Scale dependence of distributions of hotspots

Michael Wilkinson[1*] and Boris Veytsman[2,3]

[1*]School of Mathematics and Statistics, The Open University, Walton Hall, Milton Keynes, MK7 6AA, UK.
[2]Chan Zuckerberg Initiative, 1180 Main St, Redwood City, 94063, CA, USA.
[3]School of Systems Biology, George Mason University, 4400 University Dr, Fairfax, 22030, VA, USA.

*Corresponding author(s). E-mail(s): m.wilkinson@open.ac.uk;
Contributing authors: bveytsman@chanzuckerberg.com;

Abstract

We consider a random field $\boldsymbol{\phi}(\mathbf{r})$ in $\boldsymbol{d}$ dimensions which is largely concentrated around small 'hotspots', with 'weights', $\boldsymbol{w_i}$. These weights may have a very broad distribution, such that their mean does not exist, or else is not a useful estimate. In such cases, the median $\overline{\boldsymbol{W}}$ of the total weight $\boldsymbol{W}$ in a region of size $\boldsymbol{R}$ is an informative characterisation of the weights. We define the function $\boldsymbol{F}$ by $\ln \overline{\boldsymbol{W}} = \boldsymbol{F}(\ln \boldsymbol{R})$. If $\boldsymbol{F'(x)} > \boldsymbol{d}$, the distribution of hotspots is dominated by the largest weights. In the case where $\boldsymbol{F'(x)} - \boldsymbol{d}$ approaches a constant positive value when $\boldsymbol{R} \to \boldsymbol{\infty}$, the hotspots distribution has a type of scale-invariance which is different from that of fractal sets, and which we term *ultradimensional*. The form of the function $\boldsymbol{F(x)}$ is determined for a model of diffusion in a random potential.

1 Introduction

In many cases two-dimensional scalar fields are largely supported on small areas, 'hotspots'. Examples can include the distribution of human populations, which are concentrated in urban settlements, the distribution of debris on the ocean, which can be concentrated in regions where cool or saline water is subducted, and deposits of mineral ores, which can be concentrated at the points where dissolved material is deposited from evaporating water. Another example is images of star fields, where the stars appear as points. Subjectively, images of the distribution of hotspots can

appear to bear a familial resemblance. This paper addresses the question of how these distributions can be characterised, and whether they have scale-invariant features.

The fields that we consider can be modelled by random processes. We consider random, non-negative scalar fields in a two-dimensional space, denoted by $\phi(\mathbf{r})$, with statistics which are homogeneous (translationally invariant) and isotropic (rotationally invariant). Extensions to higher dimensions will be obvious.

In the cases where the field is highly concentrated in the vicinity of isolated points, we can consider the following simple model. We take a uniform, independent random scatter of points on the plane, $\mathbf{r}_i$, with density ρ. Each point is assigned a random weight w_i, drawn independently from a distribution with probability density function (PDF) $p(w)$. The weights w_i represent the integral of $\phi(\mathbf{r})$ in the neighbourhood surrounding one of the points upon which it is concentrated.

The primary interest will be in the cases where $p(w)$ has a very broad distribution. Accordingly, we introduce a *power-law model*, such that for large w, the PDF is $p(w) \sim w^{-\gamma}$. In the calculations below we shall use the following specific distribution as an example:

$$p(w) = \begin{cases} (\gamma - 1)w^{-\gamma}, & w \geq 1 \\ 0, & w < 1 \end{cases} \tag{1}$$

with $1 < \gamma < 2$, so that the distribution is normalisable, but its mean is undefined.

We shall also need to consider the cumulative distribution: if $P(w_0)$ is the probability that $w > w_0$, then (1) implies that $P(w) = w^{-(\gamma-1)}$ for $w > 1$. It will be argued that this is a foundational model for the distribution of hotspots, and that distributions obtained from more general models can be approximated using this model, with a suitable choice of the parameter γ.

Figure 1 is an illustration of 12 different realisations of this model for hotspot distributions, with $\gamma = 5/3$, plotted on four different lengthscales. It will be argued that the statistics of these images has a scale-invariance property, in that it is impossible to identify the scale factors of the panels. Non-trivial scale invariance is usually associated with fractal [1, 2] (or more generally, multifractal [3, 4]) properties, which can usually be characterised by saying that the set is, in some sense, self-similar under a change of scale. The images in figure 1 are so diverse that would require a large number of realisations to demonstrate that they are drawn from the same ensemble. We shall argue below that there is a simple quantitative distinction between the scale invariance of figure 1 and that of fractal sets.

There are two complementary aspects to characterising these sets. Firstly, we might wish to know how the total weight of the hotspots increases with the size of the region. Consider the set $\mathcal{W}_R$ of approximately ρR^2 values of w_i for which $\mathbf{r}_i$ lies inside a square of side R. The set has the total weight

$$W = \sum_{w_i \in \mathcal{W}_R} w_i. \tag{2}$$

If the weights had a compact distribution, we would estimate the mean value of W as $\langle W \rangle = \rho R^2 \langle w \rangle$, but for the distribution (1), $\langle w \rangle$ is infinite ($\langle X \rangle$ denotes the expectation value of X throughout). A more promising approach is to estimate the median

