Multifractal properties of power-law time sequences; application to ricepiles

Romualdo Pastor-Satorras
Department of Earth, Atmospheric, and Planetary Sciences
Massachusetts Institute of Technology
77 Massachusetts Avenue, Cambridge, Massachusetts 02139

We study the properties of time sequences extracted from a self-organized critical system, within the framework of the mathematical multifractal analysis. To this end, we propose a fixed-mass algorithm, well suited to deal with highly inhomogeneous one dimensional multifractal measures. We find that the fixed mass (dual) spectrum of generalized dimensions depends on both the system size \( L \) and the length \( N \) of the sequence considered, being however stable when these two parameters are kept fixed. A finite-size scaling relation is proposed, allowing us to define a renormalized spectrum, independent of size effects. We interpret our results as an evidence of extremely long-range correlations induced in the sequence by the criticality of the system.

PACS: 64.60.Lx, 47.53.+n

I. INTRODUCTION

Self-organized criticality (SOC) has been the subject of a great deal of interest, since its introduction by Bak, Tang, and Wiesenfeld. The main feature of SOC systems is that they evolve, driven by means of an external force, into a critical state characterized by the absence of any characteristic time or length scale. The resulting extremely long range correlations show up through the peculiar “1/f” power spectrum and the geometrical fractal structure. SOC behavior has been observed in many cellular automata models of sandpiles, invasion percolation, biological evolution, depinning in random media, and also in some natural systems, such as earthquakes. Even though the first cellular automaton displaying SOC was conceived to represent the dynamics of a sandpile, the experimental evidence indicates that this is not actually the case; Real sandpiles are not in a self-organized critical state. Recently, however, Frette et al. reported SOC behavior in a real granular system, a one dimensional ricepile. For grains of rice with a considerable aspect ratio, the pile behaves critically, this fact being accounted for by the increased friction, which is able to counterbalance the inertia effects predominant in real sandpiles.

In a subsequent paper, Christensen et al. analyzed the transport properties of individual grains inside a stationary ricepile. They measured the transit time of individually colored grains of rice (tracers), defined as the time necessary for a grain to escape from the pile. Christensen et al. found that the distribution of transit times follows a truncated power law form, and that the average transport velocity of the grains diminishes as the system size increases. A cellular automaton model of a ricepile was proposed (the so-called Oslo model), reproducing the phenomenological behavior of the actual experiments. Boguñá and Corral have also suggested a theoretical scenario for the Oslo ricepile, based on a continuous time random walk model.

The main results of the Oslo experiments and simulations can be expressed through a single function, the probability distribution of transit times \( P(T, L) dT \), which is defined as the probability of a given tracer spending a time between \( T \) and \( T + dT \) inside a pile of linear size \( L \). It was found that

\[
P(T, L) \sim L^{-\nu}, \quad T < L^{\nu},
\]

\[
P(T, L) \sim T^{-\chi}, \quad T > L^{\nu},
\]

where \( \nu \) and \( \chi \) are certain characteristic exponents. The experiments provided the values \( \nu = 1.50 \pm 0.20 \) and \( \chi = 2.40 \pm 0.20 \), whereas the cellular automaton model rendered the exponents \( \nu = 1.30 \pm 0.10 \) and \( \chi = 2.22 \pm 0.10 \). This numerical evidence can be summarized in the finite-size scaling ansatz

\[
P(T, L) = L^{-\beta} f \left( \frac{T}{L^\nu} \right),
\]

with

\[
f(x) = \begin{cases} 
\text{const.} & \text{for } x < 1 \\
 x^{-\chi} & \text{for } x > 1 
\end{cases}
\]

Given that \( \chi > 2 \), and provided that the probability distribution is normalized, we have that \( \beta = \nu \) and the average value of \( T \) is finite, \( \langle T \rangle \sim L^\nu \) (see Appendix). The fact that \( \chi < 3 \) implies, however, that the second moment of the distribution is infinite, \( \langle T^2 \rangle = \infty \).

The finite-size scaling compacts the experimental data into a useful relationship, which on its turn allows one to extract valuable conclusions about the system. However, it is actually quite obvious that it is possible to extract more information about the ricepile from the sequence of transit times, apart from its distribution function. In order to gain a different insight into the problem, we propose to consider the output of the experiment from a different point of view. Let us define the set \( S(N, L) \) as follows: Throw a tracer grain in a stationary pile of linear size \( L \) and measure the time elapsed until it emerges outside within any avalanche. Performing
the same measurement for \( N \) different grains, consecutively thrown in the pile, we can construct the sequence \( S(N, L) = \{T_n\}_{n=1,...,N} \), where \( T_n \) is the time, measured in units of added grains (the slow time scale), spent inside the pile by the \( n \)-th grain, in the sequence of \( N \) consecutive throws. The set \( S(N, L) \) can be interpreted as a discrete time sequence, assigning to the instant \( n = 1, \ldots, N \) the value \( T_n \). In the Fig. 1 we have represented such a sequence, for the transit times recorded in the cellular automaton model of ricepile described in Ref. [10]. The system size is \( L = 100 \). Fig. 1(a) shows a record of 90000 transit times, whereas Fig. 1(b) depicts the 5000 points closer to the center of the previous graph. These plots show rather conclusively that not only the distribution of transit times is scale-free, but also that their sequence is in some sense self-similar.

![Fig. 1: a) Sequence of transit times for 90000 tracer grains in a computer simulation of the Oslo ricepile; system size \( L = 100 \). b) Zoom close to the central section of the previous picture.](image)

In this paper we will extract more information from the Oslo ricepile model studying the sequence of transit times \( S(N, L) \). The method we have employed is that of multifractal analysis (which, on the other hand, is not new in the field of SOC [14,15]). To this end, we have developed an algorithm particularly well suited to deal with one dimensional measures, like the ones under consideration. When computing the multifractal spectrum of the sequence \( S(N, L) \), we observe that it shows considerable size effects: The spectrum of generalized dimensions \( D(q) \) (to be defined later on) depends on the system size \( L \) and, even worse, on the sequence length \( N \). This fact seems to doom any effort to describe a single well-defined spectrum. However, by analyzing \( D(q) \) in the limit \( q \to \infty \), we observe a power law dependence on \( N \) and \( L \). Extending the scaling to the whole range of \( q \) allows us to define a "renormalized" spectrum, truly independent of size effects. We interpret our results as an effect of the extremely long ranged correlations present in the sequence, correlations induced by the criticality of the ricepile.

