

PAPER • OPEN ACCESS

Missing levels in intermediate spectra

To cite this article: María Hita-Pérez *et al* 2023 *J. Phys. A: Math. Theor.* **56** 505702

View the [article online](#) for updates and enhancements.

You may also like

- [Making sense of non-Hermitian Hamiltonians](#)
Carl M Bender
- [Information geometry in vapour-liquid equilibrium](#)
Dorje C Brody and Daniel W Hook
- [Quantum chaos in Feshbach resonances of the ErYb system](#)
Maciej B Kosicki, Mateusz Borkowski and Piotr S uchowski

Missing levels in intermediate spectra

María Hita-Pérez¹ , Laura Muñoz²  and Rafael A Molina^{3,*} 

¹ Institute of Fundamental Physics, IFF-CSIC, Serrano 113b, Madrid E-28006, Spain

² Grupo de Física Nuclear, Departamento de Estructura de la Materia, Física Térmica y Electrónica, and IPARCOS, Universidad Complutense de Madrid, CEI Moncloa, Madrid E-28040, Spain

³ Instituto de Estructura de la Materia, IEM-CSIC, Serrano 123, Madrid E-28006, Spain

E-mail: rafael.molina@csic.es

Received 2 June 2023; revised 17 October 2023

Accepted for publication 16 November 2023

Published 28 November 2023



CrossMark

Abstract

We derive an expression for the nearest-neighbor spacing distribution $P(s)$ of the energy levels of quantum systems with intermediate dynamics between regularity and chaos and missing levels due to random experimental errors. The expression is based on the Brody distribution, the most widely used for fitting mixed spectra as a function of one parameter. By using Monte Carlo simulations of intermediate spectra based on the β -Hermite ensemble of random matrix theory (RMT), we evaluate the quality of the formula and its suitability for fitting purposes. Estimations of the Brody parameter and the fraction of missing levels can be obtained by a least-square two-parameter fitting of the experimental $P(s)$. The results should be important to distinguish the origins of deviations from RMT in experimental spectra.

Keywords: quantum chaos, spectral statistics, missing levels

(Some figures may appear in colour only in the online journal)

1. Introduction

Statistical analysis of spectra is a very important tool for understanding the dynamics of quantum and wave systems [1–3]. Quantifying the level repulsion, for example, it is

* Author to whom any correspondence should be addressed.



Original Content from this work may be used under the terms of the [Creative Commons Attribution 4.0 licence](https://creativecommons.org/licenses/by/4.0/). Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

possible to study the transition between integrable and chaotic quantum or wave systems [4–14], between localized and extended states in disordered systems [15–18] and between ergodic and many-body localized phases in many body strongly correlated systems [19–23]. This relationship is based on the identification of complex wave or quantum spectra with random matrix theory (RMT) [24]. These ideas initially came from nuclear physics but are now backed up with a semiclassical justification and a great body of numerical evidence behind them [3, 25]. Experimental studies of the spectral statistics of different quantum and wave systems from this perspective are numerous [2, 13, 26–28]. However, there are intrinsic limitations that have prevented a more wide use of these tools in experimental spectra. For a meaningful statistical analysis one needs complete sequences with no missing levels, no mixing of symmetries and long enough to have sufficient statistics. When the spectrum to be analyzed does not fulfill these conditions the reliability of the statistical analysis can be compromised, leading to incorrect conclusions, as explained next.

In chaotic systems with time-reversal symmetry the appropriate RMT ensemble is the GOE (Gaussian orthogonal ensemble) whereas regular systems show spectral fluctuations described by Poisson statistics, corresponding to uncorrelated spectra [29]. Thus a transition from chaos to regularity manifests in the spectral fluctuations as a transition from GOE to Poisson statistics. But this kind of transition can also be induced in a GOE spectrum by the loss of correlations caused by missing levels or mixed symmetries, taking its spectral statistics also towards Poisson. Thus, when statistical analysis of a spectrum throws an intermediate result between GOE and Poisson, it is quite difficult to distinguish the actual origin of the intermediate behavior, whether it is due to a true mixed dynamics between chaos and regularity or to missing levels and/or mixed symmetries in an actual GOE spectrum. An attempt at that requires a complex analysis probably taking into account different statistics [30]. Moreover, if it is due to both reasons it is even more difficult (when not impossible, unless one has some previous theoretical or experimental information on the dynamics of the system or the completeness of the spectrum) to find out what is the main one or estimate the weight of each, that is, estimate the degree of chaos and the number of missing levels or mixed symmetries independently.

There has been a line of research that tries to circumvent some of these limitations and even take advantage of them in order to estimate the number of missing levels and the number of mixed symmetries in a particular experimental level sequence [5, 31–35]. These approaches are based on RMT and the key point in order to be able to extract reliable information about missing levels or mixed symmetries is to assume that the spectral statistics coincide with the GOE (or the corresponding RMT ensemble for each symmetry class), that is, the actual experimental spectrum is chaotic. For example, analyzing the spectral statistics of chaotic nuclei it is possible to estimate how isospin symmetry is broken [36]. The number of missing levels in experimental sequences can also be estimated using these techniques. It is possible, then, to correct the value of experimentally obtained level densities [37–40].

To sum up, there are tools available to determine the degree of chaos assuming that the spectrum is complete (no missing levels) and tools to estimate the number of missing levels assuming that the spectrum is chaotic (GOE). But to the best of our knowledge, there are no tools with which one can obtain both parameters at the same time: there has been no attempt to study the effect of missing levels on the spectral statistics of a mixed system between chaos and regularity. It is the purpose of this paper to fill the gap. We have been able to derive a two-parameter formula for the nearest-neighbor spacing distribution (NNSD) $P(s)$ that allow an independent estimation of the degree of chaos and the fraction of experimentally observed levels. Through Monte Carlo simulations of mixed spectra based on the β -Hermite ensemble we study the accuracy of our formula and perform some tests to prove its usefulness by fitting the $P(s)$ of these simulated spectra. We prove then that our formula is useful for estimating

at the same time the chaoticity and the fraction of missing levels of an experimental level sequence when this fraction is not very large.