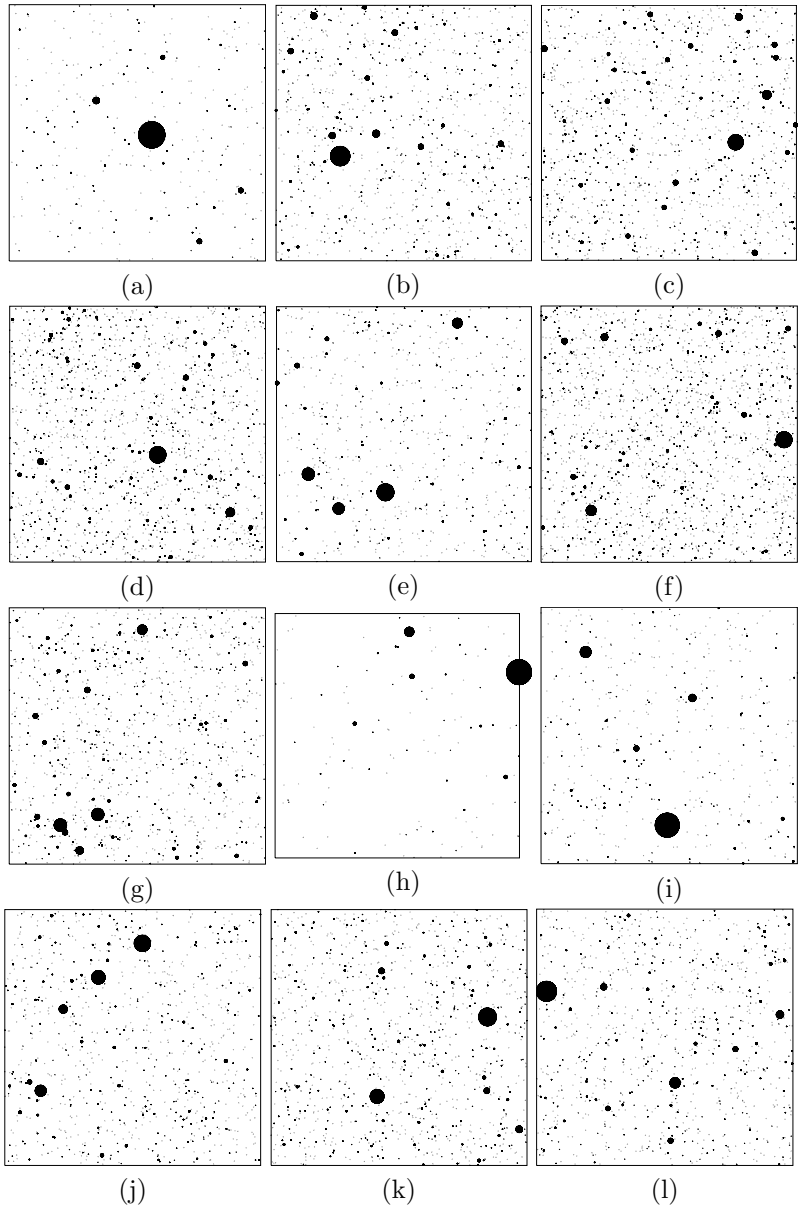


Fig. 1 Illustration of scale-independence of the 'hotspots' model. At the position of each hotspot there is a filled circle with area proportional to its weight. The total area of the circles in each image is normalised to be 1% of the total area of the image. The images use the probability distribution (1) with $\gamma = 5/3$, and the scale factors are $R = 10000$, $R = 2000$, $R = 400$, $R = 80$, (with three cases of each scale factor). The different images cannot be associated with the different values of R by any statistical test, reflecting the scale-invariance property. The scale factors are: $R = 10000$, panels (a), (g), (k), $R = 2000$, panels (b), (f), (h), $R = 400$, panels (d), (e), (l), $R = 80$, panels (c), (i), (j).

value $\overline{W}$. This is considered in Section 2 below. We anticipate that $\overline{W}$ will increase very rapidly as a function of the scale length R. Accordingly, we use a logarithmic scale. We can characterise a given hotspot distribution by means of a function F:

$$\ln(\overline{W}) = F(\ln R). \tag{3}$$

For the simple model described above, we show that $\overline{W}$ has a power law dependence upon R, so $F(x)$ is a linear function. The exponent of this power law can be thought of as a type of *dimension* of the set of hotspots, and even if F is not a linear function we can define an effective dimension D_{eff} at the length scale R as the derivative

$$D_{\text{eff}} = \frac{\mathrm{d}F(\ln R)}{\mathrm{d}(\ln R)}. \tag{4}$$

For our simplified model it will be shown that, for points distributed randomly in d dimensions with the weight distribution (1),

$$D_{\text{eff}} = \frac{d}{\gamma - 1}. \tag{5}$$

Note that, because $2 > \gamma > 1$, this effective dimension is higher than the dimension of the embedding space. This indicates that the effective dimension is different from a fractal dimension. We describe this scale invariance as *ultradimensional*.

A second aspect of describing the hotspot distribution is to consider the relative sizes of the largest values of w_i in the set $\mathcal{W}_R$. We can transform this to a filtered and normalised set, $\widetilde{\mathcal{W}}_R$, as follows. We scale the hotspot positions by dividing by R, and plot $\mathbf{r}_i/R$ inside a unit square. We eliminate the values of w_i below a chosen threshold, for example, those w_i that are less than ϵW, where ϵ is a given small positive number. We can also 'normalise' these sets by dividing every remaining w_i by W. These normalised and filtered sets are a natural representation of many types of point-set data. An example is a geographical map showing settlements using symbols with the sizes relative to of the largest settlements in the mapped region, where settlements below a certain size are not shown in order to eliminate clutter. Another example is a photograph of the night sky with the exposure adjusted so that the image saturation is normalised, and stars below a certain intensity are not registered at all. The filtered and normalised sets can be characterised by considering the relative sizes of the largest values of w_i. To this end, we can sort the weights, w_i, into a decreasing sequence $\{\overrightarrow{w}_i\}$, and consider the proportion of the total mass which is contained in the first k elements of this set

$$f_k = \frac{\sum_{j=1}^{k} \overrightarrow{w}_j}{\sum_{j=1}^{\widetilde{N}} \overrightarrow{w}_j}, \tag{6}$$

where $\widetilde{N}$ is the number of elements in the filtered set. We can consider the average of f_k over different regions of the data, and in some cases we can also average over multiple realisations of the distribution. For the model defined by equation (1), this

leads to a family of functions of γ:

$$\widetilde{f}_k(\gamma) = \langle f_k \rangle. \tag{7}$$

We shall make a hypothesis that, for a general model, the set of values of f_k at length scale R is representative of the model (1), with an effective value of γ given by rearrangement of (5):

$$\gamma_{\text{eff}} = 1 + \frac{d}{F'(\ln\, R)}. \tag{8}$$

If the derivative of the function $F(x)$ defined by equation (3) approaches a constant as $x \to \infty$, this is indicative of the sets $\widetilde{\mathcal{W}}_R$ having scale-invariant properties, such that the statistics of $\widetilde{\mathcal{W}}_R$ and $\widetilde{\mathcal{W}}_{\lambda R}$ are indistinguishable, for a wide range of values of the positive number λ. This idea can be expressed by saying that the realisation of $\widetilde{\mathcal{W}}_R$ are drawn from an ensemble which is independent of R, depending only upon γ_{eff}. In the case of the power law model, this scale invariance is manifest. This self-similarity could be trivial, or it could indicate that the hotspot distribution has fractal properties, or something different. It will be argued that it is the latter possibility which is realised. Figure 1 is an example of sets $\widetilde{\mathcal{W}}_R$ generated by this model. displayed on four different length scales (we used $\rho = 1$, $\gamma = 5/3$, and $R = 80$, 400, 2000, 10000. We show 12 sample sets, with three at each of the different scale factors. The twelve images look so different from each other that it is not evident that they are drawn from the same ensemble. The different values of the scale factor R were randomly assigned (the key is in the figure caption), and its value cannot be determined by inspection of an individual realisation. Despite the fact that the length scales vary by a factor of 125, it is not possible to distinguish which of these images corresponds to which value of R.

The results in section 2 will quantify the non-trivial scale-invariance of the power-law model under a change of the magnification of the image. We discuss the statistics of W for this simple model, leading to the relation (5) between the exponent γ and the dimension D_{eff}.

In section 3 we discuss a physical example of a hotspot distribution, namely the probability density for a particle diffusing in a two-dimensional gaussian random potential, $V(x, y)$. The equilibrium probability density is proportional to $\exp[-V(x, y)/\mathcal{D}]$ where $\mathcal{D}$ is the diffusion coefficient. In the limit where the diffusion coefficient approaches zero, this density is concentrated at 'hotspots' which are minima of the potential function, with the weights $w \approx \exp[-V_{\text{min}}/\mathcal{D}]$, where V_{min} is the value of the local minimum of the potential at the hotspot. We identity the functions $F(\ln R)$ and $\gamma_{\text{eff}}(R)$ for this model, including their dependence upon the diffusion coefficient $\mathcal{D}$. Section 4 is a brief conclusion.

