

# Chaos in QFT

Nadav Shraier<sup>\*</sup> and Jacob Sonnenschein<sup>†</sup>

*The Raymond and Beverly Sackler School of Physics and Astronomy, Tel Aviv University, Tel Aviv, Israel*



(Received 28 January 2026; accepted 27 April 2026; published 26 May 2026)

In this paper we explore the chaotic behavior of nonintegrable quantum field theories and compare them to integrable ones. We choose as prototypes the double sine-Gordon and the sine-Gordon models. We analyze their discrete spectrum determined by a truncation method. We examine the map of the corresponding energy eigenvalues to the eigenvalues of the random matrix theory Gaussian orthogonal ensemble (GOE). This is done by computing the following properties: (a) The distribution of the adjacent spacings and their ratios, (b) higher order spacings and ratios, (c) pair correlations, (d) spectral form factors, and (e) spectral rigidity. For these properties we determine the differences between the integrable and nonintegrable theories and verify that the former admits a Poisson behavior and the latter GOE (apart from the spectral rigidity, which shows mixed behavior).

DOI: [10.1103/nln-v3ny](https://doi.org/10.1103/nln-v3ny)

## I. INTRODUCTION

Chaotic behavior in classical and quantum mechanical systems both of a single body and of many bodies has been intensively researched. The study of ergodicity and chaos in continuous quantum field theories and string theories is much less mature.

Significant progress in the study of quantum chaos in quantum field theories (QFTs) in  $1+1$  dimension has been made in [1–3]. In the second a Hamiltonian truncation method was applied to the integrable sine-Gordon (SG) and nonintegrable double sine-Gordon (DSG) models to determine their discrete spectrum and in the third a light-cone conformal truncation was used in the context of the  $1+1$  dimension  $\phi^4$  theory. In these papers various aspects of the map between these systems and random matrix theory (RMT) were studied. Another contemporary approach to chaos in QFTs deals with a quantum analog of the classical Lyapunov exponent, called the out-of-time-ordering correlators (OTOCs), and a bound on the maximum value it can hold. Maximally chaotic systems are said to saturate this bound [4]. The evolution of systems from integrability to chaos has been intensively studied also using the method of “Krylov complexity” ([5–7] and references therein).

In this paper we continue the program of examining the chaotic behavior of QFTs. There are several different

directions that one can follow for that purpose: (i) Further study of the map between the discrete spectrum associated with truncated field theories and RMT; (ii) analyzing chaotic behavior via this map in different field theories in different space-time dimensions, like theories of scalar fields, spinor fields, gauge theories like QED and QCD, etc.; (iii) exploring chaotic scattering processes similar to those discovered recently and which involve high excited string states [8–12]; (iv) computing OTOC and general correlation functions; (v) using the Krylov complexity method.

Here we follow (i) and present a follow-up work to the paper [2] by studying properties related to chaos in the SG and DSG models.<sup>1</sup>

The RMT properties that we study are: (a) the distribution of the adjacent spacings and their ratios, (b) higher order spacings and ratios, (c) pair correlations, (d) spectral form factors, and (e) spectral rigidity.

The structure of this paper is as follows: In Sec. II we present several measures of chaos: (i) adjacent spacings and their ratios, (ii) higher order spacings and spacing ratios, (iii) pair correlations, (iv) spectral form factor, and (v) higher order correlations (OTOCs). In Sec. III the distributions of nearest-neighbor (NN) spacings and spacing ratios are determined. Section IV deals with the distributions of nonadjacent spacings and their ratios. In particular we discuss a probabilistic approach to higher order level spacings. In Sec. V we write down the results of the pair correlations. The scattering form factor is determined in

<sup>\*</sup>Contact author: nadavshraier@tauex.tau.ac.il

<sup>†</sup>Contact author: cobi@tauex.tau.ac.il

Published by the American Physical Society under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/). Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI. Funded by SCOAP<sup>3</sup>.

<sup>1</sup>Contradicting the common lore in [13] analytical and numerical evidence was presented that classical integrable models possessing infinitely many degrees of freedom unexpectedly exhibit some features that are typical of chaotic systems.

Sec. VI. In Sec. VII we present the  $\Delta_3$  statistics. We conclude with Sec. VIII where a summary of the results is given and future directions are discussed. We add three Appendixes to the paper. In Appendix A we review the method of the truncated conformal space approach (TCSA) and the specific parameters used to obtain the data we worked on. In Appendix B we write down the basic properties of the SG and the DSG models and Appendix C is a brief review of random matrix theory.

## II. MEASURES OF CHAOS

One possible way to define a chaotic behavior of a quantum system is if its energy eigenvalues can be mapped to the eigenvalues of a RMT. To determine the quality of the map one can compute statistical properties of the energy eigenvalues and compare them to those of RMT. These properties include (a) the distribution of the adjacent spacings and their ratios, (b) higher order spacings and ratios, (c) pair correlations, (d) spectral form factors, and (e) spectral rigidity. Prior to computing these quantities we elaborate in this section on each of them. We begin with discussing the unfolding of the spectrum, which is a necessary process in order to compare different spectra.

### A. Unfolding

The universality of chaos is often seen through local statistics of the spectrum, while longer ranged statistics are affected by the model-dependent averaged level density  $\rho_0$ , so it is required to decouple the spectrum from its averaged level density in order to make meaningful comparisons between spectra. This is done in a process called “unfolding,” where the average level density is “flattened” such that it is constant throughout the spectrum, and only the fluctuation around it remains. Another model-dependent aspect of the spectra is the scale of the energy levels, so the unfolding should also align all spectra to the same local scale, which is naturally chosen to be the mean nearest-neighbor spacing  $\langle s \rangle$  such that after unfolding it is equal to 1.

A common scheme for level unfolding is the use of the cumulative distribution of the normalized average level density, where the unfolded levels are  $N$  times the value of the cumulative distribution function

$$e_i = N \int_{-\infty}^{E_i} \rho_0(E') dE'. \quad (1)$$

To see the effect of unfolding explicitly, we will use an example from RMT. Assume  $\{\lambda_i\}$  is a set of  $N$  eigenvalues taken from a Gaussian unitary ensemble (GUE) and normalized by  $1/\sqrt{N}$ . The GUE average level density  $\rho_0$  is found by using  $\beta = 2$  in (C7)

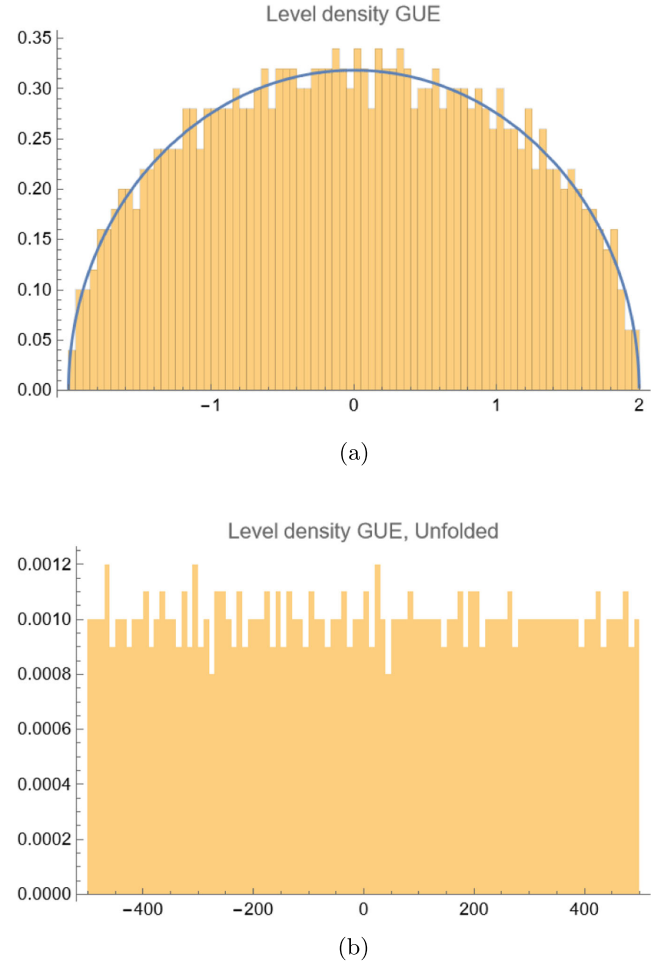


FIG. 1. Histogram of the spectrum taken from GUE,  $N = 1000$ . The blue curve in (a) is the average level density of GUE, Eq. (2) and it is used for the unfolding. (b) The spectrum after the unfolding.

$$\rho_0(\lambda) = \frac{1}{2\pi} \sqrt{4 - \lambda^2}. \quad (2)$$

If we compute the set of eigenvalues numerically with  $N = 1000$ , we can see in Fig. 1(a) that the histogram of eigenvalues (the orange bars) matches the average level density in (2) and the mean value of the nearest-neighbor spacings calculated numerically from GUE, and in Fig. 1(b) we can see the histogram of the set of levels  $\{e_i\}$  after applying the unfolding procedure in (1) using (2). As can be seen, the unfolded spectra “flattens”  $\rho_0$  to a constant, while preserving the fluctuations around it. The mean nearest-neighbor spacing turns out to be 1.00026. Note that since  $\rho_0$  is a normalized distribution (1) the range of the spectrum changes such that  $e_{\max} - e_{\min} = N$ .

When the average level density is not known, it can be calculated numerically using the histogram of the spectrum, where the bin sizes should be larger than the mean spacing but much smaller than the scale where  $\rho_0$  changes. For example, we take the high range of the DSG spectrum with

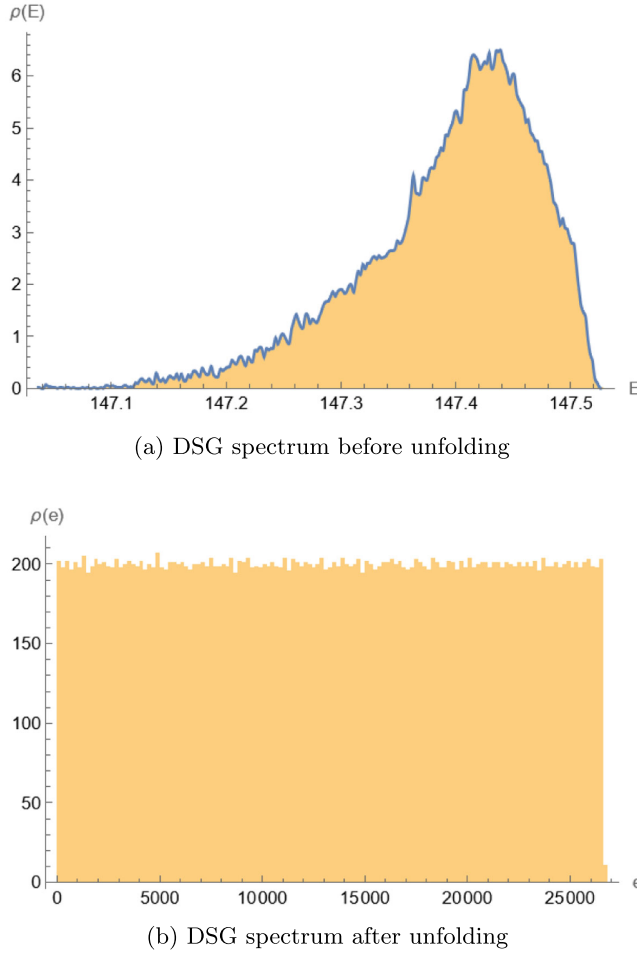


FIG. 2. The high range spectrum of DSG with  $mL = 1$ . The orange bars in (a) is the binned density and the blue curve on top is the interpolated  $\rho_0$ . (a) The same spectrum after unfolding, using the numerically evaluated  $\rho_0$ . The fluctuations are still visible in (a) as variations of the bin heights around the constant average density.

$mL = 1$  and get a histogram with properly sized bins. We use the bin counts and average energy of each bin as data points and interpolate a smooth function in between, which is our numerical  $\rho_0(E)$ , which we then use for the unfolding (see Fig. 2).

Notice that the energy scale of the unfolded spectrum changed, and the whole spectrum is now between 0 and  $N$ . In areas where the derivative of  $\rho_0$  is large, more points should be used in the interpolation, although one should bear in mind that using too many points might flatten the fluctuations as well. One should also be careful not to get negative values for the interpolated  $\rho_0$  since those could cluster the energy levels within the energy range where  $\rho_0$  is negative, once unfolded.

### B. Adjacent spacings and their ratios

An important step in the analysis of quantum chaos was made in the work of Wigner during the 1960s [14], when he

TABLE I. Constants for the three Gaussian ensembles.

| Ensemble | $\beta$ | $(A_\beta, B_\beta)$                         |
|----------|---------|----------------------------------------------|
| GOE      | 1       | $(\frac{\pi}{2}, \frac{\pi}{4})$             |
| GUE      | 2       | $(\frac{32}{\pi^2}, \frac{4}{\pi})$          |
| GSE      | 4       | $(\frac{262144}{729\pi^3}, \frac{64}{9\pi})$ |

proposed that the statistics of the eigenvalues and eigenstates of heavy nuclei and complex molecules could be understood in terms of random matrix theory, where the Hamiltonian of the system is replaced by an independent identically distributed (IID) and rotationally invariant random matrix, and the energy levels of the system are given by the eigenvalues of the matrix. In this way it is possible to compute, using random matrices, different observables seen in experiments. Wigner and Dyson [15] observed that the spacings between consecutive energy levels obey a kind of distribution that could be derived analytically from RMT by diagonalizing  $2 \times 2$  Gaussian matrices. The distribution obtained is called the Wigner surmise, or Wigner-Dyson (WD) distribution, and it is

$$P_\beta(s) = A_\beta s^\beta e^{-B_\beta s^2}, \quad (3)$$

where  $s$  is the spacing between two consecutive energy levels (or eigenvalues in terms of RMT)

$$s_n = E_{n+1} - E_n.$$

$\beta$  is the Dyson index [15] and  $A_\beta, B_\beta$  are constants explicitly shown in Table I. They are determined by the requirement that the distribution in (3) must be normalized and its mean must be equal to 1.

