Higher-order level spacings in random matrix theory based on Wigner's conjecture

Wen-Jia Rao *

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

(Received 18 May 2020; revised 30 June 2020; accepted 21 July 2020; published 4 August 2020)

The distribution of higher-order level spacings, i.e., the distribution of $\{s_i^{(n)} = E_{i+n} - E_i\}$ with $n \geqslant 1$ is derived analytically using a Wigner-like surmise for Gaussian ensembles of random matrix as well as Poisson ensemble. It is found $s^{(n)}$ in Gaussian ensembles follows a generalized Wigner-Dyson distribution with rescaled parameter $\alpha = \nu C_{n+1}^2 + n - 1$, whereas that in the Poisson ensemble follows a generalized semi-Poisson distribution with index n. Numerical evidences are provided through simulations of random spin systems as well as nontrivial zeros of the Riemann ζ function. The higher-order generalizations of gap ratios are also discussed.

DOI: 10.1103/PhysRevB.102.054202

I. INTRODUCTION

Random matrix theory (RMT) was introduced half a century ago when dealing with complex nuclei [1] and since then has found various applications in fields ranging from quantum chaos to isolated many-body systems [2,3]. This roots in the fact that RMT describes universal properties of a random matrix that depends only on its symmetry whereas independent of microscopic details. Specifically, the system with time-reversal invariance is represented by a matrix that belongs to the Gaussian orthogonal ensemble (GOE); the system with spin rotational invariance whereas breaks time-reversal symmetry belongs to the Gaussian unitary ensemble (GUE); whereas Gaussian symplectic ensemble (GSE) represents systems with time-reversal symmetry but breaks spin rotational symmetry.

Among various statistical quantities, the most widely used one is the distribution of nearest level spacings $\{s_i = E_{i+1} - E_i\}$, i.e., the gaps between adjacent energy levels, which measures the strength of level repulsion. The exact expression for $P(s)$ can be derived analytically for a random matrix with a large dimension, which is cumbersome [4,5]. Instead, for most practical purposes it is sufficient to employ the so-called Wigner surmise [6] that deals with a 2×2 matrix (this will be reviewed in Sec. II), the out-coming result for $P(s)$ has a neat expression that contains a polynomial part accounting for level repulsion and an Gaussian decaying part [see Eq. (6)].

Different models may and usually do have different density of states (DOS), hence, to compare the universal behavior of level spacings, an unfolding procedure is required to erase the model-dependent information of DOS. To overcome this obstacle, Oganesyan and Huse [7] proposed a new quantity to study the level statistics, i.e., the ratio between adjacent gaps $\{r_i = s_{i+1}/s_i\}$, whose distribution $P(r)$ is later analytically derived by Atas *et al.* [8]. The gap ratio is independent of local DOS and requires no unfolding procedure (provided the DOS does not vary in the scale of the spacings involved), hence, has found various applications especially in the context of many-body localization (MBL) [9–18] .

Both the nearest level spacing and gap ratio account for the short-range level correlations. However, long-range correlations are also important especially when studying the MBL transition phenomena. Indeed, there are several effective models describing the level distribution at the MBL transition region. For example, the Rosenzweig-Porter model [19], mean-field plasma model [20], short-range plasma model (SRPM) [21], and its generalization—so-called weighed SRPM [22], Gaussian β ensemble [23], and the generalized $\beta - h$ model [24]. All of these models more or less describe the short-range level correlations in the MBL transition region well, and their difference can only be revealed when long-range correlations are concerned. For a comparison of these models in describing the MBL transition point, see Ref. [22].

Commonly, the long-range correlations in a random matrix can be 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 and have already been a source of misleading signatures [25]. Instead, it is more direct and numerically easier to study the higher-order level spacings and gap ratios. There are existing works that generalize the level spacing and gap ratios to higher order as well as their applications in studying MBL transitions [22,24,26–32]. However, most of these works are numerical or phenomenological, and an analytical derivation for the distribution of level spacing/gap ratio is still lacking. Given the importance of higher-order level correlations, it is desirable to have an analytical formula for them, and it is then the purpose of this paper to fill in this gap.

In this paper, by using a Wigner-like surmise, we succeeded in obtaining an analytical expression for the distribution of higher-order spacing $\{s_i^{(n)} = E_{i+n} - E_i\}$ in all the Gaussian ensembles of RMT as well as the Poisson ensemble. The results show the distribution of $s_i^{(n)}$ in the former class follows a generalized Wigner-Dyson distribution with a rescaled parameter; whereas, in the Poisson ensemble, it follows a generalized semi-Poisson distribution with index n. Interestingly, the rescaling behavior of higher-order level

*wjrao@hdu.edu.cn

spacing is identical to that of the high-order gap ratio found numerically in Ref. [28], for which we will provide a heuristic explanation.

This paper is organized as follows. In Sec. II, we review the Wigner surmise for deriving the distribution of nearest level spacings and present numerical data to validate this surmise. In Sec. III A, we present the analytical derivation for higher-order level spacings using a Wigner-like surmise, and numerical fittings are given in Sec. III B . In Sec. IV, we discuss the generalization of gap ratios to higher order. Conclusion and discussion come in Sec. V.

