Scaling detection in time series: Diffusion entropy analysis

Nicola Scafetta[1,2] and Paolo Grigolini[2,3,4]
[1]*Pratt School EE Department, Duke University, P.O. Box 90291, Durham, North Carolina 27708*
[2]*Center for Nonlinear Science, University of North Texas, P.O. Box 311427, Denton, Texas 76203-1427*
[3]*Dipartimento di Fisica dell'Università di Pisa and INFM, Via Buonarroti 2, 56127 Pisa, Italy*
[4]*Istituto di Biofisica CNR, Area della Ricerca di Pisa, Via Alfieri 1, San Cataldo 56010, Ghezzano-Pisa, Italy*
(Received 31 January 2002; published 25 September 2002)

The methods currently used to determine the scaling exponent of a complex dynamic process described by a time series are based on the numerical evaluation of variance. This means that all of them can be safely applied only to the case where ordinary statistical properties hold true even if strange kinetics are involved. We illustrate a method of statistical analysis based on the Shannon entropy of the diffusion process generated by the time series, called diffusion entropy analysis (DEA). We adopt artificial Gauss and Lévy time series, as prototypes of ordinary and anomalous statistics, respectively, and we analyze them with the DEA and four ordinary methods of analysis, some of which are very popular. We show that the DEA determines the correct scaling exponent even when the statistical properties, as well as the dynamic properties, are anomalous. The other four methods produce correct results in the Gauss case but fail to detect the correct scaling in the case of Lévy statistics.

DOI: 10.1103/PhysRevE.66.036130 PACS number(s): 89.75.Da, 05.40.Fb, 05.45.Tp

I. INTRODUCTION

Scale invariance has been found to hold empirically for a number of complex systems, and the correct evaluation of the scaling exponents is of fundamental importance to assess if universality classes exist [1]. The mathematical definition of scaling is as follows. The function $\Phi(r_1,r_2,\ldots)$ is termed scaling invariant, if it fulfills the property

$$\Phi(r_1,r_2,\ldots)=\gamma^a\Phi(\gamma^b r_1,\gamma^c r_2,\ldots). \tag{1}$$

Equation (1) means that if we scale all coordinates $\{r\}$ by means of an appropriate choice of the exponents $a,b,c,\ldots,$ then we always recover the same function. The theoretical and experimental search for the correct scaling exponents is intimately related to the discovery of deviations from ordinary statistical mechanics. This aspect emerges clearly, for instance, from Ref. [2]. The author of this interesting book, with the help of dimensional analysis and regularity assumption, determines the values of the scaling exponents. These scaling exponents, however, are *trivial* in the sense that they refer to ordinary statistical mechanics.

In this paper we focus on the scaling of time series, and consequently [3] on the scaling of diffusion processes. In fact, according to the prescription of Ref. [3] we interpret the numbers of a time series as generating diffusion fluctuations, thereby shifting our attention from the time series to the probability distribution function (pdf) $p(x,t)$, where x denotes the variable collecting the fluctuations. In this case, if the time series is stationary, the scaling property of Eq. (1) takes the form

$$p(x,t)=\frac{1}{t^\delta}F\left(\frac{x}{t^\delta}\right), \tag{2}$$

where δ is the scaling exponent. Ordinary statistical mechanics is intimately related to the central limit theorem (CLT) [4], thereby implying for F the Gaussian form and for the scaling exponent the value predicted by ordinary random walk theory, namely, $\delta=0.5$.

The main purpose of this paper is to prove that a technique of statistical analysis, recently introduced to establish the thermodynamic nature of a time series of sociological, astronomical, and biological interest [5–9], affords a reliable way to evaluate the scaling exponent. This method of analysis is based on the entropy of the diffusion process and for this reason is called diffusion entropy analysis (DEA). We compare the DEA to the standard deviation analysis (SDA) [10], the detrended fluctuation analysis (DFA) [3], the rescaled range analysis (RRA) [11], and to the spectral wavelet Analysis (SWA) [12]; and we show that, while all these techniques, some of which are very popular, can yield wrong scaling exponents, the DEA always determines the correct value of the scaling exponent δ, with satisfactory precision. This important conclusion is reached by examining artificial sequences generating Gauss and Lévy statistics. The DEA is the only technique yielding the correct scaling in both the Gauss and Lévy cases. The other techniques produce correct results only in the Gauss case but fail to detect the correct scaling δ in the case of Lévy statistics.