We note that the WD distribution is an approximation of the consecutive level spacing probability, which should be derived from general  $N \times N$  random matrices. Bohigas *et al.* conjectured [16] in the 1980s that the local statistics of energy levels of chaotic systems are universal and depend only on the symmetries of the system, which correspond to Gaussian (or circular, for this matter) ensembles in RMT with the same local statistics: systems with time-reversal symmetry to the Gaussian orthogonal ensemble (GOE), systems without time-reversal symmetries to the GUE, and systems with time-reversal symmetry but spin 1/2 statistics to the Gaussian symplectic ensemble (GSE). They also posited that for integrable systems, the Poisson ensemble captures the local statistics. This follows Berry and Tabor conjecture [17] that all integrable systems have uncorrelated energy levels. From the Poisson ensemble, which is an IID diagonal random matrix where the eigenvalues are random variables. The consecutive level spacing distribution derived from the Poisson ensemble is

$$P_{\text{PE}}(s) = e^{-s}. \quad (4)$$

For systems in a mixed state of integrability and non-integrability, there are several approaches to describing the distributions of consecutive level spacings. We will describe two approaches, both phenomenological. The first example is due to Brody [18,19] which uses a heuristic argument by Wigner and an ansatz for a repulsion factor that interpolates between the Poisson ensemble (PE) and GOE to obtain the Brody distribution

$$P_q(s) = (q+1)a_q s^q \exp(-a_q s^{q+1}), \quad (5)$$

where  $q$  is an interpolating parameter that lies in the range  $0 \leq q \leq 1$  and

$$a_q = \left[ \Gamma\left(\frac{q+2}{q+1}\right) \right]^{q+1}$$

is obtained by normalizing the distribution and requiring a mean of 1.  $\Gamma$  is the gamma function. Note that when  $q = 0$  the distribution is exactly (4), and for  $q = 1$  it is the GOE nearest-neighbor spacing (3) with  $\beta = 1$ . The Brody distribution does not interpolate to the other Gaussian ensembles.

Another interpolation, this time based on localization and due to Izraelev and co-workers [20,21] is

$$P_\nu(s) = A_\nu s^\nu \exp\left(-\frac{\pi^2}{16}\nu s^2 - \left(B_\nu - \frac{\nu}{2}\right)\frac{\pi}{2}s\right), \quad (6)$$

where, again,  $A_\nu$ ,  $B_\nu$  are obtained by normalizing the distribution and its mean to 1. The parameter  $\nu$  now holds for the whole range  $0 \leq \nu \leq 4$ , and for the specific values  $\nu = 0, 1, 2$  it deviates from (4), and (3) for GOE, GUE respectively by at most 5%. We note that both  $\nu$  and  $q$  are phenomenological parameters and do not have a direct physical meaning.

In recent years there has been use of another statistical measure of local correlations in the energy levels, due to Oganessian and Huse [22]—the consecutive spacing ratios  $r_n$ , defined as

$$r_n = \frac{s_{n+1}}{s_n} = \frac{E_{n+2} - E_{n+1}}{E_{n+1} - E_n}. \quad (7)$$

Atas *et al.* [23] used a derivation similar to that which gave rise to the WD distribution, only this time using  $3 \times 3$  random matrices since (7) requires at least three eigenvalues, and got the following approximate formula for the distribution of NN level spacing ratios

$$P_\beta(r) = \frac{1}{N_\beta} \frac{(1+r)^\beta}{(1+r+r^2)^{1+3\beta/2}}, \quad (8)$$

where  $N_\beta$  is a normalization constant assuring the distribution above is normalized to unity. As for the WD distribution, a better approximation for the distribution would use bigger matrices. Atas *et al.* derive analytical expression of such distributions for up to  $5 \times 5$  matrices [24] and a general prescription for an  $N \times N$  expression, but comparison with numerical data shows that (8) is already a very good approximation.

One of the main reasons for the use of the spacing ratios rather than the spacings is the assumption that nearest-neighbor levels are still in the scale in which the mean level density  $\rho_0$  is constant and can be factored out, and so in taking the spacing ratio they cancel. This means that the level spacing ratios are independent of the mean level spacing and have no need of being unfolded. Since the order of the spacings in the definition of the spacing ratio (7) was chosen at random and could have also been  $r_n = s_n/s_{n+1}$  which is the inverse of the original definition, the whole support of the spacing ratios is sometimes restricted to being from 0 to 1 by defining

$$\tilde{r}_n = \min\left\{r_n, \frac{1}{r_n}\right\}. \quad (9)$$

A consecutive level spacing ratio distribution was also calculated in [23] for the integrable case based on (4) for  $3 \times 3$  matrices by assuming that the joint probability distribution is the product of independent probabilities. This leads to the expression

$$P_{\text{PE}}(r) = \frac{1}{(1+r)^2}. \quad (10)$$

### C. Higher order spacings and spacing ratios

It is a natural generalization to look at the next orders in the level spacings and spacing ratios in order to better understand the transition between short- and long-range correlations in the level statistics; the latter will be dealt with in the next section.

In this section we will denote  $k$ -spacing as

$$s_{n,k} = E_{n+k} - E_n, \quad (11)$$

where the nearest-neighbor spacing is the case when  $k = 1$ . In the literature the  $k$ -spacings are sometimes noted such that nearest-neighbor spacings have  $k = 0$ , so a careful read is needed. There are several approaches for analytical derivations of the  $k$ -spacing probability  $P_k(s)$  [25]. Fox and Kahn [26] derive a series expansion in the spacing  $s$  directly from RMT. Other more direct approaches involve a calculation similar to that carried by Brody [18] which is based on a statistical argument originally made by Wigner, where the spacing probability  $P(s)$  is derived from the form of the level repulsion at small  $s$  via a heuristic repulsion function  $r(s)$  [27],

$$P(s) = r(s) \exp \left( - \int_0^s r(s') ds' \right). \quad (12)$$

Abul-Magd and Simbel [28] use the ansatz  $r(k, s) \propto s^{k\beta}$  and a recursion relation to solve for the level repulsion of the  $k$ -spacing. Adding a Gaussian tail for the proper large  $s$  behavior, they obtain a WD-like expression

$$P_k(s) = A_{k,\beta} s^{\alpha_{k,\beta}} \exp(-B_{k,\beta} s^2), \quad (13)$$

where

$$\alpha_{k,\beta} = \frac{k(k+1)}{2} \beta + (k-1) \quad (14)$$

and  $A_{k,\beta}$  and  $B_{k,\beta}$  are determined by proper normalization. The same expression (13) is obtained in [29] by inserting the  $k$ -spacing as a  $\delta$  function to the Gaussian joint probability density and evaluating it directly.

As for  $k$ -spacing ratios, it is conjectured in [30] that the  $k$ -spacing ratio probability distribution is the same as (8), but replacing  $\beta$  by the expression  $\alpha_{k,\beta}$  given in (14).

There is also an explicit analytical derivation of  $k = 2$  spacing ratio distribution given in [24].

#### D. Pair correlations

The NN spacing distribution and the higher order  $k$ -spacing generalization are related to another concept—the pair correlation function and the pair cluster function, which are the  $n = 2$  cases of the general  $n$ -point correlation and cluster functions, discussed in Appendix C. We begin with the  $n$ -point correlation function (C3), which is the  $n$  marginal probability distribution of the joint probability distribution of the ensemble, and choose to work in the GUE case in the following part for simplicity, but a full derivation for all three symmetry ensembles can be found in [31]. Using determinant identities, we can insert the Gaussian measure into the Vandermonde determinant by a polynomial reparametrization of the entries of using the normalized harmonic oscillator eigenvectors,

$$\phi_n(x) = \frac{1}{\sqrt{\pi 2^n n!}} e^{-\frac{1}{2}x^2} H_n(x), \quad (15)$$

where  $H_n(x)$  are the regular Hermite polynomials

$$H_n(x) = (-1)^n e^{x^2} \frac{d^n}{dx^n} e^{-x^2}. \quad (16)$$

We can define the kernel

$$K_N(\lambda_1, \lambda_2) = \sum_{n=0}^{N-1} \phi_n(\lambda_1) \phi_n(\lambda_2) \quad (17)$$

and check, using the orthogonality of the Hermite polynomials, that it satisfies the kernel equation

$$\int K_N(\lambda_1, \lambda_2) K_N(\lambda_2, \lambda_3) d\lambda_2 = K_N(\lambda_1, \lambda_3). \quad (18)$$

With the kernel we can write the  $n$ -point correlation function as

$$R(\lambda_1, \lambda_2, \dots, \lambda_n) = \det [K_N(\lambda_i, \lambda_j)]_{i,j=1,2,\dots,n}. \quad (19)$$

Now unfolding the spectrum allows us to simplify this expression further, since unfolding turns the mean level spacing, which is the one-point correlation function, into a constant  $R(\lambda) = \rho_0(\lambda) = 1/D$  where  $D$  is the mean NN level spacing, which is usually taken to be unity in the unfolding process. From this follows that the correlation between any two levels is translational invariant, i.e. depends only on expressions like  $\lambda_i - \lambda_j$ . Using rescaled eigenvalues  $x = D\lambda$  and the large  $N$  limit, we get the sine kernel

$$\lim_{N \rightarrow \infty} DK_N(x_i, x_j) = \frac{\sin(\pi(x_i - x_j))}{\pi(x_i - x_j)}. \quad (20)$$

And for the two-point GUE correlation function we get the expression

$$R(x_1, x_2) = 1 - \left( \frac{\sin(\pi s)}{\pi s} \right)^2, \quad (21)$$

where the 1 comes from the disconnected part  $R(x_1)R(x_2)$  which were normalized by choosing  $D = 1$ , and the second term is the connected part, which is the two-point cluster function  $T(x_1, x_2)$  with  $s = x_2 - x_1$ . The two-point correlation function for GOE is derived in the same manner, though (19) is more complicated. We will write it when it is used, in Sec. V.

The pair correlation after unfolding can be seen as the probability distribution for having a spacing  $s$  between any two eigenvalues in the spectrum, and so we can relate it to the higher order spacing distributions

$$R(x_1, x_2) = \sum_{k=1}^{N-1} P_k(s). \quad (22)$$

One known use of the pair correlation is found in the surprising connection between the Riemann hypothesis (RH) and the GUE. As a reminder, the RH conjectures that the nontrivial zeros  $z_n$  of the Riemann zeta function  $\zeta(s) = \sum_n n^{-s}$  are all found along the line  $\Re\{s\} = 1/2$ ,

$$z_n = \frac{1}{2} + i\theta_n. \quad (23)$$

Montgomery conjectured [32], based on the RH, an estimate for the correlation between any two imaginary parts of zeros of the  $\zeta$  function  $\theta, \theta'$ . For fixed  $0 < \alpha < \beta < \infty$ ,

$$\frac{| \{ (\theta, \theta') : 0 < \theta, \theta' \leq T, 2\pi\alpha(\log T)^{-1} \leq \theta - \theta' \leq 2\pi\beta(\log T)^{-1} \} |}{(T \log T)/2\pi} \sim \int_{\alpha}^{\beta} 1 - \left( \frac{\sin(\pi u)}{\pi u} \right)^2 du, \quad (24)$$

where it is assumed that  $T \rightarrow \infty$ . This was later confirmed numerically in by Odlyzko in [33].

Dyson [15] pointed out that this behavior of pair correlations is identical to that found in GUE. Since the pairs in GUE are ordered eigenvalues of a random matrix, this similarity pushes forward the Hilbert and Polya conjectures (see Ref. [32]) that the nontrivial zeros of the  $\zeta$  function correspond to eigenvalues of a positive linear operator.

### E. Spectral rigidity

The pair correlation is also present in two closely related statistical measures of spectra—the level number variance  $\Sigma^2(L)$  and the spectral rigidity  $\Delta_3(L)$ , both indicators of short- and long-range phenomena between pairs of levels. We shall derive both based on [34].

Given an unfolded spectrum, the number of energy levels in a given energy interval of length  $L$  is expressed with the level density  $\rho(x)$

$$n(x, L) = \int_x^{x+L} \rho(x') dx' \quad (25)$$

and the mean number of levels  $\langle n(L) \rangle$  is the ensemble average of (25), which is independent of a specific choice of lower bound energy level  $x$  because it averages over all of them. Since the unfolding process assures that the mean spacing between NN levels is 1, it follows that the mean number of levels in an interval  $L$  is  $L$ , i.e.  $\langle n(L) \rangle = L$ . We can now define the variance in the mean number of levels as

$$\Sigma^2(L) = \langle n(L)^2 \rangle - \langle n(L) \rangle^2 \quad (26)$$

and need only to calculate the first term, which is explicitly

$$\langle n(L)^2 \rangle = \iint \langle \rho(x_1) \rho(x_2) \rangle dx_1 dx_2, \quad (27)$$

where the term  $\langle \rho(x_1) \rho(x_2) \rangle$  is the pair correlation function for  $x_1 \neq x_2$ . For  $x_1 = x_2$  we get an extra  $\delta$  function term. Overall  $\langle \rho(x_1) \rho(x_2) \rangle = 1 - T(x_1, x_2) + \delta(x_1 - x_2)$ , and inserting this into (26) gives

$$\Sigma^2(L) = L - 2 \int_0^L (L - s) T(s) ds, \quad (28)$$

where  $s = x_1 - x_2$ . Using the pair cluster functions for the GOE and GUE respectively yields the final expressions for  $\Sigma_{\beta}^2(L)$ , which asymptotically in the  $L \rightarrow \infty$  limit are

$$\begin{aligned} \Sigma_1^2(L) &= \frac{2}{\pi^2} \left( \ln(2\pi L) + \gamma + 1 - \frac{\pi^2}{8} \right), \\ \Sigma_2^2(L) &= \frac{2}{\pi^2} (\ln(2\pi L) + \gamma + 1), \end{aligned} \quad (29)$$

where  $\gamma$  is the Euler-Mascheroni constant and we use  $\beta$  as the Dyson index in the subscript.