We have organized this manuscript as follows: In Sec. II we review the multifractal analysis of general mathematical measures, stressing the difference between fixed-size and fixed-mass formalisms. In Sec. III we develop in particular the formalism needed to deal with a discrete time sequence. Sec. IV analyses different synthetic uncorrelated random time sequences. First we check the accuracy of the algorithm against sequences of known spectra. Then we study a power-law distributed random signal, mimicking the real transit time sequences. Section V deals with our final goal, actual sequences of transit times from numerical simulations of the Oslo model. Finally, our conclusions are discussed in Sec. VI.

II. MULTIFRACTAL ANALYSIS: FIXED-SIZE VS. FIXED-MASS FORMALISM

Loosely speaking, we call multifractals the mathematical sets which can be decomposed into an infinite set of interwoven subfractals, labelled with an index \( \alpha \), each of them characterized by a different fractal dimension \( f \). The collection of these dimensions form the so-called multifractal spectrum \( f(\alpha) \). Strictly speaking, however, it is only possible to assign mathematically meaningful multifractal properties to a measure (mathematical or physical) defined over a given support. A multifractal measure is completely specified either by its multifractal spectrum \( f(\alpha) \) or by its spectrum of generalized dimensions \( D(q) \).

In this section we review the main mathematical definitions and properties of multifractal analysis.

A. General definitions

Following Ref. [19] (see also [20]), consider a normalized measure \( \mu \) defined on a support \( K \subset \mathbb{R}^d \), \( \mu(K) = 1 \). Let \( \Delta \) be an arbitrary partition of \( K \) in non-intersecting elements \( \Delta_i \), that is,

\[
K \subseteq \bigcup_i \Delta_i \quad \text{and} \quad \Delta_i \cap \Delta_j = \emptyset, \ i \neq j, \quad (2.1)
\]
and let $p_i$ and $\varepsilon_i$, $i = 1, \ldots, N$ be the variables that represent the weight factor and the size factor corresponding to the element $\Delta_i$, respectively. We define the function

$$\Phi_\Delta(q, \tau) = \left( \sum_{i=1}^{N} p_i^q \varepsilon_i^{-\tau} \right), \quad (2.2)$$

where $q$ and $\tau$ are any real numbers. The sum runs over all the $N$ disconnected parts in which we decompose the support of the measure, and the brackets stand for an average over different realizations of the measure. For any measure, either deterministic or experimental (non deterministic), we will assume that, for fine enough partitions, the function $\Phi_\Delta(q, \tau)$ collapses onto a single constant value, that is,

$$\left( \sum_{i=1}^{N} p_i^q \varepsilon_i^{-\tau} \right) = \text{const.} \quad (2.3)$$

Expression (2.3) is an implicit equation, allowing one to determine $\tau(q)$ for a given $q$, or conversely, $q(\tau)$ for a given $\tau$. If we assume a partition $\Delta$ in which $\varepsilon_i = \varepsilon = \text{const.}$, then the size factor $\varepsilon$ can be factorized from the former expression, yielding

$$\left( \sum_{i=1}^{N(\varepsilon)} p_i^q \right) \sim \varepsilon^\tau, \quad (2.4)$$

where $N(\varepsilon)$ is the number of parts of size $\varepsilon$, containing a certain measure $p_i$ different from zero. From this last expression we can compute the function $\tau(q)$ and the generalized dimensions $D(q)$ [21,22], defined by $D(q) = \tau(q)/(q - 1)$. The $f(\alpha)$ spectrum is given by the Legendre transformation $f(\alpha) = \min_q (qa - (q - 1)D(q))$ [19,23]. This approach corresponds to the so-called fixed-size multifractal formalism (FSF).

On the other hand, we can select a partition $\Delta$ in which $p_i = p = \text{const.}$, which yields to

$$\left( \sum_{i=1}^{N(p)} \varepsilon_i^{-\tau} \right) \sim p^{-q}, \quad (2.5)$$

where $N(p)$ is the number of parts of measure $p$, with a certain size $\varepsilon_i$ different from zero. From this expression we can calculate the function $q(\tau)$, and then, inverting it, compute the spectrum $D(q)$. This second approach corresponds to the so-called fixed-mass multifractal formalism (FMF).

Both FSF and FMF are completely equivalent. In order to stress this correspondence, we define the new parameters $q^* \equiv -\tau$ and $\tau^* \equiv -(q - 1)$ and substitute into Eq. (2.4). Now both equations (2.4) and (2.3) read the same, the only difference being the change of role of $p_i$ and $\varepsilon_i$. The equivalence between both formalisms is explicitly illustrated by the identities

$$q^* = -(q - 1)D(q), \quad (2.6)$$

$$D^*(q^*) = \frac{q}{1 + (q - 1)D(q)}, \quad$$

with $D^*(q^*) = \tau^*(q^*)/(q^* - 1)$.

### B. Box-counting algorithms

The most common operative numerical implementations of multifractal analysis are the so-called fixed-size box-counting algorithms [18]. For a given measure $\mu$ with support $K \subset \mathbb{R}^d$, they consider the partition sum

$$Z_\varepsilon(q) = \sum_{\mu(B) \neq 0} \left( \mu(B) \right)^q, \quad (2.7)$$

$q \in \mathbb{R}$, where the sum runs over all the different non-empty boxes $B$ of a given side $\varepsilon$ in a $\varepsilon$-grid covering the support $K$, that is

$$B = \prod_{k=1}^d \lceil l_k \varepsilon, (l_k + 1)\varepsilon \rceil , \quad (2.8)$$

$l_k$ being integer numbers. The generalized fractal dimensions of the measure are defined by the limit

$$D(q) = \frac{1}{q - 1} \lim_{\varepsilon \to 0} \frac{\log Z_\varepsilon(q)}{\log \varepsilon} \quad (2.9)$$

and numerically estimated through a linear regression of

$$\frac{1}{q - 1} \log Z_\varepsilon(q) \quad (2.10)$$

against $\log \varepsilon$.

Within this formalism, the mathematical definition (2.3) is strictly valid for positive $q$ [24]. Numerical estimates work well for $q > 1$ in $d > 2$, and render usually incorrect results for $q < 0$ [26,27]. This fact is obviously due to the presence of boxes $B$ with an unnaturally small measure, which contribute to the function $Z$ with diverging terms. In those cases, one is forced to apply different prescriptions [27,28].