II. DIFFUSION ENTROPY ANALYSIS

It is remarkably simple to determine the scaling parameter δ using the DEA. First of all, we transform the time series into a diffusion process whose pdf $p(x,t)$ is estimated (Sec. IV illustrates an algorithm to do that). Then, we measure the Shannon entropy of the diffusion process,

$$S(t)=-\int_{-\infty}^{+\infty}dx\;\; p(x,t)\ln[p(x,t)]. \tag{3}$$

Let us suppose that $p(x,t)$ fits the scaling condition of Eq.

(2) and let us plug Eq. (2) into Eq. (3). After an easy algebra, based on changing the integration variable from x into $y = x/t^{\delta}$, we obtain

$$S(t) = A + \delta \ln(t), \tag{4}$$

where

$$A \equiv -\int_{-\infty}^{\infty} dy F(y) \ln[F(y)]. \tag{5}$$

Equation (4) means that the entropy $S(t)$ increases linearly with $\ln(t)$ and the slope δ of the resulting straight line is the scaling coefficient. The numerical search for the scaling coefficient δ is done with this property in mind. Actually, the numerical results are expressed in a linear-log scale that is equivalent to transform the fitting curves with the form $K + \delta \ln(t)$ into straight lines. It is evident that the diffusion time t depends on the time unit adopted. However, this arbitrary choice in no way affects the scaling parameter. The adoption of a different time unit would change the curve (4) into a new one, parallel to the original, and thus bearing the same scaling parameter δ.

III. GAUSS AND LÉVY DIFFUSION

This section is devoted to illustrating why, in spite of the unambiguous definition of diffusion scaling of Eq. (2), in the literature on time series analysis a misleading perspective is frequently adopted. The scaling property is usually expressed by means of

$$x \propto t^{\delta}. \tag{6}$$

The next step, adopted by many authors, rests on evaluating the second moment of the pdf $p(x,t)$, $\langle x^2(t) \rangle$. Let us see why this procedure is correct only in the Gaussian case. In the long-time limit, the variable $x(t)$ collecting the fluctuations $\xi(t)$ has a time evolution equivalent to

$$\dot{x}(t) = \xi(t). \tag{7}$$

By time integration we get

$$x(t) = x(0) + \int_0^t dt' \xi(t'). \tag{8}$$

Let us imagine a set of infinitely many trajectories of the type of that of Eq. (8). As to the second moment $\langle x^2(t) \rangle$, we evaluate its time evolution by squaring Eq. (8) and averaging over all the trajectories of this set. Under the assumption that the process is stationary and that $\langle \xi(t) \rangle = 0$, it is straightforward to obtain

$$\langle x^2(t) \rangle = \langle x^2(0) \rangle + 2\langle \xi^2 \rangle \int_0^t d\tau_1 \int_0^{\tau_1} d\tau_2 \Phi_{\xi}(\tau_2). \tag{9}$$

Note that to get this result we use the equilibrium correlation function

$$\Phi_{\xi}(t_1, t_2) \equiv \frac{\langle \xi(t_1) \xi(t_2) \rangle}{\langle \xi^2 \rangle}. \tag{10}$$

Under the stationary condition, this correlation function depends only on the time difference, namely, $\Phi_{\xi}(t_1, t_2) = \Phi_{\xi}(|t_1 - t_2|)$; and this property, with the help of a suitable change of integration variables, yields Eq. (9).

What is the connection between second moment and scaling? Having in mind Eq. (6), one would be tempted to make the conjecture that

$$\langle x^2(t) \rangle \propto t^{2\delta}. \tag{11}$$

However, this conjecture is not correct in general, and to be more rigorous let us replace Eq. (11) with

$$\langle x^2(t) \rangle \propto t^{2H}. \tag{12}$$

The adoption of the symbol H rather than the symbol δ is dictated by the following reasons. In general the second moment does not yield the correct scaling δ. Therefore, it is convenient to adopt for the information afforded by the second moment a symbol different from that used in this paper for the scaling coefficient. The authors in the field of time series analysis use the symbol H to denote scaling, having in mind the popular method of Hurst. The Hurst coefficient, in the special case of fractional Brownian motion (FBM) [13], is, quite correctly, identified with the scaling parameter. However, there is no guarantee that the exponent H is in general identical to the scaling coefficient δ. Thus, the adoption of the symbol H is also a way of warning the reader that in some cases H might be significantly different from δ. To prove under which conditions the equality $\delta = H$ applies, and consequently Eq. (11) [as well as Eq. (12)] is correct, let us notice that under the assumption that the fluctuation $\xi(t)$ is Gaussian (and with no other assumption), we can prove that the pdf $p(x,t)$ fulfills the following diffusion equation:

$$\frac{\partial p(x,t)}{\partial t} = D(t) \frac{\partial^2}{\partial x^2} p(x,t), \tag{13}$$

