



Scaling of the reduced energy spectrum of random matrix ensemble

Wen-Jia Rao^a, M. N. Chen

School of Science, Hangzhou Dianzi University, Hangzhou 310027, China

Received: 8 October 2020 / Accepted: 31 December 2020

© The Author(s), under exclusive licence to Società Italiana di Fisica and Springer-Verlag GmbH Germany, part of Springer Nature 2021

Abstract We study the reduced energy spectrum $\{E_i^{(n)}\}$, which is constructed by picking one level from every n levels of the original spectrum $\{E_i\}$, in a Gaussian ensemble of random matrix with Dyson index $\beta \in (0, \infty)$. It is shown the joint probability distribution of $\{E_i^{(n)}\}$ bears the same form as $\{E_i\}$ with a rescaled parameter $\gamma = \frac{n(n+1)}{2}\beta + n - 1$. Notably, the n th-order level spacing and gap ratio in $\{E_i\}$ become the lowest order ones in $\{E_i^{(n)}\}$, which explains their distributions found separately by recent studies in a consistent way. Our results also establish the higher-order spacing distributions in random matrix ensembles beyond GOE, GUE and GSE and reveal a hierarchy of structures hidden in the energy spectrum.

1 Introduction

Random matrix theory (RMT) is a powerful mathematical tool for studying complex quantum systems; it describes the universal properties of random matrix that are determined only by the system's symmetry while independent of microscopic details. For this reason, RMT has been applied to various fields ranging from disordered nuclei to isolated quantum many-body systems [1–3].

Among various statistical quantities of RMT, the most widely used ones are the distribution of nearest level spacings $\{s_i = E_{i+1} - E_i\}$ and gap ratios $\{r_i = s_{i+1}/s_i\}$. It is well established that in a chaotic system, the distribution of level spacing $P(s)$ will follow a Wigner–Dyson distribution [4, 5] (see Eq. (3) in Sect. 3), which reveals the level repulsion in a direct way. However, when accounting $P(s)$, an unfolding procedure is required to erase the model-dependent information about local density of states (DOS). In contrast, gap ratios distribution $P(r)$ is independent of DOS and requires no unfolding procedure [6, 7] and has found various applications especially in the context of many-body localization (MBL) [7–18].

Both the nearest level spacing and gap ratio account for the short-range level correlations. However, long-range correlations are also important, especially in the study of MBL transition phenomena. Actually, there're a number of RMT models accounting for the intermediate level statistics between Wigner–Dyson and Poisson ensembles [19–24], and some of them are suggested to describe the level statistics in MBL transition regime [25–29], all of which more or less describe the short-range level correlations well, while their difference can only be

^a e-mail: wjrao@hdu.edu.cn (corresponding author)

revealed when long-range correlations are concerned. Commonly, the long-range correlation in a random matrix ensemble is described by the number variance Σ^2 or the Dyson–Mehta Δ_3 statistics [5]; however, both of them are very sensitive to the concrete unfolding strategy which has already been a source of misleading signatures [30]. Instead, it's more direct and numerically easier to study the higher-order level spacings and gap ratios, as have been done in a number of recent works [28, 29, 31–40].

Formally, the n th-order level spacing and gap ratio are defined as $\{s_i^{(n)} = E_{i+n} - E_i\}$ and $\{r_i^{(n)} = s_{i+n}^{(n)}/s_i^{(n)}\}$, respectively. The higher-order gap ratios are first studied in Ref. [33], where the authors provide strong numerical evidences that $P(r^{(n)})$ in a random matrix ensemble with Dyson index β bears the same form as $P(r)$ with a rescaled parameter

$$\gamma = \frac{n(n+1)}{2}\beta + n - 1, \quad (1)$$

where $\beta = 1, 2, 4$ correspond to the Gaussian orthogonal ensemble (GOE), Gaussian unitary ensemble (GUE) and Gaussian symplectic ensemble (GSE), respectively. On the other hand, it is later proved in Ref. [39] that $P(s^{(n)})$ has the same form as Wigner–Dyson distribution with a rescaled parameter γ that is identical to the one in Eq. (1) for $\beta \in (0, \infty)$. This strongly hints a homogeneous relation between the higher-order level spacing and gap ratio, but an explanation is still lacking, which gives the first motivation of this work.

The second motivation comes from the recent works that encounter physical systems that go beyond the three standard ensembles with $\beta = 1, 2, 4$. For example, the $\beta = 3$ behavior has been found in a 2D lattice with non-Hermitian disorder [41], and the ensembles with non-integer β are suggested to describe the level statistics in the whole region along the MBL transition in 1D random spin system [27], while their efficiency in describing long-range spectral correlations is controversial [29]. Therefore, it's beneficial to have an expression for the higher-order spacing distributions in these ensembles, which will also be offered in this study.