2 Statistics of a simple model

Consider how the statistics of the total weight W depends upon R for the *power-law model*, with weight distribution (1). The mean value of w is undefined, so calculating the expectation value $\langle W \rangle$ is not a good approach. The mean value is dominated by rare realisations where one or more of the w_i takes a very large value. Estimating the

median of W, which will be denoted by $\overline{W}$, appears to be more promising. If $\hat{w}$ is the largest of the $N \sim \rho R^2$ samples of w in the square, then we might hypothesise that $\overline{W}$ is approximated by $\overline{\hat{w}}$, that is of the median of the largest value in the sample. Here it will be argued that this multiplier $\overline{W}/\overline{\hat{w}}$ is independent of both R and ϵ.

It is easy to calculate $w^* \equiv \overline{\hat{w}}$. The probability that none of the N independent values of w exceeds $\hat{w}$ is $[1 - P(\hat{w})]^N$, so that $w^* = \overline{\hat{w}}$ satisfies $[1 - P(w^*)]^N = 1/2$. This gives

$$w^* = \left(\frac{N}{\ln 2}\right)^{1/(\gamma-1)}. \tag{9}$$

Next we estimate the number of points $\widetilde{N}$ in the filtered set, and the value of $\overline{W}$. The number of values of w_i in the range from $w^* \equiv \overline{\hat{w}}$ (upper limit) to ϵw^* (lower limit) is

$$\begin{aligned} \widetilde{N} &\sim N \int_{\epsilon w^*}^{w^*} \mathrm{d}w\, p(w) \\ &= N \left[w^{1-\gamma}\right]_{\epsilon w^*}^{w^*} \\ &\sim N (w^*)^{-(\gamma-1)} \epsilon^{-(\gamma-1)} \\ &\sim \ln 2\, \epsilon^{-(\gamma-1)} \end{aligned} \tag{10}$$

so that the number of points $\widetilde{N}$ in the filtered set is independent of R, although it does depend upon ϵ.

The median of the sum W of a large number of values of w_i is estimated by noting that $W = \hat{w} + \widetilde{W}$, where $\hat{w}$ is the largest of the w_i, and $\widetilde{W}$ is the sum excluding the largest of the w_i. The value of $\widetilde{W}$ will be approximated by its mean value, which depends upon $\hat{w}$. Writing $\hat{w} = aw^*$, and taking the leading order as $N \to \infty$, $\epsilon \to 0$

$$\begin{aligned} \langle \widetilde{W} \rangle &\sim (N-1) \int_{\epsilon w^*}^{\hat{w}} \mathrm{d}w\, w p(w) \\ &\sim (N-1) \left[\frac{\gamma-1}{2-\gamma} w^{2-\gamma}\right]_{\epsilon \hat{w}}^{\hat{w}} \\ &\sim \frac{\gamma-1}{2-\gamma}(N-1)\hat{w}^{2-\gamma} \\ &\sim \frac{(\gamma-1)\ln 2}{2-\gamma} a^{2-\gamma} w^*. \end{aligned} \tag{11}$$

This gives the following estimate for W, in terms of $a = \hat{w}/w^*$:

$$W \approx w^* \left[a + a^{2-\gamma} \frac{(\gamma-1)\ln 2}{2-\gamma}\right]. \tag{12}$$

The value of W depends upon a random quantity, a. The median value of $W(a)$ is $\overline{W} = W(\overline{a})$. And because we define $w^* = \overline{\hat{w}} = w^* \overline{a}$, we have $\overline{a} = 1$. This gives the

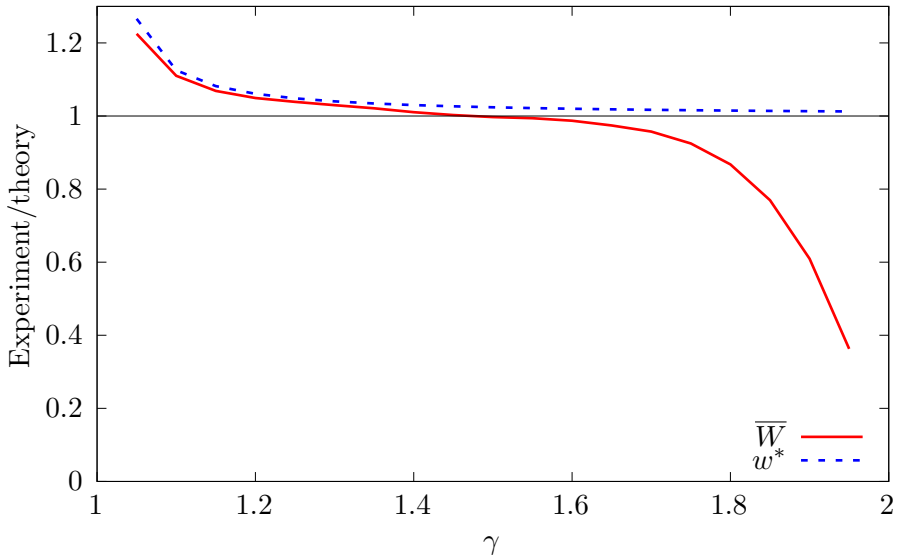


Fig. 2 Plot of ratios of values of w^* (median of the largest element) and $\overline{W}$ (median of sum of N samples) obtained from simulation, divided by their theoretical estimates, equations (9) and (13) respectively, as a function of γ. The figure shows data for $N = 1000$, averaged over $M = 10^4$ iterations. The values of $\overline{W}$ used to generate this figure span more that 58 decades.

following estimate for $\overline{W}$:

$$\overline{W} \approx \left(1 + \frac{(\gamma - 1)\ln 2}{2 - \gamma}\right)\left(\frac{N}{\ln 2}\right)^{1/(\gamma-1)}. \tag{13}$$

This indicates that $\overline{W}$ exceeds the median of the largest term by a factor which is independent of both ϵ and N (and which is therefore therefore independent of R). The independence of $\overline{W}/w^*$ upon R indicates that the filtered images are scale-invariant. The fact that this ratio does not depend upon ϵ reflects the fact that the images are dominated by the largest values of w_i. Equation (13) implies that the number of w_i, including the largest one, that make a significant contribution to $\overline{W}$ is $1 + \frac{\gamma-1}{2-\gamma}\ln 2$. When $\gamma \to 1$, there is likely to be only one w_i that dominates the filtered image. This is in accord with the *large jump principle*, discussed in [5].

The prediction for $\overline{W}$, equation (13), was tested numerically. Figure 2 shows the ratio of the empirically determined values of w^* and $\overline{W}$ to the theoretical estimates, equations (9) and (13), for $N = 1000$ with $M = 10^4$ realisations. This verifies equation (9), and shows that the N-dependence of $\overline{W}$ is the same as that of w^*. The values of $\overline{W}$ used to create figure 2 span many decades: theoretical values of $\overline{W}$ (with $N = 1000$) range from $1.58\ldots\times 10^{63}$ for $\gamma = 1.05$ to $2.99\ldots\times 10^4$ at $\gamma = 1.95$. Given this very wide range of values, figure 2 demonstrates that equation (13) is a useful approximation.