The box-counting version of the fixed-mass formalism is in general harder to implement in $d > 1$ spatial dimensions. The difficulties reside in the proper selection of boxes with a given fixed measure. (For an application in $d = 2$ see Ref. [21].) From a numerical point of view, it is well known that the FMF is a good estimator of generalized dimensions for $q < 0$ (that is, $q^* > 0$, see Eq. (2.4)) and bad for $q > 0$ ($q^* < 0$). The explanation of this behavior is related to the space distribution of the measure. The FSF operates well in the dense regions of the support, whereas the FMF is specially appropriate to deal with its sparse regions.

As we will see in the next section, however, a fixed-mass algorithm is particularly simple to implement for one dimensional measures, such as time sequences.
III. MULTIFRACTAL FORMALISM FOR DISCRETE TIME SEQUENCES

Fractal geometry and multifractal analysis are well-known tools for the study of complex time signals (see for instance [28] and references therein). In this section we will specialize the box-counting multifractal analysis sketched above for the particular case of a discrete one dimensional time sequence.

We define a general discrete time sequence $T(N)$ as any set of $N$ positive real numbers, $T(N) = \{t_n\}_{n=1,\ldots,N}$, $t_n \in \mathbb{R}^+$. At this level we will not make any assumption about the possible correlations of the sequence. However, we will assume that it is the outcome of some physical process in a stationary state, and that we can obtain sequences as long as it might be required.

A. Fixed-size algorithm

In order to study the multifractal properties of a sequence $T(N)$, we must first provide a meaningful physical measure on it. As a first ansatz, we define the naive measure $\mu$ on the support $[0,N] \subset \mathbb{R}$ over which the sequence is defined. This measure assigns to a given box in $[0,N]$ a weight proportional to the sum of the value $t_n$ of all the points $n$ inside the box $B$. Namely, if $B(x,\varepsilon)$ is a ball with center in $x$ and diameter $\varepsilon$, then

$$\mu(B(x,\varepsilon)) = \frac{1}{Q(N)} \sum_{x - \frac{\varepsilon}{2} < n \leq x + \frac{\varepsilon}{2}} t_n,$$  \hspace{1cm} (3.1)

where $Q(N) = \sum_{n=1}^{N} t_n$ is a normalization factor such that $\mu([0,N]) = 1$. In order to compute the generalized dimensions $D(q)$ of $\mu$, consider a partition of $[0,N]$ into boxes of diameter $r$, in a number $N/r$, defined by

$$B_{k,r} = [\{(k-1)r, kr\}], \hspace{1cm} k = 1,\ldots,N/r, \hspace{1cm} (3.2)$$

The partition sum will then read

$$Z_r(q) = \sum_{k=1}^{N/r} (\mu(B_{k,r}))^q = \frac{1}{Q(N)^q} \sum_{k=1}^{N/r} \left( \sum_{(k-1)r < n \leq kr} t_n \right)^q. \hspace{1cm} (3.3)$$

The generalized dimensions are defined through

$$D(q) = \frac{1}{q-1} \lim_{r/N \to 0} \frac{\log \left( \frac{1}{Q(N)^q} \sum_{k=1}^{N/r} \left( \sum_{(k-1)r < n \leq kr} t_n \right)^q \right)}{\log \frac{1}{N}} \hspace{1cm} (3.4)$$

The role of $\varepsilon$ is now played by the reduced diameter of the boxes $r/N$. Numerically, we will obtain an estimate of $D(q)$ as the slope of a linear regression of

$$\frac{1}{q-1} \log \sum_{k=1}^{N/r} \left( \sum_{(k-1)r < n \leq kr} t_n \right)^q$$

against $\log(r/N)$. Note that we have dropped the normalization factor $Q(N)^q$, since it does not depend on $r$ and therefore plays no role in the regression. Moreover, the elimination of this factor results in general in a better performance of the numerical algorithm, except for those values of $q$ very close to 1.

B. Fixed-mass algorithm

In order to define a fixed-mass algorithm for a discrete sequence $T(N)$, we start by constructing an approximate Cantor set $C_T(N)$, composed by a collection of $N$ discrete points on the interval $[0,1]$. We define the dual measure $\mu^*$ by associating a mass distribution to this approximate Cantor set. The distribution corresponds to just assigning a mass unity to each one of its points. Consider thus the sequence $T(N) = \{t_n\}_{n=1,\ldots,N}$, with $Q(N) = \sum_{n=1}^{N} t_n$, and let us define the Cantor set $C_T(N)$ by

$$C_T(N) = \{x_n \mid 0 < x_n \leq 1, n = 1,\ldots,N\} \hspace{1cm} (3.6)$$

with

$$x_n = \frac{1}{Q(N)} \sum_{k=1}^{n} t_k. \hspace{1cm} (3.7)$$

We define a measure on $C_T(N)$ through the density function

$$\rho_C(x) = \frac{1}{N} \sum_{n=1}^{N} \delta(x - x_n), \hspace{1cm} (3.8)$$

where $\delta$ is the usual Dirac delta function. The measure of a ball of center $x$ and diameter $\varepsilon$, $B(x,\varepsilon) = [x - \frac{\varepsilon}{2}, x + \frac{\varepsilon}{2}]$ is given by the integral

$$\mu^*(B(x,\varepsilon)) = \int_{x - \frac{\varepsilon}{2}}^{x + \frac{\varepsilon}{2}} \rho_C(x) dx \hspace{1cm} (3.9)$$

and is equal to the number of points from $C_T(N)$ contained in the interval $B(x,\varepsilon)$. It is easy to verify that the dual measure $\mu^*$ has holes of finite size: Consider a given $t_p$ and

$$\bar{\varepsilon} < \frac{t_p}{Q(N)}, \hspace{1cm} (3.10)$$

and define

$$\bar{x} = x_{p-1} + \frac{t_p}{2Q(N)}. \hspace{1cm} (3.11)$$
Then $\mu^*(B(\bar{x}, \bar{\varepsilon})) = 0$. If $t_p$ is very large, then it will correspond to a large hole in $C_T(N)$, with a diameter $\frac{t_p}{Q(N)}$. This implies that the fractal dimension of the support of $\mu^*$ would be in general lesser than 1. These regions of zero dual measure are related to the regions of large naive measure.