II. NEAREST LEVEL SPACINGS

We begin with the discussion about nearest level spacings, our starting point is the probability distribution of energy levels $P(\{E_i\})$ in three Gaussian ensembles, whose expression can be found in any textbook on RMT (e.g., Ref. [5]),

$$P(\{E_i\}) \propto \prod_{i<j} |E_i - E_j|^{\nu} \exp\left(-A\sum_i E_i^2\right), \tag{1}$$

where $\nu = 1, 2, 4$ for the GOE, GUE, and GSE, respectively. The distribution of nearest level spacing can then be written as

$$P(s) = \int \prod_{i=1}^{N} dE_i P(\{E_i\})\delta(s - |E_1 - E_2|), \tag{2}$$

where N is the number of levels in $\{E_i\}$ and the analytical result is quite complicated for general N. Instead, Wigner proposes a surmise that we can focus on the $N = 2$ case, the distribution then reduces to

$$P(s) \propto \int_{-\infty}^{\infty} |E_1 - E_2|^{\nu}\delta(s - |E_1 - E_2|) \times \exp\left(-A\sum_i E_i^2\right) dE_1 dE_2. \tag{3}$$

By introducing $x_1 = E_1 - E_2,\ x_2 = E_1 + E_2$, we have

$$P(s) \propto 2\int_{-\infty}^{\infty} |x_1|^{\nu}\delta(s - |x_1|)\exp\left(-\frac{A}{2}\sum_i x_i^2\right) dx_1 dx_2 = Cs^{\nu}e^{-As^2/2}. \tag{4}$$

The constants A, C can be determined by working out the integral about x_2, but it is more convenient to obtain by imposing the normalization condition,

$$\int_0^{\infty} P(s)ds = 1, \qquad \int_0^{\infty} sP(s)ds = 1. \tag{5}$$

From which we can reach to the celebrated Wigner-Dyson distribution,

$$P(s) = \begin{cases} \frac{\pi}{2}s\exp\left(-\frac{\pi}{4}s^2\right), & \nu = 1 \quad \text{GOE}, \\ \frac{32}{\pi^2}s^2\exp\left(-\frac{4}{\pi}s^2\right), & \nu = 2 \quad \text{GUE}, \\ \frac{2^{18}}{3^6\pi^3}s^4\exp\left(-\frac{64}{9\pi}s^2\right), & \nu = 4 \quad \text{GSE}. \end{cases} \tag{6}$$

On the other hand, the levels are independent in the Poisson ensemble, which means the occurrence of the next level is independent of the previous level, the nearest level spacings then follow a Poisson distribution $P(s) = \exp(-s)$.

Although the Wigner surmise is for the 2×2 matrix, it works fairly well when the matrix dimension is large. To demonstrate this, we present numerical evidence from a quantum many-body system—the spin-1/2 Heisenberg model with a random external field, which is the canonical model in the study of MBL, whose Hamiltonian in a length-L chain is

$$H = \sum_{i=1}^{L} \mathbf{S}_i \cdot \mathbf{S}_{i+1} + \sum_{i=1}^{L}\sum_{\alpha=x,y,z} h^{\alpha}\varepsilon_i^{\alpha} S_i^{\alpha}, \tag{7}$$

where we set coupling strength to be 1 and assume the periodic boundary condition in the Heisenberg term. The ε_i^{α}'s are random numbers within range $[-1, 1]$, and h^{α} is referred to as the randomness strength. We focus on two choices of h^{α}: (i) $h^x = h^z = h \neq 0$ and $h^y = 0$, the Hamiltonian matrix is orthogonal; (ii) $h^x = h^y = h^z = h \neq 0$, the model being unitary. This model undergos a thermal-MBL transition at roughly $h_c \simeq 3$ (2.5) in the orthogonal (unitary) model where the level spacing distribution evolves from the GOE (GUE) to Poisson [17].

We choose a $L = 12$ system to present a numerical simulation and prepare 500 samples at $h = 1$ and $h = 5$ for both the orthogonal and the unitary models. In Fig. 1(a), we plot the DOS for the $h = 1$ case in the orthogonal model. We can see the DOS is much more uniform in the middle part of the spectrum, which is also the case for $h = 5$ and the unitary model. Therefore, we choose the middle half of energy levels to do the spacing counting, and the results are shown in Fig. 1(b). We observe a clear GOE/GUE distribution for $h = 1$ in the orthogonal/unitary model and a Poisson distribution for $h = 5$ in the orthogonal model as expected, the fitting result for $h = 5$ in the unitary model is not shown since it almost coincides with that in the orthogonal model. It is noted the fitting for the Poisson distribution has minor deviations around the region $s \sim 0$, and this is due to the finite-size effect since there will always remain exponentially decaying but finite correlation between levels in a finite system. As we will demonstrate in a subsequent section, the fitting for higher-order level spacings will be better since the overlap between levels decays exponentially with their distance in the MBL phase.

A technique issue is, when counting the level spacings, we choose to take the middle half levels of the spectrum, whereas we can also employ a unfolding procedure using a spline interpolation that incorporates all energy levels [16], and the fitting results are almost the same [18,33].

