quad \text{for} \quad j = 2, \dots, M, \tag{130}$$

and independently distributed X_j drawn from (10). The factor 2^{l-j} exhibits an exponential decay of the correlation of the new matrix multiplied with respect to the past ones.

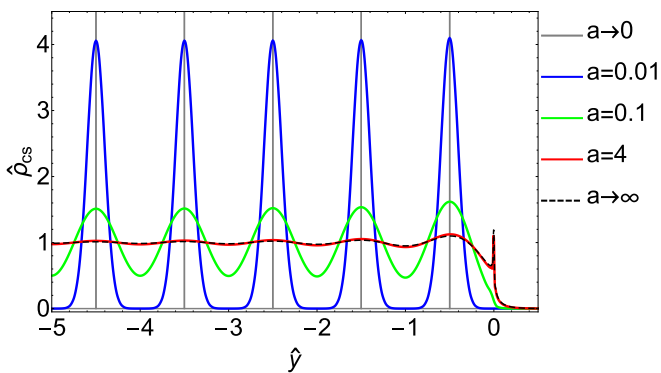


FIG. 2. The properly unfolded level density (118) at the soft edge for various ratios $a = N/M$. The bulk of eigenvalues is to the left. The peak of the distributions at the origin is the price we have to pay for unfolding, which generates a coordinate singularity since the mesoscopic level density vanishes there. Note that the individual eigenvalues are shifted by $\frac{1}{2}$ and do not lie on the integers.

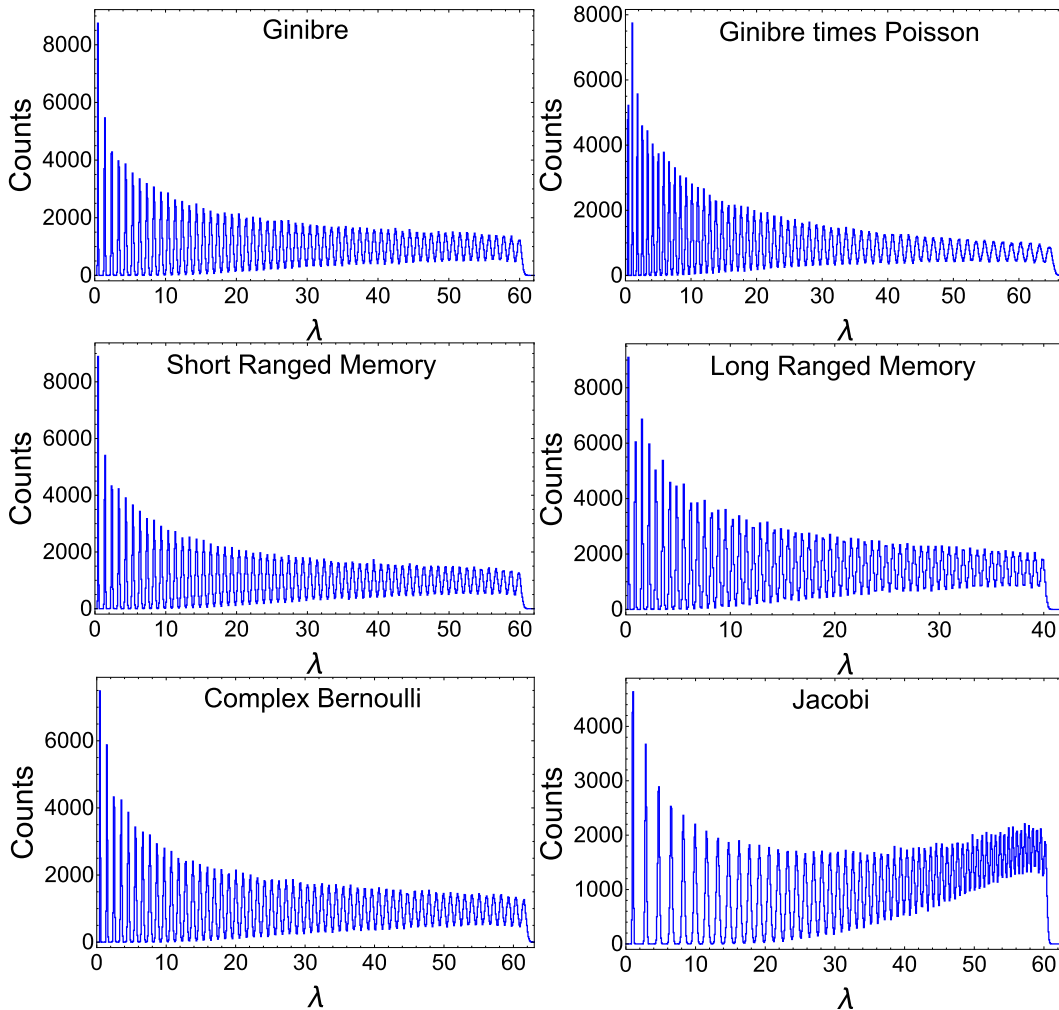


FIG. 3. The total eigenvalue counts (histogram values are proportional to the level density) of the six ensembles. The bin size is equal to 0.1 and has been chosen to be significantly smaller than the peaks' widths in order to see local fluctuations of the level density. Each ensemble consists of 10^4 configurations of a product of $M = 500$ matrices of size 60×60 . Therefore, each histogram comprises 6×10^5 eigenvalues. We have slightly rescaled the eigenvalues to be able to apply the same bin size.

(5) To emphasize that also non-Gaussian ensembles share the same limiting statistics, we employed the recursion (127), but now with independent complex Bernoulli matrices X_j , meaning each matrix entry of X_j is independently and uniformly drawn from the set $\{0, \pm 1\} + i\{0, \pm 1\}$.