We define the FSF multifractal spectrum of $\mu^*$, $D^*(q^*)$, through the partition function $Z^*_\varepsilon(q^*)$, which on its turn is defined onto the basis of a set of disjoint intervals covering $[0,1]$:

$$B_{k,\varepsilon} = [(k-1)\varepsilon, k\varepsilon], \quad k = 1, \ldots, 1/\varepsilon, \quad (3.12)$$

that is,

$$Z^*_\varepsilon(q^*) = \sum_{k=1}^{1/\varepsilon} \mu^*(B_{k,\varepsilon}) q^{q^*} = \frac{1}{Nq^*} \sum_{k=1}^{1/\varepsilon} \left[ \int_{(k-1)\varepsilon}^{k\varepsilon} \rho_C(x) dx \right] q^{q^*}. \quad (3.13)$$

The generalized dimensions are mathematically defined by the limit

$$D^*(q^*) = \frac{1}{q^* - 1} \lim_{\varepsilon \to 0} \frac{\log Z^*_\varepsilon(q^*)}{\log \varepsilon} \quad (3.14)$$

and numerically evaluated as the slope of a linear fit of

$$\frac{1}{q^* - 1} \log Z^*_\varepsilon(q^*) \quad (3.15)$$

against $\log \varepsilon$. We will drop again the normalization factor $N^*$. From a mathematical point of view, this construction represents a practical implementation of the notion of inverse multifractal measure discussed by Mandelbrot and Riedi in Ref. [31]. Let us show that $\mu^*$ indeed corresponds to the inverse of the naive measure defined on the original sequence. Consider a box $B_k$ of size $\varepsilon_k$, which contains $n_k$ points from the $C_T(N)$, and therefore has an associated dual measure

$$\mu^*(B_k) = \frac{n_k}{N}. \quad (3.16)$$

Consider that those $n_k$ points are the consecutive points $x_l, x_{l+1}, \ldots, x_{l+n_k-1}$. Assuming that the extreme points coincide with the extremes of the interval, then we have $x_{l+n_k-1} - x_l \sim \varepsilon_k$. If we recover the former definition of $x_n$, then

$$\varepsilon_k \sim x_{l+n_k-1} - x_l = \frac{1}{Q(N)} \sum_{s=1}^{l+n_k-1} t_s - \frac{1}{Q(N)} \sum_{s=1}^{l} t_s = \frac{1}{Q(N)} \sum_{s=l+1}^{l+n_k-1} t_s \simeq \mu(\hat{B}_k), \quad (3.17)$$

where $\hat{B}_k$ is a certain box, associated with the naive measure, with diameter $\hat{\varepsilon}_k \sim n_k/N$. Then we have

$$\sum_k \mu^*(B_k) q^{q^*} \varepsilon_k^{-q^*} \sim \sum_k (\frac{n_k}{N}) q^{q^*} \varepsilon_k^{-q^*} \mu(\hat{B}_k)^{-q^*} \sim \sum_k \mu(\hat{B}_k) q^{q^*} \varepsilon_k^{-q^*}. \quad (3.18)$$

**IV. NUMERICAL RESULTS FOR SYNTHETIC TIME SEQUENCES**

In this section we present our estimates for the multifractal spectrum of some synthetic (computer generated) time sequences. First we check our algorithm with two measures of known multifractal spectrum. Finally, we study the special case of a random signal whose values are distributed according to a truncated power law.

The numerical procedure for computing estimates of dimensions $D(q)$ implies the quenched average of the partition sum over an ensemble of statistically independent realizations of the signal, each one with the same length $N$. By quenched averages we refer to the mean value of the logarithm of the partition sum, $\langle \log Z_\varepsilon(q) \rangle$. As it is well-known, this kind of average is more stable and less subject to a particular sampling of scarce significance than the annealed average, which would consider the logarithm of the mean value of the partition sum, $\log (\langle Z_\varepsilon(q) \rangle)$. In order to obtain results comparable in a
straight forward way for any value of \( q \), the linear regressions to estimate \( D(q) \) are always performed over the same scaling interval, independently of the particular value of \( q \) considered.

**A. Uniform random sequence**

Firstly, we analyze a uniform random sequence \( \mathcal{R}(N, m, \sigma) \), where the different values \( t_n \) are uniform uncorrelated random variables with mean value \( m \) and standard deviation \( \sigma \). For our numerical experiments we choose \( m = 100 \) and \( \sigma = 10 \). For a smooth signal like \( \mathcal{R}(N, m, \sigma) \) we expect to obtain a flat multifractal spectrum, that is, generalized dimensions equal to unity for both naive and dual measures. This expectation is confirmed by our computations, which yield generalized dimensions very close to 1 for \( 10 < |q| \leq 40 \), independently of the particular regression to estimate \( D \). The value \( D \) is rather less complex. (see text). In our computations we averaged over 10 different approximations of the sequence. Linear regressions were performed over an interval of 2.5 decades. Error bars correspond to statistical errors from the regression algorithm. In Fig 2(b) we have plotted our numerical estimates of the dual spectrum for sequences of length \( N = 10000 \), together with the analytic spectrum \([4.1]\). The figure shows an excellent agreement between our estimates and the expected analytic result, in the whole interval of values of \( q \) considered, both positive and negative. The accuracy of the fit can be slightly improved by increasing the sequence length, but the estimates are already quite stable and correct for the value of \( N \) showed in the figure. Computations performed for the naive spectrum yielded an equally good agreement with the analytical result.

**B. Self-similar deterministic sequence**

We can construct a fully multifractal sequence starting from any self-similar deterministic multifractal measure on \( \mathbb{R} \). We considered a fixed-size measure with contraction factor \( r = 1/2 \) and probabilities \( p_1 = 0.3 \) and \( p_2 = 0.7 \) and constructed a non-normalized approximation of the measure composed by \( 1.1 \times 10^7 \) points by means of the ‘Chaos Game’. The multifractal sequence was eventually constructed by binning the sample points in \( 5 \times 10^4 \) boxes covering the interval \([0,1]\) over which the original measure was defined. The value \( t_n \) of the sequence is then given by the occupation number of the \( n \)-th box. Fig 2(a) depicts such a sequence. Its self-similarity seems obvious even to the naked eye.