where

$$D(t) \equiv \langle \xi^2 \rangle \int_0^t \Phi_{\xi}(t') dt'. \tag{14}$$

The proof of this important result rests on the cumulant theory of Ref. [14], and the reader can derive it from a more general case discussed in Ref. [15]. It is straightforward to show that the general solution of Eq. (13), for a set of particles initially located at $x = 0$, is

$$p(x,t) = \frac{1}{\sqrt{2\pi \langle x^2(t) \rangle}} \exp\left(-\frac{x^2}{2\langle x^2(t) \rangle} \right), \tag{15}$$

where $\langle x^2(t) \rangle$ is the second moment with the time evolution described by Eq. (9). It is easy to show that the time

asymptotic properties of the second moment are compatible with Eq. (12), with H ranging from 0 to 1. Let us consider the case

$$\lim_{t\to\infty}\Phi_{\xi}(t)\propto\frac{\Im}{t^{\beta}}. \tag{16}$$

It is straightforward [16] to prove that $0\leqslant\beta\leqslant 1$ and $\Im=1$ yields

$$H=1-\frac{\beta}{2}. \tag{17}$$

If $\beta>1$ and $\Im=1$, we get $H=0.5$. Note that the case when the decay of the correlation function is exponential, it corresponds to $\beta=\infty$, and so again to $H=0.5$. In principle, this picture is compatible with a diffusion slower than the Brownian diffusion. This is possible when $\Im=-1$. This means that the correlation function undergoes one or more oscillations allowing it to get negative values moving from the initial positive value $\Phi_{\xi}(0)=1$. In this case Eq. (17) holds true with $1\leqslant\beta\leqslant 2$.

Thus, we see that in the asymptotic time limit the solution of Eq. (13) can be written under the form

$$p(x,t)=\frac{1}{\sqrt{4\pi\kappa t^{2H}}}\exp\left(-\frac{x^2}{4\kappa t^{2H}}\right), \tag{18}$$

with κ being a constant depending on the time series under study. The expression of Eq. (18) shows that the scaling definition of Eq. (2) is fulfilled, with $\delta=H$, while the function $F(y)$ keeps the Gaussian form of ordinary statistical mechanics. This expression coincides with the FBM prescription [13], the only important difference being that the FBM implies that the form of Eq. (18) holds true at all time scales, while the dynamical derivation, from Eq. (13), makes it true only in the asymptotic time limit.

The condition $\delta=H$, correct in the case of FBM, is violated in general. A very popular example is given by the Lévy flight [17,18]. Let us illustrate here the special case of symmetric Lévy flight. Let us consider that at any time step a one-dimensional random walker can make jumps by a distance ξ, whose probability density $\lambda(\xi)$ has the Fourier transform of the form $\hat{\lambda}(k)=exp(-b|k|^{\mu-1})$. Here $1\leqslant\mu\leqslant 3$ and b denotes the strength of the resulting diffusion process. According to the generalized central limit theorem (GCLT) [19], the resulting diffusion process yields a pdf $p(x,t)$, whose Fourier transform $\hat{p}(k,t)$ reads

$$\hat{p}(k,t)=\exp(-b|k|^{\mu-1}t). \tag{19}$$

Note that $|k|\propto 1/|x|$. Thus, Eq. (19) shows that in this case the scaling δ is given by

$$\delta=\frac{1}{\mu-1}. \tag{20}$$

On the other hand, it is known [17] that the pdf $p(x,t)$ yields slow tails proportional to $1/|x|^{\mu}$, thereby implying diverging second moments for $\mu<3$. It is evident that in this case the property $\delta=H$ is broken, and that the numerical determination itself of H is an ill-posed problem.

Instantaneous jumps of arbitrarily large intensity are somewhat unrealistic. For this reason the authors of Refs. [20–22] made the assumption that it takes the random walker a time proportional to the jump intensity to make a given jump. This process is called Lévy walk. Furthermore, the condition that the distribution of jumps intensities has already the Lévy stable form is released and replaced by an inverse-power-law distribution. To generate Lévy walk we refer here to the algorithm of Ref. [6]. This algorithm is based on drawing first the random numbers τ_i's, $i=1,2,\ldots,$ with the probability density $\psi(\tau)$ given by

$$\psi(\tau)=(\mu-1)\frac{T^{\mu-1}}{(T+\tau)^{\mu}}, \tag{21}$$

where T is a positive constant. This is the simplest analytical form ensuring at large times an inverse power law and fitting, at the same time, the normalization condition that is necessary to interpret it as a probability density. Note that we also set the physical condition $\mu>2$, ensuring the existence of a finite first moment, namely, the mean time of this distribution, $\langle\tau\rangle$. The meaning of the parameter T is made clear by the relation