(6) Another ensemble which now destroys the independence of the matrix entries, but not the one between the matrices X_j , is the Jacobi ensemble. Such a Jacobi ensemble can be generated by taking a sub-block, which is X_j , of a Haar-distributed unitary random matrix $U_j \in U(L)$ with $L > N$. In the present case, we have chosen $L = 2N$. We have anew exploited the recursion (127) and have drawn M independent $U_j \in U(2N)$ so that also the sub-blocks X_j , that are multiplied, are independent. Recently, this product has been analytically studied in Ref. [26].

The level density (times the number of matrices generated) of the six ensembles is drawn in Fig. 3. For all ensembles we have set $M = 500$ and $N = 60$, and generated 10^4 product matrices. In some ensembles, we have rescaled the eigenvalues by a constant factor. In this way, we could choose the same bin size.

The peaks of each single eigenvalue can be nicely seen in each plot in Fig. 3. This picture is very natural and shows that the macroscopic and mesoscopic level densities are not approached uniformly, when being in the critical [$a = O(1)$] or in the subcritical ($a \ll 1$) regime. The oscillations are persistent.

Additionally, it is evident that the hard edge about the origin always converges to picket fence statistics, regardless whether the matrices are Gaussian or non-Gaussian, and uncorrelated or correlated. The distributions of the individual eigenvalues only start to overlap significantly when departing from the origin and have the strongest overlap at the soft edge.

It is the soft edge result (118) that we fit to the numerics, with an effective a that has not necessarily to be N/M for the non-Ginibre ensembles. This is particularly seen for the product of complex Jacobi matrices, where $a \approx 0.2$ while M and N are still the same as the Ginibre case. Albeit we have chosen the soft edge and not the bulk, the fitting with the soft edge result highlights also the agreement with the bulk statistics (79). The transition from the soft edge to the bulk statistics is rapid, as we know from the transition between the

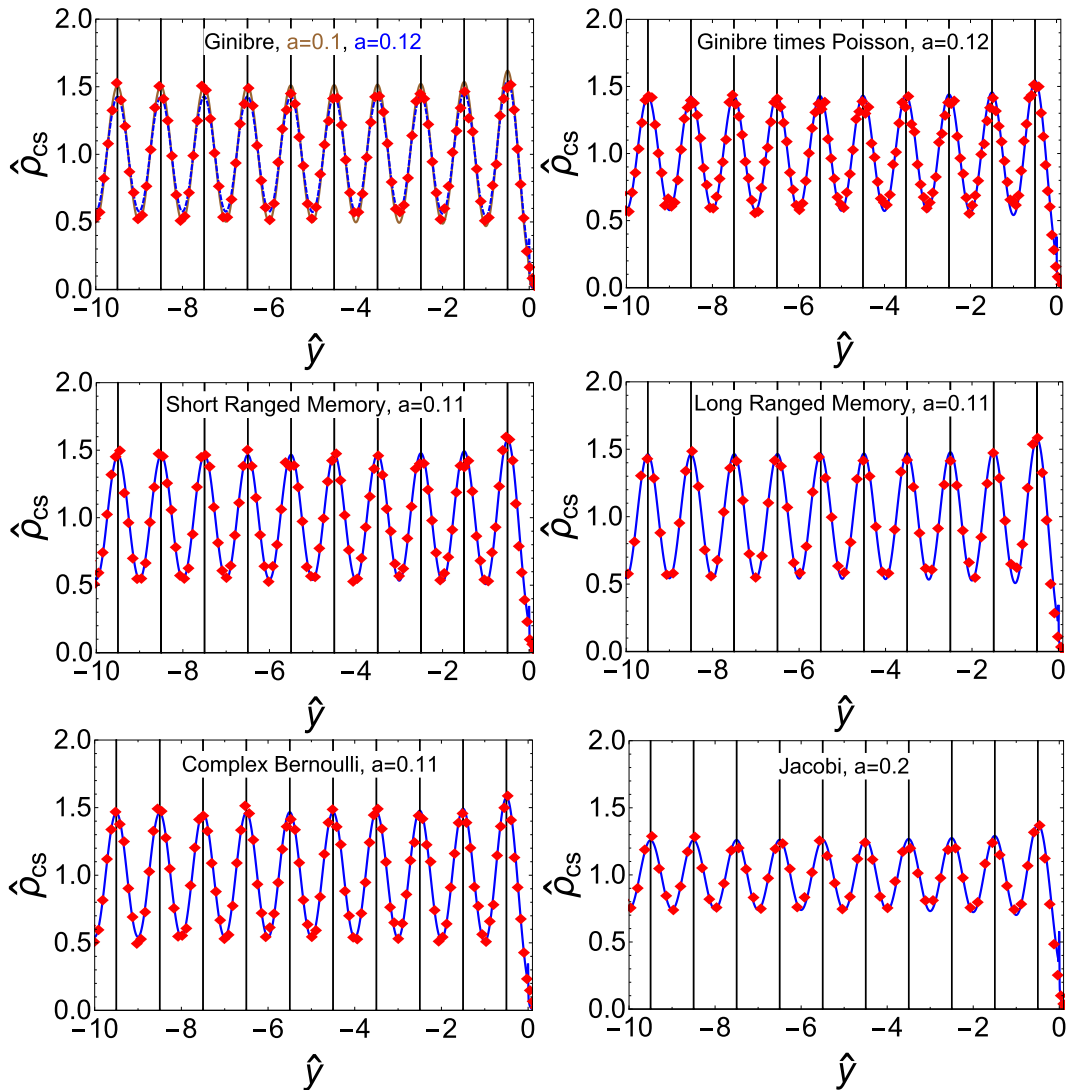


FIG. 4. The unfolded microscopic level density of the six ensembles 1–6 of product matrices at the soft edge (symbols) compared to the analytical result (118) in combination with (109) (smooth curves), with a fitted effective value a that replaces the ratio N/M . Even for the Ginibre case (upper left plot), for which we have analytically studied the double scaling limits, we cannot rely completely on the relation $a = N/M$, which would be $a = 60/500 = 0.12$ (blue dashed curve) in the present case. Due to the finite matrix size, some deviations show up for smaller eigenvalues. Therefore, we have additionally plotted the result for $a = 50/500 = 0.1$ (brown curve) which fits the 10th largest eigenvalue best. Certainly, similar deviations for the other ensembles can be explained by finite size effects, and should be always taken into account when comparing with empirical data. The matrix size is $N = 60$, the number of matrices multiplied is $M = 500$, and the ensemble size is 10^4 for all six ensembles.