The analytical dual spectrum of the sequence is given, as a function of the parameter \( s \in \mathbb{R} \), by the expression \([4.2] \)

\[
q_s = \frac{\log(p_1^s + p_2^s)}{\log(s)},
\]

\[
D(q_s) = \frac{s}{1 + \frac{\log(p_1^s + p_2^s)}{\log(s)}}.
\]

(Recall that \( D(q) \) stands now the fixed-mass spectrum of the original naive multifractal measure. The expression for its fixed-size spectrum, commonly found in the literature, is rather less complex.) In our computations we averaged over 10 different approximations of the sequence. Linear regressions were performed over an interval of 2.5 decades. Error bars correspond to statistical errors from the regression algorithm. In Fig 2(b) we have plotted our numerical estimates of the dual spectrum for sequences of\( \mathcal{R}(N, m, \sigma) \).

**C. Power-law random sequence**

The sequence of transit times seems to be distributed according to a truncated power-law of the form

\[
\rho(t, t_0) = \begin{cases} 
\text{const.} & t \in [0, t_0] \\
\frac{a}{t^{\chi}} & t \in [t_0, \infty]
\end{cases}
\]

(see Eq. \([1.1]\)). In order to explore the applicability of our algorithm to a power-law sequence, we have constructed and analyzed a synthetic random sequence \( D(N, t_0) = \{t_n\}_{n=1,...,N} \), in which each \( t_n \) is a random variable sorted according to the density \([2.2]\). In order to get results comparable to those of the transit times sequences, we will only allow values of \( \chi \) in the range...
2 < \chi < 3$. For the purposes of our computer calculations, we generate a synthetic sequence by sampling $N$ values $t_n$ according to the rule

$$
t_n = \begin{cases} 
  t_0 \frac{\eta}{\eta_0} & \text{if } \eta \leq \eta_0 \\
  t_0 \left[ \frac{1-\eta}{1-\eta_0} \right]^{-(\chi-1)} & \text{if } \eta > \eta_0
\end{cases} \quad (4.3)
$$

where $\eta$ is a uniform random number in $[0,1]$ and $\eta_0 = 1 - 1/\chi$. (See Appendix for details.) Given that each term $t_n$ of any particular realization of the sample depends linearly on $t_0$, we infer that the multifractal spectrum of the sequence will be independent of the particular cut-off $t_0$ chosen. We will report results on $D_N(q)$, the multifractal spectrum computed for an ensemble of sequences of fixed length $N$.

When computing the spectrum for any given value of $\chi \in [2,3]$, we find that for any fixed $N$, the results for different samples of the sequence do not collapse onto the same function, but are widely scattered around some average position. We explain that effect by the fact that, by construction, the signal $t_n$ has no upper bound, so that it is possible to find that, just perchance, we have generated a sample with a particular term $t_p$ extremely large, in comparison with the expected average maximum value $\langle T_M \rangle$ (that is, a rare event). Is easy to show (see Appendix) that in a sequence of $N$ random variables distributed according to a truncated power-law, the average maximum value expected scales in the limit of large $N$ as

$$
\langle T_M \rangle \sim t_0 N^{1/(\chi-1)} \quad (4.4)
$$

In order to get rid of the effect of those rare events, we proceed to compute the spectrum of a depleted sequence, in which all the values $t_n$ larger than a threshold $\overline{T}_M = t_0 N^{1/(\chi-1)}$ have been truncated to the value $\overline{T}_M$. By using this trick, we obtain stable results for all sequence lengths, collapsing onto the same average curve, within the error bars. In order to check that our particular selection of the threshold does not have an exceedingly strong effect on the computed spectra, we have repeated our calculations for different values of $\overline{T}_M$, finding always the same behavior for the generalized dimensions, even for a threshold as large as $t_0 N$. In the computations reported here, we average for each sequence length over an ensemble of 25 different realizations. Linear regressions were performed on intervals of 2 decades. Statistical error bars are all smaller than 0.01.

First of all, we observe that for $q < 0$, the dual spectra are always ill-defined, suffering from unacceptable correlation coefficients and therefore being meaningless. This fact seems to be very natural, since, as it is well-known, fixed-size algorithms render bad results for negative $q$. However, recall that what we are actually measuring is the fixed-mass spectrum of the naive measure defined in Sec. IIIA so that the fixed-size spectrum of that very measure turns out to be well behaved for negative $q$, and ill-defined for positive $q$, against all previous intuition.

The reason of this fact is the following: For negative $q$ the partition function is dominated by the sparse regions of the measure and, for positive $q$, for the dense regions. The bad behavior for $q < 0$ is a reflection of the presence on holes in the support of the dual measure, the only source of boxes with abnormally small measure. Going back to the naive measure, this means that this measure is dominated by a background of a few points with an extremely large measure (corresponding to the holes in the dual measure), which cause the break down of the algorithm for positive $q$. We claim therefore that the dual measure as defined in Sec. IIIA is the most appropriate to characterize extremely non-homogeneous series, like the power-law distribution under consideration.

In the range $q \geq 0$, for every value of $\chi$ analyzed we observe stable dual spectra, dependent on $N$, for $N > 1000$. When increasing the value of $N$, the spectrum becomes flatter and flatter. That is to say, the “multifractality” of the sequence becomes smaller and smaller, with $D_N(q) \rightarrow 1$ for any $q$, when $N \rightarrow \infty$. This result is shown in Fig. 3. A measure of the degree of multifractality (multifractality strength), of the sequence could be the expression

$$
D_N(\infty) = \lim_{q \rightarrow +\infty} D_N(q). \quad (4.5)
$$

We have computed $D_N(\infty)$ from linear regressions of the partition function computed for a value of $q$ large enough to ensure the stability of the estimates. Numerically we find that the multifractality strength is a power-law function of $N$, with an exponent dependent on $\chi$

$$
1 - D_N(\infty) \sim N^{-\gamma(\chi)}. \quad (4.6)
$$

In Fig. 4 we have plotted $1 - D_N(\infty)$ versus $N$ in log-log scale, for different values of $\chi$. The change in the slope
is evident. In Fig. 5 we represent the estimated values of \( \gamma \) as a function of \( \chi \). It is very well approximated by a linear relationship \( \gamma(\chi) \sim \chi \). Our numerical estimates of the coefficients of this relation are

\[
\gamma(\chi) = (0.48 \pm 0.01)\chi - (0.82 \pm 0.02)
\]

In the limit of infinite \( N \) we will find a flat spectrum (uniform measure); however, for any finite value of \( N \) the deterministic sequences are fully multifractal.

\[
\begin{align*}
\frac{1 - D_N(q)}{N^{-\gamma(\chi)}} &= 1 - D_R(q).
\end{align*}
\]

The renormalized spectrum \( D_R(q) \) is a universal function, independent of the length \( N \). It is an intrinsic property of the initial time sequence, independent of any particular sample, and it can be therefore regarded as its true spectrum.