III. HIGHER-ORDER LEVEL SPACINGS

Now, we proceed to consider the distribution of higher-order level spacings $\{s_i^{(n)} = E_{i+n} - E_i\}$, using a Wigner-like surmise. We first give the analytical derivation, then provide numerical evidence from simulation of the spin model in Eq. (7) as well as the nontrivial zeros of the Riemann ζ function.

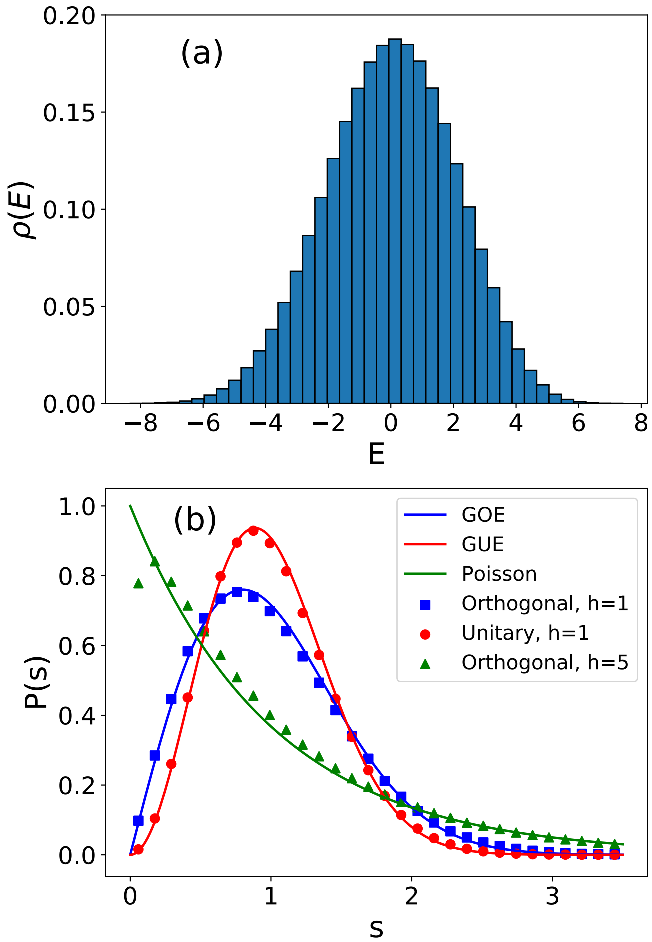

FIG. 1. (a) The DOS $\rho(E)$ of the random field Heisenberg model at $L=12$ and $h=1$ in the orthogonal case, the DOS is more uniform in the middle part, we, therefore, choose the middle half levels to perform level statistics. (b) Distribution of nearest level spacings $P(E_{i+1}-E_i)$, we see a GOE/GUE distribution for $h=1$ in the orthogonal/unitary model, whereas a Poisson distribution is found for $h=5$ in the orthogonal model, the result for $h=5$ in the unitary model is not displayed since it coincides with that in the orthogonal model.

A. Analytical derivation

Introduce $P_n(s)=P(s^{(n)}=s)\equiv P(|E_{i+n}-E_i|=s)$, to apply the Wigner surmise, we are now considering $(n+1)\times(n+1)$ matrices, the distribution $P_n(s)$ then goes to

$$P_n(s)\propto\int_{-\infty}^{\infty}\prod_{i<j}|E_i-E_j|^{\nu}\delta(s-|E_1-E_{n+1}|)\times\exp\left(-A\sum_{i=1}^{n+1}E_i^2\right)\prod_{i=1}^{n+1}dE_i. \tag{8}$$

We first change the variables to

$$x_i=E_i-E_{i+1},\qquad i=1,2,\dots,n,\qquad x_{n+1}=\sum_{i=1}^{n+1}E_i, \tag{9}$$

the $P_n(s)$ then evolves into

$$P_n(s)\propto\int_{-\infty}^{\infty}\frac{\partial(E_1,E_2,\dots,E_{n+1})}{\partial(x_1,x_2,\dots,x_{n+1})}\left(\prod_{i=1}^{n}\prod_{j=i}^{n}\left|\sum_{k=i}^{j}x_k\right|^{\nu}\right)\times\delta\left(s-\left|\sum_{i=1}^{n}x_i\right|\right)\times\exp\left\{-\frac{A}{n}\left[\sum_{i=1}^{n}\sum_{j=i}^{n}\left(\sum_{k=i}^{j}x_k\right)^2+x_{n+1}^2\right]\right\}\prod_{i=1}^{n+1}dx_i. \tag{10}$$

In this expression, the Jacobian $\frac{\partial(E_1,E_2,\dots,E_{n+1})}{\partial(x_1,x_2,\dots,x_{n+1})}$ and integral for x_{n+1} are all constants that can be absorbed into the normalization factor, hence, we can simplify $P_n(s)$ to

$$P_n(s)\propto\int_{-\infty}^{\infty}\left(\prod_{i=1}^{n}\prod_{j=i}^{n}\left|\sum_{k=i}^{j}x_k\right|^{\nu}\right)\delta\left(s-\left|\sum_{i=1}^{n}x_i\right|\right)\times\exp\left[-\frac{A}{n}\sum_{i=1}^{n}\sum_{j=i}^{n}\left(\sum_{k=i}^{j}x_k\right)^2\right]\prod_{i=1}^{n}dx_i. \tag{11}$$