$$\langle\tau\rangle=\frac{T}{\mu-2}. \tag{22}$$

This means that the parameter T serves the purpose of keeping under control the mean time $\langle\tau\rangle$, which can be made as large as we wish in two different ways, the first being making μ as close as possible to 2, and the second being assigning to T very large values that compensate an index μ significantly greater than 2, but smaller than 3. The second requirement is due to the fact that $\mu=3$ is the border with the Gaussian basin of attraction [6]. To generate Lévy walk we cannot cross this border. Note that the artificial sequences that we generate to show the different techniques of analysis in action are obtained, as we shall see in Sec. VI, from the discrete version of this algorithm. We associate each time interval τ_i to a number s_i, equal to either $+1$ or -1, according to a coin tossing prescription. We call *event* the random drawing of the pair $\{\tau_i,s_i\}$. The first event takes place at $t=0$. The random walker moves with constant velocity W, ahead or backwards, according to whether s_1 is equal to $+1$ or to -1. At time $t=\tau_1$ the random walker can change direction or keep moving in the same direction, according to whether s_2 has a sign opposite or equal to that of s_1. We keep using the same prescription at times $\tau_1+\tau_2$, $\tau_1+\tau_2+\tau_3$, and so on. We consider a time scale characterized by the property

$$t\gg\langle\tau\rangle. \tag{23}$$

It is evident that the number of events that occurred prior to a given time t is given by

$$n=\left[\frac{t}{\langle\tau\rangle}\right], \tag{24}$$

with $[a]$ denoting the integer part of a. Consequently, at a given time $t\gg\langle\tau\rangle$, the position x occupied by the random walker is given by the superposition of many highly correlated fluctuations ξ_i, of intensity W, or of n uncorrelated variables $\xi_i=\tau_i s_i$. Using the latter perspective, we have that the probability distribution function $\lambda(\xi)$ given by

$$\lambda(\xi)=\frac{1}{2W}\ \psi\left(\frac{\xi}{W}\right), \tag{25}$$

the analytical form of the function ψ being given by Eq. (21). By applying again the GCLT [19], we obtain the Lévy statistics, and, of course, the same scaling prescription as that of Eq. (20).

Lévy walk serves the very useful purpose of explaining why the emergence of the Lévy statistics does not imply a total failure of the methods relating scaling to variance. In this case, in fact, the second moment is finite, and this property does not depend on the lack of sufficient statistics. It depends on the fact that no jump can occur with a length of intensity larger than Wt. In this specific case, the renewal theory [23] prescribes that the correlation function $\Phi_\xi(t)$ is related to the waiting time distribution $\psi(\tau)$ by the important equation

$$\Phi_\xi(t)=\frac{1}{\langle\tau\rangle}\int_t^{+\infty}(t'-t)\psi(t')dt'. \tag{26}$$

From this important relation, using Eq. (21), we derive the following analytical expression for $\Phi_\xi(t)$:

$$\Phi_\xi(t)=\left(\frac{T}{t+T}\right)^\beta, \tag{27}$$

with

$$\beta=\mu-2. \tag{28}$$

At this stage we are equipped to derive the asymptotic properties of the pdf second moment. The existence of the correlation function of Eq. (27) allows us to use again Eq. (9) so as to reach quickly the conclusion that

$$H=\frac{4-\mu}{2}. \tag{29}$$

This result is, in fact, obtained by plugging Eq. (28) into Eq. (17). There is no reason to identify H with δ, in this case. Rather, if we trust the GCLT and, consequently, the scaling prescription of Eq. (20), we see that δ is related to H by

$$\delta=\frac{1}{3-2H}. \tag{30}$$

We shall prove that the DEA detects this correct scaling; the methods resting on variance cannot, even if the exponent H they detect has an interesting physical meaning. This physical meaning changes from case to case and depends on both the statistics of the process and the walking rule adopted to change the time series into a diffusion process.

IV. THE DIFFUSION ALGORITHM

In Sec. II, the DEA was illustrated adopting a continuous picture. The analysis of time series implies the adoption of a discrete picture. This is so because, in practice, we have to analyze a sequence of N numbers of ξ_i, with $i=1,\ldots,N$. We derive from this sequence the largest possible number of diffusing trajectories with the method of a mobile window of (integer) length t. In fact, we select an integer number t, fitting the condition $1\leqslant t\leqslant N$. This integer number plays the rule of the diffusion time. Therefore, for the only purpose of simplifying the notation, we adopt the same symbol "t" adopted for the continuous diffusion time in the previous sections. For any given diffusion time t, we can find $N-t+1$ subsequences defined by