Airy kernel and the sine kernel. Already after three or four eigenvalues the statistical error of the empirical data is larger than the actual deviation between the two statistics.

When fitting empirical data with our analytical result in Fig. 4, one has to be aware of two things. First, the finite size effects can be significant and visible. This can be easily observed in the Ginibre case (upper left plot in Fig. 4) where the parameter a is for the largest eigenvalue $a = N/M = 60/500 = 0.12$ while for the 10th largest eigenvalue we have $a = j/M = 50/500 = 0.1$. These deviations have to be taken into account, especially when there are no analytical formulas at hand for the macroscopic or mesoscopic level density.

The next problem to solve is to fix the position of the largest eigenvalue. While the distributions of the other eigenvalues can be relatively easily fixed by enforcing that the

mean level spacing of each pair of consecutive eigenvalues is equal to one, we have no straightforward information about the position of the largest eigenvalue. We have solved this problem by fixing the tail, by rescaling the distribution of the largest eigenvalue with a fixed constant. This very crude method works rather well when considering the results shown in Fig. 4.

IX. SUMMARY

In this work we have delivered the details of the derivation for the results for the double scaling limits of a product of Ginibre matrices, presented by us in Ref. [23]. Furthermore, we gained more insights in the mechanisms behind the transition of the local spectral statistics from picket fence

(equidistant eigenvalues) to GUE statistics. One of these insights is the exact equivalence of the kernel with those that can be obtained from Dyson's Brownian motion, with the picket fence statistics as its initial condition, which we have proven in this work for the bulk as well as the soft edge statistics. The main difference of the additive process of Dyson's Brownian motion and of its multiplicative counterpart considered here, is that the transition of the statistics is not uniform for the whole spectrum. In general, different parts of the spectrum of such a product matrix correspond to different times in the Dyson Brownian motion. This time parameter is equal to $a = j/M$, where j stands for the j th smallest eigenvalue, or in general equal to the squared width to spacing ratio $a = \text{WSR}_j^2$ [cf. Eq. (5)].

This kind of universality between additive and multiplicative stochastic processes on matrix spaces was substantiated with Monte Carlo simulations of product matrices, that also comprise correlations between the matrices multiplied and non-Gaussian ensembles. We have fitted the analytical results at the soft edge to the empirical data, finding very good agreement. In doing so, we have exploited another insight which concerns an emerging mesoscopic spectral scale interpolating between the bulk and the soft edge. Albeit the microscopic statistics in this very narrow part close to the largest eigenvalue is the same as the bulk statistics, its averaged level density does not follow the one from the bulk. It is exactly the mesoscopic level density that is needed to properly unfold the spectrum at the soft edge since it always exhibits a square root behavior. This square root edge is not seen any more, regardless whether and how the matrix dimension N and the number M of matrices multiplied are related. An open question is whether the mesoscopic wide correlators will be also very different from those of the macroscopic scale.

Several generalizations of products of Ginibre matrices have been studied such as products of non-Gaussian ensembles [9–11,26] as well as of products of rectangular matrices

[7,12,13]. Regarding the first, there was recently a study [26] on products of Jacobi ensembles (truncated unitary matrices) where the authors found similar effects as we have seen (see also Sec. VIII). We also expect that this carries over to a product over rectangular matrices where in the end only the average of the rectangularity ν (difference of the two matrix sizes) of the matrices multiplied may enter the game. We have not considered those two generalizations in this work so as to keep the technical level of our results as transparent as possible.

Another direction in which one can try to extend the ideas presented in our work is to consider products of real and quaternionic Ginibre matrices. Do those ensembles yield local spectral statistics that still follow those of Dyson's Brownian motion but now for the Gaussian orthogonal and symplectic ensemble, respectively? It is also interesting to study complex eigenvalues (instead of singular values) of the product of Ginibre matrices using the tools developed in Ref. [36], where we refer to [25] for first results. Due to its two-dimensional nature one should expect a different but related behavior due to the equivalence of the eigenvalue and singular value statistics for biunitarily invariant random matrix ensembles that was proven in Ref. [37]. Products of Ginibre ensembles surely satisfy the requirements that are needed for this kind of equivalence.

ACKNOWLEDGMENTS

This work was supported by the German Science Foundation (DFG), through Grant No. CRC1283 "Taming uncertainty and profiting from randomness and low regularity in analysis, stochastics and their applications" (G.A. and M.K.). The Faculty of Physics and Applied Computer Science at AGH University of Science and Technology is thanked for hospitality (G.A. and M.K.) as well as the School of Mathematics and Statistics of the University of Melbourne (G.A.).