Figure 3 shows the expectation value of the fraction f_k of the contribution to W from the largest k samples, as defined by (6). Figure 3 shows the expectation values $\langle f_k \rangle$ (again using $N = 1000$ elements in the sum, and $M = 10^4$ realisations) as a

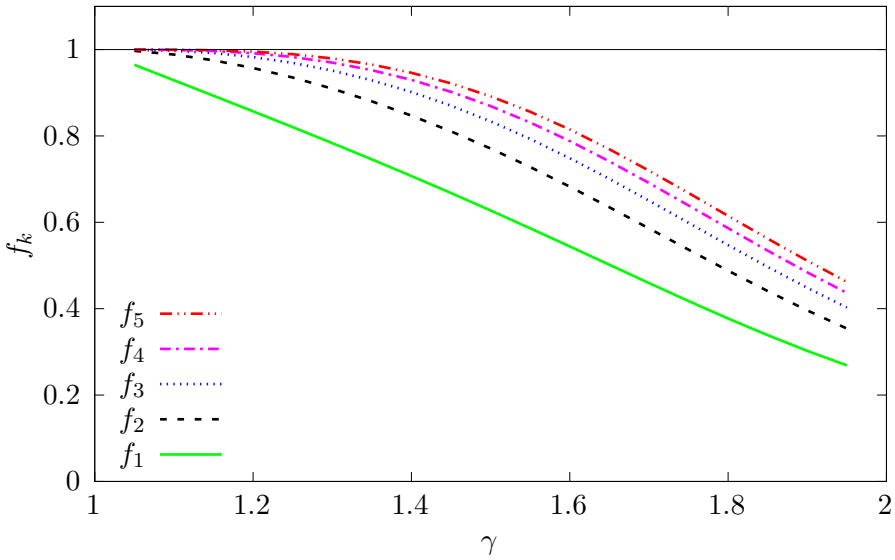


Fig. 3 Mean values of fraction of the sum W contained in its largest k elements (defined in equation (6)), as a function of γ. The number of elements of the sum was $N = 1000$, and there were $M = 10^4$ realisations.

function of γ. This verifies that, in a typical realisation, most of the contribution to W comes from a small number of the largest w_i. The fractional contribution approaches unity, in accord with the large jump principle [5], as $\gamma \to 1$

We remark that there is a further level of self-similarity in our power law model, which is concerned with varying the exponent γ. Because equation (1) implies that $y = (\gamma - 1) \ln w$ has a PDF proportional to $\exp(y)$, the ensembles for different values of γ are equivalent, if we replace w by $(\gamma - 1) \ln w$.

Our most general conclusion from this calculation follows from equation (13). When extended to d dimensions, we infer that

$$\overline{W} \sim R^{d/(\gamma-1)} \tag{14}$$

so that the apparent dimension D_{eff} which characterises the scale-invariance is given by equation (5). Note that $D_{\text{eff}} > d$. We say that this scale-invariance is *ultradimensional*. It is clearly distinguished from the self-similarity of fractal sets, where the dimension D satisfies $D < d$.

3 Diffusion model

3.1 Defining the model

We now consider a physically motivated example of a distribution of hotspots: the equilibrium probability density for diffusion in a random potential, $V(\mathbf{x})$. This example will exhibit an approximate, rather than exact, scale-invariance. Motion of a particle

is determined by a stochastic differential equation:

$$\delta x_i = -\mu \frac{\partial V}{\partial x_i} \delta t + \sqrt{2\mathcal{D}} \delta \eta_i(t), \tag{15}$$

where μ is the mobility and $\delta\eta_i(t)$ are white noise signals, independent at each timestep, satisfying $\langle \delta\eta_i \rangle = 0$ and $\langle \delta\eta_i \delta\eta_j \rangle = \delta_{ij}\delta t$. In the following we set $\mu = 1$ throughout. When $V = \text{const}$, the motion is simple diffusion with the diffusion coefficient $\mathcal{D}$. The equilibrium probability density function for the stochastic process (15) is

$$P(\mathbf{x}) = \frac{1}{\mathcal{Z}} \exp\left[-V(\mathbf{x})/\mathcal{D}\right], \tag{16}$$

where $\mathcal{Z}$ is the partition function. We shall assume that motion is confined to a finite but very large region (which we take to be a square with the side $\mathcal{R}$). When $\mathcal{D}$ is small, this density is very strongly concentrated in minima of the potential $V(\mathbf{x})$. Our aim will be to characterise the function F, defined by equation (3), for this model.

Consider the equilibrium measure when the potential $V(\mathbf{x})$ is itself a smoothly varying random function, with a Gaussian PDF, and statistics which are homogeneous and isotropic. We shall assume that $V(x, y)$ has the following statistical properties:

$$\langle V \rangle = 0, \quad \langle V^2 \rangle = 1, \quad \langle V_x^2 \rangle = \langle V_y^2 \rangle = 1, \tag{17}$$

where $V_x = \partial V/\partial x$, etc. These requirements can be satisfied by re-scaling the coordinates and the potential. Also define c by writing

$$c = \langle V_{xy}^2 \rangle. \tag{18}$$

This parameter satisfies $c \geq 1/2$, with the lower limit realised if the spectral function $S(k)$ of $V(\mathbf{x})$ (the modulus squared of the Fourier transform of its autocorrelation) has a ring spectrum, $S(k) \propto \delta(k - k_0)$. If the correlation function of V is a Gaussian, then $c = 1$.

In general, the value of $\mathcal{Z}$ depends upon the realisation of the potential $V(\mathbf{x})$, but if the scale size $\mathcal{R}$ of the region is sufficiently large, we can use ergodicity and approximate $\mathcal{Z}$ by its expectation value, $\langle \mathcal{Z} \rangle$. The partition function is then approximated by

$$\begin{aligned} \langle \mathcal{Z} \rangle &= \mathcal{R}^2 \left\langle \exp\left[-\frac{V}{\mathcal{D}}\right] \right\rangle \\ &= \frac{\mathcal{R}^2}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \mathrm{d}V \, \exp\left[-\frac{V^2}{2} - \frac{V}{\mathcal{D}}\right] \\ &= \exp\left[\frac{1}{2\mathcal{D}^2}\right] \mathcal{R}^2. \end{aligned} \tag{19}$$

