Spacing distributions for point processes on a regular fractal

Jamal Sakhr[1] and John M. Nieminen[2]
[1]*Department of Physics, Harvard University, Cambridge, Massachusetts 02138, USA*
[2]*NDI (Northern Digital Inc.), 103 Randall Drive, Waterloo, Ontario, Canada N2V 1C5*
(Received 18 November 2005; published 1 March 2006)

The homogeneous Poisson point process in $\mathbb{R}^d$ (denoted by $\mathbf{P}_d$) is a basic model of stochastic geometry and modern statistical physics. Using ideas from fractal geometry, geometrical statistics, and random matrix theory, we introduce the model of random points on a self-similar fractal as a model of intermediate statistics, in the sense that the interpoint spacing statistics of the model are intermediate between those of $\mathbf{P}_1$ and $\mathbf{P}_2$ when the fractal dimension is in between 1 and 2, and intermediate between those of $\mathbf{P}_2$ and $\mathbf{P}_3$ when the fractal dimension is in between 2 and 3, and so on. We also introduce the idea of using a continuous family of such models to interpolate between $\mathbf{P}_1$ and $\mathbf{P}_2$ and thereby effectuate crossover transitions between $\mathbf{P}_1$ statistics and $\mathbf{P}_2$ statistics. We first derive the kth-nearest-neighbor spacing distribution for the general model, and then study the interpoint spacing statistics of several realizations of the model involving Sierpinski fractals in $\mathbb{R}^2$ and $\mathbb{R}^3$. We also study a realization of a continuous interpolation between $\mathbf{P}_1$ and $\mathbf{P}_2$, in particular a continuous interpolation between a point process on a line and a point process on a plane-filling curve, using the continuous family of self-similar Koch curves in $\mathbb{R}^2$. In the latter study, we specifically analyze the second-nearest-neighbor interpoint spacing statistics, which undergo a crossover transition between semi-Poisson and Ginibre statistics.

DOI: 10.1103/PhysRevE.73.036201 PACS number(s): 05.45.Df, 02.50.Ey, 05.40.−a, 05.45.Mt

I. INTRODUCTION

There is an interesting and deep connection between the spacing statistics of eigenvalues from the classical Gaussian random matrix ensembles and the interpoint spacing statistics of the Poisson point process in $\mathbb{R}^2$ [1,2]. The nearest-neighbor spacing statistics of eigenvalues from Gaussian ensembles of 2×2 random matrices constitute the most elementary results of classical random matrix theory (RMT), and are well-known and have been exploited in many fields of physics [3–6]. The Poisson point process is a stochastic model that is often encountered in statistical physics [7] and astrophysics [8,9], but the spacing statistics of Poisson point processes are perhaps not well-known to most physicists, and therefore deserve some introduction. The homogeneous Poisson point process in $\mathbb{R}^d$ (henceforth denoted by $\mathbf{P}_d$) can be understood as the limit of a simpler stochastic model: the binomial point process in $\mathbb{R}^d$ [10]. The binomial model consists of N independent uniformly distributed random points in a compact subset W of $\mathbb{R}^d$. A convenient choice for W is a d-dimensional ball of radius R and center at the origin. Let us denote the d-dimensional volume of this ball by $b_d R^d$ (where b_d is the volume of the unit ball in $\mathbb{R}^d$). If we take the limits $N\to\infty$ and $R\to\infty$ in such a way that $N/b_d R^d \equiv \rho$ remains constant, then the limiting stochastic point process is $\mathbf{P}_d$ (with intensity ρ).The kth-nearest-neighbor spacing distribution (kth-NNSD) for $\mathbf{P}_d$, which we shall denote by $D(s;k,d,\rho)$, gives the probability $D(s;k,d,\rho)dS$ of finding the kth nearest neighbor to a given point of $\mathbf{P}_d$ at a distance between s and $s+ds$. It can be shown that the NNSD for $\mathbf{P}_d$ is given by [10]

$$D(s;1,d,\rho) = \rho b_d d s^{d-1} \exp(-\rho b_d s^d). \tag{1}$$

It is easy to verify that this distribution is normalized [i.e., $\int_0^\infty D(s;1,d,\rho)ds=1$], and that the mean nearest-neighbor spacing is given by

$$\bar{s} = \int_0^\infty s D(s;1,d,\rho)ds = (\rho b_d)^{-1/d}\Gamma(1+1/d). \tag{2}$$

If we transform to the random variable $S=s/\bar{s}$, then the distribution (1) becomes

$$D(S;1,d) = \alpha d S^{d-1}\exp(-\alpha S^d), \tag{3a}$$

where the coefficient α is a constant that depends only on d,

$$\alpha = \left[\Gamma\left(1+\frac{1}{d}\right)\right]^d. \tag{3b}$$

[Equation (3a) is also given in Ref. [1].] The NNSD $D(S;1,d)$ gives the probability $D(S;1,d)dS$ of finding the nearest neighbor to a given point of $\mathbf{P}_d$ at the (dimensionless) scaled distance between S and $S+dS$. In the context of RMT, there are two interesting special cases: the NNSD for $\mathbf{P}_1$ is the well-known Poisson distribution,

$$D(S;1,1) = P_P(S) = \exp(-S), \tag{4}$$

and the NNSD for $\mathbf{P}_2$ is the Wigner distribution,

$$D(S;1,2) = P_W(S) = \frac{\pi}{2} S \exp\left(-\frac{\pi}{4}S^2\right). \tag{5}$$

For later reference, the so-called Wigner surmises for the NNSD of eigenvalues from the GXE (X=O,U,S) are given by (see Ref. [5])

$$P_W(S;1,\beta) = \mathcal{A}(\beta) S^\beta \exp[-\mathcal{B}(\beta)S^2], \tag{6a}$$

where

$$\mathcal{A}(\beta) = 2\frac{\left[\Gamma\left(\frac{\beta}{2}+1\right)\right]^{\beta+1}}{\left[\Gamma\left(\frac{\beta+1}{2}\right)\right]^{\beta+2}}, \quad \mathcal{B}(\beta) = \frac{\left[\Gamma\left(\frac{\beta}{2}+1\right)\right]^{2}}{\left[\Gamma\left(\frac{\beta+1}{2}\right)\right]^{2}}, \tag{6b}$$

and the parameter $\beta=1,2$, and 4 labels the symmetry classes (i.e., the classical Gaussian ensembles), which are orthogonal (GOE), unitary (GUE), and symplectic (GSE), respectively. The Wigner surmises are exact for Gaussian ensembles of 2×2 random matrices, and are excellent analytical approximations for the nearest-neighbor spacing statistics of eigenvalues from Gaussian ensembles of arbitrarily large random matrices [1]. Note that $P_W(S;1,1)\equiv P_W(S)$.

Although the NNSD for $\mathbf{P}_d$ [Eq. (3)] depends only parametrically on the Euclidean space dimension d, it is, in actuality, the discrete parameter d that characterizes the distribution and determines the behavior at both small and large values of S [i.e., $D(S;1,d)\sim S^{d-1}$ as $S\to 0$ and $D(S;1,d)\sim\exp(-\alpha S^d)$ when $S>1$]. Most importantly, notice that $D(0;1,d)=0$ when $d\neq 1$. The Wigner surmises for the NNSDs of eigenvalues have the same property [see Eq. (6) and note that $P_W(0;1,\beta)=0$], and this property is (to use the conventional term from RMT) referred to as "level repulsion." In RMT, the degree of level repulsion depends on the symmetries of the Hamiltonian, which are conveniently symbolized by the so-called level-repulsion parameter β. Analogously, we can think of the discrete parameter d in Eq. (3) as a "point-repulsion" parameter. This interpretation of the dimension as a measure of the amount of "repulsion" between points of a set is consistent with elementary notions of the more general fractal dimension [66], which suggests that if we wanted to consider random points uniformly distributed on fractal subsets of $\mathbb{R}^d$, then the nearest-neighbor interpoint spacing statistics would still be described by Eq. (3) but with the dimension d as a *continuous* (i.e., noninteger) parameter equal to the fractal dimension of the set. (In fact, for strictly self-similar fractals, this assertion is correct, and we shall prove it in the next section.) The motivation for this perhaps obscure proposition is the following. If we consider point processes on a family of fractal sets whose dimension changes continuously from 1 to 2 as some set parameter is varied, then such a family of point processes would constitute a new model for a crossover transition between Poisson and Wigner statistics. Point processes on fractals [67] would therefore serve as examples of stochastic models that do not involve random matrices but whose nearest-neighbor spacing statistics nevertheless exhibit the full range of statistics between Poisson and Wigner typified by many random matrix models [13–22]. It is well-known that the nearest-neighbor energy-level spacing statistics of a typical time-reversal invariant quantum Hamiltonian undergoes the same transition as the underlying classical dynamics change from being completely integrable to completely chaotic [23–25]. There indeed is an interesting connection between point processes on fractals and quantum chaos. The connection is based on a result which we have only recently reported [26] and which we will again encounter later in this paper.