Next, we introduce the n-dimensional spherical coordinate,

$$x_1=r\cos\theta_1,\qquad x_n=r\prod_{i=1}^{n-1}\sin\theta_i,$$
$$x_i=r\left(\prod_{j=1}^{i-1}\sin\theta_j\right)\cos\theta_i,\qquad i=2,3,\dots,n-1,$$
$$0\leqslant\theta_i\leqslant\pi,\qquad i=1,2,\dots,n-2,\qquad 0\leqslant\theta_{n-1}\leqslant 2\pi, \tag{12}$$

whose Jacobian is

$$\frac{\partial(x_1,x_2,\dots,x_n)}{\partial(r,\theta_1,\theta_2,\dots,\theta_{n-1})}=r^{n-1}\prod_{i=1}^{n-2}\sin^{n-1-i}\theta_i, \tag{13}$$

which reduces to the normal spherical coordinate when $n=3$. The resulting expression of $P_n(s)$ is complicated, whereas we are mostly interested in the scaling behavior about s, hence, we can write the formula as

$$P_n(s)\propto\int_0^{\infty}r^{n-1}\int r^{\nu C_{n+1}^2}\delta[s-r|G(\boldsymbol{\theta})|]\times H(\boldsymbol{\theta})\exp\left[-\frac{A}{n}r^2J(\boldsymbol{\theta})\right]dr\,d\boldsymbol{\theta}, \tag{14}$$

where $C_{n+1}^2=n(n+1)/2$ and $d\boldsymbol{\theta}=\prod_{i=1}^{n-1}d\theta_i$, the explanation goes as follows: (i) the first term r^{n-1} comes from the radial part of the Jacobian in Eq. (13); (ii) the second $r^{\nu C_{n+1}^2}$ comes from the number of terms in $\prod_{i=1}^{n}\prod_{j=i}^{n}|\sum_{k=i}^{j}x_i|^{\nu}$ where each term contributes a factor of r^{ν}; (iii) the auxiliary function $G(\boldsymbol{\theta})=\sum_{i=1}^{n}x_i/r$; (iv) the second auxiliary function $H(\boldsymbol{\theta})$ is composed of the angular part of the Jacobian and the angular part of $\prod_{i=1}^{n}\prod_{j=i}^{n}|\sum_{k=i}^{j}x_i|^{\nu}$; (v) $J(\boldsymbol{\theta})$ is the angular part of $\sum_{i=1}^{n}\sum_{j=i}^{n}(\sum_{k=i}^{j}x_k)^2$. The key observation is that $G(\boldsymbol{\theta}), H(\boldsymbol{\theta})$, and $J(\boldsymbol{\theta})$ all depend only on $\boldsymbol{\theta}$ whereas independent of r. Since we are only interested in the scaling behavior about s, we can work out the δ function and get

$$P_n(s) \propto s^{\nu C_{n+1}^2+n-1} \int H(\boldsymbol{\theta}) \exp\left[-\frac{AJ(\boldsymbol{\theta})}{n|G(\boldsymbol{\theta})|^2}s^2\right]d\boldsymbol{\theta}. \quad (15)$$

Although the integral for $\boldsymbol{\theta}$ is tedious and difficult to handle, it will only perform a correction to the Gaussian factor whereas not influence the scaling behavior about s. Therefore, we can write $P_n(s)$ into a generalized Wigner-Dyson distribution,

$$P_n(s) = C(\alpha)s^\alpha e^{-A(\alpha)s^2}, \quad (16)$$

$$\alpha = \frac{n(n+1)}{2}\nu + n - 1. \quad (17)$$

The normalization factors $C(\alpha)$ and $A(\alpha)$ can be determined by the normalization condition in Eq. (5) for which we obtain

$$A(\alpha) = \left(\frac{\Gamma(\alpha/2+1)}{\Gamma(\alpha/2+1/2)}\right)^2, \qquad C(\alpha) = \frac{2\Gamma^{\alpha+1}(\alpha/2+1)}{\Gamma^{\alpha+2}(\alpha/2+1/2)}, \quad (18)$$

where $\Gamma(z) = \int_0^\infty t^{z-1}e^{-t}dt$ is the Γ function. When $n = 1$, $P_n(s)$ reduces to the conventional Wigner-Dyson distribution in Eq. (6).

Interestingly, there exists coincidences between distributions in different ensembles. For example, as has been known for a long time [4,34], $P_k(s)$ in the GSE coincides with $P_{2k}(s)$ in the GOE for arbitrary integer k. And $P_7(s)$ in the GOE coincides with $P_5(s)$ in the GUE, and so on. Actually, our derivations are purely mathematical that work for arbitrary positive values of ν (not limited to integer values), although the three standard Gaussian ensembles are of most physical interest.