The spectral rigidity  $\Delta_3(L)$  is the minimal variance of the difference between the number of levels in an interval  $L$ ,  $n(x, L)$ , and straight line, averaged over the spectrum,

$$\Delta_3(L) = \frac{1}{L} \left\langle \min_{A,B} \int_x^{x+L} (n(x, x') - Ax' - B)^2 dx' \right\rangle, \quad (30)$$

which can be expressed explicitly as

$$\Delta_3(L) = \frac{L}{15} - \frac{1}{15L^4} \int_0^L (L - s)^3 (2L^2 - 9Ls - 3s^3) T(s) ds. \quad (31)$$

For integrable cases, the two-point cluster function vanishes and the spectral rigidity is  $\Delta_{3,0}(L) = \frac{L}{15}$ . For the other ensembles, inserting the right two-point cluster functions in (31) and taking the large  $L$  limit gives

$$\begin{aligned} \Delta_{3,1}(L) &= \frac{1}{\pi^2} \left( \ln(2\pi L) + \gamma - \frac{5}{4} - \frac{\pi^2}{8} \right), \\ \Delta_{3,2}(L) &= \frac{1}{2\pi^2} \left( \ln(2\pi L) + \gamma - \frac{5}{4} \right), \end{aligned} \quad (32)$$

which is very similar to the  $\Sigma_{\beta}(L)^2$  expressions in (29).

### F. Spectral form factor

To conclude this part, we note another popular measure related to the pair correlation—the spectral form factor (SFF). It arises naturally when dealing with the time dependence of the expectation value of two local operators in a finite temperature  $\beta$  (not to be confused with  $\beta$  as the Dyson index) and has been of recent interest in showing the discreteness of the spectrum of black holes using holography from the bulk perspective [4] and in studying the thermalization of high energy states in the eigenstate thermalization hypothesis, see Ref. [35] for a review on the subject. Consider the two-point correlation function of a Hermitian operator  $O(t)$

$$G(t) = \frac{1}{Z(\beta)} \text{Tr}(e^{-\beta H} O(t) O(0))$$

$$= \frac{1}{Z(\beta)} \sum_{m,n} e^{-\beta E_m} |\langle m | O | n \rangle|^2 e^{-i(E_m - E_n)t}, \quad (33)$$

where  $Z(\beta) = \text{Tr}(e^{-\beta H})$  is the partition function and  $|m\rangle$  is the energy eigenstate with energy  $E_m$ . The time dependence is mostly characterized by the energy levels, which can be expressed better using the analytically continued partition function  $Z(\beta, t) = \text{Tr}(e^{-(\beta + it)H})$  and writing

$$g(t) = \left| \frac{Z(\beta, t)}{Z(\beta)} \right|^2 = \frac{1}{Z(\beta)^2} \sum_{m,n} e^{-(E_m + E_n)\beta} e^{-i(E_m - E_n)t}, \quad (34)$$

which reduces, for infinite temperature ( $\beta = 0$ ) and a finite spectrum of size  $N$  to the expression

$$g(t) = \frac{1}{N^2} \sum_{m,n} e^{-i(E_m - E_n)t}. \quad (35)$$

In the continuum limit, taking an ensemble mean (35) gives

$$\langle g(t) \rangle = \frac{1}{N^2} \iint \langle \rho(E_1) \rho(E_2) \rangle e^{-i(E_1 - E_2)t} dE_1 dE_2, \quad (36)$$

where in general we can separate the level density into  $\rho(E) = \rho_0(E) + \delta\rho(E)$  such that taking the ensemble average is the mean level density  $\langle \rho(E) \rangle = \rho_0(E)$ . This gives

$$\rho_0(E_1) \rho_0(E_2) + \langle \delta\rho(E_1) \delta\rho(E_2) \rangle,$$

which can be inserted into (36), resulting in two terms—the first one is the square of the Fourier transform of  $\rho_0(E)$ , which is a model-dependent term that decays for  $t \rightarrow \infty$  if the spectrum is finite. This term is usually referred to as the “descent.” After unfolding, the mean level density is unity, and the second term is the normalized cluster function  $\langle \delta\rho(E_1) \delta\rho(E_2) \rangle = \delta(x_1 - x_2) - T(x_1, x_2)$ . Inserting the proper  $T(x_1, x_2)$  terms for the GOE and GUE ensembles gives, after Fourier transformation, the connected part of the SFF [36]

$$\langle g_c(t) \rangle_{\text{GOE}} = \begin{cases} \delta(t) + 2|t| - |t| \log(1 + 2|t|) & \text{for } |t| \leq 1 \\ 2 - |t| \log\left(\frac{2|t|+1}{2|t|-1}\right) & \text{for } |t| > 1 \end{cases} \quad (37)$$

$$\langle g_c(t) \rangle_{\text{GUE}} = \begin{cases} \delta(t) + |t| & \text{for } |t| \leq 1 \\ 1 & \text{for } |t| > 1. \end{cases} \quad (38)$$

These terms in the SFF are usually referred to as the “ramp” for  $t < 1$  and “plateau” for  $t > 1$ , as seen for the

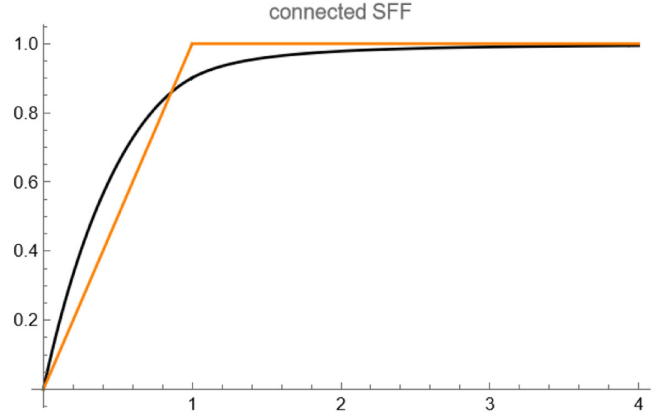


FIG. 3. RMT expressions for the connected parts of the SFF, after unfolding and normalization. The black curve corresponds to the GOE (37) and the orange curve to the GUE (38).

GOE and GUE SFF in Fig. 3. The plateau part can be thought of as the late time average of the SFF  $\lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T g(t) dt$  and it is due to the discreteness of the spectrum. The ramp is associated with the level repulsion, and indeed for integrable systems where the pair cluster function  $T(x_1, x_2)$  vanishes there is no ramp in the SFF, and the descent part goes directly to the plateau. It was suggested in the past years that the indication of a ramp in the SFF is a measure for quantum chaos, though there are examples for spectra that give rise to ramps in the SFF even though they could be written as analytical expressions and are therefore not chaotic [37].

### III. DISTRIBUTIONS OF NN SPACINGS AND SPACINGS RATIOS

In this section we apply the characteristics of quantum chaos mentioned in the previous section to the numerically calculated spectrum of the nonintegrable DSG model and compare the results with those of the spectrum of SG.

In particular, we would like to affirm that the level statistics behave like those predicted by RMT with the corresponding ensemble being the GOE [or circular orthogonal ensemble (COE), since their local level statistics are the same and they differ only in the average level density], since DSG is symmetric under time-reversal transformation. This goal can be translated into finding that the value of  $\beta$  is indeed  $\beta = 1$  in the RMT  $\beta$  ensemble. If some measurements of the value of  $\beta$  deviates from 1, taking into account statistical errors, this might mean that the system is in a transition between two analogous ensembles of RMT. Specifically, if  $\beta$  is between 0 and 1 the system might be in a transitional phase between the Poisson ensemble, which is analogous to integrable systems, and GOE, which is non-integrable with time-reversal symmetry. Such mix states can be thought of as the quantum version of the classical Kolmogorov-Arnold-Moser theory in which integrable and chaotic regions coexist in the phase space.

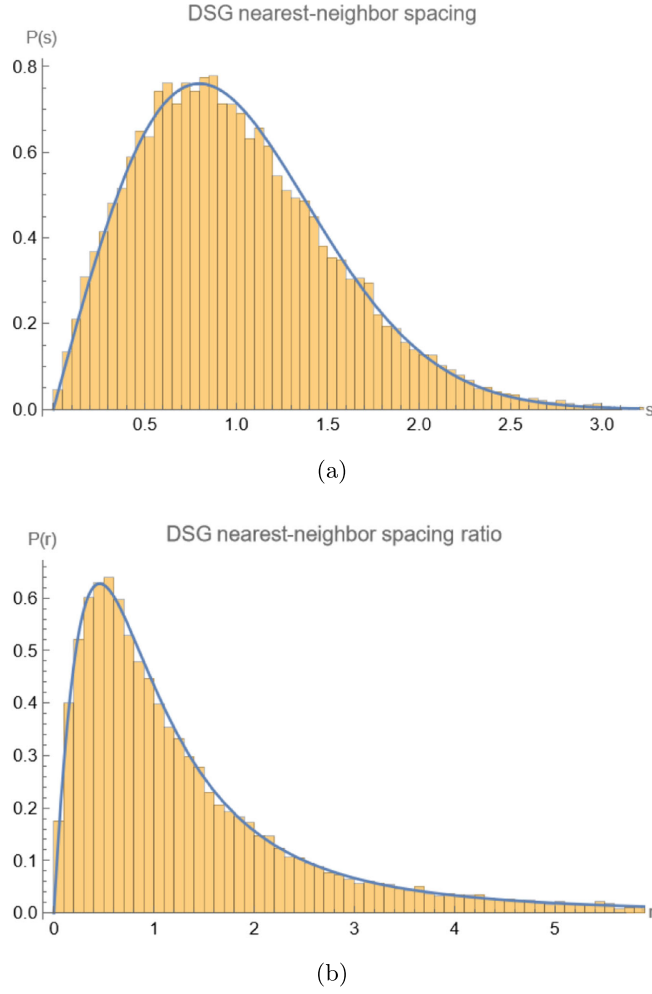


FIG. 4. (a) The unfolded DSG nearest-neighbor spacing (orange bars) with the GOE Wigner surmise (blue curve). (b) The unfolded DSG nearest-neighbor spacing ratio (orange bars) with the expression for nearest-neighbor spacing ratio distribution (8), with  $\beta = 1$  (blue curve).

The data used were calculated by Prosen and co-workers in their paper [2], where they used an energy cutoff via a Hamiltonian truncation method called the “truncated conformal space approach” in order to calculate the spectrum. The coupling constant  $\lambda$  is chosen such that the dimensionless parameter  $l = m_\beta L$  will be of order 1, where  $m_\beta$  is the mass scale corresponding to the first breather’s mass for a given frequency parameter  $\beta$  in the SG potential, and  $L$  is the size of the system (i.e. the model has the spatial identification  $x \sim x + L$ ). The two datasets we use are SG and DSG, both with  $l = 1$  and containing around 85,000 energy levels. The SG data with the frequency  $\beta = 1$  and the DSG with  $(\beta_1, \beta_2) = (1, 2.5)$ , a combination of frequencies that the authors found to introduce the most GOE-like behavior, at least in terms of the mean consecutive spacing ratio  $\langle \tilde{r} \rangle$ . In the following calculations, we use only the highest (around 25,000) energy levels out of the whole dataset, which we unfold by interpolation.

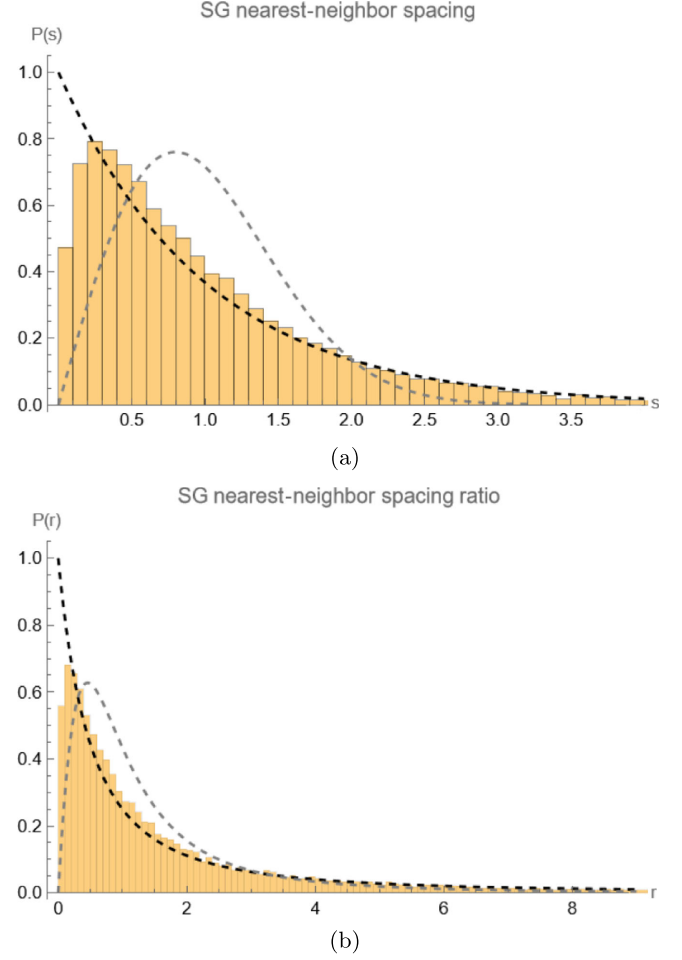


FIG. 5. Level spacing and spacing ratio statistics from the SG data. (a) The consecutive level spacing distribution, where the black dashed line is the RMT prediction for a Poisson distribution (4) and the gray dashed line is the GOE WD distribution (3). (b) The consecutive level spacing ratio distribution, with the black dashed line being the Poisson ensemble prediction for level spacing ratios (10) and the gray dashed line being the GOE level spacing ratio from RMT, (39).

To begin with, we compute the nearest-neighbor spacing distribution, compared with the analytical Wigner-Dyson distribution for  $\beta = 1$ . The results seem to fit nicely [see Fig. 4(a)]. We also compare the nearest-neighbor spacing ratio distribution with the analytical expression, see Fig. 4(b), a calculation not done in the original paper.

For a numerical comparison, we compute the mean value of the nearest-neighbor spacing ratio  $\langle \tilde{r} \rangle = 0.529808$  which is compared with the mean of the RMT-based spacing ratio “surmise,” given in (8)  $\langle \tilde{r} \rangle_{\text{GOE}} = 0.53590$ , which gives a deviation of less than 1% between the two values. We note that this calculation is the basis for the choice of the frequency parameters in DSG, in such a way that the combination of the frequencies  $(\beta_1, \beta_2)$  has the value of  $\langle \tilde{r} \rangle$  closest to the RMT prediction, but there always seems to be a minimal deviation of the same order from the

prediction, indicating that the DSG data are not completely chaotic.