2. $P(s)$ for missing levels in intermediate systems

The two extremes (regular-Poisson and chaotic-GOE) in the NNSD are universal. However, there is not a universal transition for the intermediate dynamics. Two general behaviors in this transition can be distinguished in terms of level repulsion: fractional level repulsion and level repulsion of only a fraction of levels [41, 42]. Fractional level repulsion when the NNSD behaves as $P(s) \sim s^q$ for small values of s is the most common. It can be described phenomenologically quite well by the Brody [4] and the Izrailev [6] distributions. Both describe this kind of repulsion at small spacings but are slightly different at large spacings. They are phenomenological distributions, although the Brody distribution can be derived from a power-law ansatz for the level repulsion function [43]. We choose the Brody distribution in this paper as it is the more extensively used for fitting results in intermediate systems. The Brody's formula is given by:

$$P_B(s) = as^q \exp(-bs^{q+1}), \tag{1}$$

where $a = (q + 1)b$, $b = \left[\Gamma\left(\frac{q+2}{q+1}\right) \right]^{q+1}$, and q is the so called Brody's parameter or mixing parameter, which allows to interpolate between Poisson distribution for $q = 0$ (regularity):

$$P_P(s) = \exp(-s), \tag{2}$$

and Wigner distribution for GOE for $q = 1$ (chaos):

$$P_W(s) = \frac{\pi}{2} s \exp\left(-\frac{\pi}{4} s^2\right). \tag{3}$$

The NNSD is the first of a series of spacing distributions between neighbors of n -order, $p(n, s)$, $n = 0$ being the NNSD. These distributions are defined after the unfolding of the spectrum where the level density is locally normalized to unity [3]. A correct unfolding is not an easy task [44] but, for our purposes, we assume that the spectrum has been properly unfolded. This is standard in the statistical analysis of spacing distributions. Moreover, for our calculations here with β -ensemble spectra the unfolding is straightforward as the level densities follow the semicircle law and are known from RMT [45]. For unfolded levels these distributions need to be normalized and centered with the following conditions:

$$\int_0^\infty p(n, s) ds = 1, \tag{4}$$

$$\langle s \rangle = \int_0^\infty sp(n, s) ds = n + 1. \tag{5}$$

When levels are missing from a sequence the higher order spacing distributions start to play a role. Using simple statistical considerations, Bohigas and Pato derived the general formula for the NNSD of a sequence with only a fraction of observed levels f (the fraction of missing levels would be then $1 - f$) as a function of the $p(n, s)$ distributions and assuming that the missing levels are randomly distributed [32]:

$$p(s) = \sum_{k=0}^\infty (1 - f)^k p(k, s/f). \tag{6}$$

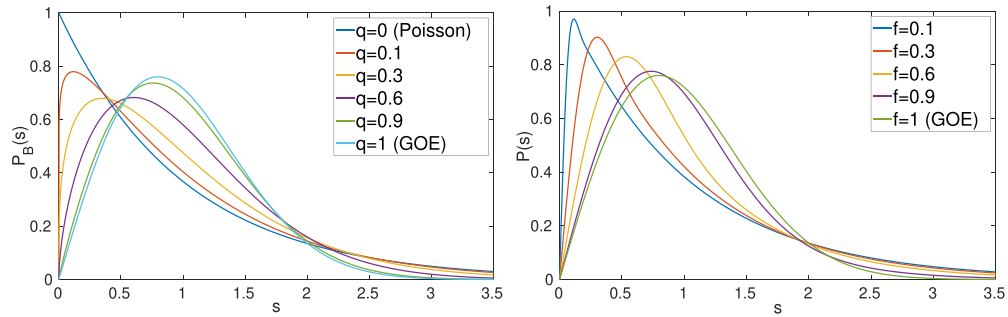


Figure 1. Comparison of the evolution of $P(s)$ from GOE to Poisson with the mixing parameter q of equation (1) for the transition with the type of dynamics (a) and with the fraction f of observed levels of equation (6) (b).

The interpretation of this formula is easy as the prefactor $(1 - f)^k$ for $k \geq 1$ is the probability that there are k missing levels between two given levels. An observed nearest neighbor spacing (left hand side) can really be a 2nd neighbor spacing with 1 missing level in between, a 3rd neighbor spacing with 2 missing levels in between, ... a $(k + 1)$ th neighbor spacing with k missing levels in between. ... in the original spectrum (right hand side). One, then, needs to use s/f instead of s , given that the level sequence with a fraction f of observed levels is unfolded to $\langle s \rangle = 1$, so the average spacing in the original sequence without missing levels is a factor $1/f$ higher.

This equation implies that missing levels are randomly sampled. This assumption is central to our work and can be justified depending on the experimental details for obtaining the particular spectra under analysis. For example, in measurements with microwave or acoustic resonators where the missing levels depend on the position of an antenna or detector with respect to the nodes of the wave function [46–48] or when observing compound nucleus resonances with limited detection sensitivity, the assumption can be justified [31, 32]. However, if missing levels come from the spectral resolution of the experimental device such that levels close enough are mistaken as one, the assumption fails as level clustering may be left undetected. Without the assumption of random sampling, the analytical treatment is much more complicated and outside the scope of this work. We discuss further this problem in the conclusions.

From equation (6) and the form of the higher-order spacing distribution for GOE one can see that the small spacing behavior of the transition from GOE to Poisson when we have fractional behavior ($P(s) \sim s^q$) is very different from the transition to Poisson due to missing levels where we still have $P(s) \sim s$. This can be seen in figure 1, where we show the evolution with the type of dynamics of equation (1) for different values of q compared to the evolution with missing levels of equation (6) for different values of f . This gives us hope that we can derive a two-parameter distribution that can distinguish the origin of the transition through a fit to experimental spectra.