In this work, we find that the key to link the higher-order level spacing and gap ratio is the reduced energy spectrum $\{E_i^{(n)} \equiv E_{in}\}$, i.e. the spectrum constructed by picking one level from every n levels in the original spectrum $\{E_i\}$. By this construction, the n th-order level spacing and gap ratio in $\{E_i\}$ become the lowest order ones in $\{E_i^{(n)}\}$. It will be verified that the joint probability distribution of $\{E_i^{(n)}\}$ (to leading order) bears the same form as $\{E_i\}$ with a rescaled parameter γ expressed in Eq. (1), which hence explains their distributions by virtue of Wigner surmise. Furthermore, this rescaling relation holds for general β beyond $\beta = 1, 2, 4$; therefore, the higher-order level spacing and gap ratio distributions in these ensembles can be obtained accordingly, which is thus a natural extension of Ref. [33].

This paper is organized as follows. In Sect. 2 we summarize the formulas regarding the higher-order level spacings and gap ratios, which motivates the construction of reduced energy spectrum. In Sect. 3 we focus on the cases that $\{E_i^{(n)}\}$ has two and three levels and provide compelling numerical evidence for the scaling of its joint probability distribution in ensembles with general β . In Sect. 4 we provide numerical simulations for the distributions of nearest level spacing and gap ratio in $\{E_i^{(n)}\}$ with a large number of levels and confirm their coincidence with the higher-order ones in $\{E_i\}$. In Sect. 5 we briefly discuss the higher-order reduced energy spectrums. Conclusion and discussion come in Sect. 6.

2 Motivating the reduced energy spectrum

The starting point to study the spectral statistics in the Gaussian ensemble of random matrix is the joint probability distribution (JPDT) of its energy levels, whose form is [4,5]

$$P(\beta, \{E_i\}) = C \prod_{i < j} |E_i - E_j|^\beta e^{-A \sum_{i=1}^N E_i^2}, \tag{2}$$

where the Dyson index $\beta \in (0, \infty)$ is a continuous parameter, C and A are coefficients correlated by the normalization condition $\int \prod_i dE_i P(\beta, \{E_i\}) = 1$. It's worth noting that only the subset $\beta = 1, 2, 4$ are physically invariant ensembles, that is, each matrix element is allowed to draw from a Gaussian distribution provided the matrix being invariant under orthogonal, unitary or symplectic transformations, while for other values of β , the JPDT stems from a special tridiagonal random matrix (see Eq. (21) in Sect. 4).

From the JPDT in Eq. (2), we can in principle calculate any statistical quantity we want, in particular the distribution of nearest level spacings and gap ratios. The general distributions for them in large dimension N are complicated, while for most practical purpose it is sufficient to adopt the so-called Wigner surmise that deals with smallest matrix that holds the quantity of interest. For example, to study nearest level spacing it's sufficient to consider a 2×2 matrix, which gives the celebrated Wigner–Dyson distribution [5]

$$P(\beta, s) = C(\beta) s^\beta e^{-A(\beta)s^2}, \tag{3}$$

where the coefficients $C(\beta)$, $A(\beta)$ are determined by the normalization conditions

$$\int_0^\infty P(\beta, s) ds = 1, \int_0^\infty s P(\beta, s) ds = 1. \tag{4}$$

It's easy to see $P(\beta, s \rightarrow 0) \sim s^\beta$; hence, β is the parameter that controls the strength of level repulsion.

As for the nearest gap ratios $\{r_i = s_{i+1}/s_i\}$, a Wigner-like surmise is applicable by studying 3×3 matrices [6], which gives

$$P(\beta, r) = \frac{1}{Z_\beta} \frac{(r + r^2)^\beta}{(1 + r + r^2)^{1+3\beta/2}} \tag{5}$$

where Z_β is the normalization factor determined by requiring $\int_0^\infty P(\beta, r) dr = 1$. It is crucial to note that the derivations for Eqs. (3) and (5) are purely mathematical, that is, applicable for arbitrary positive β .

For the higher-order level spacings $\{s_i^{(n)} = E_{i+n} - E_i\}$, its distribution is studied in Ref. [39] using a Wigner-like surmise that deals with $(n + 1) \times (n + 1)$ matrix, and the result shows they follow a generalized Wigner–Dyson distribution that bears the same form as Eq. (3) with the parameter β rescaled to γ as expressed in Eq. (1).

On the other hand, higher-order gap ratios come in two different ways, i.e. the ‘‘overlapping’’ way [31] and ‘‘non-overlapping’’ way [33]. In the former case, we are dealing with

$$\tilde{r}_i^{(n)} = \frac{E_{i+n+1} - E_{i+1}}{E_{i+n} - E_i} = \frac{s_{i+n} + s_{i+n-1} + \dots + s_{i+1}}{s_{i+n-1} + s_{i+n-2} + \dots + s_i}, \tag{6}$$

which is termed ‘‘overlapping’’ gap ratio since there are shared spacings between the numerator and denominator, while the n th-order ‘‘non-overlapping’’ gap ratio is defined as

$$r_i^{(n)} = \frac{E_{i+2n} - E_{i+n}}{E_{i+n} - E_i} = \frac{s_{i+2n-1} + s_{i+2n-2} + \dots + s_{i+n}}{s_{i+n-1} + s_{i+n-2} + \dots + s_i}. \tag{7}$$

Both these two generalizations reduce to the nearest gap ratio when $n = 1$, but they are very different when studying their distributions using Wigner surmise: for non-overlapping ratio $r^{(n)}$, the smallest matrix dimension is $(2n + 1) \times (2n + 1)$, while it is $(n + 2) \times (n + 2)$ for overlapping ratios. Naively, it's expected that $P(\tilde{r}^{(n)})$ is more complicated due to the overlapping spacings. Indeed, the analytical form of $P(\tilde{r}^{(2)})$ was worked out in Ref. [31] and the result is quite involving. For the non-overlapping ratios, Ref. [33] provides compelling numerical evidence that in cases with $\beta = 1, 2, 4$, $P(r^{(n)})$ bears the same form as $P(r)$ with the same rescaling parameter as higher-order level spacing, that is, Eq. (1).