For the uncorrelated energy levels in the Poisson class, the distribution for higher-order spacing can also be obtained. Let us start with $n = 2$, we can write $\widetilde{s} = E_{i+2} - E_i = (E_{i+2} - E_{i+1}) + (E_{i+1} - E_i) = s_{i+1} + s_i$, where s_{i+1} and s_i can be treated as independent variables that both follow Poisson distribution, therefore, the distribution $P_2(\widetilde{s})$ for unnormalized $\widetilde{s}$ is

$$P(\widetilde{s}) \propto \int_0^{\widetilde{s}} P_1(\widetilde{s} - s_1)P_1(s_1)ds_1 = \widetilde{s}e^{-\widetilde{s}}. \quad (19)$$

Then, by requiring the normalization condition, we arrive at $P_2(s) = 4se^{-2s}$—the semi-Poisson distribution [35], which is suggested to be the distribution for nearest level spacing at the thermal-MBL transition point in the orthogonal model [20]. This interesting fact indicates the (leading-order) universality of this transition point is more affected by the MBL phase rather than the thermal phase, which is already noted by previous studies [10,20].

For higher-order level spacing in the Poisson ensemble, by repeating the procedure in Eq. (19) $n-1$ times, we reach

$$P_n(s) = \frac{n^n}{(n-1)!}s^{n-1}e^{-ns}, \quad (20)$$

which is a generalized semi-Poisson distribution with index n. Compared to the Poisson distribution for nearest level spacings, it is crucial to note that $P_n(0) = 0$ for $n \geqslant 2$, this is not a result of level repulsion as in the Gaussian ensembles, rather, it simply states that $n+1$ ($n \geqslant 2$) consecutive levels do not coincide.

We note every $P_n(s)$'s in the Gaussian and Poisson ensembles tend to be the Dirac δ function $\delta(s-1)$ in the limit $n \to \infty$, which is easily understood since, in that limit, only one spacing remains in the spectrum. Finally, we want to emphasize that the levels are well correlated in the Gaussian ensembles, hence, the derivation of $P_n(s)$ for the Poisson ensemble in Eq. (19) does not hold, otherwise the result will deviate dramatically [32].

For convenience, we list the order of the polynomial part in $P_n(s)$ for the three Gaussian ensembles as well as the Poisson ensemble up to $n = 8$ in Table I, note that the exponential parts in the former class are Gaussian type and that in the Poisson ensemble is an exponential decay.

TABLE I. The order of the polynomial term in $P_n(s)$ for the three Gaussian ensembles as well as the Poisson ensemble, the decaying term is a Gaussian type for the former class and an exponential decay for the latter.

n	1	2	3	4	5	6	7	8
GOE	1	4	8	13	19	26	34	43
GUE	2	7	14	23	34	47	62	79
GSE	4	13	26	43	64	89	118	151
Poisson	0	1	2	3	4	5	6	7

B. Numerical simulation

To show how well the distributions in Eqs. (16) and (20) work for the matrix with large dimensions, we now perform numerical simulations for the random spin model in Eq. (7) where we also pick the middle half levels to do statistics. We have tested the formula up to $n = 5$, and, in Fig. 2, we display the fitting results for $n = 2$ and $n = 3$.

As expected, the fittings are quite accurate for both GOE and GUE as well as the Poisson ensemble. In fact, the fittings for higher-order spacings in the Poisson ensemble are better than that for the nearest spacing in Fig. 1(b). This is because, in the MBL phase, the overlap between levels decays exponentially with their distances, hence, the fitting for higher-order level spacings is less affected by the finite-size effect.

For another example, we consider the nontrivial zeros of the Riemann ζ function [36],

$$\zeta(z) = \sum_{n=1}^{\infty} \frac{1}{n^z}, \quad (21)$$

it was established that statistical properties of nontrivial Riemann zeros $\{\gamma_i\}$ are well described by the GUE distribution [37]. Therefore, we expect the gaps $\{s_i^{(n)} = \gamma_{i+n} - \gamma_i\}$ follow the same distribution as those in the GUE. The numerical results for $n = 1$–3 are presented in Fig. 3 as can be seen, the fittings are perfect.

IV. HIGHER-ORDER GAP RATIOS

As mentioned in Sec. I, besides the level spacings, another quantity is also widely used in the study of random