Apart from these particular considerations, it is conceptually useful in its own right to treat and study the model of random points on a self-similar fractal with dimension in-between 1 and 2 as a model that is (in terms of spacing statistics) intermediate between $\mathbf{P}_1$ and $\mathbf{P}_2$. It will be convenient to symbolize the model of random points (uniformly distributed) on a self-similar fractal with similarity dimension $d_s\geqslant 1$ by $\mathbf{R}_{d_s}$ [68]. We have specifically considered the model $\mathbf{R}_{d_s}$ when $d_s\in(1,2)$. The reason is that the nearest-neighbor spacing statistics of $\mathbf{R}_{d_s}$ will be in between Poisson and Wigner when $d_s\in(1,2)$. This case is of special interest in relation to RMT as we have already discussed above and as we shall further discuss below in the context of the higher-order spacing statistics. However, there is no formal restriction to fractals with dimension in between 1 and 2. For instance, we can (and shall in Sec. III) consider random points on a self-similar fractal embedded in $\mathbb{R}^3$, whose similarity dimension $d_s\in(2,3)$, and we will see that the nearest-neighbor spacing statistics are in between those of $\mathbf{P}_2$ and $\mathbf{P}_3$.

So far, we have discussed only the nearest-neighbor spacing statistics, and we now move on to the higher-order spacing statistics. We again begin with $\mathbf{P}_d$. The kth-nearest-neighbor spacing distribution (kth-NNSD) $D(S;k,d)$ for $\mathbf{P}_d$ gives the probability $D(S;k,d)dS$ of finding the kth nearest neighbor to a given point of $\mathbf{P}_d$ at a distance between S and $S+dS$. It can be shown that the kth-NNSD for $\mathbf{P}_1$ is given by

$$D(S;k,1) = \frac{k^k}{\Gamma(k)}S^{k-1}\exp(-kS). \tag{7}$$

Interestingly, as pointed out in Ref. [1], the *second*-NNSD for $\mathbf{P}_1$ is the so-called semi-Poisson distribution,

$$D(S;2,1) = P_{\mathrm{sP}}(S) = 4S\exp(-2S). \tag{8}$$

The semi-Poisson distribution has been used to describe the spectral statistics of various models (see Refs. [27,28]), whose *nearest-neighbor* energy-level spacing distributions exhibit linear repulsion at small S and exponential falloff at large S. The semi-Poisson distribution has been particularly useful in helping to understand the spectral statistics of pseudointegrable billiards [29,30]. There is, in fact, a more general correspondence between the higher-order spacing statistics of $\mathbf{P}_1$ and what we shall refer to as the "generalized semi-Poisson statistics" [69]

$$P_{\mathrm{sP}}(S;n,\beta) = \frac{(\beta+1)^{n(\beta+1)}}{\Gamma[n(\beta+1)]}S^{[n(\beta+1)-1]}\exp[-(\beta+1)S]. \tag{9}$$

Note that $P_{\mathrm{sP}}(S;1,1)\equiv P_{\mathrm{sP}}(S)$. Comparison of Eqs. (7) and (9) reveals that, formally,

$$D(S;\beta+1,1) = P_{\mathrm{sP}}(S;1,\beta). \tag{10}$$

The kth-NNSD for $\mathbf{P}_1$ is relevant to the subject of integrable Hamiltonian systems. It is well known that the NNSD of the energy levels of a typical quantum Hamiltonian, whose classical limit is integrable, is Poissonian. We are not aware of any work that specifically considers the higher-order energy-level spacing distributions of quantized integrable systems, but the kth-NNSD should be (by virtue of the fact that the

long-range "gap function" statistics of typical integrable systems are Poissonian [35]) the distribution $D(S;k,1)=P_{sP}(S;1,k-1)$. Further discussions of Poisson processes in one dimension and their relevance to statistical analyses of energy-level spectra can be found in Ref. [36]. We next turn to $\mathbf{P}_2$ statistics. It can be shown that the kth-NNSD for $\mathbf{P}_2$ is given by [2]

$$D(S;k,2)=\frac{2\alpha^k}{\Gamma(k)}S^{2k-1}\exp(-\alpha S^2), \tag{11a}$$

where

$$\alpha=\left[\frac{\Gamma\left(k+\frac{1}{2}\right)}{\Gamma(k)}\right]^2. \tag{11b}$$

The Wigner surmises are usually defined only for three integer values ($\beta=1,2,4$), which correspond to the three classical Gaussian ensembles. It is interesting though that if β in Eq. (6) were such that $\beta\in\mathbb{N}^{(-)}$ (the set of all postive odd integers), then Eq. (6) would exactly reproduce the set of all higher-order spacing distributions for $\mathbf{P}_2$. There is, in fact, an intricate set of relations between the higher-order spacing statistics of $\mathbf{P}_2$ and the Wigner surmises for the kth-NNSDs of eigenvalues from the GXE (X=O,U,S) [denoted by $P_W(S;k,\beta)$] (see Ref. [2]). Two examples are

$$D(S;4,2)=P_W(S;2,2) \tag{12}$$

and

$$D(S;7,2)=P_W(S;4,1)=P_W(S;2,4). \tag{13}$$

There is also a direct relation between $\mathbf{P}_2$ and non-Hermitian RMT. The *second*-NNSD for $\mathbf{P}_2$ is identical to the NNSD of complex eigenvalues from Ginibre's ensemble of 2×2 general complex non-Hermitian random matrices [37],

$$D(S;2,2)=P_G(S)=\frac{3^4\pi^2}{2^7}S^3\exp\left(-\frac{3^2\pi}{2^4}S^2\right). \tag{14}$$

Note the commonly overlooked fact that $P_W(S;1,3)=P_G(S)$. This ensemble yields *cubic* level repulsion [i.e., $P_G(S)\sim S^3$ for small values of S]. Cubic level repulsion in quantum spectra was found to be a universal property of dissipative quantum systems with a chaotic classical limit [38].

We now come back to $\mathbf{R}_{d_s}$. If $d_s\in(1,2)$, we expect the spacing statistics of $\mathbf{R}_{d_s}$ to be in between those of $\mathbf{P}_1$ [Eq. (7)] and $\mathbf{P}_2$ [Eq. (11)], and as discussed above, we can think of a family of models $\{\mathbf{R}_{d_s}:1\leqslant d_s\leqslant 2\}\equiv\mathbf{R}_{[1,2]}$ as interpolating between $\mathbf{P}_1$ and $\mathbf{P}_2$, and the correspondng spacing statistics executing a crossover transition between $\mathbf{P}_1$ statistics and $\mathbf{P}_2$ statistics. So, for instance, the *third*-NNSD for $\mathbf{R}_{d_s}$ should be in between $D(S;3,1)$ $[=P_{sP}(S;1,2)]$ and $D(S;3,2)$, and the family of *third*-NNSDs for $\mathbf{R}_{[1,2]}$ should sweep through the region bounded by the distributions $D(S;3,1)$ and $D(S;3,2)$. Given the results we have quoted above, there are a number of cases that are of particular interest in relation to RMT. For example, the *second*-NNSD for $\mathbf{R}_{d_s}$ should be an intermediate distribution in between the semi-Poisson distribution $P_{sP}(S)$ and the Ginibre distribution $P_G(S)$, and the second-nearest-neighbor spacing statistics of $\mathbf{R}_{[1,2]}$ should execute a crossover transition between semi-Poisson and Ginibre statistics as d_s ranges between 1 and 2. In Sec. III of the paper, we demonstrate and clarify these ideas with examples. We begin, however, in the next section with a derivation of the kth-NNSD for $\mathbf{R}_{d_s}$.

II. DERIVATION OF THE *k*th-NNSD

Suppose that N points of a self-similar set $K\subset\mathbb{R}^d$ are chosen randomly and uniformly. The probability $P(s)ds$ of finding the kth nearest neighbor to a given point at a distance between s and $s+ds$ is equal to the probability of finding *one* of the $(N-1)$ points at a distance between s and $s+ds$ from the given point *and* $(k-1)$ points within a distance s of the given point *and* the $[(N-2)-(k-1)]=(N-k-1)$ remaining points at a distance greater than s. Let $\mathcal{P}(s)$ denote the probability of finding a point within a distance s of a given point. The probability of finding $(k-1)$ points would require that $\mathcal{P}(s)$ be multiplied $(k-1)$ times since the points are selected independently. The probability of finding one point at a distance greater than$s$ is then $[1-\mathcal{P}(s)]$, and for $(N-k-1)$ points, the probabilities are again multiplicative due to point independence. Thus,

$$P(s)ds=\eta[1-\mathcal{P}(s)]^{(N-k-1)}[\mathcal{P}(s)]^{(k-1)}d\mathcal{P}(s), \tag{15a}$$

where the prefactor

$$\eta=\frac{(N-1)!}{(N-k-1)!\,(k-1)!} \tag{15b}$$

is a combinatorial factor that accounts for all possible point configurations (see Appendix A), and $d\mathcal{P}(s)=\mathcal{P}'(s)ds$ is the probability of finding a point in a shell with inner and outer radii s and $s+ds$ centered about the given point.