APPENDIX A: DERIVATION OF THE KERNEL REPRESENTATION

In this Appendix, we derive the kernel given in Eqs. (14) and (15). As a starting point we use the representation of the kernel in terms of Meijer G functions given in Ref. [7]:

$$K_Y(x, y) = \sum_{j=1}^N G_{1, M+1}^{1, 0} \left(0, \dots, 0 \middle| x \right) G_{1, M+1}^{M, 1} \left(0, \dots, 0 \middle| y \right). \quad (\text{A1})$$

While for the general definition of Meijer G functions we refer to [32], in this particular case the two functions in the last equation are given by the following complex contour integral representations:

$$K_Y(x, y) = \sum_{j=1}^N \int_{\gamma'_s} \frac{ds}{2\pi i} x^s \frac{\Gamma(-s)}{\Gamma^M(1+s)\Gamma(j-s)} \int_{\gamma_t} \frac{dt}{2\pi i} y^t \frac{\Gamma(j+t)\Gamma^M(-t)}{\Gamma(1+t)}. \quad (\text{A2})$$

The contour γ'_s in the first integral encloses all poles of the integrand, given by the non-negative integers, in clockwise direction. The second contour γ_t is a straight line parallel to the imaginary axis $c + i\mathbb{R}$, with $-j < c < 0$ in-between the poles of the integrand for the j th term. In order to make the contour j independent we have chosen $-1 < c < 0$.

It is convenient to first evaluate the sum over j . To that end, we regroup the terms as follows:

$$K_Y(x, y) = \int_{\gamma'_s} \frac{ds}{2\pi i} x^s \frac{\Gamma(-s)}{\Gamma^M(1+s)} \int_{\gamma_t} \frac{dt}{2\pi i} y^t \frac{\Gamma^M(-t)}{\Gamma(1+t)} \sum_{j=1}^N \frac{\Gamma(j+t)}{\Gamma(j-s)}, \quad (\text{A3})$$

and apply the following telescopic property from [13], valid for general integers $n_+ \geq n_-$:

$$\sum_{n=n_-}^{n_+} \frac{\Gamma(z+n)}{\Gamma(w+n)} = \frac{1}{z-w+1} \left(\frac{\Gamma(z+1+n_+)}{\Gamma(w+n_+)} - \frac{\Gamma(z+n_-)}{\Gamma(w-1+n_-)} \right). \tag{A4}$$

It enables us to carry out the sum over j in Eq. (A3):

$$K_Y(x, y) = \int_{\gamma_t} \frac{dt}{2\pi i} y^t \frac{\Gamma^M(-t)}{\Gamma(1+t)} \int_{\gamma'_s} \frac{ds}{2\pi i} x^s \frac{\Gamma(-s)}{\Gamma^M(1+s)} \frac{1}{t+s+1} \left\{ \frac{\Gamma(N+t+1)}{\Gamma(N-s)} - \frac{\Gamma(1+t)}{\Gamma(-s)} \right\}. \tag{A5}$$

The second term in the curly brackets cancels all residua coming from $\Gamma(-s)$. Hence, the corresponding integral over s vanishes if the pole at $t+s+1$ is not included in the contour γ'_s . Under this condition, the remaining part can be written as

$$K_Y(x, y) = \int_{\gamma_t} \frac{dt}{2\pi i} y^t \frac{\Gamma^M(-t)}{\Gamma(1+t)} \int_{\gamma'_s} \frac{ds}{2\pi i} x^s \frac{\Gamma(-s)}{\Gamma^M(1+s)} \frac{1}{t+s+1} \frac{\Gamma(N+t+1)}{\Gamma(N-s)}. \tag{A6}$$

This is the starting point for the two integral representations in Eqs. (14) and (15).

The first representation is obtained as follows. The integral over s picks up the contributions from the residua $\text{Res } \Gamma(z)|_{z=-j} = (-1)^j/\Gamma(1+j)$ of the poles at $s=j$ for $j=0, 1, \dots, N-1$, yielding

$$K_Y(x, y) = \int_{\gamma_t} \frac{dt}{2\pi i} y^t \frac{\Gamma^M(-t)}{\Gamma(1+t)} \sum_{j=0}^{N-1} \frac{(-1)^j x^j}{\Gamma(1+j)^{M+1}} \frac{1}{t+j+1} \frac{\Gamma(N+t+1)}{\Gamma(N-j)}. \tag{A7}$$

We employ Euler’s reflection formula [32]

$$\Gamma(z)\Gamma(1-z) = \frac{\pi}{\sin(\pi z)}, \quad z \notin \mathbb{Z} \tag{A8}$$

to replace $1/\Gamma(1+t)$ by $-\Gamma(-t)\sin(\pi t)/\pi$. Finally, we change the variable $-t \leftrightarrow t+1$ and we arrive at

$$K_Y(x, y) = \frac{1}{y} \sum_{j=0}^{N-1} x^j \int_{\gamma_t} \frac{dt}{2\pi i} y^{-t} \frac{\sin[\pi(j-t)]}{\pi(j-t)} \left(\frac{\Gamma(1+t)}{\Gamma(1+j)} \right)^{M+1} \frac{\Gamma(N-t)}{\Gamma(N-j)}, \tag{A9}$$

after making the contours j independent by choosing γ_t again. The last equation can be cast into (14),

$$K_Y(x, y) = \frac{1}{y} \sum_{j=0}^{N-1} \int_{\gamma_t} \frac{dt}{2\pi i} \frac{\sin[\pi(j-t)]}{\pi(j-t)} e^{i\pi \text{sign}[\text{Im}(t)]t} \exp[-\mathcal{S}(j; x) + \mathcal{S}(t; y)], \tag{A10}$$

when defining the action as

$$\mathcal{S}(z; \alpha) = -i\pi \text{sign}[\text{Im}(z)]z - \ln[\alpha]z + (M+1)\ln[\Gamma(1+z)] + \ln[\Gamma(N-z)]. \tag{A11}$$

This representation (A10) is a good starting point when the local statistics show either picket fence statistics or are in the interpolating regime.