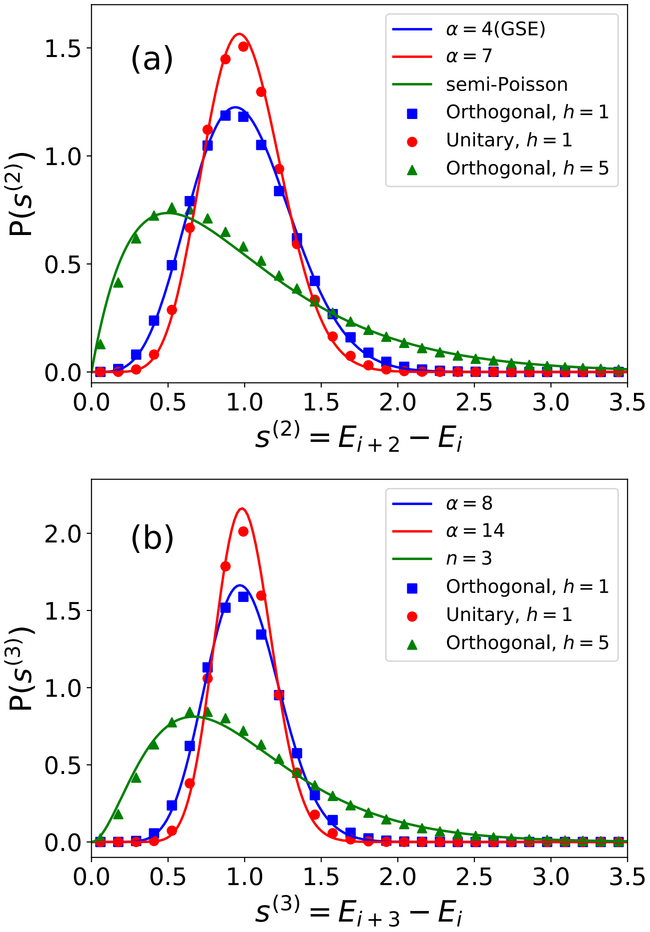

FIG. 2. Distribution of next-nearest level spacings $P(s^{(2)})$ in (a) and next-next-nearest level spacings $P(s^{(3)})$ in (b), where α and n are the indices in Eqs. (16) and (20), respectively.

matrices, namely, the ratio between adjacent gaps $\{r_i = s_{i+1}/s_i\}$, which is independent of the local DOS. The distribution of the nearest gap ratios $P(\nu, r)$ is given in Ref. [8], whose

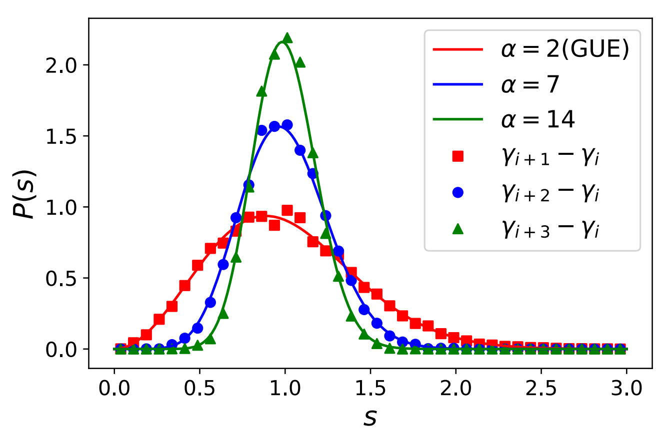

FIG. 3. The distribution of nth order spacings of the nontrivial zeros $\{\gamma_i\}$ of the Riemann ζ function, where α is the index in generalized Wigner-Dyson distribution in Eq. (16). The data come from 10^4 levels starting from the 10^{22}th zero, taken from Ref. [38].

result is

$$P(\nu, r) = \frac{1}{Z_\nu} \frac{(r + r^2)^\nu}{(1 + r + r^2)^{1+3\nu/2}}, \tag{22}$$

where $\nu = 1, 2, 4$ for the GOE, GUE, GSE, and Z_ν is the normalization factor determined by requiring $\int_0^\infty P(\nu, r)\, dr = 1$.

This gap ratio can also be generalized to higher order, but in different ways, i.e., the "overlapping" [8,26] and "nonoverlapping" [28,29] ways. In the former case, we are dealing with

$$\tilde{r}_i^{(n)} = \frac{E_{i+n} - E_i}{E_{i+n-1} - E_{i-1}} = \frac{s_{i+n} + s_{i+n-1} + \cdots + s_{i+1}}{s_{i+n-1} + s_{i+n-2} + \cdots + s_i}, \tag{23}$$

which is named the overlapping ratio since there is shared spacings between the numerator and the denominator. Whereas the nonoverlapping ratio is defined as

$$r_i^{(n)} = \frac{E_{i+2n} - E_{i+n}}{E_{i+n} - E_i} = \frac{s_{i+2n} + s_{i+2n-1} + \cdots + s_{i+n+1}}{s_{i+n} + s_{i+n-1} + \cdots + s_i}. \tag{24}$$

Both these two generalizations reduce to the nearest gap ratio when $n = 1$, but they are quite different when studying their distributions using Wigner surmise: For the overlapping ratio $\tilde{r}_i^{(n)}$, the smallest matrix dimension is $(n+2) \times (n+2)$; whereas it is $(1+2n) \times (1+2n)$ for the nonoverlapping ratio. Naively, we can expect the distribution for $\tilde{r}^{(n)}$ is more involved due to the overlapping spacings. Indeed, the $n = 2$ case for $P(\tilde{r}^{(n)})$ has been worked out in Ref. [26], and the result is very complicated. Instead, for the nonoverlapping ratio, Ref. [28] provides compelling numerical evidence for its distribution to follow:

$$P(\nu, r^{(n)}) = P(\nu', r), \tag{25}$$

$$\nu' = \frac{n(n+1)}{2}\nu + n - 1, \tag{26}$$

Surprisingly, the rescaling relation Eq. (26) coincides with that for higher-order level spacing in Eq. (17). We have also confirmed this formula by numerical simulations in our spin model Eq. (7), and the results for $n = 2$ in the GOE ($\nu = 1$) case is presented in Fig. 4 where we also draw the distribution of overlapping ratio $\tilde{r}^{(2)}$ for comparison. As can be seen, they differ dramatically, and the fitting for the nonoverlapping ratio is quite accurate. This result strongly suggests the nonoverlapping ratio is more universal than the overlapping ratio, and its distribution $P(r^{(n)})$ is homogeneously related with that for the nth-order level spacing, at least in the sense of Wigner, surmise for which we provide a heuristic explanation as follows.