In summary, we have

$$P(\beta, s^{(n)}) = P(\gamma, s), \tag{8}$$

$$P(\beta, r^{(n)}) = P(\gamma, r), \tag{9}$$

where γ is expressed in Eq. (1). For the rest of this paper, ‘‘gap ratio’’ will always refer to the non-overlapping one.

The identical rescaling behavior for higher level spacing and gap ratio hints they may be attributed to one single reason, which is found to be the reduced energy spectrum $\{E_i^{(n)}\}$. Formally, a reduced energy spectrum $\{E_i^{(n)}\}$ is constructed by picking one level from every n levels in the original spectrum $\{E_i\}$, which is mathematically achieved by tracing out every $n - 1$ levels in between. This construction is very similar to that of the reduced density matrix there we trace out the degrees of freedom in a subsystem; hence, $\{E_i^{(n)}\}$ is named ‘‘reduced energy spectrum’’. By this construction, the n th-order level spacing and gap ratio in $\{E_i\}$ are mapped to the lowest order counterparts in $\{E_i^{(n)}\}$. It's then natural to conjecture that $\{E_i\}$ and $\{E_i^{(n)}\}$ (to leading order) bear the same form for their probability distributions, with the Dyson index β for the former rescaled to γ for the latter according to Eq. (1). Therefore, by applying Wigner surmise to $\{E_i^{(n)}\}$, the scaling behaviors in Eqs. (8) and (9) can be explained simultaneously. This is the main task of current work.

Before proceeding, we want to mention the relatively trivial case of Poisson ensemble. The reduced energy spectrum in Poisson ensemble has been studied in Ref. [42] (which is named ‘‘Daisy model’’ by the authors), where n th-order level spacing is shown to follow the generalized semi-Poisson distribution

$$P(s^{(n)} = s) = \frac{n^n}{(n - 1)!} s^{n-1} e^{-ns}, \tag{10}$$

which reduces to the conventional Poisson distribution $P(s) = \exp(-s)$ when $n = 1$. For the n th-order gap ratios, we can derive their distribution from the results in Ref. [42], that is

$$P(r^{(n)} = r) = \frac{r^{n-1}}{(1 + r)^{2n}}, \tag{11}$$

which reduces to the one given in Ref. [6] when $n = 1$. The formulas in Eqs. (8), (9), (10) and (11) will be used for later numerical simulations.

3 Scaling of $P(\{E_i^{(n)}\})$

By the construction of reduced energy spectrum $\{E_i^{(n)}\}$, its formal joint probability distribution is

$$P(\{E_i^{(n)}\}) = \prod_i \int_{E_{in}}^{E_{(i+1)n}} \prod_{j=in+1}^{(i+1)n-1} dE_j P(\beta, \{E_i\}). \tag{12}$$

For reasons described in the previous section, we conjecture $P(\{E_i^{(n)}\})$ (to leading order) bear the same form as $P(\{E_i\})$ with the rescaled parameter γ as in Eq. (1), that is,

$$P(\{E_i^{(n)}\}) \sim \prod_{i < j} |E_i^{(n)} - E_j^{(n)}|^\gamma e^{-A' \sum_{i=1}^{N/n} (E_i^{(n)})^2}. \tag{13}$$

An analytical derivation from Eqs. (12) to (13) for arbitrary matrix dimension N is mathematically challenging, only the special case with $\beta = 2/k$ (k being positive integer) has been proven rigorously in Ref. [43]. It is not our purpose to give a general proof; instead, to explain the behaviors of the $P(s^{(n)})$ and $P(r^{(n)})$ in Eqs. (8) and (9), we only need to verify Eq. (13) in the sense of Wigner surmise, that is, in the cases that $\{E_i^{(n)}\}$ has only two and three levels, for which we will provide strong numerical evidence in the following.

First of all, the constant A' is not important since it is only a decay rate parameter, whose value can be tuned by global rescaling of the energy levels without affecting the distribution of level spacing or gap ratio. Therefore, we will focus on the parameter γ that controls the strength of level repulsion.

The ensembles with general positive β can be divided into four typical categories: (i) $\beta = 1, 2, 4$, corresponding to the three standard Gaussian ensembles, which are of most physical interest; (ii) β is an integer that goes beyond the three standard ensembles, for which we choose $\beta = 3$; (iii) β is a fraction, for which we choose $\beta = 1/3$; (iv) β is an irrational value, for which we choose $\beta = (\sqrt{5} - 1)/2$ (the golden ratio). We will verify the rescaling relation Eq. (13) in these cases.