The second representation (15) can be obtained from (A6), by applying the reflection formula (A8) to the two gamma functions $\Gamma(1+t)$ and $\Gamma(-s)$. After substituting $t+1 \rightarrow -t$ and $s \rightarrow -s$ we arrive at

$$K_Y(x, y) = \frac{1}{y} \int_{\gamma_t} \frac{dt}{2\pi i} \int_{\gamma_s} \frac{ds}{2\pi i} \frac{1}{s-t} \frac{\sin(\pi t)}{\sin(\pi s)} \frac{e^{i\pi \text{sign}[\text{Im}(t)]t}}{e^{i\pi \text{sign}[\text{Im}(s)]s}} \exp[-\mathcal{S}(s; x) + \mathcal{S}(t; y)]. \tag{A12}$$

Notice that the contour γ_s now runs in counterclockwise direction and has been contracted, to only enclose the interval $[0, N-1]$. This is because the remaining poles at larger positive integers of the sine function in the denominator are canceled by $\Gamma(N-z)$ from the action.

A similar representation has been derived in Ref. [24]. It will become useful in the case where the local statistics follow the sine or Airy kernel. But, we employ it also to obtain the macroscopic level density that is required to properly unfold the spectrum in Sec. III.

APPENDIX B: SADDLE POINT ANALYSIS

The digamma function $\psi(z) = \Gamma'(z)/\Gamma(z)$ has the following series representation [32, 5.76] away from its poles $z=0, -1, -2, \dots$:

$$\psi(z) = -\gamma + \sum_{l=0}^{\infty} \left(\frac{1}{l+1} - \frac{1}{l+z} \right). \tag{B1}$$

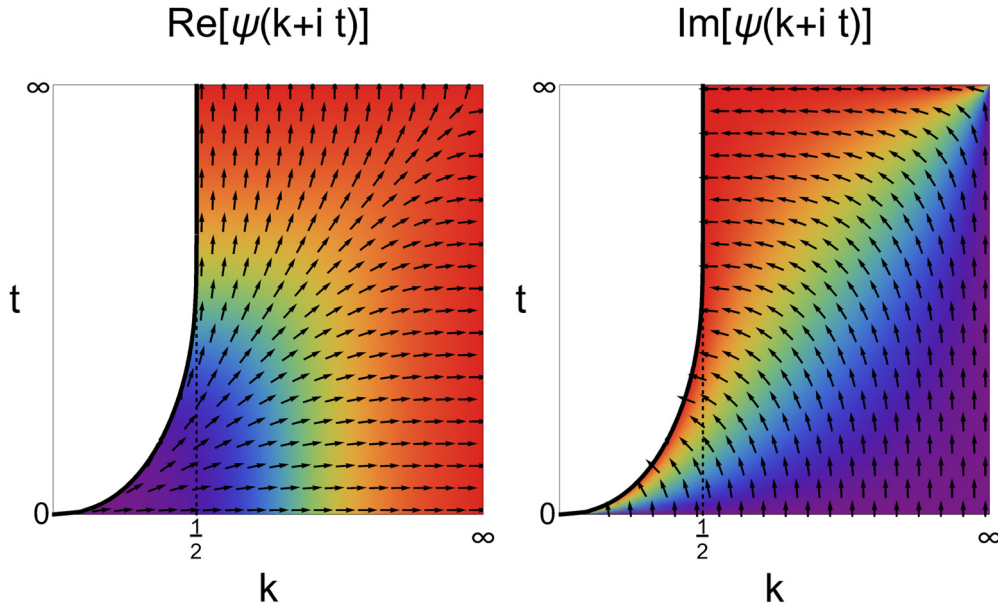


FIG. 5. The real (left plot) and imaginary (right plot) parts of the digamma function (B2) in the upper right quadrant of the complex plane. Note that we have mapped the whole quarter plane to a square by the monotonicity preserving map $(k, t) \mapsto (k/[1+k], t/[1+t])$. The black curve is defined by $\text{Im}[\psi(k+it)] = \pi/2$ and the dashed black line at $k = \frac{1}{2}$ indicates when the $\text{Im}[\psi(k+it)] < 0$ is satisfied for all positive $t > 0$. The height of the real and imaginary parts is color coded in a density plot where purple and red are either $-\infty$ and $+\infty$ for $\text{Re}[\psi(k+it)]$ or 0 and $\pi/2$ for $\text{Im}[\psi(k+it)]$. The vectors highlight the local steepest ascent of the function, in particular, they show our claim that $\text{Im}[\psi(k+it)]$ is increasing in t and $\text{Re}[\psi(k+it)]$ is increasing in k in this regime.

Here, γ is the Euler-Mascheroni constant. The solutions z_s of (17) obviously come in complex conjugate pairs, due to $\psi(z)^* = \psi(z^*)$, unless $z_s \in \mathbb{R}$. We will now show that we have a unique solution in the upper half-plane of the complex plane, and thus also in the lower half-plane.

Let us split $z = k + it$ into its real and imaginary parts. First, we fix the real part $\text{Re}(z) = k > 0$ of a complex solution and assume from now on that $\text{Im}(z) = t \geq 0$. From (B1), it follows for the imaginary part of the digamma function

$$\text{Im}[\psi(k+it)] = \sum_{l=0}^{\infty} \frac{t}{(k+l)^2 + t^2}, \quad k, t \in \mathbb{R}. \quad (\text{B2})$$