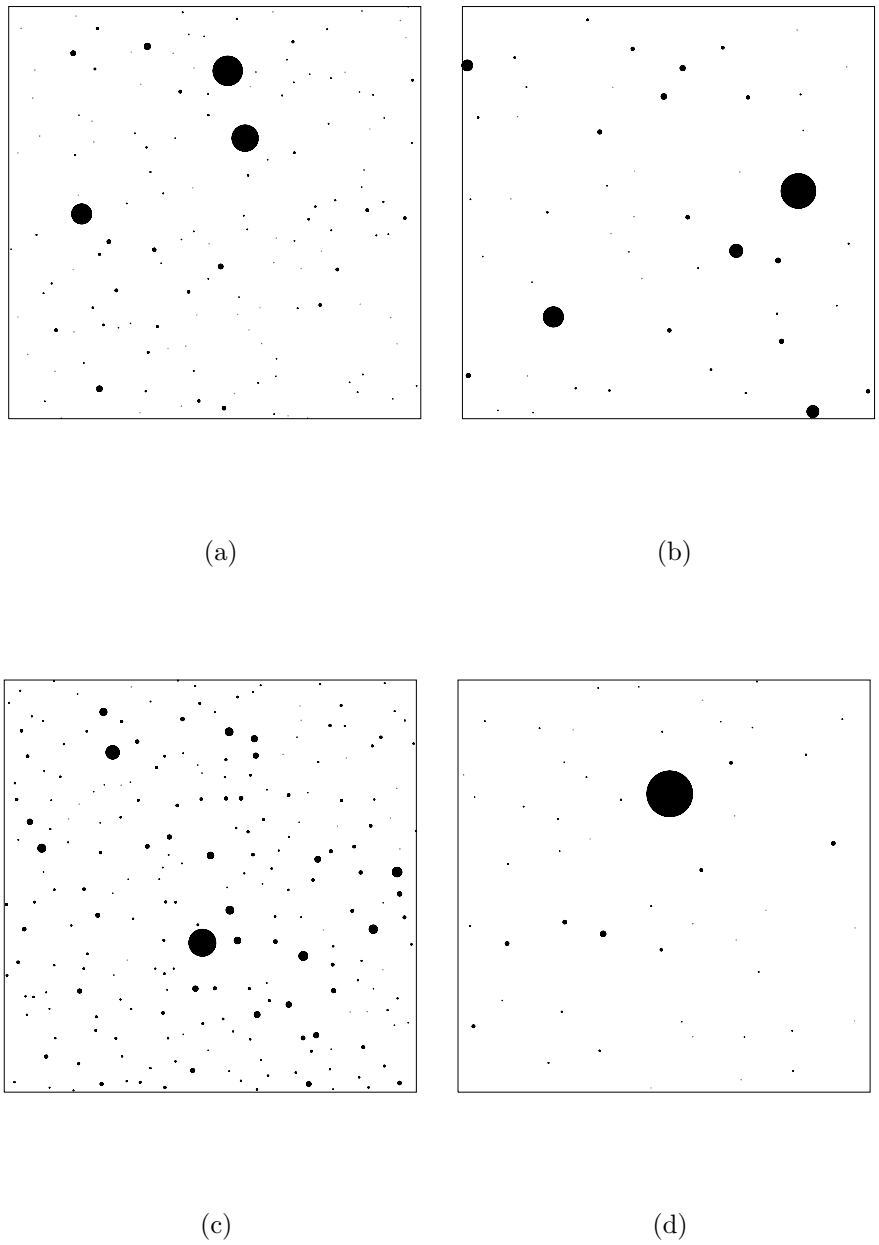


Fig. 4 Equilibrium probability density for diffusion in a two-dimensional random potential, as specified by (17) and (18) with a Gaussian correlation function, $\langle V(\mathbf{x})V(\mathbf{x}')\rangle = \exp[-(\mathbf{x}-\mathbf{x}')^2/2]$, so that $c = 1$ in (18). The presentation is the same as in figure 1: hotspots are represented by circles by the areas proportional to their weights, with 1% of the image covered. Panel (a): $\mathcal{D} = 0.25$, $R = 50$. Panel (b): $\mathcal{D} = 0.25$, $R = 25$. Panel (c): $\mathcal{D} = 0.4$, $R = 50$. Panel (d): $\mathcal{D} = 0.4$, $R = 25$.

When $\mathcal{D}$ is sufficiently small that the measure (16) is concentrated at the minima of $V(\mathbf{x})$, the weight of a hotspot is approximated by

$$w = \frac{1}{\mathcal{Z}} \int \mathrm{d}x \int \mathrm{d}y \ \exp\left[-V(x,y)/\mathcal{D}\right] \sim \frac{2\pi\mathcal{D}}{\mathcal{Z}} \Delta^{-1/2} \exp[-V^*/\mathcal{D}], \tag{20}$$

where V^* is the height of the minimum, and $\Delta = V_{xx}V_{yy} - V_{xy}^2$ is the determinant of the Hessian matrix at the minimum.

Figure 4 illustrates the distribution of the weights of the hotspots of this diffusion model, using the same presentation as figure 1 (hotspots are represented by a filled circle with the area proportional to its weight, equation (20), and the total area of circles is normalised to 1%). We used two different diffusion coefficients $\mathcal{D}$ and length-scales R. The distributions are qualitatively similar to those of the simplified model, shown in figure 1.

When $\mathcal{D}$ is small, the weights of the hotspots have a very broad distribution. The expectation value $\langle w \rangle$ is dominated by extremely rare events, which are unlikely to be realised, and it is more useful to estimate the median $\overline{W}$ of the total weight inside a region of area R^2. The growth of $\overline{W}$ as a function of R is characterised by calculating the function F defined by equation (3): $\ln \overline{W} = F(\ln R)$. It will be argued that, for this model, the large-jump principle [5] is applicable, so that $\overline{W}$ is well approximated by the median of its largest contributor, denoted by w^*.

We shall consider the following scenario. The potential $V(\mathbf{x})$ is evaluated, and the weights (20) calculated, in a region of size $\mathcal{R}$. While $\mathcal{R}$ is assumed to be large, we assume that $\mathcal{D}$ is sufficiently small that $\mathcal{Z} \ll \langle \mathcal{Z} \rangle$, so that the largest weight is $\hat{w} \approx 1$. This implies that when we estimate $\overline{W}(R)$, our estimate should satisfy $\overline{W}(\mathcal{R}) = 1$. We assume that the density of minima of $V(\mathbf{x})$ is ρ. According to equation (20), a large value of w is associated with a minimum of the potential V, which has an approximate depth $V \approx -\mathcal{D}[\ln w + \ln \mathcal{Z}]$, and we find it convenient to use a variable

$$\widetilde{V} \equiv -\mathcal{D}\left[\ln w + \ln \mathcal{Z}\right] \tag{21}$$

instead of w, because the distribution of weights has a narrow support when expressed in terms of $\widetilde{V}$. The largest values of w are observed very rarely, so we shall characterise the density of hotspots with very large values of w as follows: the probability $P(\widetilde{V}_0)$ that $\widetilde{V}$ is less than $\widetilde{V}_0$ is written in the form

$$P(\widetilde{V}) = \exp\left[-J(\widetilde{V})\right]. \tag{22}$$

(In order to unambiguously normalise this distribution we regard any minimum of $V(\mathbf{x})$ as being a hotspot). The function $J(V)$ corresponds to a 'rate function' or 'entropy function' of large deviation theory [6]. We can then estimate the median of the smallest value of $\widetilde{V}$, denoted by $\widetilde{V}^*$, by writing $1/2 = [1 - P(\widetilde{V}^*)]^{\rho R^2}$, where ρ is the density of minima. This yields:

$$J(\widetilde{V}^*) = 2\ln R + \ln \rho - \ln \ln 2. \tag{23}$$

If the inverse function of J is K (that is $K(J(V)) = V$), then the required relation between $w^* = \overline{W}$ and $\ln R$ is

$$\ln \overline{W} = -\ln \mathcal{Z} - \frac{1}{\mathcal{D}} K(2 \ln R + \ln \rho - \ln \ln 2) \equiv F(\ln R). \tag{24}$$

In order to use this expression to determine the function $F(\ln R)$ which appears in equation (3), we must determine the large-deviation rate function $J(V)$ which was introduced in equation (22).