Given Brody’s NNSD (1), Abul-Magd and Simbel [49] derived a general expression for the level-spacing distributions for higher order neighbors $p(n, s)$ using a statistic treatment proposed by Engel *et al* [43]. They obtained the following generalization of Brody’s formula:

$$p_B(n, s) = a_n s^{(n+1)q} \exp\left(-b_n s^{(n+1)q+1}\right) \int_0^s p(n-1, x) \exp\left(b_n x^{(n+1)q+1}\right) dx, \quad (7)$$

for $n \geq 1$. Here $a_n = [(n + 1)q + 1]b_n$ and both constants, a_n and b_n , are determined by $p(n, s)$ normalization conditions (4) and (5).

This expression presents many complications when trying to calculate high-order spacings. This is due to the fact that an exact solution for b_n can only be found in the case in which the system is regular ($q = 0$) where $b_n = 1$, and in all the other cases it has to be calculated by solving the integrals numerically. Those calculations are too heavy for practical purposes but we have found that Gaussian approximations for $n \geq 3$ work very well. This trick was already used in the original work of Bohigas and Pato for missing levels in GOE [32]. Then, we have proceed as follows.

For $n = 1$, a generalization of the Brody formula for the next-nearest-neighbor distribution is obtained by substituting equation (1) for $p_B(0, s)$ into equation (7).

$$p_B(1, s) = aa_1 \int_0^s s^{2q} x^q \exp \left[-bx^{q+1} - b_1 (s^{2q+1} - x^{2q+1}) \right] dx, \tag{8}$$

where $a_1 = (2q + 1)b_1$. The parameter b_1 was parametrized by Abul-Magd and Simbel and has the form:

$$b_1(q) = \frac{1}{1 + 2.7q + 3.5q^2}. \tag{9}$$

For $n = 2$, following the same steps and substituting equation (8) into equation (7), we find that

$$p_B(2, s) = aa_1 a_2 \int_0^s \int_0^x s^{3q} x^{2q} y^q \exp \left[-by^{q+1} - b_1 (x^{2q+1} - y^{2q+1}) - b_2 (s^{3q+1} - x^{3q+1}) \right] dx dy. \tag{10}$$

Here $a_2 = (3q + 1)b_2$ and b_2 has to be obtained numerically from the normalization conditions. As Abul-Magd and Simbel did for b_1 , we have calculated using Monte Carlo simulations the values of b_2 for a scan of values of q ($q = 0.1, 0.2, \dots, 0.9, 1$) and parametrized the result in the form:

$$b_2(q) = \frac{1}{1 + 6.7q + 1.3q^2 + 51q^3}. \tag{11}$$

For $n \geq 3$, we found that the n th-order level-spacing distribution follows mostly a Gaussian distribution making the calculation of b_n not worth the time it takes, especially taking into account that one should keep many terms to have a good approximation to the infinite sum of expression (6). In this work we have kept up to 150 terms. Thus instead of using the expression (7) to describe the $p(n, s)$ distribution, we use now a Gaussian approximation given by

$$p(n, s) = \frac{1}{\sqrt{2\pi}\sigma(n, q)} \exp \left[-\frac{(s - \mu)^2}{2\sigma(n, q)^2} \right], \quad n \geq 3, \tag{12}$$

where μ is the mean of the distribution, $\mu = \langle s \rangle = n + 1$, and $\sigma(n, q)$ is its standard deviation. From now on, $\sigma(n, q)$ will be calculated using spline interpolation from a battery of $\sigma(n, q)$ values previously obtained from spectra with different q values ($q = 0.1, 0.2, \dots, 0.9, 1$). The argument for the Gaussian approximation is given by the central limit theorem as a higher order spacing of order n is a sum of n nearest neighbor spacings. The higher n is, the better the Gaussian approximation becomes. We show in figure 2 the results for $P(3, s_3)$ for three different values of q . The histograms were obtained using calculations with the β -ensemble as explained in the next section. Although the approximation by a Gaussian is better for larger

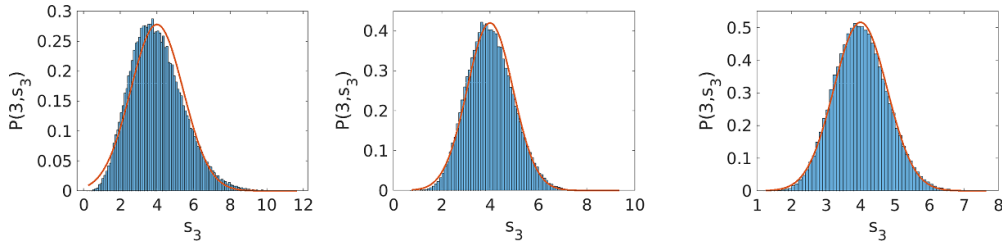


Figure 2. Three examples of $P(3, s_3)$ for $q = 0.1$ (left panel), $q = 0.5$ (central panel) and $q = 0.9$ (right panel) obtained numerically with the β -ensemble and comparison with the Gaussian approximation centered in $s_3 = 4$.

values of q , all cases with $n > 3$ are certainly good enough for our purposes. The worst case that we actually use is shown in the left panel of the figure $P(3, s)$ for $q = 0.1$.