It is a strictly increasing function in $t > 0$ as long as $0 \leq \text{Im}[\psi(k+it)] \leq \pi/2$ [see Fig. 5 (right plot)]. For $k \geq \frac{1}{2}$ the restriction $\text{Im}[\psi(k+it)] \leq \pi/2$ holds for all $t \in \mathbb{R}_+$. Unfortunately, we do not have any analytical proof for the monotonicity of the imaginary and real parts of the digamma function in this regime but only a strong numerical evidence. The condition $\text{Im}[\psi(k+it)] \leq \pi/2$ is important because the imaginary part of (17) at the saddle point $z_s = k + it$,

$$0 = \text{Im}[\partial_{z_s} \mathcal{S}(z_s; y)] = (M+1) \sum_{l=0}^{\infty} \frac{t}{(1+k+l)^2 + t^2} + \sum_{l=0}^{\infty} \frac{t}{(N-k+l)^2 + t^2} - \pi \quad \text{for } 0 < k < N, \quad (\text{B3})$$

can only be satisfied in this regime. Indeed, both sums, representing the imaginary part of the two digamma functions involved, are positive for $t > 0$. Hence, there is a unique solution $t_0(k) > 0$ for each fixed $k \in]0, N[$ as both sums are strictly increasing in $t > 0$. Let us insert this function $t_0(k)$ into the real part of the saddle point equation (17):

$$0 = \text{Re}[\partial_{z_s} \mathcal{S}(z_s; y)] = -\ln(y) + (M+1) \sum_{l=0}^{\infty} \frac{(l+k+1)(k+2) + t^2}{(l+k+1)^2 + t^2} - \sum_{l=0}^{\infty} \frac{(l+N-k)(N-k+1) + t^2}{(l+N-k)^2 + t^2} \quad (\text{B4})$$

using

$$\text{Re}[\psi(k+it)] = -\gamma + \sum_{l=0}^{\infty} \left[\frac{1}{l+1} - \frac{l+k}{(l+k)^2 + t^2} \right] = -\gamma + \sum_{l=0}^{\infty} \frac{(l+k)(k-1) + t^2}{(l+1)[(l+k)^2 + t^2]}. \quad (\text{B5})$$

The latter is also strictly increasing for a strictly increasing function in $k > \frac{1}{2}$ for $t \geq 0$ (see left plot in Fig. 5), and thus we obtain also a unique solution for k_0 along $t_0(k)$.

Equation (B2) is also ideal to get an idea of what order the imaginary part of z_s is in N and M . Since the summand in Eq. (B2) is a strictly decreasing function in the summation index l for $t > 0$, we can give the upper and lower limits for the sum

$$\int_0^\infty du \frac{t}{(k+u)^2+t^2} \leq \sum_{l=0}^\infty \frac{t}{(k+l)^2+t^2} \leq \int_{-1}^\infty du \frac{t}{(k+u)^2+t^2}. \quad (\text{B6})$$

After evaluating these integrals and using $\arctan(1/x) = \operatorname{arccot}(x) = \frac{\pi}{2} - \arctan(x)$ for $x > 0$, we have (B2),

$$\arctan\left(\frac{t}{k}\right) \leq \operatorname{Im}[\psi(k+it)] \leq \arctan\left(\frac{t}{k-1}\right), \quad (\text{B7})$$

with $k > 1$ for the upper bound and $k > 0$ for the lower one. Inserting the lower and the upper bounds of (B7) into (B3), we obtain

$$(M+1) \arctan\left(\frac{t}{1+k}\right) + \arctan\left(\frac{t}{N-k}\right) - \pi \leq 0 \quad (\text{B8})$$

and

$$(M+1) \arctan\left(\frac{t}{k}\right) + \arctan\left(\frac{t}{N-k-1}\right) - \pi \geq 0, \quad (\text{B9})$$

respectively. Simple manipulations of (B8) and (B9) lead to

$$\begin{aligned} \operatorname{Im}[z_s] &\leq (1 + \operatorname{Re}[z_s]) \tan \left\{ \frac{1}{M+1} \left[\pi - \arctan\left(\frac{\operatorname{Im}(z_s)}{N - \operatorname{Re}[z_s]}\right) \right] \right\} \leq [1 + \operatorname{Re}(z_s)] \tan \left[\frac{\pi}{M+1} \right], \\ \operatorname{Im}[z_s] &\geq \operatorname{Re}[z_s] \tan \left\{ \frac{1}{M+1} \left[\pi - \arctan\left(\frac{\operatorname{Im}(z_s)}{N-1-\operatorname{Re}[z_s]}\right) \right] \right\} \geq \operatorname{Re}[z_s] \tan \left[\frac{\pi}{2(M+1)} \right], \end{aligned} \quad (\text{B10})$$

where we have reinserted $k = \operatorname{Re}[z_s]$ and $t = \operatorname{Im}(z_s) \geq 0$. For the inequalities on the right-hand sides we have used the monotonicity of \arctan , and the fact that it is bounded by 0 and $\pi/2$ for positive arguments.

APPENDIX C: RESOLVENT: SADDLE POINT RELATION FOR $\operatorname{Re}[z_s] \gg M$

The saddle point z_s can be related to the limiting resolvent G_Y using (31). It is defined as an integral of the limiting normalized density ρ_Y over its support σ :

$$G_Y(w) = \lim_{N \rightarrow \infty} \int_{w \notin \sigma} \frac{du \rho_Y(u)}{w-u} \stackrel{|w| \gg 1}{\approx} \frac{1}{w}. \quad (\text{C1})$$

The large argument asymptotic follows from the normalization of the density. The limiting density is recovered via

$$\rho_Y(y) = -\frac{1}{2\pi i} \lim_{\epsilon \searrow 0} [G_Y(y+i\epsilon) - G_Y(y-i\epsilon)] = \frac{1}{\pi} \lim_{\epsilon \searrow 0} \operatorname{Im} G_Y(y-i\epsilon). \quad (\text{C2})$$

Comparing with the relation (31), this suggests to identify

$$z_s(y) \stackrel{M, N \gg 1}{\approx} N \frac{y}{N^M} \lim_{\epsilon \searrow 0} G_Y\left(\frac{y}{N^M} - i\epsilon\right) \quad (\text{C3})$$

with the resolvent for the matrix Y/N^M , where we have inserted the scaling implied by (32). The large argument behavior of $z_s(y)$ found above is consistent with that of the resolvent in Eq. (C1) and fixes all constants.