We start with the case that $\{E_i^{(n)}\}$ has only two levels, where the rescaling in Eq. (13) becomes

$$\begin{aligned} I(E_0, E_n) &= \int_{E_0}^{E_n} \prod_{i=1}^{n-1} dE_i P(\beta, \{E_i\}) \\ &\sim |E_0 - E_n|^\gamma e^{-A'(E_0^2 + E_n^2)}. \end{aligned} \tag{14}$$

Denote $E_0 = R \cos \theta$ and $E_n = R \sin \theta$, and keeping R constant, we reach to

$$\log I(\theta) = \gamma \log |\cos \theta - \sin \theta| + \text{const.} \tag{15}$$

Without loss of generality, we take $A = 1$ and $R = 1$ and randomly generate 200 sets of θ in the range $[0, 2\pi)$. We then numerically calculate $\log I(\theta)$ and $\log |\cos \theta - \sin \theta|$, and collect the results for $n = 2, 3, 4$, which are presented in Fig. 1. As can be seen, the $\log I(\theta)$ and $\log |\cos \theta - \sin \theta|$ shows a perfect linear dependence in all cases, with the fitted values of γ quite close to the expected ones in Eq. (1), as summarized in Table 1.

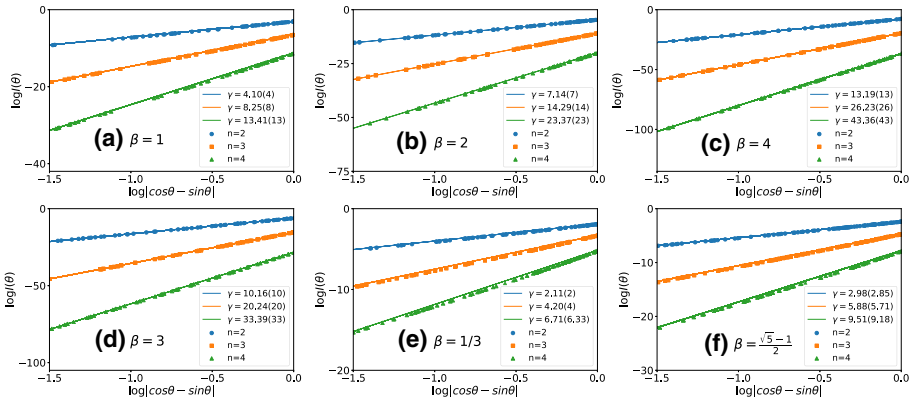


Fig. 1 The fitting of Eq. (15) for $\beta = 1, 2, 3, 4, \frac{1}{3}, \frac{\sqrt{5}-1}{2}$ with $n = 2, 3, 4$. The fitted slopes are shown in the figure legends, where the numbers in the bracket are the expected values according to Eq. (1)

Table 1 The values of γ with different β and order n , where γ_e is the expected value according to Eq. (1). $\gamma_{num.}^{(1)}$ and $\gamma_{num.}^{(2)}$ are the values fitted from Eqs. (15) to (18); their relative errors with respect to γ_e are denoted by error⁽¹⁾ and error⁽²⁾, respectively

β	n	γ_e	$\gamma_{num.}^{(1)}$	Error ⁽¹⁾ (%)	$\gamma_{num.}^{(2)}$	Error ⁽²⁾ (%)
1	2	4	4.10	2.50	4.03	0.75
	3	8	8.25	3.13	8.27	3.38
	4	13	13.41	3.15	13.49	3.77
2	2	7	7.14	2.00	7.20	2.86
	3	14	14.29	2.07	15.10	7.86
	4	23	23.37	1.61	24.75	7.29
4	2	13	13.19	1.46	13.42	3.23
	3	26	26.23	0.88	27.66	6.38
	4	43	43.36	0.84	46.48	8.09
3	2	10	10.16	1.60	10.42	4.20
	3	20	20.24	1.20	21.35	6.75
	4	33	33.39	1.18	35.46	7.45
1/3	2	2	2.11	5.5	1.91	-4.5
	3	4	4.20	5.0	3.90	-2.5
	4	19/3	6.71	6.0	6.25	-1.3
$\frac{\sqrt{5}-1}{2}$	2	$\frac{3\sqrt{5}-1}{2}$	2.98	4.4	2.82	-1.2
	3	$3\sqrt{5}-1$	5.88	3.0	5.75	0.7
	4	$5\sqrt{5}-2$	9.51	3.6	9.43	2.7