Next we would like to compare the results above with the corresponding integrable SG model. We use data calculated numerically by the same authors, with the conditions used in DSG that  $m_\beta L = 1$ , and a frequency  $\beta = 2.5$ . As for the DSG spectrum, we take only the highest  $\sim 25,000$  energy level as the dataset and unfold it. From the unfolded data we compute the consecutive level spacing and spacing ratio and plot it using (4) and (10) as RMT predictions.

From Fig. 5 we can see in both the level spacing and level spacing ratios that the data do not fit the Poisson distributions neatly—there is a clear level repulsion for  $s, r \rightarrow 0$ , which should not exist for pure integrable models and a following “excess” of levels for  $s, r$  a bit larger. The tails of both the spacing and the ratios do seem to fit the Poisson prediction. Both cases are far from the GOE distributions which are denoted by the gray curves, in the language of a mixing parameter it seems closer to the integrable value. SG is an integrable model, so these results are attributed to the numerical truncation in the calculation of the spectrum. We also calculate the mean spacing ratio  $\langle \tilde{r} \rangle = 0.4361$ , which is greater than the RMT predicted  $\langle \tilde{r} \rangle_{\text{PE}} = 0.3863$ .

Taking SG data with  $mL < 1$  yields better results. Using the unfolded data of SG with  $mL = 0.01$ , which corresponds to taking a higher cutoff energy, gives results that fit the RMT predictions much better, with  $\langle \tilde{r} \rangle_{mL=0.01} = 0.3819$ . Thus we use  $mL = 0.01$  for all further calculations of the SG data.

#### IV. DISTRIBUTIONS OF NONADJACENT SPACINGS AND THEIR RATIOS

We also measure the mean  $k$ -spacing and spacing ratio, as elaborated in Sec. II C. Figure 6(a) shows that the mean  $k$  behaves linearly in  $k$  with a slope of 1. We suggest an explanation for this behavior in Sec. IV A. The deviations of  $\langle s^k \rangle$  from linearity are very small, as can be seen from Fig. 6(b).

The pattern in the residual plot looks random at first, but in closer inspection there seems to be a reflection symmetry around the value of  $k \approx 12,500$ , which happens to be  $k = N/2$ . This symmetry is masked by the distortion of the values for  $k > N/2$ , but that distortion is due to the shrinking size of the samples of spacings, which is a linear function in  $k$  as well,

$$N_{\text{samples}}(k) = N - k - 1.$$

To see the symmetry and the distortion, we take the residual plot of Fig. 6(b) and reflect the plot around  $k = N/2$  in Fig. 6(c) such that the x axis is now the distance from the reflection point,  $|k - N/2|$  and the orange curve in the reflected part of the plot. The blue curve is the part unchanged where  $k > N/2$ . We find that the two curves

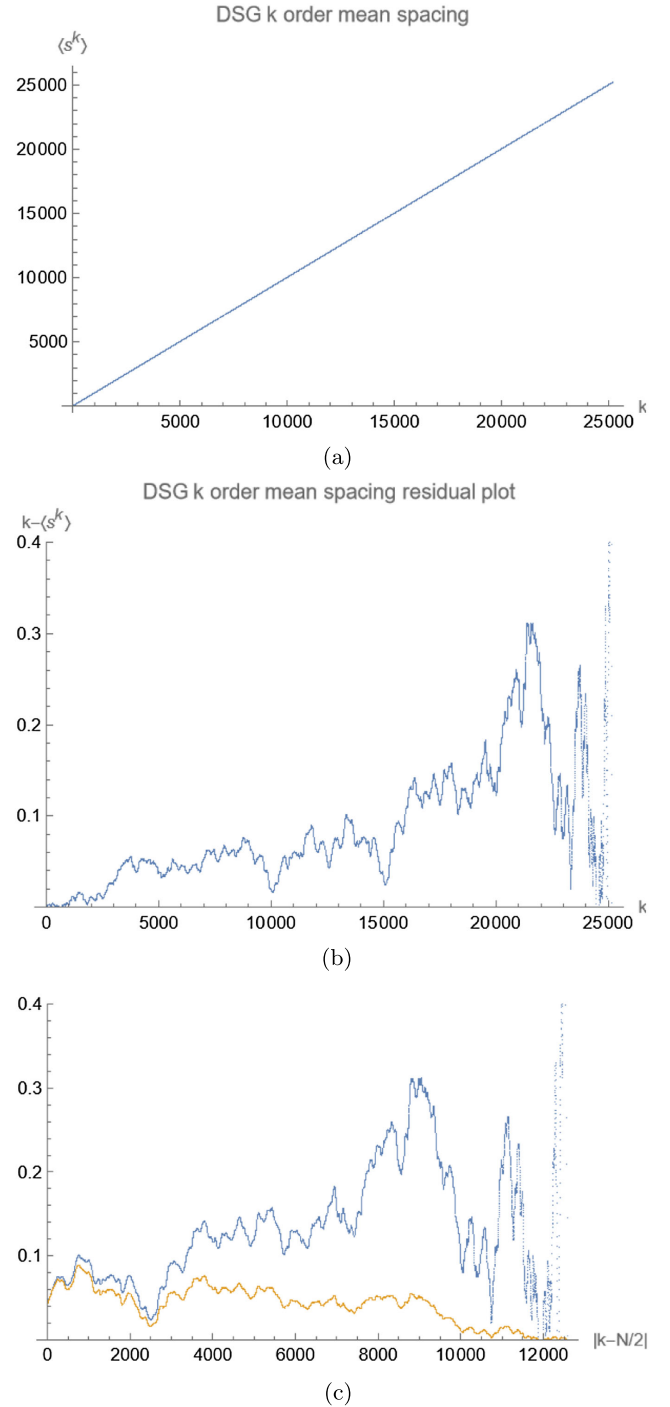
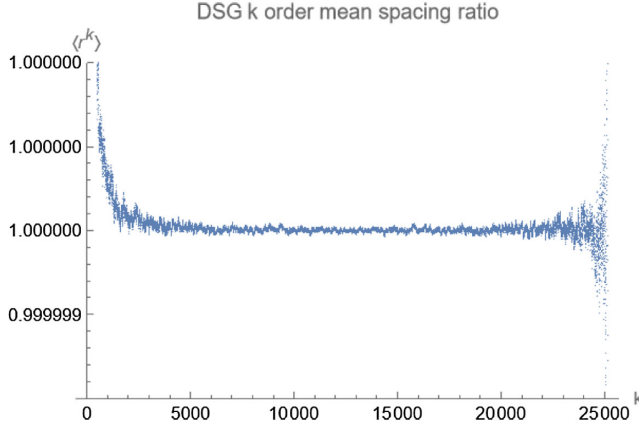


FIG. 6. (a) The mean  $k$ -spacing ratio  $\langle r^k \rangle$  from the unfolded DSG high energy levels. It shows a linear behavior with a slope of 1. The deviation from the predicted linear behavior is seen in (b). (c) The similarity of the residuals under reflection around  $k = N/2$ .

indeed correspond, though the distortion shifts their values apart.

We also compute the mean  $k$ -spacing ratios, seen in Fig. 7. Since there is no prediction to compare the results with, we will note that the mean is predominantly 1, with

FIG. 7. Mean  $k$ -spacing ratio, plotted against  $k$ .

fluctuations 7 orders of magnitude smaller, as can be seen from the main part of Fig. 7. Again, though it is harder to see in this case, we find the reflection symmetry around  $k = N/2$  and a distortion of the mean values around the value of 1, which grows with  $k$  as the sample size shrinks, though the distortion seems to be prominent for  $k$ 's much larger than in the spacing case.

For small  $k$ , we can check the validity of the results with the conjecture in [30] where the probability distribution of the  $k$ -spacing ratios is given by (39),

$$P(r, \beta') = \frac{1}{N'_\beta} \frac{(1+r)^{\beta'}}{(1+r+r^2)^{1+3\beta'/2}} \quad (39)$$

with

$$\beta' = \frac{k(k+1)}{2} \beta + (k-1)$$

and compare the DSG results with RMT results and (39).

The DSG and RMT results match even for small  $k$ , as can be seen in Fig. 8(b) where the difference is of order  $10^{-4}$  already for  $k = 3$ .

For higher order spacings and spacing ratios of integrable systems, there is an expression due to [24] for the  $k$ -spacing ratio distribution

$$P_k(r) = \begin{cases} \frac{r^k(1+k+kr)}{(1+r)^2} & \text{for } r \leq 1 \\ \frac{k+r(k+1)}{r^{k+1}(1+r)^2} & \text{for } r > 1, \end{cases} \quad (40)$$

from which we can derive the mean  $k$ -spacing ratio

$$\langle r_k \rangle = \frac{1}{2} \left[ 2 + H\left(\frac{k}{2}\right) - H\left(\frac{k+1}{2}\right) - H\left(\frac{k-2}{2}\right) + H\left(\frac{k-1}{2}\right) \right], \quad (41)$$

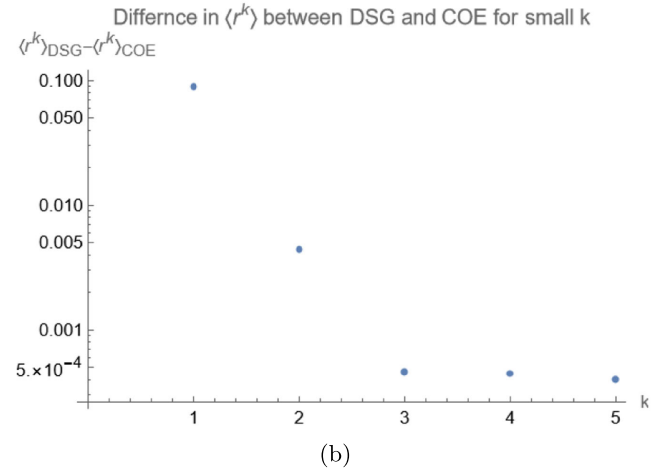
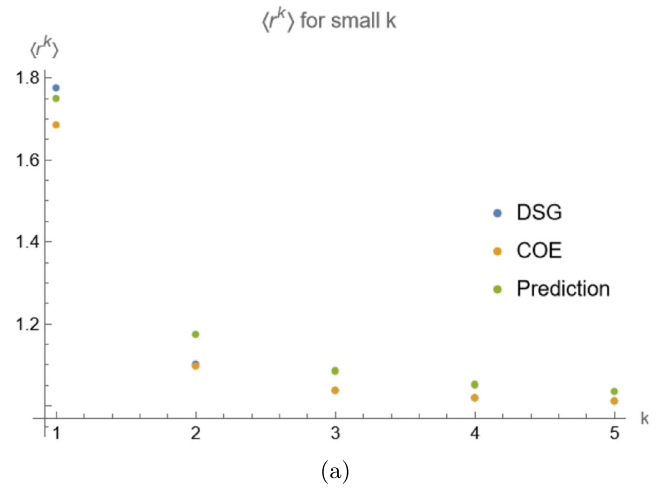


FIG. 8. (a)  $\langle r^k \rangle$  for the first five spacing ratios calculated from the DSG unfolded data, from the COE with  $N = 5000$  and from the analytical expression given in (39). Aside from  $k = 1$ , the DSG and COE results are indistinguishable. To get a better understanding of the comparison of DSG and COE, in (b) we plot in a log scale the difference between the mean ratios for each  $k$ .

where  $H(k)$  are the harmonic numbers, defined as  $H(k) = \sum_{n=1}^k \frac{1}{n}$  for a finite integer  $k$ . Asymptotically, (41) can be written as

$$\langle r_k \rangle \approx 1 + \frac{1}{2} \log \left( \frac{k(k-1)}{(k+1)(k-2)} \right), \quad (42)$$

where it is clearer that  $\langle r_k \rangle$  converges to 1 as  $k \rightarrow \infty$ . We can compare (41) with the mean  $k$ -spacing ratio calculated from the data, as seen in Fig. 9(a). The prediction agrees with the calculation from the data, and taking the difference of the two shows that the  $k = N/2$  reflection symmetry seems to hold for integrable cases as well.

In conclusion, we have seen that the mean  $k$ -spacing and spacing ratios show fluctuation around the predictions. For example

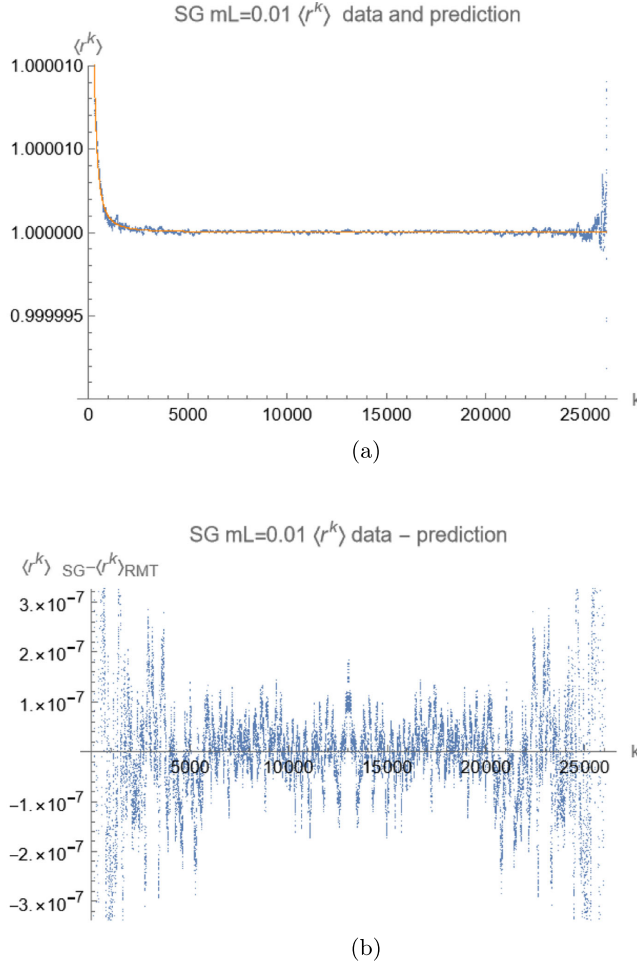


FIG. 9. (a)  $\langle r_k \rangle$  calculated from the SG data (blue) together with the prediction (42) (orange), which is easier to compute than (41) for large  $k$ . (b) The residual plot, the difference between the calculated means and the prediction. Notice the small scale of  $10^{-7}$  of the fluctuations of the data and the appearance of the  $k = N/2$  reflection symmetry.