In Fig. 6 we have tested the scaling ansatz (4.9) for actual computations. The best collapse is achieved for sequences with length in between \( 10^4 \) and \( 10^6 \), and for an exponent \( \gamma' = 0.265 \). The power-law sequence considered has a distribution exponent \( \chi = 2.22 \) and a predicted value \( \gamma = 0.25 \) according to Eq. (4.7), quite close to the actual value.

\[ \gamma = 0.48 \chi - 0.82 \]

**V. NUMERICAL RESULTS FOR TRANSIT TIMES SEQUENCES**

We now turn to the numerical analysis of the sequence of SOC transit times \( S(N, L) \). By construction, the value \( T_n \) is the time spent into the pile by the \( n \)-th grain in a series of \( N \) consecutive throws. It is conceivable that the landing of a tracer may provoke an avalanche which would eventually evacuate out of the pile the very tracer that caused it. In such a case, we assign a value \( T = 1 \) to the transit time of that particular tracer. We have therefore \( T_n \in [1, \infty] \). Since the computer time devoted
to any simulation is always a finite amount, one has to stop the run at some point, leaving inside the pile, with nonzero probability, some of the tracers thrown at intermediate stages of the simulation. These tracers which did not emerge at the end of the run would represent a gap in the sequence $S(N, L)$. We fill these gaps by shifting the sequence one site to the left at the points $n$ when a tracer did not come out. We have also considered sequences in which each gap was filled with a lower bound of its corresponding transit time, estimated by subtracting the time of addition of that gap to the total time that the simulation was running. The results obtained with both procedures were identical, within the error bars.

We work with sequences of total length $M = 10^6$ points from simulations of the one dimensional Oslo model of size $L = 25, 50, 100, 200, 400, 800$, and $1600$. In order to average our partition sum, we proceed to decompose the sequences into subsequences of length $N \ll M$, and perform the averages over the sample of the resulting $M/N$ subsequences. When computing the spectra, however, we find that they do not stabilize well. This is again due to the presence of rare events: In a subsequence of length $N$ there are some points with extremely large relative measure, corresponding to tracers which spent a long time inside the pile. In order to correct this effect, we proceed in the same way as we did in the random power-law signal above: We truncate the largest events up to a maximum cut-off $T_M$. In view of Eqs. (1.1) and (4.2), the SOC signal is akin to a truncated power-law distributed sequence with cut-off $t_0 \sim L^\nu$; comparing with Eq. (4.4), we select $T_M = L^{\nu} N^{1/(\chi-1)}$, with $\nu = 1.30$ and $\chi = 2.22$, according to the simulations. Our results are the spectra $D_{N,L}(q)$, computed for an ensemble of sequences of fixed length $N$, coming from a ricepile of size $L$.

With the expertise we gained from the analysis of the random power-law signal, we would expect the multifractal spectrum of any SOC sequence to be ill-defined for $q < 0$, to depend on the length $N$, and to be independent of the cut-off, that is, of the system size $L$. The first prediction turns out to be true; for $q < 0$ the poor correlation coefficients yield meaningless estimations. However, for $q > 0$ we obtain stable spectra depending on both $N$ and $L$. They show an even more striking property: the spectra decrease monotonically (become flatter) with $N$, and increase (become steeper) with $L$. This behavior is shown in Figs. 7 and 8.

![Fig. 7: Multifractal dual spectrum for SOC sequences from a ricepile of size $L = 100$. The different plots correspond to different sequence lengths; from top to bottom, $N = 10^5$, $3 \times 10^4$, $10^4$, $3 \times 10^3$, and $10^3$.](image1)

![Fig. 8: Multifractal dual spectrum for SOC sequences of length $N = 10^4$. The different plots correspond to different system sizes; from top to bottom, $L = 25$, $50$, $100$, $200$, $400$, $800$, and $1600$.](image2)

We work with sequences of total length $M = 10^6$ points

![Fig. 9: Finite-size scaling of the multifractal dual spectrum for SOC sequences.](image3)

In a similar way as we did for the synthetic signal, we proceed to investigate the degree of multifractality of the SOC sequence. Studying the same strength parameter, we find that the magnitude $1 - D_{N,L}(\infty)$ can be fitted
as a double power-law, both in $N$ and $L$, that is,

$$1 - D_{N,L}(\infty) \sim N^{-\gamma_1} L^{\gamma_2}. \tag{5.1}$$

Our estimates are $\gamma_1 = 0.27 \pm 0.02$ and $\gamma_2 = 0.32 \pm 0.02$. These results are valid in the range $N \geq 10000$ and $L \leq 400$.

The previous formula suggests again the possibility of constructing a renormalized spectrum, universal for all values of $q$, and independent of $N$ and $L$. This is done by plotting the finite-size relationship

$$1 - D_{N,L}(q) \sim N^{-\gamma_1} L^{\gamma_2} = 1 - D_R(q). \tag{5.2}$$

The validity of this scaling is checked up in Fig. 9. The plotted spectra correspond to the smaller values of $L$ and larger values of $N$ for which the relation (5.2) holds. The best collapse is obtained for effective exponents $\gamma_1' = 0.29$ and $\gamma_2' = 0.34$, very close to the ones predicted in the limit $q \to \infty$. The rescaled spectra collapse onto a unique function, which is interpreted again as a renormalized spectrum, in the sense that it is a property of the intrinsic dynamics of the ricepile where the data came from, and independent of particular samples considered when computing it.

This scaling behavior can be accounted for by the effect of the correlations inside the SOC sequence. No dependence whatsoever on the system size (the cut-off) was observed in the synthetic power-law distributed signal in Sec. IV C. The only difference between that signal and the SOC one resides in the correlations. While the different points in the synthetic sequence are completely uncorrelated by construction, the SOC transit times suffer obviously from long-range correlations. This fact is easy to realize when one considers that grains introduced into the pile at widely scattered initial times can emerge at the same instant in a single gigantic avalanche.