Next, we consider the case that $\{E_i^{(n)}\}$ has three levels; now the rescaling in Eq. (13) becomes

$$Q(E_0, E_n, E_{2n})$$

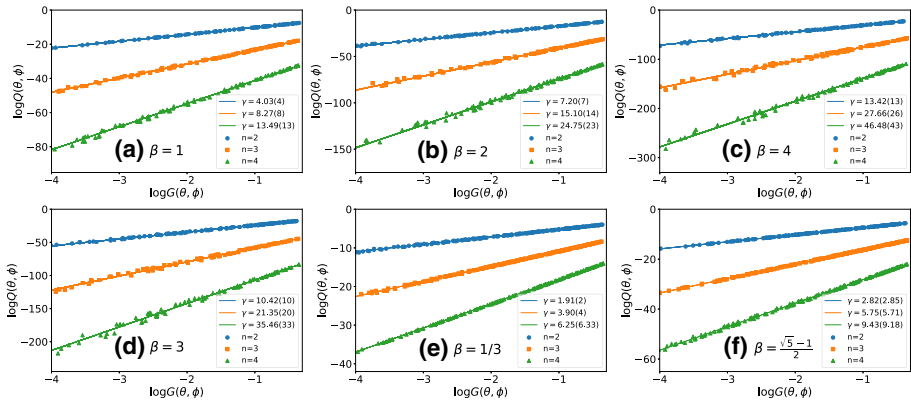


Fig. 2 The fitting of Eq. (18) for $\beta = 1, 2, 3, 4, \frac{1}{3}, \frac{\sqrt{5}-1}{2}$ with $n = 2, 3, 4$. The fitted γ values are shown in the figure legends, with the expected values according to Eq. (1) in the brackets

$$\begin{aligned} &\equiv \int_{E_0}^{E_n} \prod_{i=1}^{n-1} dE_i \int_{E_n}^{E_{2n}} \prod_{j=n+1}^{2n-1} dE_j P(\beta, \{E_i\}) \\ &\sim |E_0 - E_n|^\gamma |E_n - E_{2n}|^\gamma |E_0 - E_{2n}|^\gamma e^{-A'(E_0^2 + E_n^2 + E_{2n}^2)}. \end{aligned} \tag{16}$$

With the transformation to spherical coordinates

$$E_0 = R \sin \theta \cos \varphi, E_n = R \sin \theta \sin \varphi, E_{2n} = R \cos \theta \tag{17}$$

and keeping R constant, we can reach to

$$\log Q(R, \theta, \varphi) = \gamma \log G(\theta, \varphi) + \text{const.} \tag{18}$$

where

$$\begin{aligned} G(\theta, \varphi) = &|(\sin \theta \cos \varphi - \sin \theta \sin \varphi) \\ &\times (\sin \theta \cos \varphi - \cos \theta) \\ &\times (\sin \theta \sin \varphi - \cos \theta)|. \end{aligned} \tag{19}$$

We then perform numerical checks, where we fix $R = 1$ and $A = 1$ as before. We randomly generate 200 pairs of (θ, φ) and numerically determine $\log Q(\theta, \varphi) \equiv \log Q(1, \theta, \varphi)$ and $\log G(\theta, \varphi)$; the results are displayed in Fig. 2. As can be seen, the linear dependence between $\log Q(\theta, \varphi)$ and $\log G(\theta, \varphi)$ is still perfect in all cases, and the fitted values of γ are close to the expected ones in Eq. (1).

For convenience, we collected the theoretical and numerical values of γ in Table 1, where γ_c refers to the expected value according to Eq. (1), $\gamma_{\text{num.}}^{(1)}$ and $\gamma_{\text{num.}}^{(2)}$ are the numerical values fitted from Eqs. (15) and (18), and $\text{error}^{(1)}$ and $\text{error}^{(2)}$ are their relative deviations from γ_c respectively. In general, the numerical errors increase with larger β , but they're controlled within a satisfactory level in all cases.

Up to now, we have verified the scaling behavior for the probability distribution of $\{E_i^{(n)}\}$ with two and three levels, the cases with more levels can be verified with the same method, but it will become more and more tedious when the number of integrals increases. Nevertheless, the present results are sufficient to justify the scaling behaviors of higher-order level spacing/gap ratio by applying Wigner surmise to $\{E_i^{(n)}\}$. More importantly, we have verified

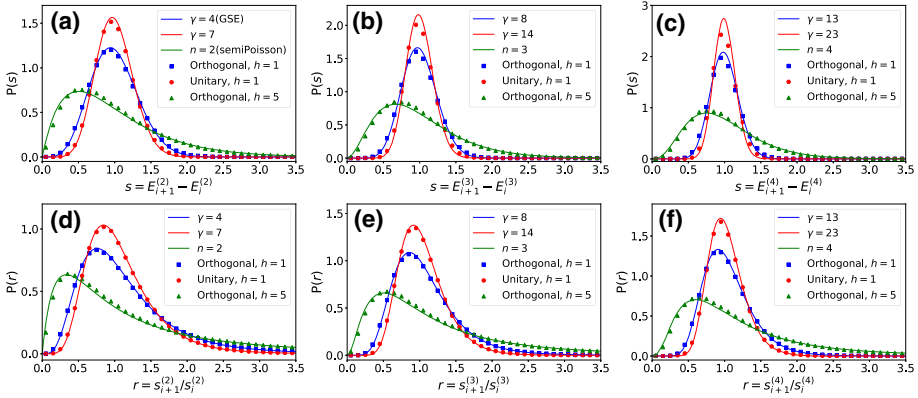


Fig. 3 Distribution of nearest level spacing and gap ratio in the reduced energy spectrum $\{E_i^{(n)}\}$ of model Eq. (20) in a $L = 12$ chain with $n = 2$ **a** and **d**, $n = 3$ **b** and **e** and $n = 4$ **c** and **f**. The data from $h = 1$ in the orthogonal (unitary) model represent GOE (GUE), and those from $h = 5$ in orthogonal model represent Poisson, respectively. The reference curves correspond to the higher-order spacing distributions in $\{E_i\}$ according to Eqs. (8), (9) for GOE and GUE, and Eqs. (10), (11) for Poisson; the parameter γ for the former is calculated by Eq. (1). The perfect fittings in all cases confirm the coincidence between higher-order spacing distributions in $\{E_i\}$ and the lowest order counterparts in $\{E_i^{(n)}\}$