$$\langle s_k \rangle = k + \langle s_k \rangle_{fl},$$

where  $\langle s_k \rangle_{fl}$  is of  $O(1)$  regardless of the size of  $k$  and obeys a reflection symmetry around  $k = N/2$ , as can be seen in Figs. 6(c), 9(b), and 10(b). We notice that it is present for means on the whole spectrum, as in the COE example of Sec. IV A, and for means on partial parts of the spectrum, as in the SG and DSG calculations, and that it is present in integrable (SG) and nonintegrable (DSG) cases, so it seems to be robust phenomena that are unrelated to the data themselves. We also notice that such a symmetry is not present in the  $k$ -standard deviation, only in the fluctuations from the mean. We suggest that this symmetry is related to the autocorrelation structure of the mean: given a discrete spectrum  $\{e_i\}$  of size  $N$ , the mean  $k$ -spacing is explicitly

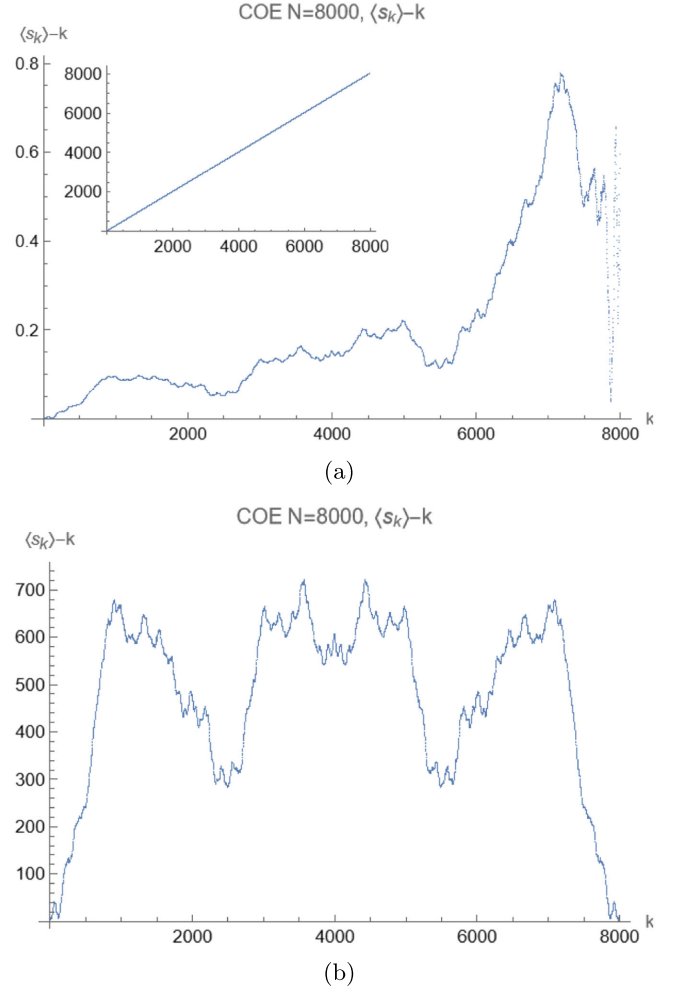


FIG. 10. Data taken from COE with  $N = 8000$ . The inset in (a) shows the mean  $k$ -spacing of the data as a function of  $k$ , which in this scale is indistinguishable from (45). The rest of (a) is the deviation of the mean  $k$ -spacing from (45). (b) The same as (a) but multiplied by a factor of  $N_{\text{sample}}$  to compensate for an apparent  $1/N_{\text{sample}}$  distortion. The result shows a clear reflection symmetry around the  $k = N/2$  value.

$$\langle s_k \rangle = \frac{1}{N-k} \sum_{n=1}^{N-k} (e_{n+k} - e_n).$$

And in the continuum limit we add a  $\delta$  function to fix the distance  $s$  between the two levels  $e, e'$ ,

$$\begin{aligned} & \frac{1}{N-s} \int de \int de' \rho(e) \rho(e') (e - e') \delta(e - e' - s) \\ &= \frac{1}{N-s} \int de \rho(e) \rho(e-s) s = \frac{s}{N-s} \int de \rho(e) \rho(e-s), \end{aligned} \quad (43)$$

where in the last term, the integral is the autocorrelation of  $\rho(e)$  with lag  $s$ ,  $(\rho \star \rho)(s)$ , which has a reflection symmetry

around  $s = N/2$ . In the last calculation we should not confuse the continuous lag  $s$ , which is analogous to the discrete spacing  $k$ , with the symbol  $s$  for spacing in the brackets of  $\langle s_k \rangle$ .

### A. Probabilistic approach to higher order level spacings

In light of Sec. II C, we suggest a naive approach to the linear behavior of the mean  $k$ -spacing: given an energy level  $E_n$ , WD statistics suggests that the consecutive energy level  $E_{n+1}$  can be constructed as  $E_{n+1} = E_n + S$  where  $S \sim P(s)$  is a random variable drawn from the NN spacing probability distribution  $P(s)$ . This means that for the  $k$  separated energy levels, the  $k$ -spacing is

$$s_{n,k} = E_{n+k} - E_n = \sum_i^k S_i, \quad (44)$$

which is the convolution of  $k$  NN probability density functions (PDFs). It follows that the mean is

$$\langle s_k \rangle = \left\langle \sum_i^k S_i \right\rangle = k \langle s \rangle \quad (45)$$

and for normalized NN distributions, where  $\langle s \rangle = 1$ , we get that  $\langle s_k \rangle = k$ . We can also see that, for large  $k$ , using the central limit theorem

$$\lim_{k \rightarrow \infty} s_{n,k} \sim N(k, \sigma_{NN} \sqrt{k}) \quad (46)$$

where  $N$  is a normal probability distribution with mean  $\mu = k$  and standard deviation  $\sigma = \sigma_{NN} \sqrt{k}$ . The contribution of the choice of NN distribution comes only in  $\sigma_{NN}$  when  $k$  is large. Note that like other approaches that are based on local statistics, we assume a constant mean level density, either by construction (as in a normalized circular ensemble) or by unfolding.

We check this approach numerically using randomly generated eigenvalues of the COE with  $N = 5000$  and  $N = 8000$ . The mean does seem to behave linearly in  $k$  as suggested by (45). When looking at the deviation from linearity, we find the deviations to be of order 1, even for large  $k$ . If we multiply the deviations in the mean with a factor of the size sample for each  $k$ , which is  $N_{\text{sample}} = N - k$ , we cancel a  $1/N_{\text{sample}}$  factor and get a picture that shows a reflection symmetry around  $k = N/2$ .

On the other hand we can calculate the expression  $P_k(s)$  for  $k = 2$  analytically by convolving two WD distributions with  $\beta = 1$ . We get

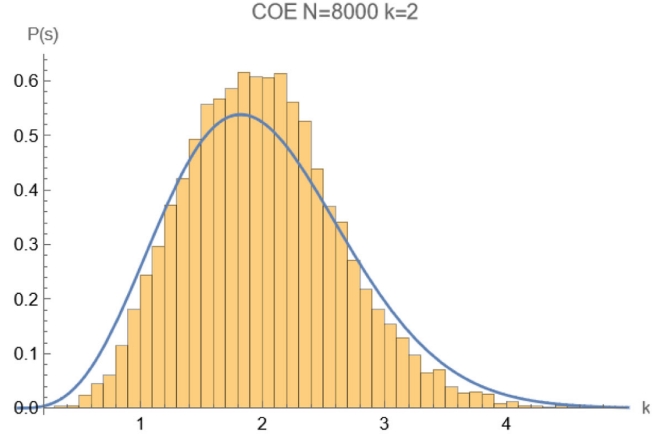


FIG. 11. COE = 8000  $k = 2$  spacing distribution (orange bar) with prediction from the convolution of two WD  $\beta = 1$  probability distributions [(47), blue curve]. The spacing data, as opposed to (47), have an apparent surplus of spacings around the mean, at the expense of less spacing in the higher and lower values. This is also indicated by a smaller standard deviation of the data, which is around 0.633, as opposed to the prediction, which is  $\sqrt{\frac{8}{\pi} - 2} \sim 0.739$ .

$$P_{k=2}(s) = \frac{\pi}{16} \exp\left(-\frac{\pi s^2}{4}\right) \left[ 4s + \sqrt{2} \exp\left(\frac{\pi s^2}{8}\right) (\pi s^2 - 4) \operatorname{erf}\left(\frac{1}{2} \sqrt{\frac{\pi}{2}} s\right) \right], \quad (47)$$

where  $\operatorname{Erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x \exp(-t^2) dt$  is the error function. When comparing (47) with the numerical  $k = 2$  spacing distribution (Fig. 11) we get the mean is 2 for both, but the numerical data have a smaller standard deviation.

When we go to larger  $k$ , we see that, even for  $k$  as small as 4 or 5, the central limit theorem is a good enough approximation and the numerical distribution fits nicely with a normal distribution of mean  $k$ . The standard deviation, however, differs from  $\sigma_{NN} \sqrt{k}$  and remains in same order of magnitude as  $\sigma_{NN}$  for all  $k$ , as can be seen in Fig. 12(a). We emphasize these results by showing how the  $k$ -spacing distributions of the data fit nicely with normal distributions with mean  $k$  and standard deviations taken from the data,  $N(k, \sigma_{k,\text{data}})$  [Figs. 12(b)–12(d)]. We suggest that the discrepancy between the naive probabilistic approach and the data comes from the assumption that the random variables in (44) are independent, whereas in practice they are highly correlated due to the diagonalization of the initial random matrix. Taking into account the covariance would likely improve the results.

## V. PAIR CORRELATION

We now turn to long-range correlations using higher order spacings and calculate the pair correlation. Using the unfolded data of DSG we calculate the level spacing for all

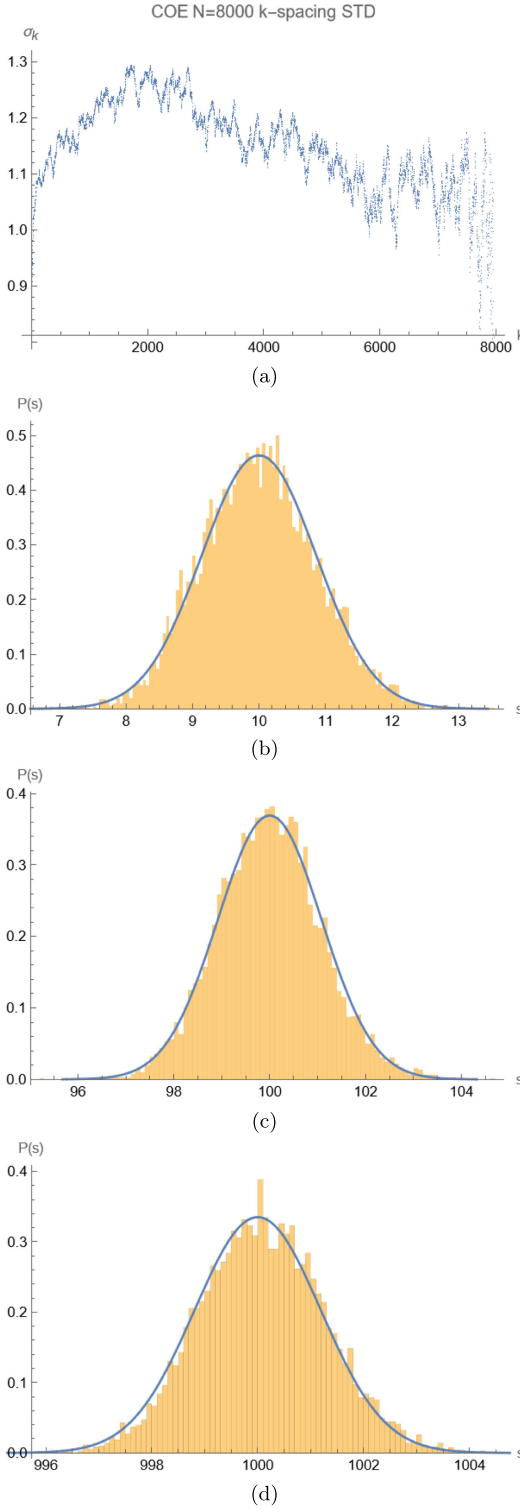


FIG. 12. (a) The standard deviation of  $k$ -spacing distribution calculated from COE with  $N = 8000$ . Note that it does not follow the  $\sim\sqrt{k}$  prediction of (46). (b)–(d) Different  $k$ -spacing distributions from the COE data, together with normal distributions with mean  $k$  and standard deviations from the corresponding data, which are seen for all  $k$  in (a) (blue curves). (d) The data seem to be bulged slightly to the right, since the mean is actually 1000.09, as can be seen from (a).

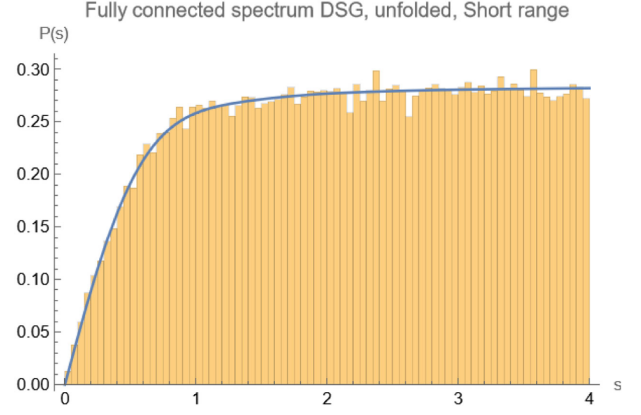


FIG. 13. DSG all order (not only nearest-neighbor) spacings (orange bars) with the predicted  $1 - Y_{\text{GOE}}$  (blue curve), normalized to the range shown.  $Y_{\text{GOE}}$  is given by (48).

orders and show the normalized probability distribution of the short spacings (around order of 1). (Note the unfolding rescaled the energy levels such that the mean consecutive level spacing is 1.) Comparing that with the two-point cluster function calculated from the GOE, which reads [38]

$$Y_{\text{GOE}}(r) = \left( \frac{1}{2} - \int_0^r S(r') dr' \right) \left( \frac{dS}{dr} \right) + S(r)^2, \quad (48)$$

where

$$S(r) = \frac{\sin(\pi r)}{\pi r}$$

and  $r$  is the spacing between any two energy levels, we find a high level of agreement between the data and the prediction (Fig. 13).