$$\xi_i^{(s)}\equiv\xi_{i+s}\quad\text{with}\quad s=0,\ldots,N-t. \tag{31}$$

For any of these subsequences, we build up the sth diffusion trajectory, defined by the position

$$x^{(s)}(t)=\sum_{i=1}^{t}\xi_i^{(s)}=\sum_{i=1}^{t}\xi_{i+s}. \tag{32}$$

We can also imagine a collection of $N-t+1$ particles, with the value $x^{(s)}(t)$ denoting the position of the sth particle at the time t. In other words, we imagine $x^{(s)}(t)$ as the position of a particle that, at regular intervals of time, has been jumping forward or backward according to the values of the corresponding subsequences of Eq. (31). This means that the particle, before reaching the position that it holds at time t, has been making t jumps. The jump made at the ith step has the intensity $|\xi_i^{(s)}|$, and is forward or backward according to whether the number $\xi_i^{(s)}$ is positive or negative. We adopt a perspective inspired to Brownian motion (random walker) for the tutorial purpose of illustrating how the diffusion algorithm works. Actually, the ultimate task of this algorithm is to express in a quantitative way the departure of the observed process from the statistical properties of the ordinary Brownian motion.

We are now ready to evaluate the entropy of this diffusion process. To do that we have to partition the x axis into cells of size ϵ. We count how many particles are found in the same jth cell at a given time t. We denote this number by $N_j(t)$. Then we use this number to determine the probability that a particle can be found in the jth cell at time t, $p_j(t)$, by means of

$$p_j(t)\equiv\frac{N_j(t)}{(N-t+1)}. \tag{33}$$

At this stage the entropy of the diffusion process at time t is determined, and reads

$$S(t) = -\sum_j p_j(t)\ln[p_j(t)]. \tag{34}$$

The easiest way to proceed with the choice of the cell size ϵ is to assume it to be a fraction of the square root of the variance of the fluctuation, $\xi(i)$, and consequently, independent of t.

V. THE METHODS OF ANALYSIS BASED ON VARIANCE

In this section, to make easier for the reader to appreciate the benefits afforded by the adoption of the DEA, we review four alternative methods of time series analysis. The last three methods are very popular, and are all related (to a somewhat different extent) to the first one, based on the direct evaluation of variance.

Standard deviation analysis. SDA is probably the most natural method of variance detection. This method was used, for instance, in Ref. [10]. The starting point is given by the diffusion algorithm of Sec. IV, Eq. (32). All trajectories start from the origin $x(t=0)=0$. With increasing time t, the subsequences generate a diffusion process. At each discrete time t, it is possible to calculate the standard deviation on the trajectory position with the well-known expression

$$D(t) = \sqrt{\frac{\sum_{s=0}^{N-t}[x^{(s)}(t)-\overline{x}(t)]^2}{N-t}}, \tag{35}$$

where $\overline{x}(t)$ is the average on the positions of the $N-t+1$ particles at time t. We note that the prescription illustrated in Sec. IV to define that the trajectories of this diffusion process are based on overlapping windows. This means that the trajectories are not totally independent of one another. In principle, we can also adopt a nonoverlapping window method. In this case the largest trajectory that we can build up with the whole sequence is divided into $L=[N/t]$ smaller trajectories of size t (as done earlier, with the symbol $[a]$ denoting the integer part of a). Thus, the quantity $D(t)$ can be referred to many trajectories totally independent of each other. This is the advantage of using many nonoverlapping windows. However, nonoverlapping windows generate a number of trajectories much smaller than when using the overlapping window method, and, consequently, statistics is poorer than that with the method of overlapping windows. We prefer to work with rich statistics; therefore, in this paper we use the method of overlapping windows.

According to the traditional wisdom of the methods based on variance, the existence of scaling is assessed by observing, with numerical methods, the following property:

$$D(t) \propto t^H. \tag{36}$$

The exponent H is interpreted as the scaling exponent. As discussed in Sec. III, there is no guarantee that this exponent coincides with the genuine scaling δ. This is the reason why with all the methods of analysis of this section we shall use the symbol H to denote the result of the statistical analysis. To assess whether this is the true scaling or not, it is necessary to also use the DEA.