the scaling behavior to hold for general β s that go beyond GOE, GUE and GSE ($\beta = 1, 2, 4$), even when β is non-integer or irrational. This indicates the distribution for higher order level spacing and gap ratio in Eqs. (8) and (9) also hold for general β ensembles, for which we will present numerical evidences in the next section.

4 Numerical simulations

In this section, we numerically check the distribution of nearest level spacing/gap ratio in $\{E_i^{(n)}\}$ with many levels, and show they indeed coincide with the n th-order counterparts in $\{E_i\}$. Before that, a technical issue needs to be pointed out, that is, the nearest level spacings in $\{E_i^{(n)}\}$ is actually $\{E_{i+1}^{(n)} - E_i^{(n)} = E_{(i+1)n} - E_{in}\}$, whose total number is $\lfloor \frac{N}{n} \rfloor - 1$, while the n th-order level spacing in original energy spectrum $\{E_i\}$ are $\{E_{i+n} - E_i\}$ with total number $N - n$; therefore, the mapping does not strictly hold, and the same thing happens to gap ratios. However, since the distribution is extracted from a large number of level spacings (gap ratios), it's natural to suspect the difference is negligible when the number of samples and matrix dimension are large, which we will soon justify.

For GOE, GUE and Poisson ensemble, we perform simulations from a real physical system, that is, the 1D Heisenberg chain with random external fields, which is the canonical model to study many-body localization [44]. The Hamiltonian reads

$$H = J \sum_{i=1}^L \mathbf{S}_i \cdot \mathbf{S}_{i+1} + \sum_{i=1}^L \sum_{\alpha=x,y,z} h^\alpha \varepsilon_i^\alpha S_i^\alpha, \tag{20}$$

where \mathbf{S}_i is spin-1/2 operator. The anti-ferromagnetic coupling strength J is set to be unity, and ε_i^α s are random numbers within range $[-1, 1]$. The h^α is referred as randomness strength. We consider two sets of h^α : (i) $h^x = h^z = h \neq 0$ and $h^y = 0$, the model is orthogonal

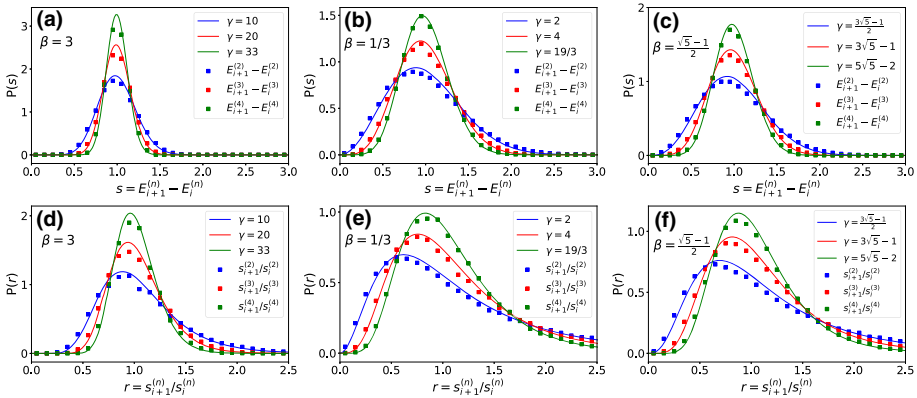


Fig. 4 Distribution of nearest level spacing and gap ratio in the reduced energy spectrum $\{E_i^{(n)}\}$ with $n = 2, 3, 4$ from ensembles with $\beta = 3$ **a, d**, $\frac{1}{3}$ **b, e** and $\frac{\sqrt{5}-1}{2}$ **c, f**, and the reference curves with index γ correspond to the ones in Eqs. (8) and (9) for level spacing and gap ratio, respectively. These fittings confirm the higher-order spacing distributions in general β ensembles

5 Higher-order reduced energy spectrums

Results in Sect. 3 verified the scaling of $\{E_i^{(n)}\}$ in cases with two/three levels, which is sufficient for applying Wigner surmise to $\{E_i^{(n)}\}$ regarding level spacing/gap ratio, and the numerical simulations in Sect. 4 provide strong support for the cases where $\{E_i^{(n)}\}$ has more levels. Actually, we can continue to construct the 2nd-order reduced energy spectrums, i.e. the “reduced energy spectrum of reduced energy spectrum”, whose nearest level spacing/gap ratio will correspond to the higher-order ones in $\{E_i^{(n)}\}$, and will rescale in a similar manner. Denote the m th-order level spacing and gap ratio in $\{E_i^{(n)}\}$ as $s_i^{(n,m)}$ and $r_i^{(n,m)}$, that is,

$$s_i^{(n,m)} = E_{i+m}^{(n)} - E_i^{(n)}, r_i^{(n,m)} = \frac{E_{i+2m}^{(n)} - E_{i+m}^{(n)}}{E_{i+m}^{(n)} - E_i^{(n)}}, \tag{24}$$

it’s straightforward to write down their expected probability distributions

$$P(\beta, s^{(n,m)}) = P(\delta, s), P(\beta, r^{(n,m)}) = P(\delta, r)$$

$$\delta = \frac{m(m+1)}{2}\gamma + m - 1, \gamma = \frac{n(n+1)}{2}\beta + n - 1. \tag{25}$$

The same procedure can continue for higher-order reduced energy spectrums.