Note that, according to equations (20) and (21),

$$\widetilde{V} = V^* + \mathcal{D} \ln \left(\frac{\sqrt{\Delta}}{2\pi \mathcal{D}} \right), \tag{25}$$

so that, in the limit as $\mathcal{D} \to 0$, $\widetilde{V} \to V$, and it is sufficient for our purposes to determine the PDF of the heights of minima of the function $V(x, y)$. We can, therefore, use the cumulative probability of the heights of local minima as the function P in equation (22).

Equations (8) and (24) imply that

$$\gamma_{\text{eff}} \sim 1 - \mathcal{D} J'(\widetilde{V}^*), \tag{26}$$

so that $\gamma_{\text{eff}} \sim 1$ when $\mathcal{D} \to 0$. This observation justifies the claim that $\hat{W} \sim w^*$.

3.2 Distribution of weights

We now turn to evaluating the distribution of heights of minima. The two-dimensional case is quite technical, so we shall start by discussing the estimate of $\overline{W}(R)$ in one dimension.

Here we require the density of local minima, ρ, and the probability $P(V)$ that the height of a local minimum is less than V. These are readily obtained using the approach developed by Rice [7], following pioneering work by Kac [8]. The density of minima is

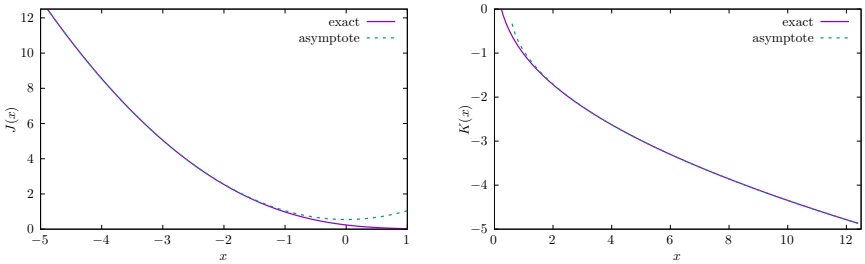
$$\rho = \int_{-\infty}^{\infty} \mathrm{d}V \int_0^{\infty} \mathrm{d}V''\, V'' P(V, 0, V''), \tag{27}$$

where $P(V, V', V'')$ is the joint PDF of V and its first two derivatives, evaluated at the same point. We consider the case where $V(x)$ is Gaussian, with correlation function

$$\langle V(x) V(x') \rangle = \exp[-(x - x')^2/2]. \tag{28}$$

We find the following non-zero statistics of the potential and its derivatives at a given point: $\langle V^2 \rangle = \langle V'^2 \rangle = 1$, $\langle V''^2 \rangle = 3$, $\langle V V'' \rangle = -1$. Using the standard formula for multivariate Gaussian distribution, we find

$$P(V, V', V'') = \frac{1}{4\pi^{3/2}} \exp \left[-\frac{1}{4} \left(3V^2 + 2VV'' + V''^2 \right) \right] \exp(-V'^2/2), \tag{29}$$



(a) (b)

Fig. 5 (a) The large deviation function for the distribution of minima, $J(V)$ (defined by equation (22)), for a one-dimensional random potential with Gaussian correlation function, $\langle V(x)V(x')\rangle = \exp[-(x-x')^2/2]$. The large deviation rate function $J(V)$ derived from the cumulative distribution of minima $P(V)$, as defined by equation (32), is shown as a solid line. Its asymptotic approximation, (33), is shown as a dashed line. (b) The inverse function, $K(J)$: exact shown as solid line, asymptote (33) is dashed line.

and hence the density of minima is

$$\rho = \frac{\sqrt{3}}{2\pi}. \tag{30}$$

The PDF of the heights of minima is

$$\begin{aligned} p(V) &= \frac{1}{\rho}\int_0^\infty \mathrm{d}V''\, V'' P(V,0,V'') \\ &= \frac{1}{\sqrt{3\pi}}\exp(-3V^2/4) - \frac{1}{2\sqrt{3}} V \exp(-V^2/2)\,\mathrm{erfc}\left(\frac{V}{2}\right) \end{aligned} \tag{31}$$

and the cumulative probability for the minimum being at a level less that V is

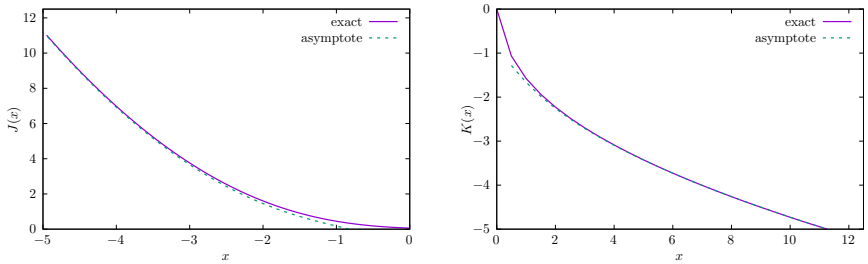
$$P(V) = \int_{-\infty}^{V} \mathrm{d}v\, p(v) = \frac{1}{2}\left[1 + \mathrm{erf}\left(\frac{\sqrt{3}V}{2}\right)\right] + \frac{1}{2\sqrt{3}}\exp(-V^2/2)\,\mathrm{erfc}\left(\frac{V}{2}\right) \tag{32}$$

which also allows us to obtain $J(V) = -\ln[P(V)]$ explicitly. The asymptote for $J(V)$ as $V \to -\infty$, and the corresponding asymptote for its inverse function $K(J)$ are:

$$J(V) \sim \frac{V^2}{2} + \frac{\ln 3}{2}, \quad K(J) \sim \sqrt{2J - \ln 3}. \tag{33}$$

The functions $J(V)$ and $K(J)$ for the one-dimensional model with Gaussian correlation function ($c = 1$) are plotted in figure 5.

In the two-dimensional case the calculation of the distribution is more difficult, but the result is already known: for the case where the correlation function is a Gaussian, the PDF of the distribution of minima is [9] (see also erratum, [10]):



(a) (b)

Fig. 6 (a) The large deviation function for the distribution of minima, $J(V)$ (defined by equaiton (22)), for a two-dimensional random potential with Gaussian correlation function, $\langle V(\mathbf{x})V(\mathbf{x}')\rangle = \exp[-(\mathbf{x}-\mathbf{x}')^2/2]$. The large deviation rate function $J(V)$ derived from the cumulative distribution of minima $P(V)$, obtained by numerical integration of equation (34), is shown as a solid line. Its asymptotic approximation, (36), is shown as a dashed line. (b) The inverse function, $K(J)$: exact shown as solid line, asymptote (36) is dashed line.

$$p(V) = \frac{\sqrt{3}}{2\pi}\left[\sqrt{\pi}\exp\left(-\frac{3}{4}V^2\right)\operatorname{erfc}\left(\frac{V}{2}\right) - V\exp\left(-V^2\right) + \sqrt{\frac{\pi}{2}}(V^2-1)\exp\left(-\frac{V^2}{2}\right)\operatorname{erfc}\left(\frac{V}{\sqrt{2}}\right)\right] \tag{34}$$

and the density of minima in two dimensions is [9]

$$\rho = \frac{1}{2\pi\sqrt{3}}. \tag{35}$$

The corresponding cumulative distribution cannot be expressed in terms of familiar special functions, so we obtained $P(V)$ by numerical integration. The asymptote can, however, be determined analytically:

$$\begin{aligned} J(V) &\sim \frac{V^2}{2} - \ln\left(\frac{V^2+1}{|V|}\right) + \frac{1}{2}\ln\left(\frac{2\pi}{3}\right) \\ K(J) &\sim \sqrt{2J + \ln\left(2J + 1\sqrt{2J}\right) + \frac{1}{2}\ln\left(\frac{3}{2\pi}\right)}. \end{aligned} \tag{36}$$

Figure 6 shows the corresponding function $J(V)$ and its inverse, compared with (36).