Rescaled range analysis. RRA was introduced by Hurst in 1965 [11], mainly for the purpose of studying the water storage of the Nile River. In t years, the average influx is

$$\langle \xi \rangle_t = \frac{1}{t}\sum_{i=1}^{t} \xi_i . \tag{37}$$

The amount of water accumulated in the reservoir in τ years is

$$x(t,\tau) = \sum_{i=1}^{\tau} (\xi_i - \langle \xi \rangle_t). \tag{38}$$

The reservoir neither overflows nor empties during the period of t years if its storage capacity is larger than the difference, $R(t)$, between the maximum and minimum amounts of water contained in the reservoir. $R(t)$ is

$$R(t) = \max_{1\le\tau\le t} x(t,\tau) - \min_{1\le\tau\le t} x(t,\tau). \tag{39}$$

For getting a dimensionless value, Hurst divided $R(t)$ by the standard deviation $S(t)$ of the data during the t years:

$$S(t) = \sqrt{\frac{1}{t}\sum_{i=1}^{t} (\xi_i - \langle \xi \rangle_t)^2}. \tag{40}$$

Hurst observed that many phenomena are very well described by the following scaling relation:

$$\frac{R(t)}{S(t)} \propto t^H. \tag{41}$$

The exponent H (denoted by the letter K by Hurst) was called Hurst exponent, and consequently denoted by the letter H, by Mandelbrot [13].

Detrended fluctuation analysis. DFA was introduced in 1994 by the authors of Ref. [3]. Since 1994, hundreds of papers analyzing fractal properties of time series with the DFA have been published. In summary, DFA works as follows. Given a time sequence $\{\xi_i\}$ $(i=1,\ldots,N)$, the DFA is based upon the following steps. First, the entire sequence of length N is integrated, thereby leading to

$$x_l = \sum_{i=1}^{l} (\xi_i - \langle \xi \rangle), \tag{42}$$

where

$$\langle \xi \rangle = \frac{1}{N}\sum_{i=1}^{N} \xi_i , \tag{43}$$

with l being an integer number whose maximum value is N. Second, the resulting time series is divided into $[N/t]$ nonoverlapping boxes. The number t, which indicates the size of the box and plays a rule analogous to the diffusion time, is an integer much smaller than N. A local trend is defined for each

box by fitting the data in the box. The linear least-squares fit may be done with a polynomial function of order $n \geqslant 0$ [24]. Let $x_l(t)$ be the local trend built with boxes of size t. Third, a detrended walk is defined as the difference between the original walk and the local trend given by the linear least-squares fit according to the following relation:

$$X_l(t) = x_l - x_l(t). \tag{44}$$

Finally, the mean squared displacement of the detrended walk is calculated, that is,

$$F_D^2(t) = \frac{1}{N} \sum_{l=1}^{N} [X_l(t)]^2. \tag{45}$$

The application of this method of analysis to a number of complex systems (see, for instance, Refs. [3,24]) shows that

$$F_D(t) \propto t^H. \tag{46}$$

Again, according to the traditional wisdom of the methods based on variance, the exponent H is considered to be a scaling exponent. Thus, the extent of the departure from the ordinary condition of Brownian motion is given by $|H - 0.5| > 0$.

Spectral wavelet analysis. SWA is a new and powerful method for studying the fractal properties of variance [12]. SWA decomposes the sample variance of a time series on a *scale-by-scale* basis. Wavelet transform makes use of scaling wavelets $\tilde{\psi}_{\tau,t}(u)$, localized in time and frequency. The wavelet frequency is given by a scaling coefficient τ that measures the width of the wavelet. The position of the wavelet in space is given by t, with u being the space variable. Two typical wavelets widely used are the Haar wavelet and the sombrero wavelet [12].

Given a signal $\xi(u)$, the continuous wavelet transform is defined by

$$W(\tau,t) = \int_{-\infty}^{\infty} \tilde{\psi}_{\tau,t}(u)\xi(u)du. \tag{47}$$

The original signal can be recovered from its continuous wavelet transform via

$$\xi(u) = \frac{1}{C_{\tilde{\psi}}} \int_0^{\infty} \left[\int_{-\infty}^{\infty} W(\tau,t)\tilde{\psi}_{\tau,t}(u)dt \right] \frac{d\tau}{\tau^2}, \tag{48}$$

where $C_{\tilde{\psi}}$ is a constant that depends on the wavelet, see Ref. [12] for details. Finally, it is possible to prove that

$$\int_{-\infty}^{\infty} \xi^2(u)du = \frac{1}{C_{\tilde{\psi}}} \int_0^{\infty} \left[\int_{-\infty}^{\infty} W^2(\tau,t)dt \right] \frac{d\tau}{\tau^2} \equiv \int_0^{\infty} S_W(\tau)d\tau. \tag{49}$$

The function $W^2(\tau,t)/\tau^2$ defines an energy density function that decomposes the energy across different scales and times. Equation (49) is the wavelet equivalent to the Parseval's theorem in the traditional Fourier analysis. The function $S_W(\tau)$, defined by Eq. (49), is the wavelet spectral density function that gives the contribution to energy at the scale τ.