Of course, such a construction is artificial; meanwhile, it reveals a hierarchy of energy spectrums can emerge from single spectrum, which (to lowest order) bear the same form of probability distributions. Moreover, by taking a closer look at the scaling expression $\gamma = \frac{n(n+1)}{2}\beta + n - 1$, we immediately recognize an infinite number of coincident relations between $\{E_i^{(n)}\}$ from different ensembles. For example, $\{E_i^{(2)}\}$ in $\beta = 1$ (GOE) has the same structure as $\{E_i\}$ in $\beta = 4$ (GSE)—a result known before [43,47], and both of them coincide with $\{E_i^{(3)}\}$ in $\beta = \frac{1}{3}$. Actually, it’s easy to verify that for $\gamma \in (k, k+1]$ there exists k different sets of (β, n) that have equal γ , and their lower-order level statistics are expected to be identical.

6 Conclusion and discussion

We studied the reduced energy spectrum $\{E_i^{(n)}\}$, constructed by picking one level from every n levels in original spectrum $\{E_i\}$. It is verified the distribution of $\{E_i^{(n)}\}$ (to leading order) bears the same form as $\{E_i\}$, with the Dyson index rescaled from β to $\gamma = \frac{n(n+1)}{2}\beta + n - 1$. It's then demonstrated the nearest level spacing and gap ratio in $\{E_i^{(n)}\}$ corresponds to the n th-order ones in $\{E_i\}$, which explains the distributions of the latter found recently in Ref. [39] and Ref. [33] simultaneously.

Moreover, we find the rescaling of reduced energy spectrum holds for Gaussian ensembles that go beyond the standard GOE, GUE and GSE and establish the distributions of higher-order level spacings and gap ratios in these ensembles. We also confirmed such correspondences in the Poisson ensemble and discovered the distribution of n th-order gap ratios, as expressed in Eq. (11).

It's noted the reduced energy spectrum has been studied for the Poisson ensemble [42] and some Gaussian ensembles, that is, the special ones with $\beta = 2/k$ (k being positive integer) [43], which contains the well-known coincidence between $\{E_i^{(2)}\}$ in GOE and $\{E_i\}$ in GSE. Our work is thus a natural extension of these studies.

The significance of our work is three-folded. First, we explained the distributions of higher-order level spacings and gap ratios—both of which are widely used in the study of MBL and whose distributions are found separately in recent studies—by a single common mechanism: the reduced energy spectrum. Second, we generalized the higher-order spacing distributions in Eqs. (8) and (9) to general β ensembles, which may be beneficial for studying systems that go beyond the standard Gaussian ensembles [27, 41, 48–50]. Third, our results reveal a rich set of structures hidden in the energy spectrum, by constructing the reduced energy spectrums.

Last but not least, in our numerical simulations in Sect. 4, we were employing the modelling matrix of general β ensemble as expressed in Eq. (21). It's thus natural and interesting to ask whether this “parent matrix” corresponds to a real quantum system, and what's the property of such a system if it does. Given the physically relevant GOE, GUE and GSE are incorporated in Eq. (21), we conjecture such a “parent Hamiltonian” does exist, whose construction is left for a future study.

Acknowledgements This work is supported by the National Natural Science Foundation of China through Grant No. 11904069 and No. 11847005 and No. 11804070.

Data Availability Statement This manuscript has associated data in a data repository [Authors comment: The data that support the figures within this paper are available from the corresponding author upon reasonable request.]

References

1. C.E. Porter, *Statistical Theories of Spectra: Fluctuations* (Academic Press, New York, 1965)
2. T.A. Brody et al., *Rev. Mod.* **53**, 385 (1981)
3. T. Guhr, A. Müller-Groeling, H.A. Weidenmuller, *Phys. Rep.* **299**, 189 (1998)
4. M.L. Mehta, *Random Matrix Theory* (Springer, New York, 1990)
5. F. Haake, *Quantum Signatures of Chaos* (Springer, 2001)
6. Y.Y. Atas, E. Bogomolny, O. Giraud, G. Roux, *Phys. Rev. Lett.* **110**, 084101 (2013)
7. V. Oganessian, D.A. Huse, *Phys. Rev. B* **75**, 155111 (2007)
8. Y. Avishai, J. Richert, R. Berkovits, *Phys. Rev. B* **66**, 052416 (2002)
9. N. Regnault, R. Nandkishore, *Phys. Rev. B* **93**, 104203 (2016)