As a numerical experiment, we have estimated the correlation length in our SOC sequences as the minimum length $\hat{N}$ above which an $R/S$ analysis [64] provides a Hurst exponent close to 0.5. Our estimates show that for $L < 400$ the sequences become roughly uncorrelated for lengths larger that $\hat{N} = 10^4$, whereas no serious estimate can be done for $L > 400$. This result seems to be in contradiction with our multifractal scaling, since in the range of validity of Eq. (5.1), the $R/S$ analysis predicts a complete decorrelation and, hence, an independence on the system size. We interpret our results as a hint towards the existence of more deep intrinsic correlations that those revealed by a simple $R/S$ analysis.

VI. CONCLUSIONS

In this paper we have investigated the multifractal properties of sequences of transit times of individual grains inside an Oslo ricepile. To this purpose, we have developed a fixed-mass multifractal algorithm, yielding the so-called dual spectrum, particularly well-suited to deal with highly inhomogeneous one dimensional measures (in our case, time series). This is particularly for the transit time sequences, which are power-law distributed and are hence constituted at any length scale by a more or less average flat background, interspersed by relatively infrequent huge peaks.

The main result of our analysis is the finite-size scaling relation (5.2). This scaling shows a particular behavior: The dual spectrum tends to decrease when increasing the sequence length $N$, whereas it tends to increase with the systems size $L$. While the first statement is in complete agreement with numerical experiments on synthetic uncorrelated power-law sequences, the second constitutes a completely unexpected result: As we show in Sec. IV C, the spectra of an uncorrelated random power law signal do not depend on the distribution's cut-off. Since the cut-off is related to the system size of the ricepile, we should expect in the SOC case to obtain results independent of $L$. That is not the case, however, in our computations.

The renormalized spectrum defined in (5.2) allows one to get rid of those finite-size effects, and constitutes a magnitude that can be associated to the very ricepile dynamics, not influenced by the hazards of the samples used in its estimation.

We interpret the initial $L$ dependence as an effect of the extremely long correlations in the transit time sequence. As the authors point out in Ref. [10], the fact that the average speed of the tracers decreases with the system size proves that there are correlations all along the system. These correlations show up even more spectacularly when analyzing the multifractal properties of the sequences. A simple $R/S$ analysis seems to show an absence of correlations for $L < 400$ and $N > 10^4$. Hence, it could seem reasonable that, for these values of the parameters, the spectra should become independent of $L$. This is not the case, however. We conclude, therefore, that the transit time sequences indeed possess correlations of a range far larger than that possibly revealed by the $R/S$ analysis, correlations which are made evident only in our more sophisticated multifractal analysis.

ACKNOWLEDGMENTS

I am very indebted to Alvaro Corral for providing the simulational data analyzed in this paper and for many clarifying discussions at the first stages of this work. The manuscript benefited greatly from critic readings by Jens Feder, Jordi Mach, Rudolf Riedi, and Daniel H. Rothman. This work has been financially supported by a scholarship grant from the Ministerio de Educación y Cultura (Spain).
APPENDIX:

In this appendix we derive some useful properties of a truncated power-law random variable. Consider a random variable $t$ distributed according to the density $\rho(t) = \left\{ \begin{array}{ll} \alpha t_0^{-1} \frac{t}{t_0}^{-\chi} & t \in [0, t_0] \\ \alpha t_0^{-1} \left( \frac{t}{t_0} \right)^{-\chi} & t \in [t_0, \infty] \end{array} \right.$ (A1)

Continuity of the density at $t = t_0$ imposes the actual form

$$\rho(t, t_0) = \left\{ \begin{array}{ll} \alpha t_0^{-1} \frac{t}{t_0}^{-\chi} & t \in [0, t_0] \\ \alpha t_0^{-1} \left( \frac{t}{t_0} \right)^{-\chi} & t \in [t_0, \infty] \end{array} \right.$$

If $\chi > 1$, then the density is normalizable, with a normalization constant

$$a^{-1} = \int_0^{t_0} \frac{t}{t_0} dt + \int_{t_0}^{\infty} \left( \frac{t}{t_0} \right)^{-\chi} dt = \frac{\chi}{\chi - 1}.$$ (A2)

If we demand that the first moment of the distribution does exist, we have to impose $\chi > 2$ to obtain

$$\langle t \rangle = \int_0^{\infty} t \rho(t, t_0) dt = \frac{1}{2\lambda} t_0 \sim t_0,$$ (A3)

where $\lambda = (\chi - 2)/(\chi - 1)$.

The distribution function $P(t, t_0) = \int_0^t \rho(t, t_0) dt$ has the form

$$P(t, t_0) = \left\{ \begin{array}{ll} \frac{1}{\chi^{\chi-1} t_0^{\chi-2}} & t \in [0, t_0] \\ 1 - \frac{1}{\chi} \left( \frac{t}{t_0} \right)^{-\chi+1} & t \in [t_0, \infty] \end{array} \right.$$ (A4)

In order to sample a sequence according to this distribution, we use the inversion method: We equate the distribution function to a uniform random number $\eta$ and obtain the corresponding value of $t$ by inverting $P(t, t_0) = \eta$. It is easy to check that the resulting sample is given by Eq. (A5).

Consider now that we sort $N$ independent random variables according to the distribution $\rho(t, t_0)$, obtaining the sample $\{t_1, \ldots, t_N\}$. Define $T_M$ as the maximum value in this particular sample, $T_M = \max\{t_1, \ldots, t_N\}$. We want to compute the average value $\langle T_M \rangle$, weighted with the distribution (A3). It is easy to see that the probability of this maximum value being lesser or equal than $T_M$ is just equal to the probability of all the individual values $t_n$ being on their turn lesser or equal than $T_M$. This means that the distribution function of the maximum value $T_M$ is just

$$\Pi(T_M, N) = P(T_M, t_0)^N.$$ (A5)

By differentiating Eq. (A5) we get the probability density of maximum values

$$\pi(T_M, N) = \frac{d\Pi(T_M, N)}{dT_M} = \left\{ \begin{array}{ll} \frac{N}{\chi^2 t_0^{\chi-1}} \left( \frac{T_M}{t_0} \right)^{N-1} & T_M \in [0, t_0] \\ N \chi^{-1} \left( \frac{T_M}{t_0} \right)^{-\chi} \left[ 1 - \frac{1}{\chi} \left( \frac{T_M}{t_0} \right)^{-\chi+1} \right]^{N-1} & T_M \in [t_0, \infty] \end{array} \right.$$ (A6)