For a given energy spectrum $\{E_i\}$ from a Gaussian ensemble with index ν, we can make up a new spectrum $\{E_i'\}$ by picking one level from every n levels in $\{E_i\}$, then, the nth-order level spacing $s^{(n)}$ in $\{E_i\}$ becomes the nearest level spacing in $\{E_i'\}$, and the nth-order nonoverlapping ratio in $\{E_i\}$ becomes the nearest gap ratio in $\{E_i'\}$. Since we have analytically proven the rescaling relation in Eq. (17), we conjecture the probability density for $\{E_i'\}$ (to leading order) bear the same form as $\{E_i\}$ in Eq. (1) with the rescaled parameter α in Eq. (17). Therefore, the higher-order nonoverlapping gap ratios also follow the same rescaling as expressed in Eqs. (25)

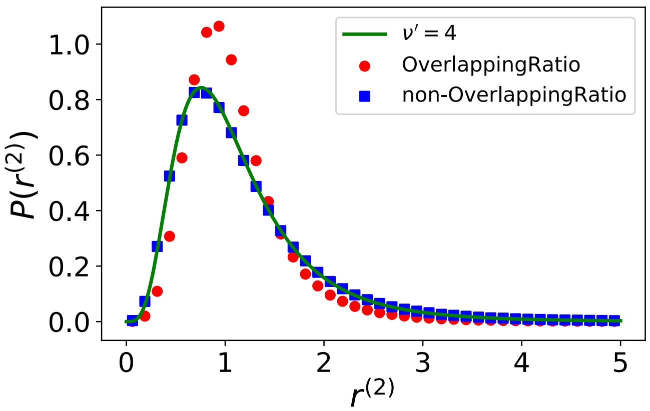

FIG. 4. The distribution of the second-order gap ratio in the orthogonal model where red and blue dots correspond to overlapping and nonoverlapping ratios, respectively, the latter fits perfectly with the formula in Eq. (25) with $\nu' = 4$. Note the data are taken from the whole energy spectrum without unfolding.

and (26). For this point of view, numerical evidences are provided in a recent work of the author [39].

V. CONCLUSION AND DISCUSSION

We have analytically studied the distribution of higher-order level spacings $\{s_i^{(n)} = E_{i+n} - E_i\}$ which describes the level correlations on long range. It is shown $s^{(n)}$ in the Gaussian ensemble with index ν follows a generalized Wigner-Dyson distribution with index $\alpha = \nu C_{n+1}^2 + n - 1$, where $\nu = 1, 2, 4$ for the GOE, GUE, and GSE, respectively. This results in a large number of coincident relations for distributions of level spacings of different orders in different ensembles. Whereas $s^{(n)}$ in the Poisson ensemble follows a generalized semi-Poisson distribution with index n. Our derivation is rigorous based on a Wigner-like surmise, and the results have been confirmed by numerical simulations from random spin system and nontrivial zeros of the Riemann ζ function.

We also discussed the higher-order generalization of gap ratios, which come in two different ways—the overlapping and nonoverlapping way—and point out their difference in studying their distributions using the Wigner-like surmise. Notably, the distribution for the nonoverlapping gap ratio has been studied numerically in Ref. [28] in which the authors find a scaling relation Eq. (26) that is identical to the one we find analytically for higher-order level spacings. This strongly indicates the distribution of higher-order spacing and the nonoverlapping gap ratio are correlated in a homogeneous way, for which we provided a heuristic explanation.

It is noted that the higher-order level spacings have played an important role in the study of the spacing distribution in a spectrum with missing levels [40], where the second-order level spacing distribution in the GOE is derived by a method different from this paper. Our derivations for $P(s^{(n)})$ in Gaussian ensembles are purely mathematical that work for arbitrary positive values of ν, although $\nu = 1, 2, 4$ for the GOE, GUE, and GSE are of most physical interest. Therefore, it is possible for our results to find applications in models that go beyond the three standard Gaussian ensembles. For example, the $\nu = 3$ behavior for level spacing has been found in a two-dimensional lattice with non-Hermitian disorder [41].

It is also interesting to note the distribution of next-nearest level spacing in Poisson class is semi-Poisson $P_2(s) \propto s\exp(-2s)$, which is suggested to be the distribution for nearest level spacing at the thermal-MBL transition point in the orthogonal model [20]. This indicates—to leading order—the universality property of this transition point is more affected by the MBL phase than the thermal phase, a fact already noted by previous studies [10,20]. This observation, thus, motivates a natural question: How will the thermal phase affect the universality of the MBL transition point? To answer this question, a comparison between the GOE-Poisson and GUE-Poisson transition points is suggested, which is left for a future study.

Last but not least, in this paper, the distribution of higher-order level spacing is derived only in $(n+1) \times (n+1)$ matrix, its exact value in a large matrix as well as the difference between them can in principle be estimated using the method in Ref. [8], this is also left for a future study.

ACKNOWLEDGMENTS

The author acknowledges the helpful discussions with X. Wan and R. Kausar. This work was supported by the National Natural Science Foundation of China through Grants No. 11904069 and No. 11847005.

[1] C. E. Porter, *Statistical Theories of Spectra: Fluctuations* (Academic, New York, 1965).