It remains to specify the probability $\mathcal{P}(s)$. To do this, we need to determine the number of points within a volume defined by a d-dimensional ball of radius s (centered about the given point). In order to derive this "number-radius" relation, there is a crucial fact that must be observed: a uniformly sampled subset of K has the same self-similar structure as K itself. Suppose that the d-dimensional ball of radius $\mathcal{R}$ centered about the given point contains all N points. Then, a ball of radius $r=(1/\mathrm{r})\mathcal{R}$ contains $(1/\mathrm{n})N$ points, where the scaling factors r and n depend on the scaling structure of K. As a conceptual example, consider the well-known triadic Koch fractal on the interval $[0,\mathcal{R}]$, and consider disks of various radii centered about the origin in $\mathbb{R}^2$. The disk of radius $\mathcal{R}$ contains all N points, and if the random points are uniformly selected, then the disk of radius $r=\mathcal{R}/3$ contains $N/4$ points. We can then immediately specify the probability $\mathcal{P}(s)$ based on the scaling behavior of the set, but we shall first give the following more formal argument. Suppose that within radius $\mathcal{R}$ of the given point, there are N points, and within radius $r_1=(1/\mathrm{r})\mathcal{R}$ there are $N_1=(1/\mathrm{n})N$ points, within radius $r_2=(1/\mathrm{r})^2\mathcal{R}$ there are $N_2=(1/\mathrm{n})^2N$ points, and more generally within radius $r_n=(1/\mathrm{r})^n\mathcal{R}$ there are $N_n=(1/\mathrm{n})^nN$ points. It follows that there is a discrete scaling relation N_n

$=\varrho r_n^{d_m}$, where the prefactor ϱ depends on the upper cutoffs N and $\mathcal{R}$, that is, $\varrho=N/\mathcal{R}^{d_m}$, and the scaling exponent (i.e., the fractal dimension) $d_m=\ln\mathtt{n}/\ln\mathtt{r}$ depends only on the scaling factors $\mathtt{r}$ and $\mathtt{n}$. The continuum limit of the above discrete scaling relation (appropriate when $N\gg 1$) gives the following relation between N and r:

$$N(r)=\left[\frac{N(\mathcal{R})}{\mathcal{R}^{d_m}}\right]r^{d_m}=\left[\frac{N}{\mathcal{R}^{d_m}}\right]r^{d_m}. \tag{16}$$

The definition of the fractal dimension d_m given above coincides with all other definitions of the dimension for strictly self-similar fractals, and in particular, $d_m=d_s$. Number-radius scaling relations are common in the theory of fractals and in physics [39]. In various applications, number-radius relations are often referred to as "mass-length" relations and the fractal dimension d_m is often called the mass dimension. This terminology is no doubt in large part due to the various mass-radius relations given in Mandelbrot's influential book [40]. The number-radius relation (16) can be understood as a mass-radius relation when an individual sampling point becomes the unit of mass. However, it is important to be aware that there are other (slightly more technical) definitions of the mass dimension [41]. Although it is of no consequence here, we should also make casual reference to the fact that, in most applications where mass-length relations are employed, the mass dimension is extracted from the slope of a $\ln\mathtt{n}$ versus $\ln\mathtt{r}$ plot. Coming back now to $\mathcal{P}(s)$, the probability of finding a point within a distance s of a given point is governed by the power law

$$\mathcal{P}(s)=\frac{N(s)}{N}=As^{d_s}, \tag{17}$$

where the coefficient $A=1/\mathcal{R}^{d_s}$, and $d_s\geqslant 1$ is the similarity dimension of K [70]. Therefore, Eq. (15) becomes

$$P(s)ds=\eta[1-As^{d_s}]^{(N-k-1)}[As^{d_s}]^{(k-1)}Ad_s s^{d_s-1}ds. \tag{18}$$

Recall that $d\mathcal{P}(s)=\mathcal{P}'(s)ds=Ad_s s^{d_s-1}ds$. It is straightforward to verify that $P(s)$ is normalized [i.e., $\int_0^{\mathcal{R}}P(s)ds=1$].

Next, we calculate the mean kth-nearest-neighbor spacing $\bar{s}=\int_0^{\mathcal{R}}sP(s)ds$,

$$\begin{aligned}\bar{s}&=\frac{\eta d_s}{(\mathcal{R}^{d_s})^k}\int_0^{\mathcal{R}}\left[1-\left(\frac{s}{\mathcal{R}}\right)^{d_s}\right]^{(N-k-1)}s^{(k-1)d_s}s^{d_s}ds\\&=\eta\mathcal{R}d_s\int_0^1(u^{d_s})^k(1-u^{d_s})^{(N-k-1)}du\\&=\eta\mathcal{R}\int_0^1 v^{(k-1)+1/d_s}(1-v)^{(N-k-1)}dv\\&=\eta\mathcal{R}B(k+1/d_s,N-k)\\&=\eta\mathcal{R}\Gamma(N-k)\Gamma(k+1/d_s)/\Gamma(N+1/d_s).\end{aligned}$$

In the second line, we have made a change of variables $u=s/\mathcal{R}$, and in the third line, we have made one further change of variables $v=u^{d_s}$. The integral in the third line we recognize as the Beta function $B(\mu,\nu)$ with parameters $\mu=k+1/d_s$ and $\nu=N-k$, and this then gives the fourth line. We then used the usual relation between the Gamma and Beta functions to arrive at the fifth line. Finally, using Eq. (15b), the mean spacing can be written as

$$\bar{s}=\mathcal{R}\frac{\Gamma(N)}{\Gamma(N+1/d_s)}\frac{\Gamma(k+1/d_s)}{\Gamma(k)}. \tag{19}$$

It can be shown that the term

$$\frac{\Gamma(N)}{\Gamma(N+1/d_s)}=\frac{1}{N^{1/d_s}}\left[1+O\left(\frac{1}{N}\right)\right]\quad\text{as } N\to\infty, \tag{20}$$

and, therefore, the asymptotic mean kth-nearest-neighbor spacing, is, to leading order,

$$\bar{s}=\frac{\mathcal{R}}{N^{1/d_s}}\left[\frac{\Gamma(k+1/d_s)}{\Gamma(k)}\right]\quad\text{as } N\to\infty. \tag{21}$$

It is worth noting here that results similar to Eq. (21) have been found in the context of the Euclidean "traveling salesman problem" (TSP) in Ref. [42] (see also footnote [71]), and also in the context of point processes on closed manifolds in $\mathbb{R}^d$ [45] and on compact sets in $\mathbb{R}^d$ [46]. Introducing the rescaled spacing $S=s/\bar{s}$ and taking the limit $N\to\infty$, the distribution $P(s)$ in Eq. (18) becomes the distribution

$$P(S;k,d_s)=\frac{\alpha^k d_s}{\Gamma(k)}S^{kd_s-1}\exp(-\alpha S^{d_s}), \tag{22a}$$

where

$$\alpha=\left[\Gamma\left(k+\frac{1}{d_s}\right)/\Gamma(k)\right]^{d_s}. \tag{22b}$$

Note that $\int_0^\infty P(S;k,d_s)dS=1$ and $\bar{S}=\int_0^\infty SP(S;k,d_s)dS=1$. The distribution (22) has two important asymptotic properties: $P(S;k,d_s)\sim S^{kd_s-1}$ for small S (i.e., the point repulsion between kth nearest neighbors is kd_s-1), and $P(S;k,d_s)\sim\exp(-\alpha S^{d_s})$ for large S, where the coefficient α depends on both k and d_s. For large values of k, $\alpha\sim k$, and so $P(S;k,d_s)\sim\exp(-kS^{d_s})$ for large S (i.e.,$S>1$).

The distribution $P(S;k,d_s)$ is valid for point processes on any self-similar subset of $\mathbb{R}^d$ with similarity dimension $d_s\geqslant 1$. If K is a *classical* (nonfractal) self-similar set (i.e., a d-dimensional cube), then $d_s=d$ and Eq. (22) should then be consistent with the kth-NNSD for $\mathbf{P}_d$ (see Ref. [1], but note that there is an error in the distribution given there). We should point out that there are, in fact, a number of self-similar fractals whose similarity dimension $d_s=2$ (one example is studied in Sec. III). These special fractals play a fundamental role here because random points uniformly distributed on these fractal sets should have the same spacing statistics as $\mathbf{P}_2$.

As we have communicated in Ref. [26], the NNSD for $\mathbf{R}_{d_s}$ is

$$P(S;1,d_s)=\alpha d_s S^{d_s-1}\exp(-\alpha S^{d_s}), \tag{23a}$$

where

$$\alpha = \left[\Gamma\left(1+\frac{1}{d_s}\right)\right]^{d_s}, \tag{23b}$$

which is the Brody distribution [47] with Brody parameter $q=(d_s-1)$,

$$P(S;1,d_s) = P_B(S;q=d_s-1). \tag{24}$$

In mathematical statistics [48], the distribution $P(S;1,d_s)$ would be referred to as the (unit-mean) Weibull distribution with shape parameter d_s (see Appendix C).

If $1<d_s<2$, then the usual Brody parameter q is the fractional part of the similarity dimension. Incidentally, the similarity dimension of a self-similar curve can be expressed as (see Ref. [49]) $d_s=1+d_r$, where d_r is the exponent of the power law ($\ell=1/c^{d_r}$) that relates the measured length ℓ of a self-similar curve to the "compass setting" c (or precision $1/c$). Therefore, if K is a self-similar curve, then the usual Brody parameter $q=d_r$.

It is both interesting and important to cite here the work of Badii and Politi [50], who obtained an analytical formula for the NNSD of random points on a Cantor set with capacity dimension $D_0<1$ [see Eq. (5.7) of Ref. [50]]. The authors actually noticed the resemblence between their formula and the Brody distribution, but did not carry out the essential task of rescaling the random variable s by its mean $\bar{s}$ (as we have done here) to arrive at *the* Brody distribution. This rescaling might appear to be an almost trivial matter (after the fact), but the result is not inconsequential.