Once we have an expression for $p(n, s)$, we are ready to propose an expression to describe the NNSD, $P(s)$, associated with incomplete spectra of intermediate and chaotic systems. Substituting equations (1), (8), (10) and (12) into equation (6), we obtain:

$$\begin{aligned}
 P(s, f, q) = & a \left(\frac{s}{f}\right)^q \exp \left[-b \left(\frac{s}{f}\right)^{q+1} \right] \\
 & + (1-f) a a_1 \int_0^{\frac{s}{f}} \left(\frac{s}{f}\right)^{2q} x^q \exp \left\{ -b x^{q+1} - b_1 \left[\left(\frac{s}{f}\right)^{2q+1} - x^{2q+1} \right] \right\} dx \\
 & + (1-f)^2 a a_1 a_2 \int_0^{\frac{s}{f}} \int_0^x \left(\frac{s}{f}\right)^{3q} x^{2q} y^q \exp \left\{ -b y^{q+1} - b_1 [x^{2q+1} - y^{2q+1}] \right. \\
 & \left. - b_2 \left[\left(\frac{s}{f}\right)^{3q+1} - x^{3q+1} \right] \right\} dy dx + \sum_{n \geq 3} \frac{(1-f)^n}{\sqrt{2\pi} \sigma(n, q)} \exp \left\{ -\frac{\left[\left(\frac{s}{f}\right) - n - 1 \right]^2}{2\sigma(n, q)} \right\},
 \end{aligned}
 \tag{13}$$

where all the parameters have already been defined.

This is the central result of this work. Despite its complicated appearance it is suitable for calculations and fitting purposes, as shown in the next sections. But besides its practical interest, the two-parameter formula is interesting in its own as a first step to start to unravel the puzzle of the actual origin of intermediate behavior of fluctuations in a given spectrum.

3. Comparison with β -ensemble results

The β -Hermite random matrix ensemble was proposed as a continuous generalization for all values of the repulsion parameter $\beta > 0$ of the classical random matrix ensembles corresponding to integer values $\beta = 1$ (GOE), 2 (GUE) and 4 (GSE) [45, 50]. It allows a transition between the Poisson integrable results corresponding to $\beta = 0$ and the chaotic GOE results $\beta = 1$ with fractional level repulsion. The physical interpretation of having a continuous value of β has been explored theoretically in the early days of the application of RMT to quantum physics through the work on models of repelling quantum particles in one-dimension, like the Dyson’s Coulomb gas and the Calogero-Sutherland model [51–55]. However, it was not until

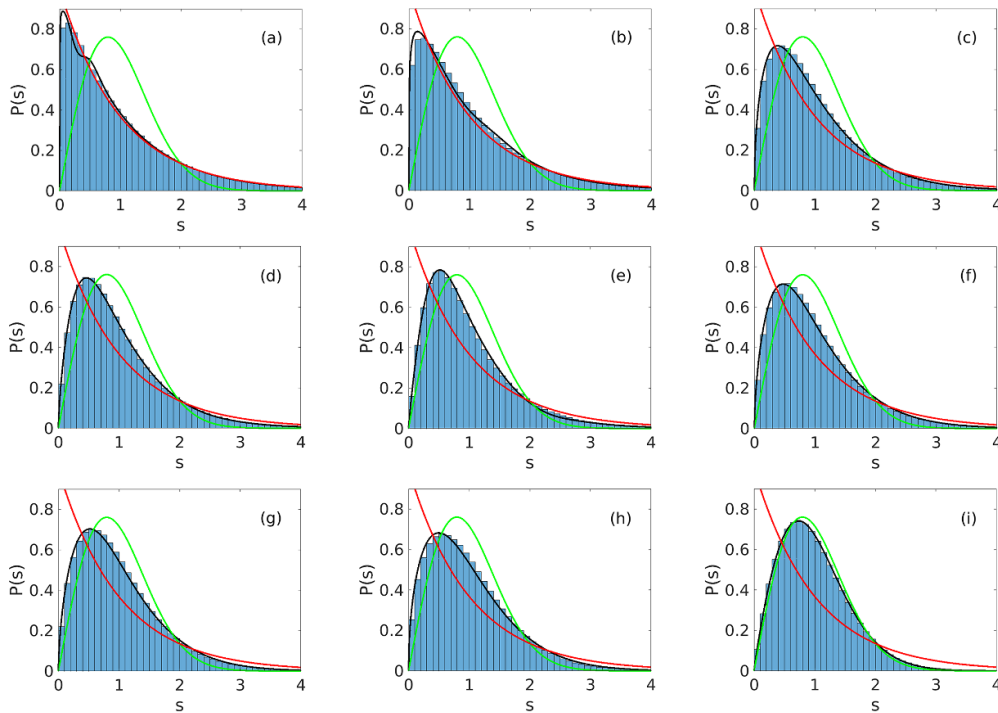


Figure 3. Comparison of equation (13). (black line) with ensemble averages of 1000 matrices of dimension $N = 10000$ from the β -Hermite ensemble with $\beta = q$ and a fraction f of observed levels (that is, the theoretical formula, not a fit). Green line is equation (3) (GOE) and red line is equation (2) (Poisson). The values of (f, q) are: (a)–(0.1, 0.1), (b)–(0.3, 0.2), (c)–(0.6, 0.5), (d)–(0.6, 0.7), (e)–(0.6, 0.9), (f)–(0.7, 0.6), (g)–(0.8, 0.6), (h)–(0.9, 0.5), (i)–(0.95, 0.9).

scanned all these values of β and f for the sake of completeness but this region is of no much practical interest. No one can expect reasonable estimations with too low fractions of observed levels in too uncorrelated spectra.

Second, looking at table 1 there is another region of values which could be strikingly higher than the rest: $f \geq 0.9, q \leq 0.3$, where χ^2 is of order tenths. However, this is also expected as this region correspond to the fitting of Brody’s formula for low values of β for a practically complete spectrum ($f \simeq 1$), and, as we have mentioned before, the $P(s)$ of the β -ensembles deviate from the Brody distribution in this case [57].

In summary, in this section we made use of the β -Hermite ensemble to show that the derivation of the two-parameter formula is correct by checking its accuracy with very high dimension. In the next section we explore if the formula could be used for realistic spectra to gain information confidently about the chaoticity and the number of missing levels simultaneously by fitting data with our expression.