Following the discussion in (21), the cluster function for integrable systems is 0 so the correlation function, after normalization, is 1. The calculation of the pair correlations shows indeed a “plateau” when looking at the short range of spacings; at the full range it has the triangular shape due to the pair correlation being the autocorrelation of the level densities, which are constant at scales larger than the scale in which the fluctuations average to that constant, which is the same order as the correlation length—the length from which level repulsion becomes significant when there is one. The average level density is of course constant due to the unfolding.

## VI. SPECTRAL FORM FACTOR

We proceed and calculate the spectral form factor for the unfolded spectra of SG with  $mL = 0.01$  and DSG with  $mL = 1$ . To do so we choose from the higher end of the spectrum energy windows of size  $N = 100$  and use (35) to calculate the SFF as a function of time. We repeat that calculation 250 times for different energy windows which

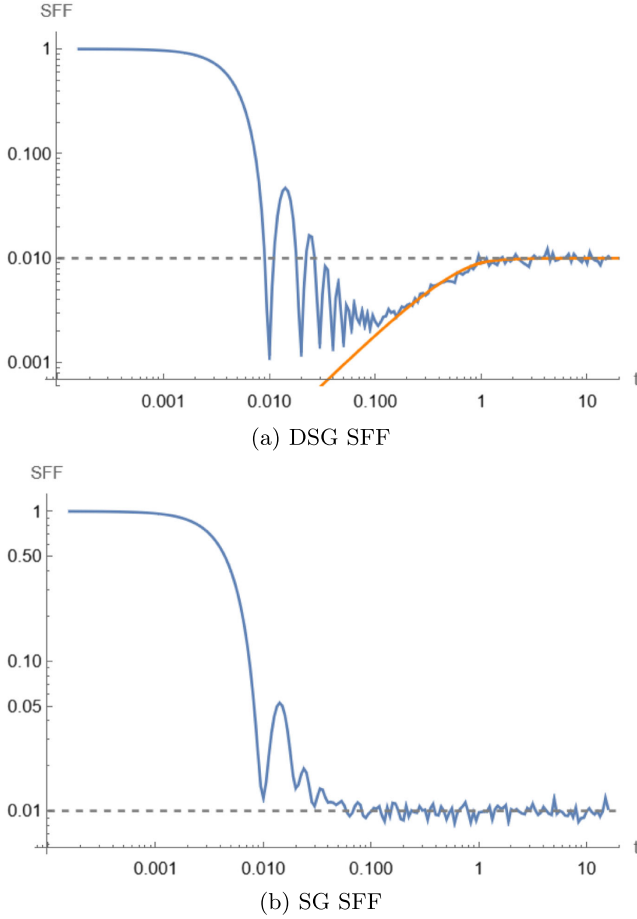


FIG. 14. Average SFF of DSG and SG as a function of time  $t$  (blue curves). The orange curve in (a) is the GOE SFF given in (37), with  $t \rightarrow t/2\pi$ . The dashed line in both (a) and (b) is  $1/N$ , which is the asymptotic value of the averaged SFF for both chaotic and integrable cases.

are mutually disjoint and finally take the average of all those calculations. The result can be seen in Fig. 14. The DSG shows structure characteristic of chaotic spectra, where there is initially an oscillating descent that goes to values smaller than  $1/N$ , then a ramp which should be linear for systems with time-reversal symmetry that correspond to the GOE, as in (37) when  $t \leq 1$ , and finally a plateau around the value of  $1/N$  when  $t \geq 1$ . The region where the value of the SFF is smaller than  $1/N$  is called the correlation hole. On the other hand, the SFF from SG exhibits no correlation hole, and the SFF is made of an oscillating descent that stops at the plateau at the value of around  $1/N$ . This is characteristic behavior of SFF of integrable systems, where the energy levels are uncorrelated.

We compare this to the SFF of the DSG spectrum that was not unfolded, only normalized such that the mean NN spacing is 1 (see Fig. 15). Though the correlation hole is still visible, the SFF from the unfolded spectrum is a better fit to the GOE curve than the one from the regular

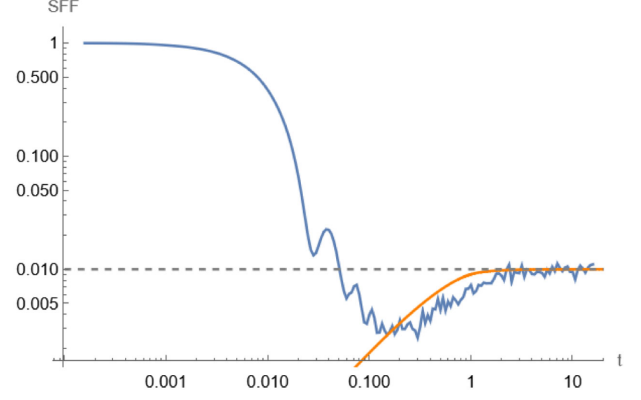


FIG. 15. SFF from the DSG spectrum which is normalized but not unfolded, together with the GOE connected SFF (orange curve).

spectrum. We note that, even without the unfolding, the SFF is not far from the GOE curve, since the change of the average mean level density is small in the energy windows chosen.

## VII. SPECTRAL RIGIDITY

Like the SFF and pair correlation, the spectral rigidity (also known as  $\Delta_3$  statistics) is a measure of long-range correlations between the energy levels. It shows distinct behavior for integrable systems, where it is linear in the energy scale  $L$ , as opposed to chaotic systems where it goes like  $\sim \ln(2\pi L)$ . Using (30), we calculate the spectral rigidity of the unfolded SG and DSG data and compare them to the RMT expressions in (32). In accordance with previous calculations, we choose  $mL = 1$  for DSG and  $mL = 0.01$  for SG, and take the high range part of the spectrum.

In general, we find that the spectral rigidity depends on the energy window it is calculated from. When the energy window was taken from the higher values of the energy shell, we find that SG fits well with  $\Delta_{3,0} = \frac{1}{15}L$  for  $L$  up to 500. When the energy window is taken from the lower part of the energy shell, one can see that the spectral rigidity saturates for some value  $L_{\max}$ . This saturation is related in the semiclassical picture to the classical shortest periodic orbit accessible by the specific model, which is nonuniversal and sets the “outer” energy scale for energy level correlations, as opposed to the “inner” scale, which is the average NN spacing [39,40]. In this regard we must add that the saturation of the spectral rigidity may also relate to the proximity of the energy window to the energy shell edge, which is also seen in deviation of other spectral properties from RMT, such as the average NN spacing ratio and the information entropy of the components of the eigenvectors (Appendix D in [2]). We see that, for energy windows further away from the lower edge of the shell, the spectral rigidity gets closer to RMT, eventually reaching Fig. 16(b).

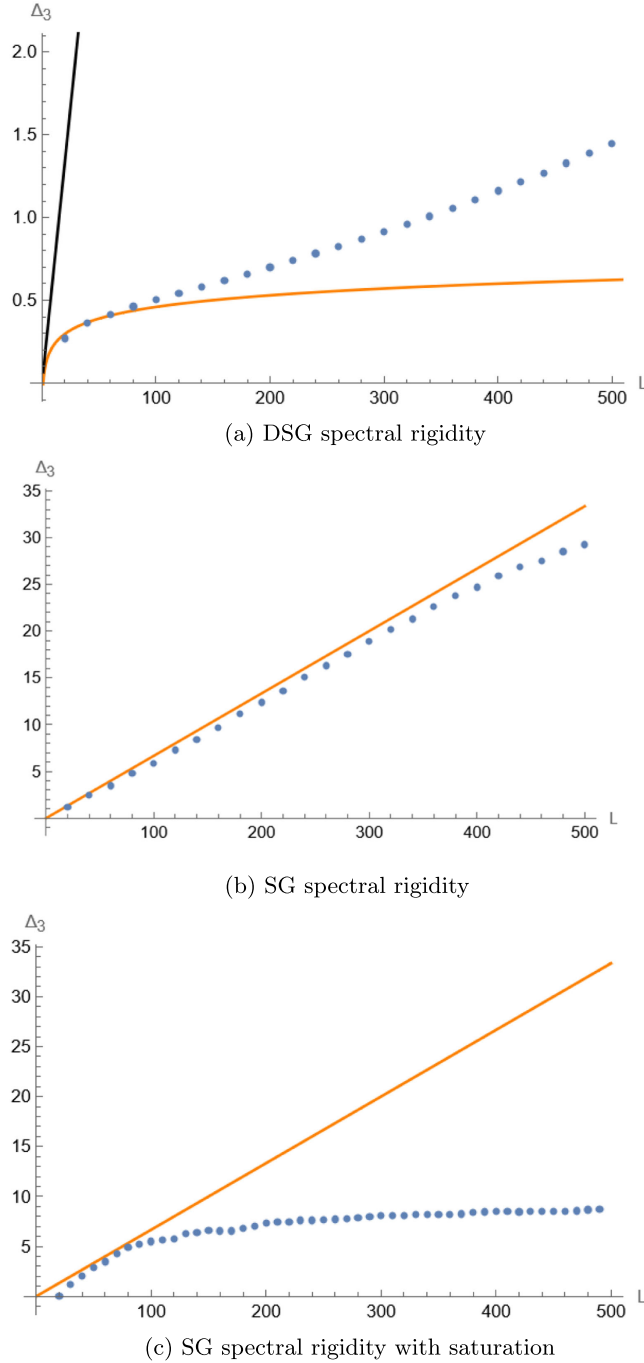


FIG. 16. The spectral rigidity of SG and DSG, calculated from the high range of the unfolded data, seen as blue dots. (a) The orange curve is  $\Delta_{3,1}$  in (32) and the black curve is  $\Delta_{3,0}$ . The numerical calculations follow  $\Delta_{3,1}$  until  $L \sim 70$ , where they detach from the curve. The orange curve in (b) is  $\Delta_{3,0}$ . (c) The SG spectral rigidity calculated from an energy window near the edge of an energy shell, showing a saturation from around  $L \sim 80$ .

For DSG, Fig. 16(a) shows that for small  $L$  the data fit with the expression (the orange curve), but around  $L \sim 70$  it diverges toward the integrable curve (black). This suggests that even though, for the local measures, the DSG data

seem to fit well with GOE, there is an energy scale below which the levels are correlated, but above the correlation ceases and Poisson statistics dominate. Similar results have been shown for the Sachdev-Ye-Kitaev model [41] and random quantum circuits [42], though the measure used for the detection of the spectral rigidity in those examples is the number variance  $\Sigma^2$  (26), which is closely related but not the same as  $\Delta_3$ . Recall that (29) and (32) are asymptotic expressions of  $1/N$  expansions, which remain small until  $L$  reaches a significant size of the energy band. As pointed out in [42], Sec. III. 1, in many-body problems the fluctuations are larger than few body or RMT, since they are suppressed by  $1/n$  ( $n$  being the particle number) rather than the Hilbert space size which goes like  $N \sim e^n$ . This makes the deviation from the asymptotic RMT expression appear at smaller  $L$  than predicted by finite RMT.

Another possible reason for the breaking of the spectral rigidity is given by [43], where the authors predict a hydrodynamic region of QFTs where long wavelength modulations act as approximately conserved charges, partitioning the spectrum into uncorrelated sectors. The hydrodynamic region is determined by the Thouless energy  $E_{th} = D/L^{3/2}$ , which for  $1+1d$  QFTs is given by the Kardar-Parisi-Zhang universality class rather than by diffusion, with  $D$  determined by the equilibrium of the equation of state.

Similar behavior can also be an indication of a mixed system [44], where the authors use a random matrix model that smoothly interpolates between Poisson-like and GOE statistics, using a mixing parameter  $\Delta \in [0, 1]$ . For mixed spectra, the spectral rigidity starts on the GOE curve and diverges around a value  $L_{max}$ , which is a function of both the mixing parameter  $\Delta$  and the size of the matrices  $N$ , such that for  $N \rightarrow \infty$  and  $\Delta > 0$ ,  $L_{max} \rightarrow \infty$  and the GOE statistics  $\Delta_{3,1}$  are recovered. It is seen physically in the framework of many-body localization transition [45,46], where the Thouless energy sets the scale below which the statistics of the energy level obey RMT and above which show Poisson statistics.

The Thouless time  $t_{Th}$  relates to the Thouless energy  $E_{Th}$  via  $t_{Th} = \hbar/E_{Th}$  and can be calculated using the dipping time  $t_p$  in the SFF and the Heisenberg time  $t_H = 2\pi/\langle s \rangle$  by  $t_{Th} = t_H t_p$  [47], where  $\langle s \rangle$  is the mean NN spacing before the unfolding. Using the calculation seen in Fig. 14(a) we get  $t_{Th} \sim 44,500$ .

We have also checked the calculation on the truncated spectrum without unfolding, to rule out a distortion of long-range correlations due to the unfolding, but the results were similar to Fig. 16(a).

## VIII. CONCLUSIONS AND FUTURE DIRECTIONS

We have computed various measures of characteristic quantum chaotic behavior on numerically calculated spectra of SG and DSG.

- (i) We have found that the DSG data when  $mL = 1$  match the behavior of the GOE in the NN spacing distribution and NN spacing ratio distributions. We found that mean value of the nearest-neighbor spacing ratio deviates by less than 1% from the value of the GOE. For the SG model with  $mL = 1$  the correspondence to the Poisson distributions is not as good as for the DSG. Using the unfolded data with  $mL = 0.01$  gives results that fit the RMT predictions much better.
- (ii) We computed the means of higher order level spacing distributions. The mean  $k$ -spacing  $\langle s_k \rangle$  shows a clear linear behavior with a slope of 1, and the mean  $k$ -spacing ratio  $\langle r_k \rangle$  matches the analytical prediction (42) from the unfolded DSG high energy levels. It shows a linear behavior with a slope of 1.
- (iii) We have suggested a probabilistic interpretation to the  $k$ -spacing distribution, which matches the Gaussian-like distributions that are obtained for large  $k$ , and the linear relation of mean  $k$ -spacing  $\langle s_k \rangle$  and  $k$ . This approach falls short when it comes to predictions of the standard deviation in  $k$ -spacing distributions, and it is less accurate in the specific distributions of small  $k$ 's, both cases due to the assumption that the distribution of each energy level is uncorrelated to the other, which is true only for the Poisson ensemble. We would like to measure the deviation of these distributions from Gaussianity in order to quantify this dependence.
- (iv) We have computed several long-range correlation measures. The pair correlation shows good agreement with the RMT prediction (48). The SFF also seems to agree with the predicted connected part (37). As for the spectral rigidity, SG matches the prediction well for all  $L$ , but in DSG we see that the numerically calculated spectral rigidity matches the prediction (32) only up to a certain value of  $L$ , for which we have also computed the corresponding Thouless time. This indicates that the GOE behavior of DSG is scale dependent, and beyond a certain energy scale the energy levels are uncorrelated.