The Brody distribution has a long history in random matrix theory and in quantum chaos. It is an indispensable tool for analyzing the nearest-neighbor energy-level spacing statistics of quantized nonintegrable Hamiltonian systems. Unfortunately, in the context of Hamiltonian systems, the Brody distribution has no physical basis (i.e., it is not derived from any physical theory). It is a purely suppositional formula and generally regarded as nothing more than a mathematical function that interpolates between the Poisson and Wigner distributions. Although the Brody parameter has served as a phenomenological "chaoticity parameter" [see, for example, Ref. [51]], most authors believe the Brody parameter (in the context of Hamiltonian systems) is only a fitting parameter with no deeper physical meaning [24]. In the context of the model $\mathbf{R}_{d_s}$, the Brody parameter is not a meaningless control parameter; it is precisely the (relative) similarity dimension of the fractal set (i.e., $q=d_s-1$). To our knowledge, the model $\mathbf{R}_{d_s}$ is the first model whose NNSD is the Brody distribution, and the first model for which the Brody parameter has been found to have a precise physical meaning.

III. NUMERICS

We now turn to numerical experiments. We have studied numerically the spacing statistics of random points on many of the well-known classical fractals in $\mathbb{R}^2$, $\mathbb{R}^3$, and $\mathbb{R}^4$, but we only give results here for a few of the Sierpinski fractals that reside in $\mathbb{R}^2$ and $\mathbb{R}^3$. See Figs. 1 and 2.

We begin with a brief overview of the numerical procedure. Random points on each fractal were selected using the random iteration algorithm (RIA), which is well-known in the subject of random iterated function systems (RIFSs). An excellent description of this algorithm and the underlying theory is given in Ref. [52] (see also Appendix D for a brief summary). MATLAB 6.5 and its default RAND function [53] were used for this part of the numerical procedure. The spacings were then computed using the usual Euclidean metric $\Delta(\mathbf{u},\mathbf{v})$ between any two points $\mathbf{u}=(u_1,u_2,\ldots,u_d)$ and $\mathbf{v}=(v_1,v_2,\ldots,v_d)$ in $\mathbb{R}^d$: $\Delta(\mathbf{u},\mathbf{v})=\sqrt{(u_1-v_1)^2+(u_2-v_2)^2+\cdots+(u_d-v_d)^2}$. The distance between any given point $\mathbf{x}_i$ and its nearest neighbor is then defined by $s_i^{(1)}=\min\{\Delta(\mathbf{x}_i,\mathbf{x}_j):i,j=1,\ldots,N\ (j\neq i)\}$, and similarly the distance between $\mathbf{x}_i$ and its furthest neighbor is defined by $s_i^{(N)}=\max\{\Delta(\mathbf{x}_i,\mathbf{x}_j):i,j=1,\ldots,N\ (j\neq i)\}$. If the spacings for each point $\mathbf{x}_i$ are sorted by size (in ascending order), then the kth-nearest-neighbor spacing is the kth element of the set $\{s_i^{(1)},s_i^{(2)},\ldots,s_i^{(k)},\ldots,s_i^{(N)}\}$. Although the set $\{s_i^{(k)}:i=1,\ldots,N\}$ defines a set of spacings, the kth-NNSD is actually defined in terms of the scaled spacings $S_i=s_i^{(k)}/\bar{s}^{(k)}$, where $\bar{s}^{(k)}=(1/N)\Sigma_{i=1}^{N}s_i^{(k)}$ is the (numerically calculated) mean kth-nearest-neighbor spacing. We constructed the histograms by binning all values of S_i and subsequently normalizing the area under the histogram to unity. Each histogram displayed in Figs. 1 and 2 was constructed from *one* sample of $N=20\,000$ random points (selected uniformly) on the fractal.

It is useful to know how accurately the histogram data follow the theoretical distribution (22). In studies of energy-level statistics, a theoretical distribution is often overlaid on a histogram, and sometimes certain statistical tests (for example, the χ^2 test) are performed in order to determine how accurately the theoretical distribution reproduces the histogram data. We could also use the χ^2 test here, but we prefer to compute two simpler figures-of-merit which more readily convey how well the theory describes the given data. We first use the Levenberg-Marquardt (LM) method [54] to find the numerical value of the parameter d_s that gives the optimal fit of $P(S;k,d_s)$ [Eq. (22)] to the histograms. This number d_f can then be immediately compared to the theoretical value d_s. We should point out that the use of methods based on distances to kth-nearest-neighbors has been proposed and developed in a number of papers [50,55–59] for the specific purpose of estimating the dimension of a fractal set (especially a strange attractor). The use of the LM method to extract a fractal dimension from fits to a kth-nearest-neighbor spacing histogram could indeed be propounded as a different method of using kth-nearest-neighbor distance information to estimate fractal dimensions, but we should make clear that the current goal is not to exemplify a new method of dimension estimation; the goal here is to produce a figure-of-merit for how accurately the distribution (22) reproduces the histogram data obtained from particular realizations of $\mathbf{R}_{d_s}$. It is important to keep in mind that each realization yields a unique histogram (and hence a unique d_f), and it is thus more informative to average over several realizations (all independent of each other) and subsequently define $d_f=\bar{d}_f\pm\sigma$, where $\bar{d}_f$ is the average d_f value obtained from n realizations, and σ is the standard deviation. The percentage error of $\bar{d}_f$ relative to d_s is denoted by ε. For the present study, we analyzed ten

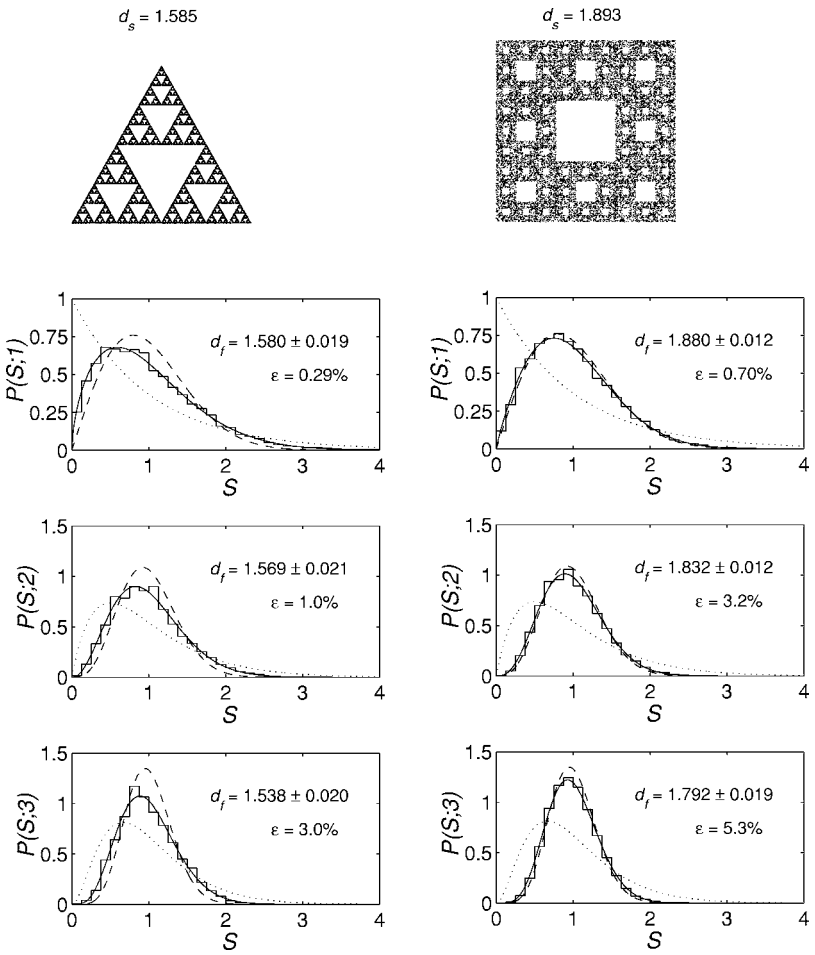

FIG. 1. Particular realizations of $\mathbf{R}_{d_s}$ consisting of 20 000 random points (selected uniformly) on the Sierpinski gasket (top left) and the Sierpinski carpet (top right), and the corresponding spacing distributions for the Sierpinski gasket (left panel) and the Sierpinski carpet (right panel). The numerical data (i.e., $d_f=\bar{d}_f\pm\sigma$ and ε) given in each window are figures-of-merit and are described in the main text. The exact similarity dimension d_s of each fractal is also given (to three decimal places). The dotted, dashed, and solid curves are the distributions $P_P(S)$, $P_W(S)$, and $P(S;1,\bar{d}_f)$ for $k=1$; $P_{sP}(S)$, $P_G(S)$, and $P(S;2,\bar{d}_f)$ for $k=2$; and $D(S;3,1)$, $D(S;3,2)$, and $P(S;3,\bar{d}_f)$ for $k=3$.

independent realizations for each fractal, generated ten corresponding histograms, and obtained ten different fits to these histograms. Up to and including the third-nearest-neighbor statistics, we find that $\bar{d}_f$ (obtained from $n=10$ independent realizations) is typically within 5% of the exact similarity dimension d_s. Almost all of the fractals we studied gave comparable results. Most of the error is due to edge effects. We inevitably introduce errors due to finite size or edge effects by including all of the points selected by the RIA in the statistical analysis. These errors are statistically insignificant as long as N is sufficiently large [see Eq. (21)]. Nevertheless, as the numerical results indicate, when the number of points is not large ($N\sim 10^4$), errors due to edge effects become increasingly important for longer-range correlations (i.e., higher k). If the number of points is large [for instance, $N\sim O(10^{10})$], and assuming $k\leqslant 5$ and $1<d_s<2$, we would expect errors due to edge effects to be negligible. It is interesting to note here that the simple procedure of excluding points within some prescribed distance from the edge of the fractals does not substantially reduce the error. This is unlike point processes on the classical self-similar sets, where the same procedure does significantly reduce the error due to edge effects. Edge effects account for the obvious difference in the size of the error in the two examples under consideration in Fig. 1. The error for the carpet is about twice as large as that for the gasket. This can be understood easily upon recognition of the fact that edge effects are practically irrelevant for the gasket except near the corners of the gasket, whereas edge effects are important along the entire edge of the carpet.