4. Distributions of fitted results

In this section we again take advantage of the β -ensemble in order to show that the two-parameter formula is suitable for fitting purposes and the possible limitations one should be

Table 1. Values of χ^2 for the NNSD of β -ensembles of 1000 matrices of dimension $N = 10\,000$ generated with mixing parameter $q = \beta$ and fraction of observed levels f with respect to the formula of equation (13).

$q \backslash f$	0.1	0.3	0.6	0.7	0.8	0.9	0.95	1
0.1	0.021	0.040	0.052	0.056	0.063	0.073	0.078	0.082
0.2	0.017	0.035	0.054	0.063	0.076	0.091	0.10	0.11
0.3	0.013	0.029	0.048	0.054	0.066	0.082	0.092	0.10
0.4	0.013	0.022	0.035	0.041	0.049	0.061	0.071	0.083
0.5	0.013	0.016	0.023	0.027	0.031	0.039	0.050	0.057
0.6	0.013	0.0096	0.012	0.014	0.016	0.020	0.024	0.032
0.7	0.012	0.0062	0.0059	0.0065	0.0075	0.0076	0.0089	0.013
0.8	0.015	0.0079	0.0087	0.0083	0.0076	0.0054	0.0034	0.0022
0.9	0.019	0.018	0.023	0.022	0.020	0.015	0.0096	0.0014
1	0.030	0.047	0.054	0.051	0.045	0.038	0.029	0.012

careful with. We generate ensembles of matrices with certain values of the dimension N and the parameters β and f as in the previous section, but now we perform fits of these data with our two-parameter formula and we represent in a two-dimensional histogram the distribution of the results for q and f , so we can analyze the probability of obtaining the correct values of the parameters when analyzing a single spectrum of interest. Thus, the uncertainty in the estimation of the parameters is not only related to errors in the fit but to the variance of the parameter distribution in the proper ensemble [35]. This variance increases when the matrix size is reduced.

We want to stress that we do not intend to give a general recipe on how to use the formula. We have used the β -ensemble as a simple RMT model to simulate the transition from regularity to chaos. However, the transition is not universal and can, then, be different in actual experimental spectra. So we will not present here results for a fine scan of both parameters as in the previous section, but only show some representative cases to explain important aspects to take into account and give a general recommendation. That is, we make available the formula and await for experimental results to evaluate its real applicability and usefulness when approaching each case with its particular features.

For few missing levels and dynamics close to chaos (high values of f and β), the spectra are still very correlated and we can obtain reliable estimations of both parameters. We show an example in figure 4 where we have represented the results for the joint distribution of fitted parameters for an ensemble of 10 000 spectra of dimension $N = 1000$ with $f = 0.8$ and $\beta = 0.9$. Here we find a narrow peak of values of the parameters concentrated around the correct estimations. The ensemble variance is small in this case, so the probability of obtaining the correct values of the parameters q and f from a fit of the data is high.

However, it is obvious that when spectra are very uncorrelated (very close to Poisson statistics) little information can be gained. The loss of correlations can occur when there are many missing levels (low values of f) or when the dynamics is very close to regularity (low values of β). This intuition should be reflected in a very high ensemble variance of the parameters. For example, in figure 5 we show the results of the distribution of the fitted parameters of the two-parameter formula for $P(s)$ in an ensemble of 10 000 spectra of dimension $N = 1000$ with $f = 0.6$ and $\beta = 0.3$. As can be seen the results for (q, f) are spread over a wide region. One obtains with a high probability values around $(f \simeq 1, q \simeq 0.3)$ and, as expected, there are many different combinations of values (q, f) to reproduce such an uncorrelated spectra as any NNSD in which any of the two parameters is very low will be very close to Poisson.

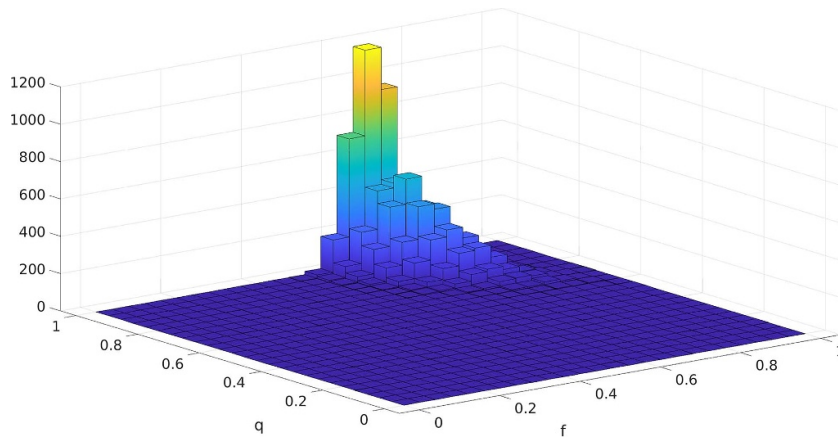


Figure 4. Histogram of the joint distribution of fitted values of the parameters (q, f) of an ensemble of 10 000 β -Hermite matrices with $f = 0.8$ and $\beta = 0.9$. The matrix sizes are $N = 1000$.

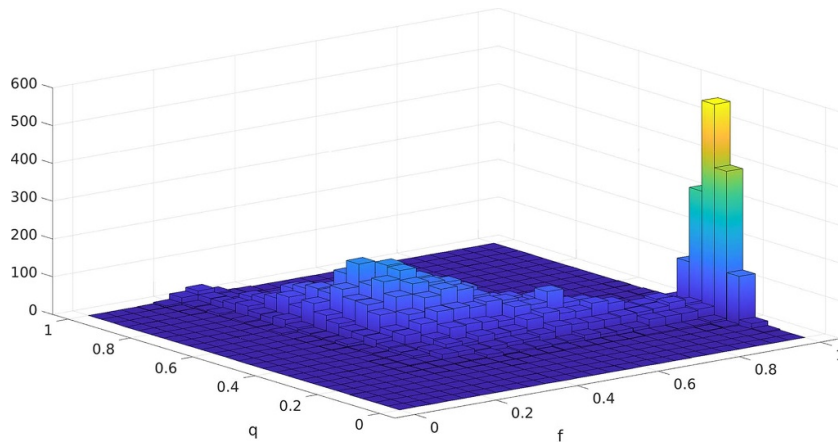


Figure 5. Histogram of the joint distribution of fitted values of the parameters (q, f) of an ensemble of 10 000 β -Hermite matrices with $f = 0.6$ and $\beta = 0.3$. The matrix sizes are $N = 1000$.