[2] T. A. Brody, J. Flores, J. B. French, P. A. Mello, A. Pandey, and S. S. M. Wong, Rev. Mod. Phys. **53**, 385 (1981).

[3] T. Guhr, A. Muller-Groeling, and 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, Berlin, 2001).

[6] E. P. Wigner, in Conference on Neutron Physics by Time-of-Flight, Gatlinburg, TN, 1957 (Oak Ridge National Laboratory, Oak Ridge, TN, 1957), p. 59.

[7] V. Oganesyan and D. A. Huse, Phys. Rev. B **75**, 155111 (2007).

[8] Y. Y. Atas, E. Bogomolny, O. Giraud, and G. Roux, Phys. Rev. Lett. **110**, 084101 (2013).

[9] V. Oganesyan, A. Pal, and D. A. Huse, Phys. Rev. B **80**, 115104 (2009).

[10] A. Pal and D. A. Huse, Phys. Rev. B **82**, 174411 (2010).

[11] S. Iyer, V. Oganesyan, G. Refael, and D. A. Huse, Phys. Rev. B **87**, 134202 (2013).

[12] X. Li, S. Ganeshan, J. H. Pixley, and S. Das Sarma, Phy. Rev. Lett. **115**, 186601 (2015).

[13] Y. Bar Lev, G. Cohen, and D. R. Reichman, Phys. Rev. Lett. **114**, 100601 (2015).

[14] K. Agarwal, S. Gopalakrishnan, M. Knap, M. Müller, and E. Demler, Phys. Rev. Lett. **114**, 160401 (2015).
[15] D. J. Luitz, N. Laflorencie, and F. Alet, Phys. Rev. B **91**, 081103(R) (2015).
[16] Y. Avishai, J. Richert, and R. Berkovits, Phys. Rev. B **66**, 052416 (2002).
[17] N. Regnault and R. Nandkishore, Phys. Rev. B **93**, 104203 (2016).
[18] S. D. Geraedts, R. Nandkishore, and N. Regnault, Phys. Rev. B **93**, 174202 (2016).
[19] P. Shukla, New J. Phys. **18**, 021004 (2016).
[20] M. Serbyn and J. E. Moore, Phys. Rev. B **93**, 041424(R) (2016).
[21] E. B. Bogomolny, U. Gerland, and C. Schmit, Eur. Phys. J. B **19**, 121 (2001).
[22] P. Sierant and J. Zakrzewski, Phys. Rev. B **99**, 104205 (2019).
[23] W. Buijsman, V. Cheianov, and V. Gritsev, Phys. Rev. Lett. **122**, 180601 (2019).
[24] P. Sierant and J. Zakrzewski, Phys. Rev. B **101**, 104201 (2020).
[25] J. M. G. Gomez, R. A. Molina, A. Relano, and J. Retamosa, Phys. Rev. E **66**, 036209 (2002).
[26] Y. Y. Atas, E. Bogomolny, O. Giraud, P. Vivo, and E. Vivo, J. Phys. A: Math. Theor. **46**, 355204 (2013).
[27] S. H. Tekur, S. Kumar, and M. S. Santhanam, Phys. Rev. E **97**, 062212 (2018).
[28] S. H. Tekur, U. T. Bhosale, and M. S. Santhanam, Phys. Rev. B **98**, 104305 (2018).
[29] P. Rao, M. Vyas, and N. D. Chavda, arXiv:1912.05664.
[30] A. Y. Abul-Magd and M. H. Simbel, Phys. Rev. E **60**, 5371 (1999).
[31] M. M. Duras and K. Sokalski, Phys. Rev. E **54**, 3142 (1996).
[32] R. Kausar, W.-J. Rao, and X. Wan, J. Phys.: Condens. Matter **32**, 415605 (2020).
[33] W.-J. Rao, J. Phys.: Condens. Matter **30**, 395902 (2018).
[34] M. L. Mehta and F. J. Dyson, J. Math. Phys. **4**, 713 (1963).
[35] E. B. Bogomolny, U. Gerland, and C. Schmit, Phys. Rev. E **59**, R1315(R) (1999).
[36] Definition of the Riemann $\zeta(z)$ function given in Eq. (21) is valid only for $\mathrm{Re}(z) > 1$. To overcome this problem, see, e.g., H. M. Edwards, *Riemann's ζ Function* (Dover, New York, 1974), Chap. 1.4.
[37] H. L. Montgomery, in Proceedings of Symposia in Pure Mathematics, edited by H. G. Diamond (AMS, Providence, 1973), Vol. 24, p. 181; E. B. Bogomolny and J. P. Keating, Nonlinearity **8**, 1115 (1995); **9**, 911 (1995); Z. Rudnick and P. Sarnak, Duke Math. J. **81**, 269 (1996); J. P. Keating and N. C. Snaith, Commun. Math. Phys. **214**, 57 (2000).
[38] A. Odlyzko, www.dtc.umn.edu/~odlyzko/zeta.
[39] W.-J. Rao and M. N. Chen, arXiv:2006.07774.
[40] O. Bohigas and M. P. Pato, Phys. Lett. B **595**, 171 (2004).
[41] A. F. Tzortzakakis, K. G. Makris, and E. N. Economou, Phys. Rev. B **101**, 014202 (2020).