Random points on the Sierpinski gasket and Sierpinski carpet (see Fig. 1) are examples of models that are intermediate between $\mathbf{P}_1$ and $\mathbf{P}_2$. The spacing statistics of these models should be in between those of $\mathbf{P}_1$ and $\mathbf{P}_2$. As we can observe from Fig. 1, the NNSDs are in between Poisson and Wigner, the second-NNSDs are in between semi-Poisson and Ginibre, and the third-NNSDs are in between $D(S;3,1)$ and $D(S;3,2)$. The Sierpinski tetrahedron is an example of a self-similar fractal whose similarity dimension $d_s=2$, and as such, random points uniformly distributed on this fractal should have the same spacing statistics as $\mathbf{P}_2$. So, the NNSD, the second-NNSD, and the fourth-NNSD should follow the Wigner surmises $P_W(S;1,1)\equiv P_W(S)$, $P_W(S;1,3)\equiv P_G(S)$, and $P_W(S;2,2)$, and so on. As we can see from the top panel of Fig. 2, the second-NNSD of the random points on the tetrahedron does follow Ginibre statistics. Random points on the Sierpinski pyramid is an example of a model that is intermediate between $\mathbf{P}_2$ and $\mathbf{P}_3$, and the spacing statistics should be intermediate between those of $\mathbf{P}_2$ and $\mathbf{P}_3$. This example is interesting since it illustrates an important point: the NNSD for this and other higher-dimensional models is the Brody distribution with Brody parameter $q=d_s-1$. In RMT, and in studies of quantum chaos, the Brody distribution is not meant to be used beyond the Wigner limit. In the context of $\mathbf{R}_{d_s}$, there is no such restriction and the Brody

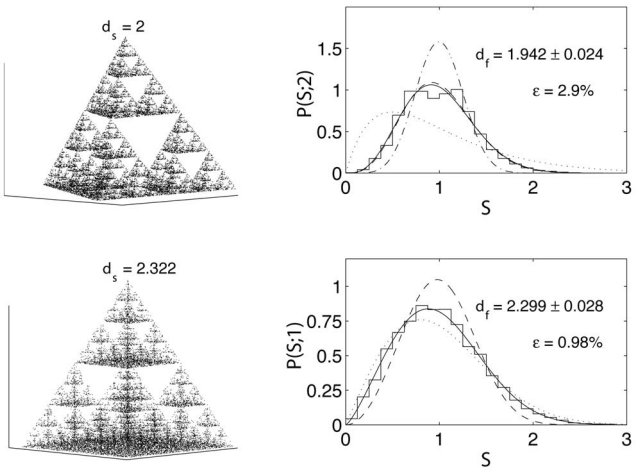

FIG. 2. (Top panel) A sample of 20 000 random points (selected uniformly) on the Sierpinski tetrahedron, and the corresponding second-NNSD. (Bottom panel) A sample of 20 000 random points (selected uniformly) on the Sierpinski pyramid, and the corresponding NNSD. (Top right) The dotted, dashed, solid, and dotted-dashed curves are the distributions $P_{sP}(S)$, $P_G(S)$, $P(S;2,\bar{d}_f)$, and $P(S;2,3)$, respectively. (Bottom right) The dotted, dashed, and solid curves are $D(S;1,2)$, $D(S;1,3)$, and $P_B(S;\bar{d}_f-1)$, respectively.

distribution continues to be valid beyond the Wigner limit. As we can see in the bottom panel of Fig. 2, the nearest-neighbor spacing histogram does indeed closely follow the Brody distribution.

A number of random matrix models have been devised whose eigenvalue statistics exhibit crossover transitions between the different classes of statistics that govern the different ensembles. In the present context, a crossover transition between say $\mathbf{P}_1$ and $\mathbf{P}_2$ statistics can be realized by considering point processes on a family of self-similar sets whose dimension ranges between 1 and 2. To demonstrate this idea explicitly, we use the continuous family of Koch curves in $\mathbb{R}^2$. These fractals can be thought of as the attractors of a one-parameter family of iterated function systems (IFSs). The similarity transformations defining the IFS involve a rotation that is conveniently parametrized by the angle θ. When $\theta=0$, the attractor is a line, and when $\theta=\pi/2$, the attractor is the famous Sierpinski-Knopp plane-filling curve, whose image is a solid isosceles triangle in $\mathbb{R}^2$ [60]. For intermediate values [i.e., $\theta\in(0,\pi/2)$], the attractors are various self-similar curves of prescribed dimension $d_s \in(1,2)$. The second-nearest-neighbor spacing statistics of the random points undergo a crossover transition from semi-Poisson to Ginibre statistics (see Fig. 3) as the self-similar set continuously deforms from a line to a plane-filling curve (i.e., as the rotation angle θ varies between 0 and $\pi/2$). Random points on these fractals were again selected using the RIA, and MATLAB 6.5 and its default RAND function were used for the numerics. Histograms were constructed from samples of $N=20\,000$ random points, and the LM method was then employed to fit Eq. (22) to the histograms (as before) from which the number d_f is determined. The percentage error ε of $\bar{d}_f$ (obtained from $n=10$ independent realizations) relative to d_s is typically under 1% when $d_s \in[1,1.5)$ and typically under 3% when $d_s\in[1.5,2)$.

We could also study the higher-order (i.e., higher k) spacing statistics of this model. Note that we studied the nearest-neighbor ($k=1$) spacing statistics of this model in Ref. [26]. For any k, the spacing statistics will always be understood as executing a crossover transition between $\mathbf{P}_1$ statistics and $\mathbf{P}_2$ statistics. Formally, the transition could also be understood as a crossover transition between Poisson statistics [Eq. (7)] and Wigner statistics [Eq. (6)] with $\beta=(2k-1)$.

IV. CONCLUSION

In this paper, we have studied the kth-nearest-neighbor spacing statistics of $\mathbf{R}_{d_s}$, the model of random points (uniformly distributed) on a self-similar set $K\subset\mathbb{R}^d$ with similarity dimension $d_s\geqslant 1$. We introduced $\mathbf{R}_{d_s}$ as a model whose spacing statistics are intermediate between those of $\mathbf{P}_1$ and $\mathbf{P}_2$ when $d_s\in(1,2)$, and intermediate between those of $\mathbf{P}_2$ and $\mathbf{P}_3$ when $d_s\in(2,3)$. This idea was inspired by a recurrent theme in RMT and in quantum chaos—models of "intermediate statistics," that is, models whose spectral statistics are intermediate between different classes of statistics. When $d_s\in(1,2)$, the spacing statistics of $\mathbf{R}_{d_s}$ are intermediate between $\mathbf{P}_1$ statistics and $\mathbf{P}_2$ statistics, and in the context of RMT, this case is especially interesting because of two remarkable correspondences: (i) the correspondence between the spacing statistics of $\mathbf{P}_1$ and the generalized semi-Poisson statistics [Eq. (10)], and (ii) the correspondence between the spacing statistics of $\mathbf{P}_2$ and the Wigner surmises of RMT [for example, Eqs. (5) and (12)–(14)]. However, the scenario when $d_s\in(2,3)$ exemplifies the more general idea. We also introduced the idea of using a continuous family of models $\{\mathbf{R}_{d_s}:1\leqslant d_s\leqslant 2\}\equiv\mathbf{R}_{[1,2]}$ to interpolate between $\mathbf{P}_1$ and $\mathbf{P}_2$, and this was motivated by another prevalent theme in RMT—random matrx models whose spectral statistics execute crossover transitions between different universality classes.

We first derived the kth-NNSD for $\mathbf{R}_{d_s}$ [Eq. (22)]. This generalizes the result for $\mathbf{P}_d$ [1]. The key to this generalization was the observation that the probability measure about a given point obeys a number-radius (or mass-length) scaling law, and that this scaling law is a consequence of the underlying self-similarity of the fractal set. Another crucial ingredient of the derivation is the rescaling of the distance between a given point and its kth-nearest-neighbor by the mean kth-nearest-neighbor distance. It is important to also keep in mind that the spacing distributions for $\mathbf{P}_2$ coincide with the Wigner surmises of RMT only when the distance variable s is rescaled by its mean $\bar{s}$.