Now let us analyze in more detail a case of more practical interest with intermediate statistics between the two extremes. Let us think of a experimental spectrum with a 30% of missing levels and dynamics between chaos and regularity with $\beta = 0.6$, which we can represent here by one individual member of a β -ensemble generated with $f = 0.7$ and $\beta = 0.6$. We can perform a fit to the two-parameter formula to obtain a first estimation of the parameters q and f . But in view of the previous figures 5 and 4 it seems clear that the best estimation of the uncertainty of the parameters deals with a simulation of an ensemble of spectra of the same dimension of the experimental one to analyze the dispersion of values of the parameters more than with any error estimation from the fit. We represent the results for this ensemble, for $N = 1000$, in the left panel of figure 6 and we can see that the spread of values is wider than the one for higher

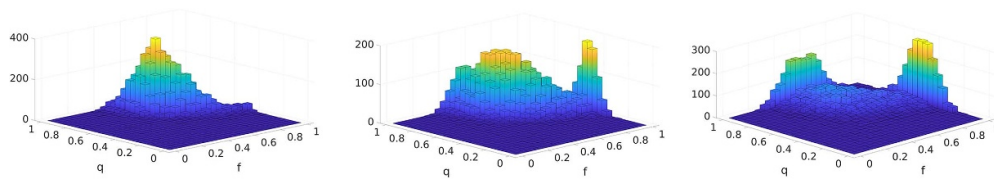


Figure 6. Histogram of the joint distribution of fitted values of the parameters (q, f) of ensembles of 10 000 β -Hermite matrices with $f = 0.7$ and $\beta = 0.6$. The matrix sizes are from left to right $N = 1000$, $N = 500$, and $N = 200$.

values shown in figure 4, and one could have a first estimation of the uncertainties of the parameters from the widths of this distribution. Moreover, for lower dimensions one obtains even wider distributions of the parameters, as shown in the central panel of figure 6 ($N = 500$) and the right panel of figure 6 ($N = 200$). That is, the probability of obtaining the correct values of the parameters from a single fit decreases, as expected, confirming that a safe estimation of the parameters and their errors cannot be obtained from a single fit but analyzing the distributions of several simulations of ensembles. In these two panels we can also observe the same effect as in figure 5 for the extreme of very uncorrelated spectra when the $P(s)$ can be well-fitted in some realizations of the ensemble with a combination of f and q such that one of them is close to unity and the other one is representative of the actual intermediate behavior of the NNSD. That is, the conclusion from the fit in these realizations would be that the intermediate statistics is due to just one reason, missing levels or intermediate dynamics.

Thus, in view of these examples, our conclusion and our recommendation when analyzing a spectrum of interest would be not only performing the fit but simulating several β -ensembles of the same dimension and observe the histograms of the distribution of the fitted results in order to try to obtain the most reliable estimation of the parameters and their uncertainties. The procedure we propose for finding an estimate of the errors in the fit if we have an experimental spectra with N_e levels is the following. We fit our formula and obtain a fitted value for f and q , f_e and q_e . We then should choose ensembles with different values of the dimension $N = N_e/f$ with an ensemble of values f around f_e and take out $N_e(1 - f)$ levels to have the correct final number of levels N_e . We should do that for different values of q around q_e . Then, to find a good estimation of the errors for this particular experiment, one should take the inverse distribution of the original values of f and q when the fitted result coincides with f_e and q_e . For lower dimensions, like the ones in the right panel of figure 6, one should take a result close to value of unity for any of the two parameters with caution and try to perform more simulations and checks to be sure of its reliability and make the best error estimation. To facilitate this task we have made available the codes and data used in the last two sections in a Zenodo repository [58]. Our codes use *MATLAB* and for the fitting we have used the standard Matlab Optimization subroutine *lsqcurvefit*.

Apart from intrinsic limitations from low values of the parameters or low dimension, here we only have simulated spectra and have assumed that no previous information on them is known, but in practical cases if one has some previous information about the type of dynamics of the system or the estimated fraction of missing levels, the limitations become less. In order to have the safest estimations of the uncertainties of the parameters we recommend to perform simulations of several ensembles in any case, but if the values can be restricted to some ranges from previous known information, much narrower distributions will be obtained with only one peak instead of two or three, even for low dimensions.

5. Conclusions

Quantum spectra with statistical properties intermediate between the GOE result of RMT and the Poisson uncorrelated spectra can be due to genuine intermediate properties between chaos and integrability or because there are missing levels destroying correlations in the level fluctuations induced by chaotic dynamics. But what happens when both things occur at the same time? Is it possible to distinguish the different causes of intermediate behavior and to quantify them?

In this work, we have tried to answer this question. We have developed a two-parameter formula for fitting short-range spectral correlations with one of the parameters based on the Brody distribution accounting for the chaoticity of the spectra, q , and the other for the fraction of observed levels, f . The formulas work perfectly well when comparing with RMT results of the β -Hermite ensemble. This proves the correctness of the formula for describing fractional level repulsion and missing levels in the spectra at the same time.