### A. Future directions

Other directions we would like to pursue in future research are as follows.

#### 1. Quantum chaos characteristics

There are several more measures for quantum chaoticity that were not studied in this work and we would like to implement in the future:

- (i) Spectral rigidity. We intend to further study the discrepancy between the expected spectral rigidity as a function of  $L$  and the observed result.
- (ii) Lyapunov exponents (OTOCs). These were mentioned in Sec. I. We would like to explore further the connection to RMT and its values for QFT.
- (iii) Krylov complexity. Also known as “ $K$ -complexity,” an increasingly popular measure in recent times, specifically in describing the spread of local operators in Hilbert space as they evolve in time. The late time Krylov complexity of a system is different in a distinguished way for integrable and nonintegrable systems [5,48].

### 2. Transient quantum chaos and scattering

Though the study of quantum chaos usually relies on the energy spectrum or eigenmodes of models, one can also study the quantum analogs of classically transient chaos, i.e. the chaos of scattering [8–12]. A canonical classical example is the erratic and self-similar behaviors of “pinballs”—the scattering of a particle of a potential of three hard disks. The chaotic behavior is seen by plotting the final deflective angle of the particle as a function of the impact parameter [49] or in the transient time the particle spends inside the potential before shooting off. In quantum systems this behavior is described by the  $S$ -matrix, and the chaotic behavior could be described by the decay time or deflective angle of a highly excited state. This approach is more suitable for QFT since the  $S$ -matrix is a more natural observable than the energy levels.

### 3. Applications to other QFT

We would like to explore chaoticity using the measures mentioned above for more  $(1+1)$ D QFT, by applying the TCSCA [50] or light-cone conformal truncation [51] methods to compute the spectrum of integrable models and nonintegrable perturbations thereof by controlling some order parameter. Some possible directions are:

- (i) Conformal field theory (CFT) ensembles. A recent work [52] proposes a CFT analog to RMT, where each member of the ensemble is an approximate dataset of a CFT.
- (ii) Wess-Zumino-Witten model. Well studied case of  $(1+1)$ D CFT with several sectors which could show chaotic behavior [53].
- (iii) 2D QCD. At  $N \rightarrow \infty$  't Hooft limit where the model is integrable and  $1/N$  perturbations may introduce chaotic behavior [54].
- (iv) Bosonized 2D QCD. Multiflavored bosonized 2D QCD in the baryonic sector [55].
- (v) Higher-dimensional CFT such as the 4D  $\mathcal{N} = 4$  Super Yang-Mills in the large  $N$  limit, by studying the spectrum from spin chains [56] and nonintegrable  $1/N$  perturbations thereof.
- (vi) Higher-dimensional QFTs in general, such as the  $\lambda\phi^4$  model [3].

#### 4. Black holes and string theory

Recent studies have shown connections between black holes and chaos by considering the OTOC [4] and also in the  $S$ -matrix of particles interacting with a black hole (BH) horizon [57]. Chaotic behavior is also observed in the decay of highly excited strings [9,11]. It might be interesting to see connections between the two subjects, based on the BH-string correspondence [58,59].

#### 5. New measures for mixed ensemble states

We would like to explore more methods for determining whether a system is in a mixed state of e.g. integrable and GOE statistics. We suspect that a linear combination of the pair correlations shown in Sec. IID for the different ensembles could indicate a mixed state with a single interpolating parameter.

#### ACKNOWLEDGMENTS

We would like to thank A. Gaikwad who took part in the early stages of this project and for useful conversations. J. S. would like to thank M. Bianchi, A. Bhattacharya, M. Firrotta, A. Jana, and D. Weissman for useful discussions. This work was supported in part by Grant No. 01034816 titled “String theory reloaded- from fundamental questions to applications” of the “Planning and budgeting committee.”

#### DATA AVAILABILITY

The data that support the findings of this article are openly available [60].

#### APPENDIX A: TRUNCATED CONFORMAL SPACE APPROACH

Aside from calculations in RMT, the data used in this paper were taken from [2], where the authors used the TCSA to obtain the spectrum of Sine-Gordon and double Sine-Gordon models. TCSA is a nonperturbative numerical method that can be used to study QFT, which takes advantage of the properties of CFT by using its eigenstates as the basis of the calculations. It is part of the larger class of Hamiltonian truncations and was first suggested by Yurov and Zamolodchikov in 1989 [61] (for a more detailed account of TCSA, see for example [62]). The basic idea is to start with a known CFT Hamiltonian and to perturb it with a relevant operator

$$H = H_{\text{CFT}} + g \int dx \mathcal{V}(x), \quad (\text{A1})$$

where  $g$  is the coupling constant and  $\mathcal{V}(x)$  is a local relevant operator. This is done by choosing the eigenstates of the CFT as the basis  $|i\rangle$  with the highest eigenvalue less than a fixed energy cutoff  $E_{\text{cut}}$  and computing the matrix elements

of the perturbations

$$V_{ij} = \langle i | \int dx \mathcal{V}(x) | j \rangle. \quad (\text{A2})$$

The result is then diagonalized numerically to obtain the new eigenvalues.

Though pretty conceptually straightforward, this method might be computationally pricey since the number of states tends to grow exponentially in the cutoff energy. This problem is tackled by exploiting the symmetries of the system so that the truncated Hamiltonian becomes block diagonal and can be made even more efficient using renormalization group (RG) flow methods (e.g. numerical RG [63]).

The validity of the method relies on the assumption that the cutoff energy is large enough so that the IR regime is unaffected by the cutoff scale, i.e.

$$E_{\text{cut}} \gg |g|^{\frac{1}{2-\Delta}}, \quad (\text{A3})$$

where  $\Delta$  is the scaling dimension of  $\mathcal{V}$ , and  $|g|^{\frac{1}{2-\Delta}}$  is the mass scale of the system.

When the spatial dimension is compact with length  $L$ ,  $E_{\text{cut}}$  can be written explicitly as

$$E_{\text{cut}} = \frac{2\pi}{L} N_{\text{cut}}, \quad (\text{A4})$$

where  $N_{\text{cut}}$  is the maximal conformal level in the CFT basis. In practice, it is more useful to work with the dimensionless terms  $L|g|^{\frac{1}{2-\Delta}}$  and  $N_{\text{cut}}$ .

For the case of SG and DSG,  $H_{\text{CFT}}$  was taken to be that of a free scalar boson in  $d = 2$  on a cylinder of length  $L$ . The perturbation is  $\mathcal{V}_i = \cos(\beta_i \phi)$  with  $i = 1$  for SG and  $i = 1, 2$  for DSG. The maximal value of  $N_{\text{cut}}$  used in the calculations was 42 (which gives a spectrum of  $\sim 85,000$  energy levels), and the dimensionless term  $l = ML$  with  $M$  being the IR mass scale which is the breather mass (B7) with  $n = 1$ , showed best convergence when  $l \sim 1$  for SG and  $l \sim 3$  for DSG.

#### APPENDIX B: SINE-GORDON AND DOUBLE SINE-GORDON MODELS

In this appendix we introduce SG) and DSG as the QFT models which will be tested with the measures from the previous sections.

##### 1. SG model

Sine-Gordon is a known  $(1+1)\text{D}$  field theory which has been well studied throughout the years, describing physical phenomena such as Josephson junctions [64] and 1D spin chains. It is a nonlinear integrable model with known nonperturbative solutions and their interactions.

The Lagrangian density is

$$\mathcal{L}_{\text{SG}} = \frac{1}{2} \partial_\mu \phi \partial^\mu \phi - \mu_0 \cos \beta \phi. \quad (\text{B1})$$

The basic soliton solutions are called kinks

$$\phi(x, t) = 4 \arctan \left[ \exp \left( \pm \frac{x - vt}{\sqrt{1 - v^2}} \right) \right] \quad (\text{B2})$$

and are thought of as centered around  $x$  for which  $\phi = 0$  with a velocity  $0 \leq v < 1$  and a soliton doublet “mass” of

$$m = \frac{8\sqrt{\mu_0}}{\beta}.$$

They correspond to an overall shift in the phase  $\phi \rightarrow \phi + 2\pi$  for the solutions with  $+$  in the exponent, and  $\phi \rightarrow \phi - 2\pi$  for those with a minus sign, called “antikinks.” These solutions are topological and carry a topological charge of  $\theta_k = +1$  for the kinks and  $\theta_k = -1$  for the antikinks. A bound state of a kink and antikink is called a “breather,”

$$\phi(x, t) = 4 \arctan \left[ \exp \left( \frac{\sqrt{1 - \omega^2} \cos \omega t}{\omega \cosh(\sqrt{1 - \omega^2} x)} \right) \right] \quad (\text{B3})$$

with a continuous mass in  $|\omega| < 1$ , that is calculated by inserting the solution to the energy-momentum tensor

$$M = 2m \sin \frac{\pi}{2} \omega. \quad (\text{B4})$$

Since we are interested in the discrete spectrum of the model, for the quantized version of SG and later DSG it will be useful to look at the free field Hamiltonian as a CFT of a free boson with  $c = 1$  and add the potential terms as a perturbation to the CFT,

$$H_{\text{SG}} = H_{\text{CFT}} + \mu \int dx : \cos \beta \phi :, \quad (\text{B5})$$

where

$$H_{\text{CFT}} = \int dx : \frac{1}{2} (\partial_t \phi)^2 + \frac{1}{2} (\partial_x \phi)^2 :$$

The quantum spectrum consists of a soliton with mass  $M$  that relates to the coupling parameter  $\mu$  [65] as

$$\mu = \frac{2\Gamma(\Delta)}{\pi\Gamma(1 - \Delta)} \left( \frac{\sqrt{\pi}\Gamma(\frac{1}{2-2\Delta})M}{2\Gamma(\frac{\Delta}{2-2\Delta})} \right)^{2-2\Delta} \quad (\text{B6})$$

with  $\Delta = \beta^2/8\pi$ . The breather mass is now quantized and reads

$$M_n = 2M \sin \left( \frac{\pi}{2} n \frac{\Delta}{1 - \Delta} \right), \quad (\text{B7})$$

where  $n = 1, 2, \dots, \Delta^{-1} - 1$ . It was discovered by Coleman [66] that for  $\Delta < 1$  SG is dual to the charge zero sector massive Thirring model, with the special case of  $\Delta = \frac{1}{2}$  corresponding to a free massive Dirac field.

## 2. DSG

The Hamiltonian of the quantum DSG can be written in the same way as that of SG (B5) with another potential term

$$H_{\text{DSG}} = H_{\text{CFT}} + \mu \int dx : \cos \beta \phi : + \lambda \int dx : \cos(\alpha \phi + \delta) : \quad (\text{B8})$$

where the new term is seen as a perturbation of (B5), from which it is possible to calculate the perturbative form factors (PFFTs) and add them to the integrable spectrum, using the dimensionless parameter

$$\eta \equiv \lambda \mu^{\frac{1-\Delta_\alpha}{1-\Delta_\beta}} = \lambda \mu^{\frac{8\pi-\alpha^2}{8\pi-\beta^2}} \quad (\text{B9})$$

in the regimes where the limits of  $\eta \rightarrow 0$  and  $\eta \rightarrow \infty$  converge. The latter obtained by swapping the roles of the two potential terms such that the term with  $\alpha$  is now part of SG and the term with  $\beta$  becomes the perturbation.

It is worth mentioning that adding the second frequency breaks the degeneracy of the vacuum state explicitly, and when the frequencies  $\beta$  and  $\alpha$  are incommensurate, the vacuum of the theory becomes unique. As for the solitons and antisolitons, breaking the  $2\pi/\beta$  periodicity and lifting the degeneracy of the minimal state leads to an attractive linear potential between the soliton-antisoliton pairs which causes their collapse into a string of bound states [67].

Explicit calculations of the PFFT corrections for the vacuum energy, first and second breather masses, and the  $S$ -matrix are found in [68]. A semiclassical calculation gives the following correction to the first breather mass:

$$M_1^k = \sqrt{\mu}\beta + \lambda \frac{\alpha^2}{2\sqrt{\mu}\beta} \cos \left( \frac{2\pi}{\beta} k + \delta \right) + O(\lambda^2), \quad (\text{B10})$$

where  $k$  is one of the  $n$  copies of the breather due to the remaining vacuum degeneracy when  $\alpha/\beta$  is rational.

## APPENDIX C: RANDOM MATRIX THEORY

The theory of random matrices, as the name suggests, deals with matrices which are composed of random variables. When these matrices are symmetric under rotations (i.e. orthogonal or unitary transformations), and

when the random variables are made of identical distributions that are independent up to the rotational symmetry, those are called Wigner matrices, and they hold a special place in quantum chaos since they were derived as an attempt to capture the complex behavior of quantum many-body problems, specifically heavy nuclei. It is conjectured in what is now called the Wigner-Dyson-Gaudin-Mehta conjecture that the spectral statistics of Wigner matrices are universal, i.e. are independent of the choice of distribution for the random variables in the matrix [69].