Perhaps the most interesting result (from the point of view of RMT) is that the Brody distribution is the NNSD for $\mathbf{R}_{d_s}$. The implications to quantum chaos are deep (see Ref. [26]). Let us not, however, bypass the result itself; the Brody distribution is a probability density that can be derived from an actual stochastic model, $\mathbf{R}_{d_s}$, and in the context of this model,

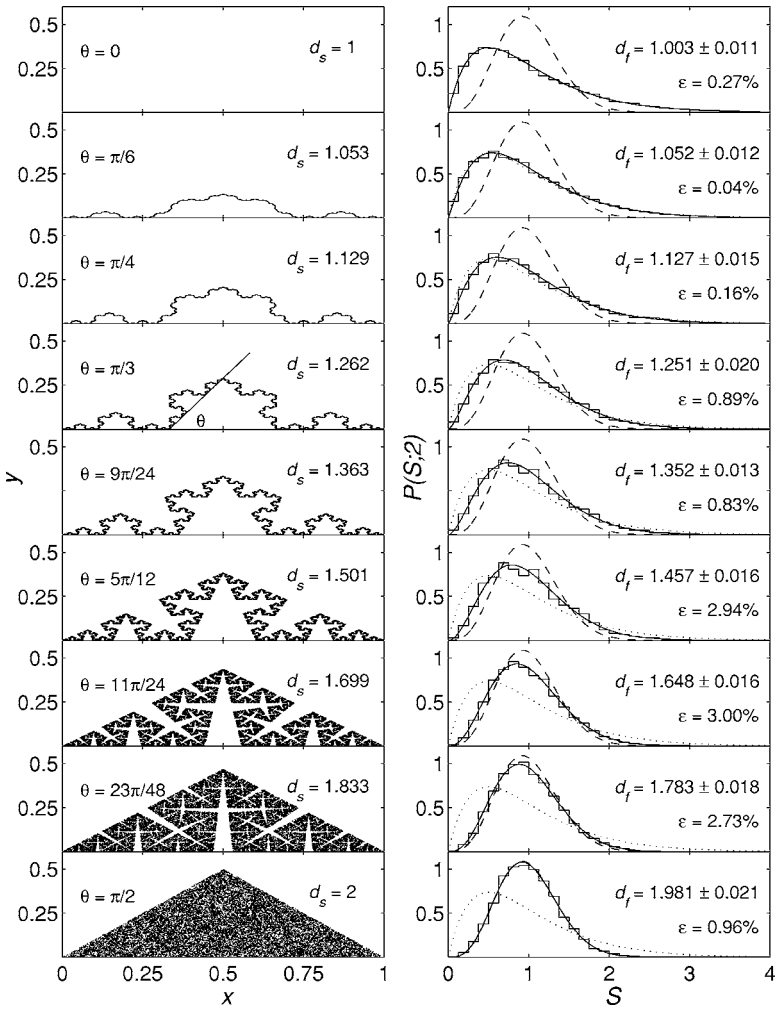

FIG. 3. A crossover transition between semi-Poisson and Ginibre statistics resulting from point processes on the family of Koch fractals in $\mathbb{R}^2$. The left panel shows random points on several fractals that belong to the family (each one specified by a particular choice of the rotation angle θ). The exact similarity dimension d_s of each of the fractals (given to three decimal places) is indicated on each window. The right panel shows the corresponding second-NNSD of the points on each fractal. The numerical data (i.e., $d_f=\bar{d}_f\pm\sigma$ and ε) indicated on each window are the figures-of-merit described in the main text. The dotted, dashed, and solid curves are the distributions $P_{sP}(S)$, $P_G(S)$, and $P(S;2,\bar{d}_f)$, respectively.

the Brody parameter $q=d_s-1$ is not a meaningless control parameter.

The ideas put forth in the Introduction were realized in several numerical demonstrations involving Sierpinski fractals in $\mathbb{R}^2$ and $\mathbb{R}^3$. Here we have only given a sample of the extensive numerical work that has been carried out using many of the well-known classical fractals. We also studied the second-nearest-neighbor spacing statistics of random points on a family of self-similar Koch curves with similarity dimension $d_s\in[1,2]$, and observed a crossover transition between semi-Poisson and Ginibre statistics. To our knowledge, this is the first example of a model that exhibits a semi-Poisson-to-Ginibre crossover transition. It is relevant to recall that, although the Ginibre ensemble of 2×2 random matrices with arbitrary complex elements belongs to the universality class of strongly non-Hermitian random matrices, there is no universal ensemble of random matrices from which semi-Poisson statistics can be derived [61]. In this sense, the semi-Poisson-to-Ginibre transition does not even have an analog in RMT, where crossover transitions occur between universality classes. There is one other point which should not be overlooked; the parameter that governs the crossover transitions manifested by the various random matrix models is usually quite vague [5], whereas the parameter that governs the semi-Poisson-to-Ginibre transition in this paper and the Poisson-to-Wigner transition in Ref. [26] is precisely the similarity dimension d_s.

"Point repulsion" is a useful theoretical construct. Recall that this concept was introduced early on in the paper and the context was that we could think of the Euclidean space dimension d in Eq. (3) (the NNSD for $\mathbf{P}_d$) as a "point-repulsion" parameter, and that this interpretation of the space dimension d actually reinforces certain notions of the more general fractal dimension. In the present picture, the fractal dimension can be interpreted as a measure of the amount of "repulsion" between points of a fractal set; a larger value indicates that points are more spread out and a smaller value indicates that points are more clustered. So, for instance, points on the Sierpinski gasket are more clustered than points on the Sierpinski carpet (see Fig. 1).

The kth-NNSD is one of the basic theoretical tools of RMT (although it is not widely used except for $k=1$). Methods based on kth-nearest-neighbor distances are among the most rudimentary tools of geometrical statistics [62], and are often used in spatial statistics [63] to characterize spatial point patterns that arise from theoretical models and physical data (again $k=1$ is most popular). The next logical step would be to consider statistics such as the two-point correlation function and geometrical analogs of the other long-range statistics common in RMT, such as the variance, skewness, and excess (kurtosis) of the number statistic [4]. As in the

present paper, efforts along these lines would bring together ideas from point processes, stochastic geometry, fractal geometry, geometrical statistics, and random matrix theory. Besides other fundamental long-range correlation studies, one important application would be to use the nearest-neighbor spacing statistic as the basis for a new method of computing the mass dimension. The idea would be to obtain numerical estimates of the mass dimension from curve fits to nearest-neighbor spacing histograms. Dimension estimation is a basic problem, especially for irregular fractals (i.e., fractals that are not *strictly* self-similar), and so the approach we are suggesting would be worthwhile if it could be aptly applied to irregular fractals. We are especially interested in applying this type of approach to problems in the subject of random fractals [49], and in particular, percolation clusters, diffusion-limited aggregation (DLA), and other fractal growth processes. We hope to explore these problems in a future project.

ACKNOWLEDGMENTS

J.S. acknowledges the Natural Sciences and Engineering Research Council of Canada (NSERC) for financial support.

APPENDIX A: THE COMBINATORIAL FACTOR η

Equation (15a) *without* the prefactor η accounts for only *one* particular configuration of points and therefore only gives the probability of having this one configuration. We must sum all of the probabilities obtained from each possible configuration. All of these probabilities are identical, but the question is how many configurations, in total, must be considered. The kth nearest neighbor can be any of the $(N-1)$ points (other than the "given point" itself). For *each* possible choice of the kth nearest neighbor, there are $(N-2)$ remaining points and $(k-1)$ of these must be closer than the kth nearest neighbor. So, we need to calculate the number of ways of choosing $(k-1)$ points (less than distance s) from a total of $(N-2)$ points. This combinatorial number is precisely the binomial coefficient ${}_{(N-2)}C_{(k-1)}$. Therefore, the total number of configurations is

$$\begin{aligned}\eta &= (N-1) \times {}_{(N-2)}C_{(k-1)} \\ &= \frac{(N-1)(N-2)!}{[(N-2)-(k-1)]\,!\,(k-1)!} \\ &= \frac{(N-1)!}{(N-k-1)\,!\,(k-1)!}. \end{aligned} \tag{A1}$$

Of course, no such combinatorial factor is needed to account for the fact that we could have specified any of the N points to be the "given point". The result would be the same because of symmetry, but we should not sum these probabilities since each of these correspond to the kth-NNSD for each of the points. In summary, specifying one configuration (which we must do initially) yields only *one* of the probability distributions, and in the end, we must add all of these separate (but equal) distributions.

APPENDIX B: ALTERNATIVE DERIVATION OF EQ. (22) FROM EQ. (18)

It is edifying to give here an alternative derivation of the kth-NNSD. We start from Eq. (18). If $N\to\infty$ and $\mathcal{P}(s)\to 0$ such that the product $N\mathcal{P}(s)$ remains fixed, then we can put to good use the so-called Poisson approximation to the binomial distribution [64]. The Poisson limit can be achieved if we let $N\to\infty$ *and* $\mathcal{R}\to\infty$ simultaneously and in such a way so as to ensure the ratio $N/\mathcal{R}^{d_s}\equiv\varrho$ (which we can heuristically think of as a "fractal intensity") remains fixed. If so, we obtain directly the asymptotic probability density

$$P(s;k,d_s,\varrho) = \frac{d_s\varrho^k}{\Gamma(k)} s^{kd_s-1}\exp(-\varrho s^{d_s}). \tag{B1}$$

The mean spacing can then be evaluated immediately,

$$\bar{s} = \int_0^\infty sP(s;k,d_s,\varrho)ds = \frac{\Gamma(k+1/d_s)}{\Gamma(k)\varrho^{1/d_s}}. \tag{B2}$$

If we then transform to the random variable $S=s/\bar{s}$, as usual, this leads directly to Eq. (22).