It is important to stress that any analysis of the spacing distribution requires a correct unfolding of the spectra. In general, lack of unfolding would tend to give results closer to Poisson for the $P(s)$ and systematic errors with lower values of both q and f will occur. Overunfolding the spectra (by including fluctuations in the average density of states) has the opposite effect giving results closer to GOE and systematic errors with higher values of both q and f . However, it is important to notice that the dependence on the unfolding is not as critical in short-range correlations, as the $P(s)$, than in long-range correlations, as the Σ_2 , Δ_3 or the power spectrum [44, 59, 60]. An advantage of calculations using the β -Hermite ensemble is that the correct unfolding is guaranteed as the average density of levels follows the semicircle law of RMT.

Another important issue of our derivation is the assumption of random sampling. For complex spectra, the position of levels and their widths are uncorrelated so the random sampling assumption is justified when the missing levels are due to sensitivity limitations of the measurement apparatus. However, this assumption cannot be justified for missing levels coming from resolution limitations of the experimental device when two close levels can be mistaken for one. It should be noted that due to level repulsion in chaotic spectra this problem is more relevant for integrable systems. An interesting extension of our work would be to study this case. As a starting point, equation (6) could be modified by taking into account the finite experimental resolution as a parameter. The easiest way to study this problem would be through numerical Monte Carlo simulations of the experimental spectra from RMT matrices. This extension is very relevant for studying experimental spectra but also very demanding and outside the scope of this work.

The theoretical interest of the formula is clear, as a first step to start to unravel the puzzle of the actual origin of intermediate behavior of fluctuations in a spectrum. Now, the next step would be to judge its applicability for actual experimental individual spectra. In this work we have shown some examples by simulating individual spectra from matrices of β -Hermite ensembles ($q = \beta$) and analyzed the accuracy and precision of the fit by computing the full distribution of the fitted results of the whole ensemble. From these distributions one can evaluate the probability to obtain reliable results depending on the values of β and f and the size of the spectrum, and estimate the uncertainties of the parameters. So this is our recommendation on how to proceed when using the two-parameter formula.

In view of the distributions of fitted results, we conclude that reliable results of both parameters at the same time can be obtained when the fraction of missing levels is not very large $f > 0.6$ and the total number of levels in the analyzed sequence is large enough $N > 1000$,

assuming no previous information on the spectrum of interest. Though, in each particular case one can have some previous information on the values of q and f of the spectrum, which can help to shorten their ranges and improve the results even for lower dimensions. The formula would certainly be applicable to estimate the fraction of missing levels if we know in advance what is the value of the repulsion parameter q or viceversa.

A similar procedure as the one we describe here for obtaining the two-parameter formula could be implemented for the Izrailev distribution, as it is also suitable to describe fractional level repulsion. We expect very similar results as the ones we present here. It should also be possible to start with the Berry–Robnik distribution that is able to fit the spectral statistics in the so-called ‘far-semiclassical regime’ which sometimes show a different behavior for small spacings, $P(0) \neq 0$ and its value is given by the fraction of the classical regular phase space. Thus, in these systems there is not fractional level repulsion but full level repulsion only for a fraction of levels of the spectrum. In order to obtain a suitable two-parameter formula for the $P(s)$ in this case one should consider the Berry–Robnik distribution as a starting point [41, 42].

We have taken advantage of the β -Hermite ensemble for testing our formula as it is a simple way to scan the transition chaos ($\beta = 1$)–regularity ($\beta = 0$) and its simple form results in an unrivalled speed-up of numerical simulations, but actual experimental spectra can be different from the β -ensemble, as the transition chaos-regularity is not universal. Thus, we make available this two-parameter formula and await results from experimental spectra to evaluate its real applicability and usefulness. The difficulties we have found in the estimation, specifically when the number of levels is limited, point also to the fact that, whenever possible, one should analyze together all the information available including short and long-range statistics and statistics of the widths in order to have more reliable conclusions regarding the chaoticity and the completeness of experimental spectra.

Data availability statement

The data that support the findings of this study are openly available at the following URL/DOI: <https://doi.org/10.5281/zenodo.8434033> [58].

Acknowledgments

This research has been supported by CSIC Research Platform on Quantum Technologies PTI-001. We also acknowledge financial support from European Union’s Horizon 2020 FET-Open project AVaQus Grant No. 899561, Project Nos. PGC2018-094180-B-I00, PID2019-106820RBC21, RTI2018-098868-B-I00, PID2021-126998OB-I00 and PID2022-136285NB funded by MCIN/AEI/10.13039/501100011033 and FEDER ‘A way of making Europe’.