From Ref. [12], we derive that SWA applied to studying a noisy sequence $\{\xi_i\}$, at the scale τ, yields

$$S_W(\tau) \propto \tau^{\alpha}. \tag{50}$$

The exponent α is related to the variance scaling exponent H in the same way as in the conventional Fourier analysis. Therefore, $\alpha = 2H - 1$ for the SWA of the noise, and $\alpha = 2H$ for the SWA of the integral of the noise. The connection with the methods of scaling detection based on variance is evident.

VI. ARTIFICIAL SEQUENCE ANALYSIS

In this section we verify the theoretical predictions of the previous sections about the pdf scaling exponent δ and the variance scaling exponent H by using artificial sequences of 5×10^6 data. With the help of artificial time series, we compare the methods of analysis based on variance with the DEA. We prove that the DEA always determines the true scaling δ, whereas the variance based methods detect the true scaling only in the Gaussian case. Thus, in the Lévy case, only the DEA reveals the true scaling.

A. Gauss statistics: fractional Brownian diffusion

Fractional Brownian diffusion is produced by fractional Gaussian noise. For historical reasons, we generate a time series $\{\xi_i\}$ of N data by using the original algorithm by Mandelbrot, which can be found in work of Feders [25]. Other more recent algorithms are suggested in Refs. [12,26]. Chosen a value of $H \in [0,1]$, let $\{\theta_i\}$ be a set of Gaussian random variables with unit variance and zero mean. The discrete fractional Brownian increment is given by

$$\xi_i = \frac{m^{-H}}{\Gamma(H+0.5)} \left\{ \sum_{j=1}^{m} j^{H-0.5} \theta_{1+m(M+i)-j} + \sum_{j=1}^{m(M-1)} [(m+j)^{H-0.5} - j^{H-0.5}] \theta_{1+m(M-1+i)-j} \right\},$$

where M is an integer that should be moderately large, and m indicates the number of the fractional steps for each unit time. In the simulation, good results are obtained with $M = 1000$ and $m = 10$. The time series $\{\xi_i\}$ is then used for generating a diffusion process with the trajectories (32).

According to the theoretical arguments of Sec. III, we expect $\delta = H$ in this case. To confirm this expectation by means of the statistical analysis, we generate five different sequences with the following values of H: (1) $H = 0.8$, (2) $H = 0.6$, (3) $H = 0.5$, (4) $H = 0.4$, and (5) $H = 0.2$. We analyze these sequences with the SDA (Fig. 1) and with the DEA (Fig. 2). The results of the numerical analysis fully confirm our expectation. Let us see all this in more detail. For illustration convenience, in Fig. 1 we plot $D(t)/D(1)$ against the diffusion time t, where $D(t)$ is the standard deviation defined by Eq. (35). With this choice, at $t = 1$ all the

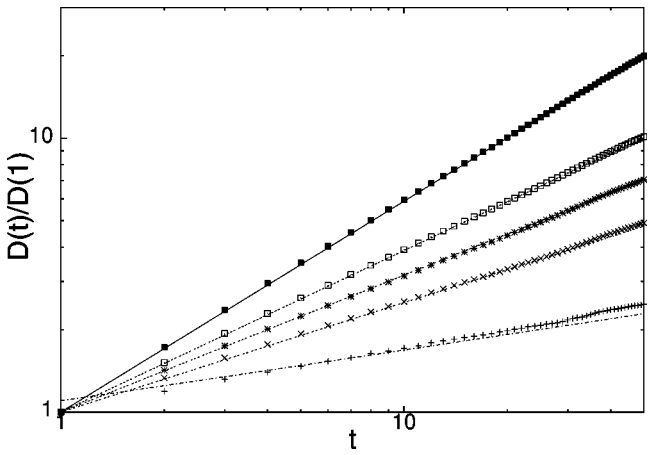

FIG. 1. SDA acting on five time series of fractional Brownian noise of Sec. VI A. We plot $D(t)/D(1)$ against the diffusion time t. The straight lines of this log-log representation are fitting functions with the form $f_D(t)=K_D t^H$. From the top to the bottom we have (1) $H=0.8$, (2) $H=0.6$, (3) $H=0.5$, (4) $H=0.4$, and (5) $H=0.2$.

numerical results yield, in the ordinate axis, the same value, equal to the unity. For the same reason, in Fig. 2, we plot the entropy difference $S(t)-S(1)$, thereby making all five numerical curves depart from the same ordinate value at $t=1$. In both figures the straight lines are the results of a fitting procedure, based on $f_D(t)=K_D t^H$ in Fig. 1, and on $f_S(t)=K_S+\delta \ln(t)$, in Fig. 2. These fitting functions become straight lines due to the log-log representation adopted in Fig. 1 and to the linear-log representation adopted in Fig. 2. The parameters H of the straight lines of Fig. 1 and the parameters δ of the straight lines of Fig. 2 coincide, curve by curve, with the actual values of H used to generate the artificial FBM sequences. The good fits of Figs. 1 and 2 prove that the condition $H=\delta$ is verified for FBM. The constants K_D and K_S are fitting parameters.