This alternative course to arrive at the formula for the kth-NNSD is of no practical significance given our program of rescaling the random variable s by $\bar{s}$, but it does bring into obvious view (albeit perhaps heuristically) the fundamental role of the intensity. The statistical description of point processes is usually accomplished through the use of intensity functions. We bypassed the potential technicalities and difficulties of properly defining and using intensity functions (in a sense) by expressing the probability density in terms of the random variable S, which is actually independent of ϱ. However, if we were to do a more rigorous analysis of the general model, then we would have to carefully consider the role of a fractal intensity as a characteristic parameter of the model, and its meaning in actual realizations of the model. It would have been necessary to do this here already if we had opted to do our analysis in terms of the random variable s. To conclude this discussion, note the one formal difference between the two derivations. In the derivation given in Sec. II, the parameter $\mathcal{R}$ is a *finite* positive constant, whereas in the derivation given here, $\mathcal{R}$ is allowed to become arbitrarily large.

APPENDIX C: A COMMENT ON THE FORMAL RELATION BETWEEN THE WEIBULL AND BRODY DISTRIBUTIONS

The Brody distribution and the unit-mean Weibull distribution are essentially the same distribution and it is only context that sets them apart. We begin with the Brody distribution, which is usually written as

$$P_B(X;q) = \lambda(q+1)X^q\exp(-\lambda X^{q+1}), \tag{C1a}$$

where

$$\lambda = \left[\Gamma\left(\frac{q+2}{q+1}\right)\right]^{q+1}, \tag{C1b}$$

and q is the Brody parameter. The two-parameter Weibull distribution is usually written as [48]

$$W(x;\alpha,\beta) = \alpha\beta x^{\beta-1}\exp(-\alpha x^{\beta}), \tag{C2}$$

where $\alpha>0$ is the scale parameter, and $\beta>0$ is the shape parameter. The mean of the Weibull distribution is

$$\bar{x} = \int_0^{\infty} xW(x;\alpha,\beta)dx = \alpha^{-1/\beta}\Gamma\left(1+\frac{1}{\beta}\right). \tag{C3}$$

If we transform to the random variable $X=x/\bar{x}$, then the unit-mean Weibull distribution with shape parameter β is

$$W(X;\beta) = \omega\beta X^{\beta-1}\exp(-\omega X^{\beta}), \tag{C4a}$$

where

$$\omega = \left[\Gamma\left(\frac{\beta+1}{\beta}\right)\right]^{\beta}. \tag{C4b}$$

So formally,

$$P_B(X;q) = W(X;\beta=q+1). \tag{C5}$$

The fact that the Brody distribution is a special case of the general two-parameter Weibull distribution was also recently noted in Ref. [65].

The Weibull and Brody distributions can be used equitably to interpolate between the Poisson and Wigner limits, and it is only a matter of what values the respective parameters assume at these limits. The Weibull distribution reduces to the Poisson and Wigner distributions when the shape parameter $\beta=1$ and $\beta=2$, respectively, whereas the Brody distribution reduces to the Poisson and Wigner distributions when the Brody parameter $q=0$ and $q=1$, respectively. The family of distributions that result as β ranges from 1 to 2 is the same family of distributions that result when the Brody parameter q ranges from 0 to 1. Nevertheless, it is the Brody distribution that is prevalent in RMT. This is not a historical accident, but rather reflects the original interpretation of the Brody parameter. The distribution $P_B(X;q)$ [Eq. (C1)] introduced by Brody has the property that $P_B(X;q)\sim X^q$ as $X\to 0$, and thus the Brody parameter q can be immediately identified as the degree of level repulsion. The distribution $P(S;1,d_s)$ [Eq. (23)] (which is really the unit-mean Weibull distribution with shape parameter d_s) rather has the property that $P(S;1,d_s)\sim S^{d_s-1}$ as $S\to 0$, and so the parameter d_s can be interpreted as the degree of point repulsion provided that $d_s\geqslant 1$.

APPENDIX D: A BRIEF ON RIFSs AND THE RIA

Fractals can be conceptualized as the attractors of random iterated function systems (RIFSs) [52]. A RIFS in $\mathbb{R}^d$ consists of a set of M affine linear transformations $\{w_1,\ldots,w_\ell,\ldots,w_M\}$ (where each $w_\ell:\mathbb{R}^d\to\mathbb{R}^d$) and a set of probabilities $\{p_1,\ldots,p_\ell,\ldots,p_M\}$, where $\Sigma_{\ell=1}^{M}p_\ell=1$. For a given RIFS, there is a unique associated geometric object K (a subset of $\mathbb{R}^d$) called the attractor of the IFS and also a unique associated measure that is related to the distribution of points on K.

The random iteration algorithm (RIA) can be summarized as follows: If $\mathbf{x}_1$ is an initial point in $\mathbb{R}^d$, then for $k=1,2,3,\ldots,N-1$, the points $\mathbf{x}_{k+1}=w_{n_k}(\mathbf{x}_k)$, where n_k is chosen randomly with probability p_ℓ from the set $\{1,\ldots,\ell,\ldots,M\}$. If $\mathbf{x}_1$ is a fixed point of one of the transformations, then all points $\mathbf{x}_1,\mathbf{x}_2,\mathbf{x}_3,\ldots,\mathbf{x}_N$ lie on the attractor and these points can be thought of as a "random orbit" on the attractor. Even if the initial point is not on the attractor, numerical convergence onto the attractor is obtained after a small number of iterations. We use the RIA *with equal probabilities* to generate random orbits on the various fractals.

For a given RIFS, the random sequence n_k together with the initial point $\mathbf{x}_1$ is a complete description of the "point process" on the fractal. The RIA effectively produces a random sampling of the fractal. The affine transformations are chosen randomly at each step, and thus there are no correlations between any two consecutively chosen points. The random points are uniformly distributed on the fractal *if* all probabilities that define the RIFS are equal. The structure of K is controlled exclusively by the affine maps of the IFS, and it is the probabilities that govern the distribution of the points on K. The derivation of $P(S;k,d_s)$ [Eq. (22)] assumes the points to be uniformly distributed, and for regular fractals, a uniform distribution of points can be obtained by using equal probabilities. For fractals that are not strictly self-similar (for example, self-affine fractals), the specification of these probabilities is an open problem.

[1] F. Haake, *Quantum Signatures of Chaos*, 2nd ed. (Springer, Berlin, 2001).

[2] J. Sakhr and J. M. Nieminen (unpublished).

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

[4] M. L. Mehta, *Random Matrices*, 3rd ed. (Elsevier, San Diego, 2004).

[5] T. Guhr, A. Müller-Groeling, and H. A. Weidenmüller, Phys. Rep. **299**, 189 (1998).

[6] P. J. Forrester, N. C. Snaith, and J. J. M. Verbaarschot, J. Phys. A **36**, R1 (2003).

[7] *Statistical Physics and Spatial Statistics: The Art of Analyzing and Modeling Spatial Structures and Pattern Formation*, Lecture Notes in Physics Vol. 554, edited by K. Mecke and D. Stoyan (Springer-Verlag, Berlin, 2000).

[8] V. J. Martinez and E. Saar, *Statistics of the Galaxy Distribution* (Chapman and Hall/CRC, New York, 2002).

[9] A. Gabrielli, F. Sylos Labini, M. Joyce, and L. Pietronero, *Statistical Physics for Cosmic Structures* (Springer-Verlag, Berlin, 2004).

[10] D. Stoyan, W. S. Kendall, and J. Mecke, *Stochastic Geometry and Its Applications*, 2nd ed. (Wiley, Chichester, 1995).

[11] F. C. Moon, *Chaotic and Fractal Dynamics* (Wiley, New York, 1992).