ORCID iDs

María Hita-Pérez  <https://orcid.org/0000-0002-8330-6237>

Laura Muñoz  <https://orcid.org/0000-0002-1952-0482>

Rafael A Molina  <https://orcid.org/0000-0001-5728-0734>

References

- [1] Mehta M 2004 *Random Matrices* 3rd edn (Academic)
- [2] Stockmann H J 2006 *Quantum Chaos* (Cambridge University Press)
- [3] Haake F 2013 *Quantum Signatures of Chaos* (Springer)
- [4] Brody T A 1973 *Lett. Nuovo Cimento* **7** 482–4
- [5] Brody T A, Flores J, French J B, Mello P A, Pandey A and Wong S S M 1981 *Rev. Mod. Phys.* **53** 385–479
- [6] Izrailev F M 1989 *J. Phys. A: Math. Gen.* **22** 865–78
- [7] Izrailev F M 1990 *Phys. Rep.* **196** 299–392
- [8] Prosen T and Robnik M 1993 *J. Phys. A: Math. Gen.* **26** 2371–87
- [9] Blecken C, Chen Y and Muttalib K A 1994 *J. Phys. A: Math. Gen.* **27** L563
- [10] Molina R A, Gómez J M G and Retamosa J 2000 *Phys. Rev. C* **63** 014311
- [11] Abul-Magd A Y, Dietz B, Friedrich T and Richter A 2008 *Phys. Rev. E* **77** 046202
- [12] Weidenmüller H A and Mitchell G E 2009 *Rev. Mod. Phys.* **81** 539–89
- [13] Frisch A, Mark M, Aikawa K, Ferlaino F, Bohn J L, Makrides C, Petrov A and Kotochigova S 2014 *Nature* **507** 475–9
- [14] Muñoz L, Molina R A, Gómez J M G and Heusler A 2017 *Phys. Rev. C* **95** 014317
- [15] Evers F and Mirlin A D 2008 *Rev. Mod. Phys.* **80** 1355–417
- [16] Molina R A, Benito-Matías E, Somoza A D, Chen L and Zhao Y 2016 *Phys. Rev. E* **93** 022414
- [17] Benito-Matías E and Molina R A 2017 *Phys. Rev. B* **96** 174202
- [18] Šuntajs J, Prosen T and Vidmar L 2021 *Ann. Phys., NY* **435** 168469
- [19] Oganesyan V and Huse D A 2007 *Phys. Rev. B* **75** 155111
- [20] Huse D A, Nandkishore R and Oganesyan V 2014 *Phys. Rev. B* **90** 174202
- [21] Luitz D J, Laflorencie N and Alet F 2015 *Phys. Rev. B* **91** 081103
- [22] Serbyn M and Moore J E 2016 *Phys. Rev. B* **93** 041424
- [23] Corps A L, Molina R A and Relaño A 2021 *SciPost Phys.* **10** 107
- [24] Bohigas O, Giannoni M J and Schmit C 1984 *Phys. Rev. Lett.* **52** 1–4
- [25] Heusler S, Müller S, Altland A, Braun P and Haake F 2007 *Phys. Rev. Lett.* **98** 044103
- [26] Zimmermann T, Köppel H, Cederbaum L S, Persch G and Demtröder W 1988 *Phys. Rev. Lett.* **61** 3–6
- [27] Stöckmann H-J and Stein J 1990 *Phys. Rev. Lett.* **64** 2215–8
- [28] Dietz B, Heine A, Richter A, Bohigas O and Leboeuf P 2006 *Phys. Rev. E* **73** 035201
- [29] Berry M V and Tabor M 1977 *Proc. R. Soc. London A* **356** 375–94
- [30] Mur-Petit J and Molina R A 2015 *Phys. Rev. E* **92** 042906
- [31] Agvaanluvsan U, Mitchell G, Shriner J and Pato M 2003 *Nucl. Instrum. Methods Phys. Res. A* **498** 459–69
- [32] Bohigas O and Pato M P 2004 *Phys. Lett. B* **595** 171–6
- [33] Molina R A, Retamosa J, Muñoz L, Relaño A and Faleiro E 2007 *Phys. Lett. B* **644** 25–32
- [34] Mulhall D 2011 *Phys. Rev. C* **83** 054321
- [35] Casal I, Muñoz L and Molina R A 2021 *Eur. Phys. J. Plus* **136** 263
- [36] Mitchell G E, Bilpuch E G, Endt P M and Shriner J F 1988 *Phys. Rev. Lett.* **61** 1473–6
- [37] Białous M, Yunko V, Bauch S, Ławniczak M, Dietz B and Sirko L 2016 *Phys. Rev. Lett.* **117** 144101
- [38] Ławniczak M, Białous M, Yunko V, Bauch S and Sirko L 2018 *Phys. Rev. E* **98** 012206
- [39] Ławniczak M, Białous M, Yunko V, Bauch S, Dietz B and Sirko L 2017 *Acta Phys. Pol. A* **132** 1672–6
- [40] Che J, Lu J, Zhang X, Dietz B and Chai G 2021 *Phys. Rev. E* **103** 042212
- [41] Berry M V and Robnik M 1984 *J. Phys. A: Math. Gen.* **17** 2413–21
- [42] Prosen T and Robnik M 1994 *J. Phys. A: Math. Gen.* **27** 8059–77
- [43] Engel D, Main J and Wunner G 1998 *J. Phys. A: Math. Gen.* **31** 6965–74
- [44] Gómez J M G, Molina R A, Relaño A and Retamosa J 2002 *Phys. Rev. E* **66** 036209
- [45] Dumitriu I and Edelman A 2002 *J. Math. Phys.* **43** 5830–47
- [46] Faleiro E, Kuhl U, Molina R, Muñoz L, Relaño A and Retamosa J 2006 *Phys. Lett. A* **358** 251–5
- [47] Nogueira T N, Sartorelli J C, Pato M P and Ellegaard C 2008 *Phys. Rev. E* **78** 055201
- [48] Białous M, Dietz B and Sirko L 2021 *Phys. Rev. E* **103** 052204
- [49] Abul-Magd A Y and Simbel M H 2000 *Phys. Rev. E* **62** 4792
- [50] Dumitriu I and Edelman A 2006 *J. Math. Phys.* **47** 063302
- [51] Dyson F J 1962 *J. Math. Phys.* **3** 140

- [52] Calogero F 1969 *J. Math. Phys.* **10** 2191
- [53] Calogero F 1969 *J. Math. Phys.* **10** 2197
- [54] Sutherland B 1971 *Phys. Rev. A* **4** 2019
- [55] Sutherland B 1971 *J. Math. Phys.* **12** 246
- [56] Buijsman W, Cheianov V and Gritsev V 2019 *Phys. Rev. Lett.* **122** 180601
- [57] Le Caër G, Male C and Delannay R 2007 *Physica A* **383** 190–208
- [58] Hita-Pérez M, Muñoz L and Molina R A 2023 Code and data for “missing levels in intermediate spectra” (available at: <https://doi.org/10.5281/zenodo.8434033>)
- [59] Relaño A, Gómez J M G, Molina R A and Retamosa J 2002 *Phys. Rev. Lett.* **89** 244102
- [60] Gómez J, Kar K, Kota V, Molina R, Relaño A and Retamosa J 2011 *Phys. Rep.* **499** 103–226