We investigated the statistics of hotspots for the model of diffusion in Gauss random potential (with Gaussian correlation function), by evaluating the function $F(x)$ defined by equation (3), for different values of the diffusion coefficient $\mathcal{D}$. Because we were able to perform numerical simulations over a wider range of R-values in the one-dimensional case, we present results for both the one- and two-dimensional models.

The results for the one-dimensional case are summarised in figures 7 and 8. For each value of $\mathcal{D}$, we generated $M = 20$ realisations of the random potential $V(x)$ on an interval of length $L = 100 \times 2^{18}$, by smoothing white noise using a Gaussian kernel. The partition function $\mathcal{Z}$ was calculated for each realisation, and the local minima x_i

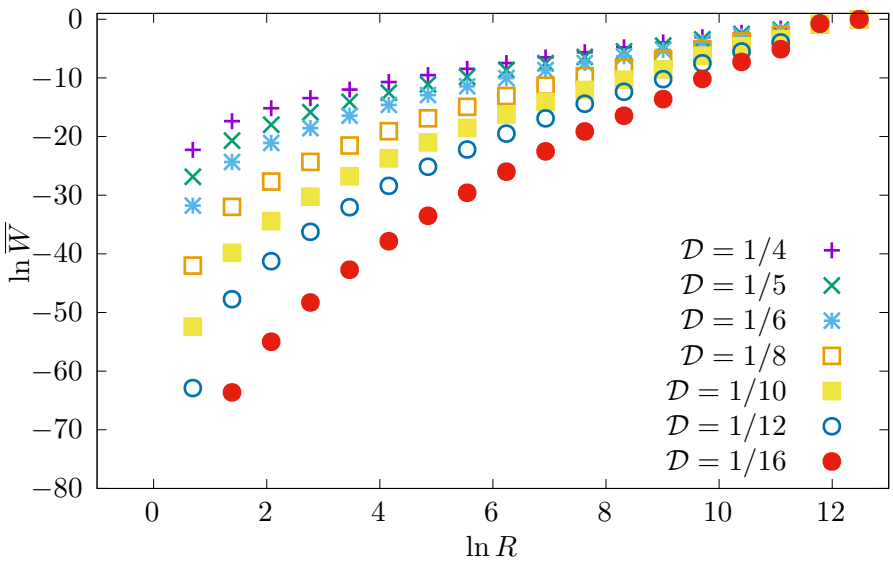


Fig. 7 Numerical investigation of the function $F(x)$ defined in equation (3) for the one-dimensional model with Gaussian correlation function: $\ln \overline{W}$ is evaluated as a function of $\ln R$, for a range of different values of $\mathcal{D}$.

of the potential were identified, together with the values of $V(x_i)$ and $V''(x_i)$. For each realisation, we divided the interval into sub-intervals, halving the length each time, for 18 generations. At generation $k = 1,\ldots,18$, for each of the $M \times 2^{k-1}$ sub-intervals of length $R = L/2^{k-1}$, we sum the weights w to determine the total weight W of each sub-interval. We then determined the median values, $\overline{W}$, of these $M \times 2^{k-1}$ weights. In figure 7 we plot the resulting 18 values of $\ln \overline{W}$ as a function of $\ln R$, for several different values of the diffusion coefficient $\mathcal{D}$.

Because the values of $\mathcal{R}$ and $\mathcal{D}$ were chosen so that the largest weights w_i were of order unity, equation (24) simplifies to

$$\ln \overline{W}(R) = \frac{1}{\mathcal{D}} K\left(\ln \mathcal{R} + \ln \rho - \ln \ln 2\right) - \frac{1}{\mathcal{D}} K\left(\ln R + \ln \rho - \ln \ln 2\right), \tag{37}$$

where ρ is the density of minima ((30) in one dimension, (35) in two dimensions). Figure 8 verifies this expression by showing a collapse of the data in figure 7 onto the inverse function of the large-deviation entropy, $K(x)$ in the one-dimensional case.

We generated $M = 4$ realisations of $V(x, y)$ on a square of size $\mathcal{R} = 256$, with toroidal boundary conditions, by convoluting a discrete representation of white noise with a Gaussian kernel. Figure 9 displays plots of $\ln \overline{W}$ as a function of $\ln R$ for the two-dimensional Gaussian potential, with different values of the diffusion coefficient $\mathcal{D}$, for $R = 2, 4, \ldots, 256$. Figure 10 illustrates the collapse of these data onto a plot of $K(x)$, the inverse of the large-deviation rate function $J(x)$. The range of values of R is much smaller than that shown in figures 7 and 8, because the two-dimensional simulations are more numerically demanding.

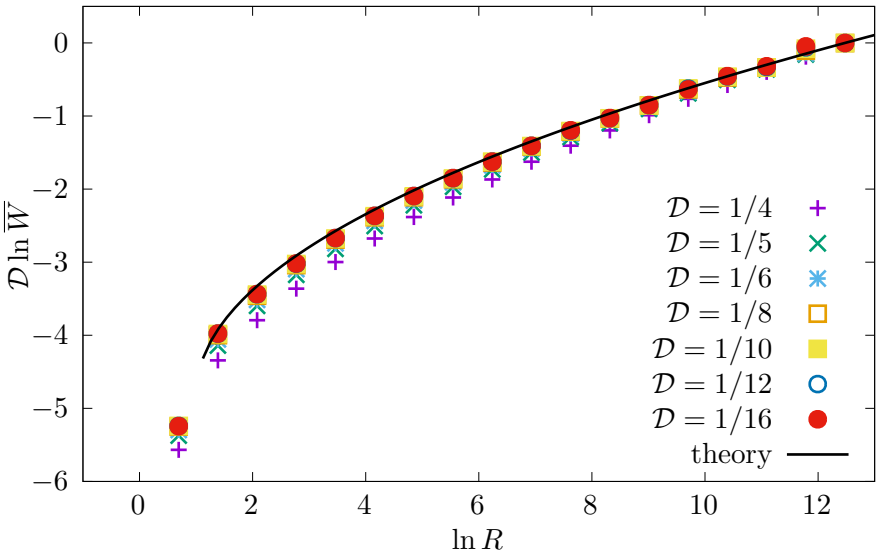


Fig. 8 The data in figure 7 collapse onto the function $K(x)$ plotted in figure 5(b), in accord with equation (37).