[12] D. J. Daley and D. Vere-Jones, *An Introduction to the Theory of Point Processes, Volume 1: Elementary Theory and Meth-*

ods, 2nd ed. (Springer-Verlag, Berlin, 2003).
[13] T. H. Seligman, J. J. M. Verbaarschot, and M. R. Zirnbauer, Phys. Rev. Lett. **53**, 215 (1984); J. Phys. A **18**, 2751 (1985).
[14] T. Cheon, Phys. Rev. Lett. **65**, 529 (1990).
[15] E. Caurier, B. Grammaticos, and A. Ramani, J. Phys. A **23**, 4903 (1990).
[16] F. M. Izrailev, J. Phys. A **22**, 865 (1989); Phys. Rep. **196**, 299 (1990).
[17] M. Feingold, D. M. Leitner, and M. Wilkinson, Phys. Rev. Lett. **66**, 986 (1991).
[18] G. Lenz and F. Haake, Phys. Rev. Lett. **67**, 1 (1991); G. Lenz, K. Życzkowski, and D. Saher, Phys. Rev. A **44**, 8043 (1991).
[19] D. M. Leitner, Phys. Rev. E **48**, 2536 (1993).
[20] P. Persson and S. Åberg, Phys. Rev. E **52**, 148 (1995).
[21] K. Życzkowski and R. Serwicki, Z. Phys. B: Condens. Matter **99**, 449 (1996).
[22] C. I. Barbosa *et al.*, Phys. Rev. E **59**, 321 (1999).
[23] O. Bohigas, in *Chaos and Quantum Physics*, edited by M.- J. Giannoni, A. Voros, and J. Zinn-Justin (Elsevier, Amsterdam, 1991).
[24] H.- J. Stöckmann, *Quantum Chaos: An Introduction* (Cambridge University Press, Cambridge, 1999).
[25] L. E. Reichl, *The Transition to Chaos in Conservative Classical Systems: Quantum Manifestations*, 2nd ed. (Springer, New York, 2004).
[26] J. Sakhr and J. M. Nieminen, Phys. Rev. E **72**, 045204(R) (2005).
[27] E. B. Bogomolny, U. Gerland, and C. Schmit, Phys. Rev. E **59**, R1315 (1999).
[28] E. B. Bogomolny, U. Gerland, and C. Schmit, Eur. Phys. J. B **19**, 121 (2001).
[29] B. Grémaud and S. R. Jain, J. Phys. A **31**, L637 (1998).
[30] J. Wiersig, Phys. Rev. E **65**, 046217 (2002).
[31] A. Pandey (unpublished).
[32] S. R. Jain and A. Khare, Phys. Lett. A **262**, 35 (1999).
[33] G. Auberson, S. R. Jain, and A. Khare, J. Phys. A **34**, 695 (2001).
[34] H. Hernández-Saldaña, J. Flores, and T. H. Seligman, Phys. Rev. E **60**, 449 (1999).
[35] M. Robnik and G. Veble, J. Phys. A **31**, 4669 (1998).
[36] P. Garbaczewski, Acta Phys. Pol. B **33**, 1001 (2002).
[37] J. Ginibre, J. Math. Phys. **6**, 440 (1965).
[38] R. Grobe, F. Haake, and H.- J. Sommers, Phys. Rev. Lett. **61**, 1899 (1988); R. Grobe and F. Haake, *ibid.* **62**, 2893 (1989).
[39] J. Feder, *Fractals* (Plenum Press, New York, 1988).
[40] B. B. Mandelbrot, *The Fractal Geometry of Nature* (W. H. Freeman and Company, San Francisco, 1982).
[41] B. B. Mandelbrot and M. Frame, *Fractals*, in Encyclopedia of Physical Science and Technology, Vol. 6, 3rd ed. (Academic Press, Boston, 2002).
[42] A. G. Percus and O. C. Martin, Phys. Rev. Lett. **76**, 1188 (1996).
[43] S. P. Lalley, Probab. Eng. Inform. Sci. **4**, 1 (1990).
[44] I. Hueter, Adv. Appl. Probab. **31**, 34 (1999).
[45] A. G. Percus and O. C. Martin, Adv. Appl. Math. **21**, 424 (1998).
[46] D. Evans, A. J. Jones, and W. M. Schmidt, Proc. R. Soc. London, Ser. A **458**, 2839 (2002).
[47] T. A. Brody Lett. Nuovo Cimento Soc. Ital. Fis. **7**, 482 (1973); T. A. Brody, *et al.* Rev. Mod. Phys. **53**, 385 (1981).
[48] I. Miller and M. Miller, *Mathematical Statistics with Applications*, 7th ed. (Prentice Hall, Englewood Cliffs, NJ, 2004); R. A. Johnson, *Probability and Statistics for Engineers*, 7th ed. (Prentice Hall, Englewood Cliffs, NJ, 2004).
[49] H.- O. Peitgen, H. Jürgens, and D. Saupe, *Chaos and Fractals*, 2nd ed. (Springer, New York, 2004).
[50] R. Badii and A. Politi, J. Stat. Phys. **40**, 725 (1985).
[51] T. Terasaka and T. Matsushita, Phys. Rev. A **32**, 538 (1985).
[52] M. F. Barnsley, *Fractals Everywhere*, 2nd ed. (Academic Press, Boston, 1993).
[53] C. B. Moler, *Numerical Computing with MATLAB* (SIAM, Philadelphia, 2004).
[54] W. H. Press *et al.*, *Numerical Recipes in C: The Art of Scientific Computing*, 2nd ed. (Cambridge University Press, Cambridge, 1992).
[55] Y. Termonia and Z. Alexandrowicz, Phys. Rev. Lett. **51**, 1265 (1983).
[56] R. Badii and A. Politi, Phys. Rev. Lett. **52**, 1661 (1984).
[57] P. Grassberger, Phys. Lett. **107A**, 101 (1985).
[58] W. van de Water and P. Schram, Phys. Rev. A **37**, 3118 (1988); Phys. Lett. A **140**, 173 (1989).
[59] M. de Rover and W. van de Water, Phys. Rev. E **51**, 4132 (1995).
[60] H. Sagan, *Space-Filling Curves* (Springer-Verlag, New York, 1994).
[61] T. Gorin, M. Müller, and P. Seba, Phys. Rev. E **63**, 068201 (2001).
[62] D. Stoyan and H. Stoyan, *Fractals, Random Shapes, and Point Fields: Methods of Geometrical Statistics* (Wiley, Chichester, 1994).
[63] N. A. C. Cressie, *Statistics for Spatial Data* (John Wiley and Sons, New York, 1993).
[64] J. G. Kalbfleisch, *Probability and Statistical Inference, Volume 1: Probability*, 2nd ed. (Springer-Verlag, New York, 1985).
[65] L. Peroncelli, G. Grossi, and V. Aquilanti, Mol. Phys. **102**, 2345 (2004).
[66] The fractal dimension can be succinctly described as (quoting Ref. [11]) "a quantitative property of a set of points in $\mathbb{R}^d$ that measures the extent to which the points fill a subspace [of $\mathbb{R}^d$] as the number of points becomes large." Suppose that, for instance, we have two fractal sets K_1 and K_2, embedded in $\mathbb{R}^2$, and consider the subset of each fractal that lies in the unit square $\mathbb{S}=[0,1]\times[0,1]$. The fractal having smaller holes (say K_2) fills out a larger portion of $\mathbb{S}$ (when the number of points is large). If the holes of K_2 are smaller, the points of K_2 are less clustered than the points of K_1 or equivalently the points of K_2 are more spread out (relative to each other) than the points of K_1, and if the points are more spread out, then there is greater "repulsion" between them. Accordingly, the dimension of K_1 is smaller than the dimension of K_2.
[67] We are aware that there are precise (albeit technical) meanings for the term "point processes" (see Ref. [12] for rigorous definitions). We use the phrase "point processes on fractals" heuristically (for lack of better terms) to denote any stochastic process whose outcome is a set of random points on a fractal. In this paper, we avoid the use of the abstract theory of point processes, random sets, and stochastic geometry.
[68] Any particular realization of this model depends on the fractal itself and any parameters involved in defining the fractal. Nevertheless, we will soon see that the spacing distribution (in

units of the mean spacing) only depends on the dimension of the fractal, and hence the dimension is the only characteristic parameter of importance for the spacing statistics. This is analogous to using the notation $\mathbf{P}_d$ for the homogeneous Poisson point process in $\mathbb{R}^d$.

[69] The distribution $P_{sP}(S;n,\beta)$ [Eq. (9)] was introduced in Ref. [27] as the nth-NNSD for the short-range plasma model (SRPM), which has been referred to as the short-range Dyson model in Ref. [29]. There is also unpublished work by A. Pandey [31], who investigated a model equivalent to the SRPM and derived Eq. (9) in the framework of random banded matrices. The distribution $P_{sP}(S;n,\beta)$ was also shown to be relevant to a class of exactly solvable models with nearest- and second-nearest-neighbor interactions [32,33]. Finally, $P_{sP}(S;n,\beta)$ was found to coincide exactly with the nth-NNSD for the so-called Poissonian "daisy model" of rank β [34]. Note that in Refs. [31–34], there is no restriction on the value of the system parameter β in Eq. (9), but in Refs. [27,28] the parameter β takes only the values $1,2$, and 4.

[70] The exclusion of self-similar sets with similarity dimension $d_s \in (0,1)$ is intentional. There are log-periodic corrections to Eq. (17) due to lacunarity effects. These corrections are not trivial to include in the derivation of the kth-NNSD. Fortunately, the corrections to Eq. (17) are quite small in amplitude for most (but not all) fractals. Log-periodic corrections are known to be very important for Cantor sets and other self-similar sets with $d_s < 1$, and thus should be included for a quantitative analysis of these sets. The purpose of the restriction $d_s \geqslant 1$ here is thus to avoid the analytical complications that would arise from including these corrections. We shall elaborate on this matter elsewhere.

[71] The Euclidean TSP is a famous combinatorial optimization problem. We should mention, at least parenthetically, a mathematical paper by Lalley [43], who considered the TSP on a self-similar set. Lalley obtained a limit law for the length of the shortest trajectory through finite samples of points chosen randomly and uniformly (in an appropriate sense) from self-similar subsets of $\mathbb{R}^2$ whose Hausdorff dimension $d_H < 2$. There is also a related paper that examines the convex hull of points sampled from a self-similar fractal set [44].