The scaling considered here is for $N \gg M$, as follows from Sec. III A. Let us therefore make contact with previous results for ρ_Y obtained in the limit $N \rightarrow \infty$ with fixed M , as here the limiting density and resolvent are known [33,38]. We denote by $G^{(M)}$ the limiting Green function of the product matrix Y/N^M when we take $N \rightarrow \infty$ for a fixed M , i.e.,

$$G^{(M)}(\hat{z}) = \int_0^{(M+1)^{M+1}/M^M} \frac{\rho^{(M)}(\zeta) d\zeta}{\hat{z} - \zeta}. \quad (\text{C4})$$

The limiting density fulfils the equation [38]

$$\hat{z}^M (G^{(M)}(\hat{z}))^{M+1} - \hat{z} G^{(M)}(\hat{z}) + 1 = 0. \quad (\text{C5})$$

The corresponding limiting level density $\rho^{(M)}(\zeta)$ was found analytically in Ref. [33], which is given in Eq. (48).

APPENDIX D: DISTANCE BETWEEN z_0 AND z_s WHEN $\text{Re}[z_s] \leq O(M)$

In a similar fashion as in Eqs. (36) and (38), one can also estimate the real part of the difference of the action at z_0 and at the original saddle point z_s in the complex plane, i.e.,

$$\tilde{\Delta} = \text{Re}[S(z_s; y) - S(z_0; y)] = \text{Re} \left[(\Delta z)^2 \int_0^1 d\lambda (1 - \lambda) [(M + 1)\psi'(1 + z_0 + \Delta z\lambda) + \psi'(N - z_0 - \Delta z\lambda)] \right], \quad (\text{D1})$$

with $\Delta z = z_s - z_0$. The value of the quantity $\tilde{\Delta}$ will tell us whether the approximation of the original saddle point by z_0 is legitimate.

In order to get a feeling whether Δz is large or small, we start with the initial saddle point equation (17) and show that Δz is maximally of order one. Making use of (18), meaning $\text{Im}[z_s] = O(\text{Re}[z_s]/M)$, we assume that we can expand the saddle point equation (17) in $\text{Im}[z_s] \leq O(1)$, and then take the real part. The first term in the Taylor expansion vanishes since it is imaginary, so that we arrive at

$$-\ln(y) + (M + 1)\psi(1 + \text{Re}[z_s]) - \psi(N - \text{Re}[z_s]) + O \left([(M + 1)\psi''(1 + \text{Re}[z_s]) - \psi''(N - \text{Re}[z_s])] \left[\frac{\text{Re}[z_s]}{M} \right]^2 \right) = 0. \quad (\text{D2})$$

Our assumption has been that we stay far away from the hard edge and soft edge, i.e., $1 + \text{Re}[z_s] \gg 1$ and $N - 1 - \text{Re}[z_s] \gg 1$, respectively [see (19)]. Thus, we are allowed to approximate the digamma functions and its derivatives by their leading terms [see (6)]. Then, the correction of the saddle point equation in Eq. (D2) is of order

$$O \left([(M + 1)\psi''(1 + \text{Re}[z_s]) - \psi''(N - \text{Re}[z_s])] \left[\frac{\text{Re}[z_s]}{M} \right]^2 \right) = O \left(\max \left\{ \frac{1}{M}, \frac{\text{Re}[z_s]^2}{M^2(N - \text{Re}[z_s])^2} \right\} \right) \ll 1. \quad (\text{D3})$$

Here, we have exploited $\text{Re}[z_s] \leq O(M)$ in the current situation, and that M is large.

Let us come back to the question of how far away we are with z_0 from the true saddle point z_s . The solution of (33) is uniquely given by z_0 , without any correction term. Expanding the first term in Eq. (D2) in the difference $\text{Re}[\Delta z] = \text{Re}[z_s] - z_0$, which is small compared to z_0 , we need to enforce that the leading correction in $\text{Re}[\Delta z]$ has to cancel the second term shown in Eq. (D2). Consequently, we get the following scaling relation from setting both orders to be equal:

$$[(M + 1)\psi'(1 + z_0) + \psi'(N - z_0)]\text{Re}[\Delta z] \approx \left[\frac{M}{z_0} + \frac{1}{N - z_0} \right] \text{Re}[\Delta z] = O \left(\max \left\{ \frac{1}{M}, \frac{\text{Re}[z_s]^2}{M^2(N - \text{Re}[z_s])^2} \right\} \right). \quad (\text{D4})$$

Equivalently, it holds

$$|\text{Re}[\Delta z]| = O \left(\max \left\{ \frac{z_0}{M^2}, \frac{z_0 \text{Re}[z_s]^2}{M^3(N - \text{Re}[z_s])^2} \right\} \right) \ll 1 \quad (\text{D5})$$

since M is large, $\text{Re}[z_s] \leq O(M)$, $z_0/M \leq O(1)$, and we are in the bulk (19). This implies $|\Delta z| = |z_s - z_0| = O(|\text{Im}[z_s]|)$ because the real part $\text{Re}[z_s]$ always dominates the imaginary part $\text{Im}[z_s]$ [cf. (18)]. Furthermore, the imaginary part $\text{Im}[z_s] = O(\text{Re}[z_s]/M)$ is always bigger than $|\text{Re}[\Delta z]|$, and thus it determines the order of Δz .