The average maximum value that we expect to observe in $N$ samples of the initial power law distribution will then be

$$\langle T_M \rangle = \int_{t_0}^{\infty} T_M \pi(T_M, t_0) dT_M.$$ (A7)

After substituting Eq. (A6), we obtain

$$\langle T_M \rangle = \frac{N}{\chi - 1} \left( \frac{\chi - 1}{\chi} \right)^N + \frac{N}{\lambda} \int_0^1 \left[ 1 - \frac{1}{\chi} \xi^{1/\lambda} \right]^{N-1} d\xi.$$ (A8)

In the limit $N \to \infty$, the only contribution in the last integral comes from values of $\xi$ very close to 0. We can therefore evaluate the leading behavior for large $N$ by expanding the integrand in Taylor series, keeping only the first order:

$$\int_0^1 \left[ 1 - \frac{1}{\chi} \xi^{1/\lambda} \right]^{N-1} d\xi = \int_0^1 \exp \left\{ (N - 1) \ln \left[ 1 - \frac{1}{\chi} \xi^{1/\lambda} \right] \right\} \approx \int_0^1 \exp \left\{ -(N - 1) \xi^{1/\lambda} \right\} d\xi \simeq \lambda \left( \frac{N - 1}{\chi} \right)^{-\lambda} \Gamma(\lambda).$$ (A9)

In estimating the last integral we have extended to infinity the upper limit, approximation allowed again in the limit of large $N$.

Collecting everything we get finally

$$\frac{\langle T_M \rangle}{t_0} \simeq \frac{N}{N - 1} \exp \left\{ -N \ln \frac{\chi}{\chi - 1} \right\} + \chi^{1-\lambda} \Gamma(\lambda) N (N - 1)^{-\lambda}.$$ (A10)
The first term decays exponentially. Hence in the limit of large $N$, the leading behavior is given by

$$ (T_M) \sim t_0 N^{1-\lambda} = t_0 N^{1/(\chi-1)}, $$

(A11)

up to a constant prefactor, depending only on $\chi$.

---

* Email address: romu@segovia.mit.edu

[1] P. Bak, C. Tang, and K. Wiesenfeld, Phys. Rev. Lett. 59, 381 (1987).
[2] S. Roux and E. Guyon, J. Phys. A 22, 3693 (1989).
[3] P. Bak and K. Sneppen, Phys. Rev. Lett. 71, 4083 (1993).
[4] K. Sneppen, Phys. Rev. Lett. 69, 3539 (1992).
[5] P. Bak and C. Tang, J. Geophys. Res. B 94, 15635 (1989).
[6] G. A. Held, D. H. Solomon, D. T. Keane, W. J. Haag, P. M. Horn, and G. Grinstein, Phys. Rev. Lett. 65, 1120 (1990).
[7] J. Rosendahl, M. Vekić, and J. E. Rutledge, Phys. Rev. Lett. 73, 537 (1994).
[8] J. Feder, Fractals 3, 431 (1995).
[9] V. Frette, K. Christensen, A. Malthe-Sørenssen, J. Feder, T. Jøssang, and P. Meakin, Nature 379, 49 (1996).
[10] K. Christensen, A. Corral, V. Frette, J. Feder, and T. Jossang, Phys. Rev. Lett. 77, 107 (1996).
[11] A. Corral (private communication).
[12] M. Boguña and A. Corral, Phys. Rev. Lett. 78, 4950 (1997).
[13] A ricepile is stationary when it is in a stage far apart from its initial transient regime and its average slope is constant in time.
[14] L. P. Kadanoff, S. R. Nagel, L. Wu, and S. Zhou, Phys. Rev. A 39, 6524 (1989).
[15] Z. Olami and K. Christensen, Phys. Rev. A 46, R1720 (1992).
[16] B. B. Mandelbrot, J. Fluid. Mech. 62, 331 (1974).
[17] R. Benzi, G. Paladin, G. Parisi, and A. Vulpiani, J. Phys. A 17, 3521 (1984).
[18] T. Halsey, M. Jensen, L. Kadanoff, I. Procaccia, and B. Schraiman, Phys. Rev. A 33, 1141 (1986).
[19] K. J. Falconer, J. Theoretical Probability 7, 681 (1994).
[20] J. Mach, F. Mas, and F. Sagüés, J. Phys. A 28, 5607 (1995).
[21] P. Grassberger, Phys. Lett. A 97, 227 (1983).
[22] H. G. E. Hentschel and I. Procaccia, Physica D 8, 435 (1983).
[23] R. Cawley and R. D. Mauldin, Adv. Math. 92, 196 (1992).
[24] R. H. Riedi, J. Math. Anal. Appl. 189, 462 (1995).
[25] P. Grassberger, R. Badii, and A. Politi, J. Stat. Phys. 51, 135 (1988).
[26] H. S. Greenside, A. Wolf, J. Swift, and T. Pignataro, Phys. Rev. A 25, 3453 (1982).
[27] R. Pastor-Satorras and R. H. Riedi, J. Phys. A 29, L391 (1996).
[28] M. Alber and J. Peinke (to be published).
[29] J. Lévy Véhel, Fractal Approaches in Signal Processing, in Fractal Geometry and Analysis, The Mandelbrot Festchrift, Curacao 1995, C. J. G. Evertsz, H.-O. Peitgen and R. F. Voss (eds), World Scientific, (1996).
[30] R. H. Riedi and J. Lévy Véhel, Multifractal properties of TCP traffic: a numerical study, INRIA research report 3129 (1997).
[31] B. B. Mandelbrot and R. H. Riedi, Adv. Appl. Math. 18, 50 (1997).
[32] K. J. Falconer, Fractal Geometry: Mathematical Foundations and Applications, (John Wiley and Sons, New York, 1990).
[33] M. Barnsley, Fractals Everywhere, (Academic, New York, 1988).
[34] J. Feder, Fractals, (Plenum, New York, 1988).
[35] P. Bratley, B. L. Fox, and L. E. Schrage, A Guide to Simulation, 2nd ed. (Springer-Verlag, New York, 1987).