Let us consider an  $N \times N$  matrix  $H$  where each entry is a random variable with the same probability distribution, with a zero mean and a variance of  $1/N$ ,

$$\mathbb{E}[H_{ij}] = 0 \quad \mathbb{E}[|H_{ij}|^2] = \frac{1}{N} \quad \text{for } i, j = 1, 2, \dots, N. \quad (\text{C1})$$

The object of interest will be the joint probability distribution of these variables  $P(H)$ , which according to Euler's lemma are weighted by  $\exp -\text{Tr}(H^{2n})$  for integer  $n$ . When rotated to the diagonalized basis the joint probability distribution is

$$P(\lambda_1, \lambda_2, \dots, \lambda_N) = C_N \Delta^\beta \exp \{-\beta N V(\lambda)\}, \quad (\text{C2})$$

where  $C_N$  is a normalization constant,  $\beta$  is the Dyson index, and it indicates rotational symmetry of  $H$ :  $\beta = 1$  for real symmetric matrices,  $\beta = 2$  for complex Hermitian matrices, and  $\beta = 4$  for quaternion self-adjoint matrices.  $\Delta = \prod_{j < i} (\lambda_i - \lambda_j)$  is the Vandermonde determinant and it is common to choose a quadratic form for the potential and call the family of such matrices the Gaussian ensemble (these correspond to Gaussian random variables).

This joint probability distribution carries the whole of the information of the system. In order to understand the system, it is useful to look at the  $n$  marginal probability distribution where  $n < N$ , also called the  $n$ -point correlation function

$$R_n(\lambda_1, \lambda_2, \dots, \lambda_n) = \frac{N!}{(N-n)!} \int P(\lambda_1, \lambda_2, \dots, \lambda_N) \times d\lambda_{n+1} d\lambda_{n+2} \cdots d\lambda_N, \quad (\text{C3})$$

where the combinatorial factor outside of the integration counts the number of possible  $N - n$  integration combinations, assuming the set of eigenvalues is ordered.

For example, the spectral level density, which is defined generally as

$$\rho(\lambda) = \frac{1}{N} \sum_i \delta(\lambda - \lambda_i), \quad (\text{C4})$$

is the one-point correlation  $R_1(\lambda)$ . There are several methods to derive the spectral density explicitly, most include for simplicity the large  $N$  limit. Wigner originally

used the moments method of computing  $\mathbb{E}[\text{Tr}(H^n)]$ . Another method is first writing the Gaussian ensemble version of (C2) in exponential form

$$P(\lambda_1, \dots, \lambda_N) = C_N \exp \left\{ -\frac{\beta N}{2} \sum_i \lambda_i^2 + \beta \sum_{i < j} \log |\lambda_i - \lambda_j| \right\}, \quad (\text{C5})$$

which now resembles the potential terms of a one-dimensional constrained 2D Coulomb gas in a harmonic potential. Taking the large  $N$  limit is now similar to taking the semiclassical limit, where sums are replaced with integrals that include the spectral density

$$\sum_i \rightarrow N \int d\lambda \rho(\lambda).$$

The requirement of a stationary phase is now translated to the equation

$$\lambda_i = \frac{1}{2} \int \frac{d\lambda \rho(\lambda)}{\lambda - \lambda_i}, \quad (\text{C6})$$

which can be solved with the resolvent method. The solution is the renowned ‘‘Wigner semicircle law’’

$$\rho_0(\lambda) = \frac{1}{\beta\pi} \sqrt{2\beta - \lambda^2}. \quad (\text{C7})$$

The two-point and higher correlation function are dealt with in Sec. II D.

One can also derive the distribution of eigenvectors, based on the rotational invariance [70]. For a real symmetric  $N \times N$  matrix, the  $N$ -dimensional eigenvectors are placed on an  $N$ -dimensional unit sphere and cover the unit sphere uniformly as we move through the ensemble. The normalized joint probability distribution for its components is

$$P(a_1, a_2, \dots, a_N) = \pi^{-N/2} \Gamma\left(\frac{N}{2}\right) \delta\left(\sum_i a_i^2 - 1\right), \quad (\text{C8})$$

which we can use to obtain the probability density of a single component  $y = a_i^2$ ,

$$P(y) = \int P(a_1, a_2, \dots, a_N) \delta(y - a_i^2) da_1 da_2 \cdots da_N \\ = \frac{1}{\sqrt{\pi}} \frac{\Gamma(N/2)}{\Gamma((N-2)/2)} \frac{(1-y)^{(N-3)/2}}{\sqrt{y}}. \quad (\text{C9})$$

Taking the large  $N$  limit and changing variables to  $\eta = Ny$ , we get

$$P(\eta) = \frac{1}{\sqrt{2\pi\eta}} e^{-\frac{\eta}{2}}, \quad (\text{C10})$$

which is known as the Porter-Thomas distribution. Similar results can be obtained for the GUE and GSE, assuming the components of the eigenvectors are complex or quaternion.

RMT can be generalized in various ways, for example beyond rotational invariance, where the joint probability distribution could be written with a general potential  $V(\lambda)$ , such as the Laguerre potential  $V(\lambda) = a \log \lambda - \lambda/2$  of Wishart ensembles. The parameter  $\beta$  can also be generalized to be continuous  $\beta > 0$ . In this case it is shown in [71] that the corresponding random matrices are tridiagonal, and for the Gaussian case (with a quadratic potential), these tridiagonal matrices are

$$H_\beta = \frac{1}{\sqrt{2}} \begin{pmatrix} N(0,2) & \chi_{(n-1)\beta} & & & \\ \chi_{(n-1)\beta} & N(0,2) & \chi_{(n-2)\beta} & & \\ & \ddots & \ddots & \ddots & \\ & & \chi_{2\beta} & N(0,2) & \chi_\beta \\ & & & \chi_\beta & N(0,2) \end{pmatrix} \quad (\text{C11})$$

where  $N(0,2)$  are random variables of a normal distribution with  $\mu = 0$  and  $\sigma = 2$  and the PDF of  $\frac{\chi_k}{\sqrt{2}}$  is

$$\frac{2}{\Gamma(k/2)} y^{k-1} e^{-y^2}.$$

- 
- [1] G. P. Brandino, R. M. Konik, and G. Mussardo, *J. Stat. Mech.* (2010) P07013.
- [2] M. Srdinsek, T. Prosen, and S. Sotiriadis, *Phys. Rev. Lett.* **126**, 121602 (2021).
- [3] L. V. Delacrétaz, A. L. Fitzpatrick, E. Katz, and M. T. Walters, *J. High Energy Phys.* **02** (2023) 045.
- [4] J. Cotler, G. Gur-Ari, M. Hanada, J. Polchinski, P. Saad, S. Shenker, D. Stanford, A. Streicher, and M. Tezuka, *J. High Energy Phys.* **05** (2015) 118.
- [5] E. Rabinovici, A. Sánchez-Garrido, R. Shir, and J. Sonner, *J. High Energy Phys.* **07** (2022) 151.
- [6] A. Avdoshkin, A. Dymarsky, and M. Smolkin, *J. High Energy Phys.* **06** (2024) 066.
- [7] V. Balasubramanian, R. N. Das, J. Erdmenger, and Z.-Y. Xian, *J. Stat. Mech.* (2025) 033202.
- [8] D. J. Gross and V. Rosenhaus, *J. High Energy Phys.* **05** (2021) 048.
- [9] V. Rosenhaus, *Phys. Rev. Lett.* **129**, 031601 (2022).
- [10] M. Bianchi, M. Firrotta, J. Sonnenschein, and D. Weissman, *Phys. Rev. Lett.* **129**, 261601 (2022).
- [11] M. Bianchi, M. Firrotta, J. Sonnenschein, and D. Weissman, *arXiv:2303.17233*.
- [12] M. Bianchi, M. Firrotta, J. Sonnenschein, and D. Weissman, *arXiv:2403.00713*.
- [13] S. Negro, F. K. Popov, and J. Sonnenschein, *Phys. Rev. D* **108**, 105024 (2023).
- [14] E. P. Wigner, *On the Statistical Distribution of the Widths and Spacings of Nuclear Resonance Levels* (1957).
- [15] F. J. Dyson, *J. Math. Phys. (N.Y.)* **3**, 157 (1962).
- [16] O. Bohigas, M. Giannoni, and C. Schmit, *Phys. Rev. Lett.* **52**, 1 (1984).
- [17] M. Berry and M. Tabor, *Proc. R. Soc. A* **356**, 375 (1977).
- [18] T. Brody, *Lett. Nuovo Cimento* **7**, 482 (1973).
- [19] T. Brody, J. Flores, J. French, P. Mello, A. Pandey, and S. Wong, *Rev. Mod. Phys.* **53**, 385 (1981).
- [20] F. Izrailev, *Phys. Rep.* **196**, 299 (1990).
- [21] B. Chirikov, F. Izrailev, and D. Shepelyansky, *Physica (Amsterdam)* **33D**, 77 (1988).
- [22] V. Oganessian and D. Huse, *Phys. Rev. B* **75**, 155111 (2007).
- [23] Y. Y. Atas, E. Bogomolny, O. Giraud, and G. Roux, *Phys. Rev. Lett.* **110**, 084101 (2013).
- [24] Y. Y. Atas, E. Bogomolny, O. Giraud, E. Vivo, and P. Vivo, *J. Phys. A* **46**, 355204 (2013).
- [25] P. Martinez-Azcona, R. Shir, and A. Chenu, *Phys. Rev. B* **111**, 165108 (2025).
- [26] D. Fox and P. Kahn, *Phys. Rev.* **134** (1964).
- [27] D. Engel, J. Main, and G. Wunner, *J. Phys. A* **31** (1998).
- [28] A. Abul-Magd and M. Simbel, *Phys. Rev. E* **60**, 5371 (1999).
- [29] W. Rao, *Phys. Rev. B* **102**, 054202 (2020).
- [30] S. H. Tekur, U. Bhosale, and M. Santhanam, *Phys. Rev. B* **98**, 104305 (2018).
- [31] T. Guhr, A. Müller-Groeling, and H. Weidenmüller, *Phys. Rep.* **299**, 189 (1998).
- [32] H. Montgomery, *Proc. Sympos. Pure Math.* **24**, 189 (1973).
- [33] A. M. Odlyzko, *Math. Comput.* **48**, 273 (1987).
- [34] H.-J. Stöckmann and M. Gutzwiller, *Am. J. Phys.* **68**, 777 (2000).
- [35] L. D'Alessio, Y. Kafri, A. Polkovnikov, and M. Rigol, *Adv. Phys.* **65**, 239 (2016).
- [36] F. Haake, S. Gnutzmann, and M. Kuś, *Quantum Signatures of Chaos* (Springer, Cham, 2018).
- [37] S. Das, S. K. Garg, C. Krishnan, and A. Kundu, *J. High Energy Phys.* **01** (2024) 172.
- [38] F. Haake, S. Gnutzmann, and M. Kuś, *Random Matrices and the Statistical Theory of Energy Levels* (Academic Press, New York, 1967).
- [39] M. V. Berry, *Proc. R. Soc. A* **400**, 229 (1985).
- [40] G. Casati, B. V. Chirikov, and I. Guarneri, *Phys. Rev. Lett.* **54**, 1350 (1985).

- [41] A. M. García-García and J. J. Verbaarschot, *Phys. Rev. D* **94**, 126010 (2016).
- [42] H. Gharibyan, M. Hanada, S. H. Shenker, and M. Tezuka, *J. High Energy Phys.* **07** (2018) 124.
- [43] L. Delacrétaz, A. L. Fitzpatrick, E. Katz, and M. Walters, *SciPost Phys.* **12**, 119 (2022).
- [44] P. Persson and S. Aberg, *Phys. Rev. E* **52**, 148 (1995).
- [45] M. Serbyn and J. E. Moore, *Phys. Rev. B* **93**, 041424 (2016).
- [46] P. Braun, D. Waltner, M. Akila, B. Gutkin, and T. Guhr, *Phys. Rev. E* **101**, 052201 (2020).
- [47] P. Sierant, D. Delande, and J. Zakrzewski, *Phys. Rev. Lett.* **124**, 186601 (2020).
- [48] A. Bhattacharya, A. Jana, and J. Sonnenschein (2026) (to be published).
- [49] V. Rosenhaus, *Phys. Rev. Lett.* **127**, 021601 (2021).
- [50] V. Yurov and A. Zamolodchikov, *Int. J. Mod. Phys. A* **5**, 3221 (1989).
- [51] N. Anand, A. Fitzpatrick, E. Katz, Z. Khandker, M. Walters, and Y. Xin, *arXiv:2005.13544*.
- [52] A. Belin, J. de Boer, D. Jafferis, P. Nayak, and J. Sonner, *arXiv:2308.03829*.
- [53] P. Goddard and D. Olive, *Int. J. Mod. Phys. A* **01**, 02 (1986).
- [54] G. 't Hooft, *Nucl. Phys.* **B75**, 461 (1974).
- [55] Y. Frishman and J. Sonnenschein, *Phys. Rep.* **223**, 309 (1993).
- [56] J. Polchinski, *arXiv:1505.08108*.
- [57] J. Polchinski, *Phys. Rep.* **223**, 309 (2015).
- [58] G. Horowitz and J. Polchinski, *Phys. Rev. D* **55**, 6189 (1997).
- [59] Y. Chen, J. Maldacena, and E. Witten, *arXiv:2109.08563*.
- [60] M. Srdinšek, T. Prosen, and S. Sotiriadis (2020), *10.5281/zenodo.4319672*.
- [61] V. P. Yurov and A. B. Zamolodchikov, *Int. J. Mod. Phys. A* **05**, 3221 (1990).
- [62] M. Hogervorst, S. Rychkov, and B. C. van Rees, *Phys. Rev. D* **91**, 025005 (2015).
- [63] R. M. Konik and Y. Adamov, *Phys. Rev. Lett.* **98**, 147205 (2007).
- [64] F. S. Cataliotti, S. Burger, C. Fort, P. Maddaloni, F. Minardi, A. Trombettoni, A. Smerzi, and M. Inguscio, *Science* **293** (2001).
- [65] R. Baxter, *Exactly Solved Models in Statistical Mechanics* (Academic Press, New York, 1982).
- [66] S. Coleman, *Phys. Rev. D* **11**, 2088 (1975).
- [67] G. Delfino and G. Mussardo, *Nucl. Phys.* **B516**, 675 (1998).
- [68] Z. Bajnok, L. Palla, G. Takacs, and F. Wagner, *Nucl. Phys.* **B601**, 503 (2001).
- [69] L. Erdos and H.-T. Yau, *arXiv:1106.4986*.
- [70] L. Reichl, *The Transition to Chaos: Conservative Classical Systems and Quantum Manifestations* (Springer, New York, 2004).
- [71] I. Dumitriu and A. Edelman, *J. Math. Phys. (N.Y.)* **43**, 5830 (2002).