10. S.D. Geraedts, R. Nandkishore, N. Regnault, *Phys. Rev. B* **93**, 174202 (2016)
11. V. Oganesyan, A. Pal, D.A. Huse, *Phys. Rev. B* **80**, 115104 (2009)
12. A. Pal, D.A. Huse, *Phys. Rev. B* **82**, 174411 (2010)
13. S. Iyer, V. Oganesyan, G. Refael, D.A. Huse, *Phys. Rev. B* **87**, 134202 (2013)
14. X. Li, S. Ganeshan, J.H. Pixley, S. Das Sarma, *Phys. Rev. Lett.* **115**, 186601 (2015)
15. D.J. Luitz, N. Laflorencie, F. Alet, *Phys. Rev. B* **91**, 081103(R) (2015)
16. W.-J. Rao, *J. Phys.: Condens. Matter* **30**, 395902 (2018)
17. A. Sarkar, M. Kothiyal, S. Kumar, *Phys. Rev. E* **101**, 012216 (2020)
18. Á.L. Corps, A. Relaño, *Phys. Rev. E* **101**, 022222 (2020)
19. G. Date, S.R. Jain, M.V.N. Murthy, *Phys. Rev. E* **51**, 198 (1995)
20. B. Grémaud, S.R. Jain, *J. Phys. A* **31**, L637 (1998)
21. G. Auberson, S.R. Jain, A. Khare, *J. Phys. A* **34**, 695 (2001)
22. E.B. Bogomolny, U. Gerland, C. Schmit, *Eur. Phys. J. B* **19**, 121 (2001)
23. M.L. Ndawana, V.E. Kravtsov, *J. Phys. A* **36**, 3639 (2003)
24. A.D. Mirlin, Y.V. Fyodorov, F.-M. Dittes, J. Quezada, T.H. Seligman, *Phys. Rev. E* **54**, 3221 (1996)
25. P. Shukla, *New J. Phys.* **18**, 021004 (2016)
26. M. Serbyn, J.E. Moore, *Phys. Rev. B* **93**, 041424(R) (2016)
27. W. Buijsman, V. Cheianov, V. Gritsev, *Phys. Rev. Lett.* **122**, 180601 (2019)
28. P. Sierant, J. Zakrzewski, *Phys. Rev. B* **99**, 104205 (2019)
29. P. Sierant, J. Zakrzewski, *Phys. Rev. B* **101**, 104201 (2020)
30. J.M.G. Gomez, R.A. Molina, A. Relano, J. Retamosa, *Phys. Rev. E* **66**, 036209 (2002)
31. Y.Y. Atas, E. Bogomolny, O. Giraud, P. Vivo, E. Vivo, *J. Phys. A: Math. Theor.* **46**, 355204 (2013)
32. S.H. Tekur, S. Kumar, M.S. Santhanam, *Phys. Rev. E* **97**, 062212 (2018)
33. S.H. Tekur, U.T. Bhosale, M.S. Santhanam, *Phys. Rev. B* **98**, 104305 (2018)
34. U.T. Bhosale, S.H. Tekur, M.S. Santhanam, *Phys. Rev. E* **98**, 052133 (2018)
35. P. Rao, M. Vyas, and N. D. Chavda, [arXiv:1912.05664v1](https://arxiv.org/abs/1912.05664v1)
36. A.Y. Abul-Magd, M.H. Simbel, *Phys. Rev. E* **60**, 5371 (1999)
37. M.M. Duras, K. Sokalski, *Phys. Rev. E* **54**, 3142 (1996)
38. R. Kausar, W.-J. Rao, X. Wan, *J. Phys.: Condens. Matter* **32**, 415605 (2020)
39. W.-J. Rao, *Phys. Rev. B* **102**, 054202 (2020)
40. L. Gong, K. Ma, *Phys. Lett. A* **384**, 126298 (2020)
41. A.F. Tzortzakakis, K.G. Makris, E.N. Economou, *Phys. Rev. B* **101**, 014202 (2020)
42. H. Hernández-Saldaña, J. Flores, T.H. Seligman, *Phys. Rev. E* **60**, 449 (1999)
43. P.J. Forrester, *Comm. Math. Phys.* **285**, 653 (2009)
44. F. Alet, N. Laflorencie, *C. R. Physique* **19**, 498–525 (2018)
45. I. Dumitriu, A. Edelman, *J. Math. Phys. (N.Y.)* **43**, 5830 (2002)
46. For standard GOE, GUE and GSE, we can model either by their physical construction or by using Eq. (21) with $\beta = 1, 2, 4$ to get the energy spectrums, although the DOS might be different. We have verified that these two constructions work almost equally well when studying spacing distributions, even when the matrix is as small as 200×200
47. M.L. Mehta, F.J. Dyson, *J. Math. Phys.* **4**, 701–712 (1963). <https://doi.org/10.1063/1.1704008>
48. L. Sa, P. Ribeiro, T. Prosen, *Phys. Rev. X* **10**, 021019 (2020)
49. Z. Ahmed, S.R. Jain, *Phys. Rev. E* **67**, 045106(R) (2003)
50. S.R. Jain, S.C.L. Srivastava, *Phys. Rev. E* **78**, 036213 (2008)