Summarizing, we have not only found that the true saddle point z_s is close to z_0 , but also that the real part $\text{Re}[z_s]$ converges to z_0 when $M, N \rightarrow \infty$. Next, we evaluate the difference $\tilde{\Delta}$ in Eq. (D1). This can be done by expanding the original difference (36) with the replacement $j \rightarrow z_s$, and inserting $z_s = \Delta z + z_0$ to expand in Δz . The logarithms just give an expansion in the digamma function and its derivatives, and after the cancellation in the first order we obtain

$$\tilde{\Delta} \approx (M + 1) \frac{1}{2} (\Delta z)^2 \psi'(1 + z_0) + \frac{1}{2} (\Delta z)^2 \psi'(N - z_0) \approx \frac{1}{2} (\Delta z)^2 \left(\frac{M}{z_0} + \frac{1}{N - z_0} \right). \quad (\text{D6})$$

Due to the condition (37), we can exploit the asymptotic formulas (6) for the derivative of the digamma function in the second step. Clearly, the first term in the brackets on the right-hand side is larger than or equal to the order $O(1)$ and the second term is much less than one so we obtain

$$|\tilde{\Delta}| = O(z_0/M) \leq O(1) \quad (\text{D7})$$

agreeing with the order of $|\text{Im}[z_s]|$. Hence, the expansion about z_0 for the t integral as well as for the summation index j in Eq. (14) about the point z_0 instead of z_s is justified because we cover all contributions from the saddle point.

[1] F. J. Dyson, *J. Math. Phys.* **3**, 1191 (1962).

[2] K. Johansson, *Commun. Math. Phys.* **252**, 111 (2004).

[3] Z. Burda, J. Grela, M. A. Nowak, W. Tarnowski, and P. Warchoř, *Phys. Rev. Lett.* **113**, 104102 (2014).

- [4] O. N. Dorokhov, Pis'ma Zh. Eksp. Teor. Fiz. **36**, 259 (1982) [*JETP Lett.* **36**, 318 (1982)].
- [5] P. A. Mello, P. Pereyra, and N. Kumar, *Ann. Phys.* **181**, 290 (1988).
- [6] J. R. Ipsen and H. Schomerus, *J. Phys. A: Math. Theor.* **49**, 385201 (2016).
- [7] G. Akemann, J. R. Ipsen, and M. Kieburg, *Phys. Rev. E* **88**, 052118 (2013).
- [8] G. Akemann, M. Kieburg, and L. Wei, *J. Phys. A: Math. Theor.* **46**, 275205 (2013).
- [9] G. Akemann and E. Strahov, *Ann. Henri Poincaré* **19**, 2599 (2018).
- [10] M. Kieburg and H. Kösters, *Ann. Inst. Henri Poincaré* **55**, 98 (2019).
- [11] M. Kieburg, A. B. J. Kuijlaars, and D. Stivigny, *Int. Math. Res. Not.* **2016**, 3392 (2016).
- [12] A. B. J. Kuijlaars and D. Stivigny, *Random Matrices: Theor. Appl.* **3**, 1450011 (2014).
- [13] A. B. J. Kuijlaars and L. Zhang, *Commun. Math. Phys.* **332**, 759 (2014).
- [14] D.-Z. Liu, D. Wang, and L. Zhang, *Ann. Inst. Henri Poincaré* **52**, 1734 (2016).
- [15] G. Akemann, Z. Burda, and M. Kieburg, *J. Phys. A: Math. Theor.* **47**, 395202 (2014).
- [16] H. Furstenberg and H. Kesten, *Ann. Math. Statist.* **31**, 457 (1960).
- [17] M. Isopi and C. M. Newman, *Commun. Math. Phys.* **143**, 591 (1992).
- [18] C. M. Newman, *Commun. Math. Phys.* **103**, 121 (1986).
- [19] D. S. P. Richards, *J. Multivariate Anal.* **29**, 326 (1989).
- [20] N. K. Reddy, *Int. Math. Res. Not.* **2019**, 606 (2019).
- [21] V. V. Sazonov and V. N. Tutubalin, *Theor. Probab. Its Appl.* **11**, 1 (1966).
- [22] G. Akemann and J. R. Ipsen, *Acta Phys. Pol. B* **46**, 1747 (2015).
- [23] G. Akemann, Z. Burda, and M. Kieburg, *Europhys. Lett.* **126**, 40001 (2019).
- [24] D. Z. Liu, D. Wang, and Y. Wang, [arXiv:1810.00433](https://arxiv.org/abs/1810.00433).
- [25] D. Z. Liu, and Y. Wang, [arXiv:1912.11910](https://arxiv.org/abs/1912.11910).
- [26] A. Ahn, [arXiv:1910.00743](https://arxiv.org/abs/1910.00743).
- [27] Z. Burda, *Acta Phys. Pol. B* **51**, 1641 (2020).
- [28] G. Akemann, F. Götze, and T. Neuschel, [arXiv:2006.15180](https://arxiv.org/abs/2006.15180).
- [29] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, National Bureau of Standards Applied Mathematics Series 55, 10th printing (Dover, New York, 1972).
- [30] J. Ginibre, *J. Math. Phys.* **6**, 440 (1965).
- [31] K. Johansson, in *Mathematical Statistical Physics*, edited by A. Bovier *et al.*, Vol. 83, Les Houches Summer School 2005, Session LXXXIII (Elsevier, Amsterdam, 2006).
- [32] F. W. J. Olver, D. W. Lozier, R. F. Boisvert, and C. W. Clark (eds.), *NIST Handbook of Mathematical Functions* (Cambridge University Press, Cambridge, 2010), <http://dlmf.nist.gov/>
- [33] T. Neuschel, *Random Matrices Theory Appl.* **3**, 1450003 (2014).
- [34] M. L. Mehta, *Random Matrices*, 3rd ed. (Elsevier, Amsterdam, 2004).
- [35] P. J. Forrester, *Nucl. Phys. B* **402**, 709 (1993).
- [36] G. Akemann and Z. Burda, *J. Phys. A: Math. Theor.* **45**, 465201 (2012).
- [37] M. Kieburg and H. Kösters, *Random Matrices: Theor. Appl.* **05**, 1650015 (2016).
- [38] Z. Burda, G. Livan, and A. Swiech, *Phys. Rev. E* **88**, 022107 (2013).