4 Concluding remarks

Images which show the distribution of 'hotspots', where a field has an unusually high intensity, appear to have a family resemblance, which may not be strongly dependent upon the size of the sample region.

The distribution of weights of hotspots was characterised by considering the median $\overline{W}$ of the total weight in a region of size R, and defining a function $F(x)$ by writing $\ln \overline{W} = F(\ln R)$ (equation (3)). The derivative of F is an effective dimension, $D_{\text{eff}} = F'(\ln R)$.

We investigated two models for hotspot distributions. Firstly, we considered a one-parameter family of weight distributions, defined by (1), which were contrived to be scale-invariant. The scale-invariance of these models is characterised by an effective dimension $D_{\text{eff}} = d/(\gamma - 1)$, where $\gamma \in (1, 2)$ is the parameter in the definition of the model. Because this dimension is greater than that of the embedding space, the scale invariance is distinct from the self-similarity which characterises fractal sets. Examples of realisations of this model are shown in figure 1. While it is a mathematical fact that the individual images are drawn from the same ensemble, the realisations do look very different from each other.

We also considered a physically motivated example, namely the equilibrium distribution for diffusion in a random potential. Here the realisations of the hotspot distribution, illustrated in figure 4, are qualitatively similar to those of the simplified model. We were able to determine the function $F(x)$ for this model. Because it is not a linear function, this system does not exhibit strict scale invariance.

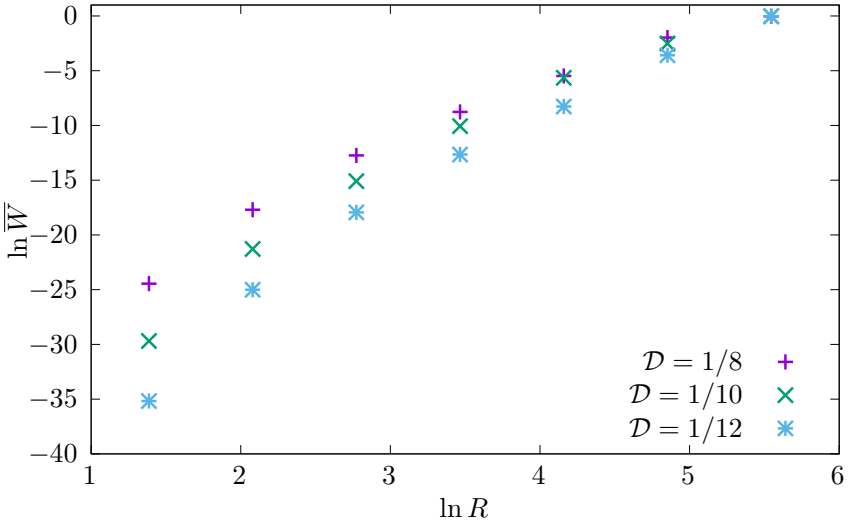


Fig. 9 Numerical investigation of the function $F(x)$ defined in equation (3) for the two-dimensional model with Gaussian correlation function: $\ln \overline{W}$ is evaluated as a function of $\ln R$, for a range of different values of $\mathcal{D}$.

Acknowledgments. This work was initiated while MW was a guest of the Chan-Zuckerberg Biohub.

Statements. No externally sourced data was processed. The programs and data used to generate the figures are available from the corresponding author. No grants were received specifically for this work, and there are no relevant financial or non-financial interests to disclose.

References

[1] Mandelbrot, B.B.: The Fractal Geometry of Nature, 3rd edn. W. H. Freeman and Comp., New York (1983)

[2] Falconer, K.: Fractal Geometry—Mathematical Foundations and Applications., Wiley, New York (1990)

[3] Halsey, T.C., Jensen, M.H., Kadanoff, L.P., Procaccia, I., Shraiman, B.I.: Fractal measures and their singularities: The characterization of strange sets. Phys. Rev. A **33**, 1141–1151 (1986) https://doi.org/10.1103/PhysRevA.33.1141

[4] Salat, H., Murcio, R., Arcaute, E.: Multifractal methodology. Physica A: Statistical Mechanics and its Applications **473**, 467–487 (2017) https://doi.org/10.1016/j.physa.2017.01.041

[5] Vezzani, A., Barkai, E., Burioni, R.: Single-big-jump principle in physical modeling. Phys. Rev. E **100**, 012108 (2019) https://doi.org/10.1103/PhysRevE.100.012108

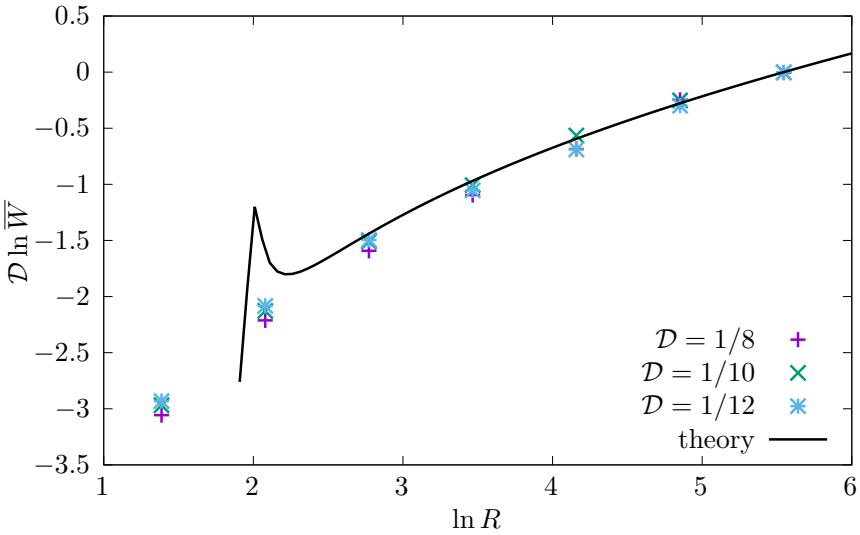


Fig. 10 The data in figure 9 collapse onto the function $K(x)$ plotted in figure 6(b), in accord with equation (37). Note that the assumptions behind asymptotics (37) break at $\ln R < 2.5$, but the data still collapse even at smaller R.

[6] Touchette, H.: The large deviation approach to statistical mechanics. Physics Reports **478**(1), 1–69 (2009) https://doi.org/10.1016/j.physrep.2009.05.002

[7] Rice, S.O.: Mathematical analysis of random noise. Bell System Technical Journal **24**(1), 46–156 (1945) https://doi.org/10.1002/j.1538-7305.1945.tb00453.x

[8] Kac, M.: On the average number of real roots of a random algebraic equation. Bulletin of the American Mathematical Society **49**, 314–320 (1943) https://doi.org/10.1090/S0002-9904-1943-07912-8

[9] Wilkinson, M., Yang, F., Austin, E.J., O'Donnell, K.P.: A statistical topographic model for exciton luminescence spectra. Journal of Physics: Condensed Matter **4**(45), 8863 (1992) https://doi.org/10.1088/0953-8984/4/45/019

[10] Wilkinson, M., Yang, F., Austin, E.J., O'Donnell, K.P.: Corrigendum. Journal of Physics: Condensed Matter **6**(16), 3123 (1994)