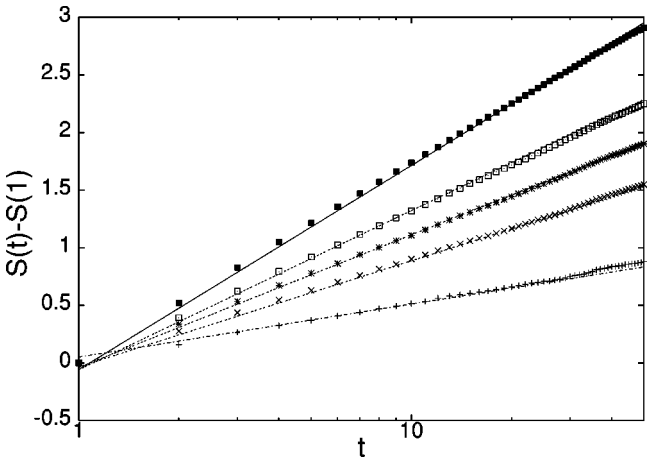

FIG. 2. DEA of the five time series of fractional Brownian noise of Sec. VI A. For illustration convenience, in ordinate we plot the entropy increment $S(t)-S(1)$ as a function of diffusion time t. The straight lines of this linear-log representation are fitting functions with the form $f_S(t)=K_S+\delta \ln(t)$. From the top to the bottom we have (1) $\delta=0.8$, (2) $\delta=0.6$, (3) $\delta=0.5$, (4) $\delta=0.4$, and (5) $\delta=0.2$.

It is remarkable that for all the values of H the parameters K_D and K_S are very close to one and zero, respectively. This is a consequence of an important property that the statistical analysis of times series should properly take into account. The short-time regime is a kind of dynamic regime and the scaling regime is a kind of thermodynamic regime. It takes time for a transition from the dynamic to the thermodynamic regime. Only in the case of a transition time equal to zero, that is, in the presence of an ideal FBM, the two fitting parameters are $K_D=1$ and $K_S=0$. Figures 1 and 2 show that for small values of H, for example $H=0.2$, the transition regime becomes more extended in time. We note that at this small value of H the diffusion process becomes significantly antipersistent. This might be a physical property where the Mandelbrot algorithm that we are adopting does not satisfactorily reproduce the ideal condition of FBM.

B. Lévy statistics: flight and walk diffusion

We generate a sequence of pairs $\{r_i,s_i\}$, with $i=1,2,\ldots$. The numbers $\{r_i\}$ are the integer part of the times τ_i generated by the algorithm of Ref. [6], yielding the probability density of Eq. (21). We select for all the artificial sequences the value $T=1$. It must be pointed out that this choice for the value of T is adopted to ensure a transition to the scaling regime as fast as possible. The asymptotic time limit predicted by the theoretical remarks of Sec. III remains unchanged. The mean time $\langle \tau \rangle$ of Eq. (22) has to be referred to a kind of effective value of T, i.e. $T_{(eff)}$, and so does the correlation function of Eq. (27). It is not worth defining the exact value of $T_{(eff)}$, since this does not have any significant consequence on the asymptotic time limit. As illustrated in Sec. III, the numbers s_i hold either $+1$ or -1, according to the coin tossing prescription. We can use the sequence $\{r_i,s_i\}$ to generate both Lévy flight and Lévy walk.

Lévy flight is obtained by changing the original sequence of pairs $\{r_i,s_i\}$ into the new sequence $\{\xi_i\}$, where $\xi_i \equiv s_i r_i$. Notice that this means that the probability density does not have the Lévy form. However, thanks to the GCLT [19], after a few time steps the resulting pdf $p(x,t)$ is expected to get the Lévy form of Eq. (19).

Lévy walk is obtained by building up the sequence $\{\xi_i\}$ as follows. We assign to the first r_1 sites of this sequence the value Ws_1, to the next r_2 sites the value Ws_2, and so on. As explained in Sec. III, the Lévy flight and the Lévy walk, in the asymptotic time limit have the same scaling, given by

TABLE I. Theoretical relation between the waiting time distribution power exponent μ, the variance scaling exponent H, and the pdf scaling exponent δ.

μ	H	δ
2.2	0.90	0.833
2.4	0.80	0.714
2.5	0.75	0.667
2.6	0.70	0.625
2.8	0.60	0.556

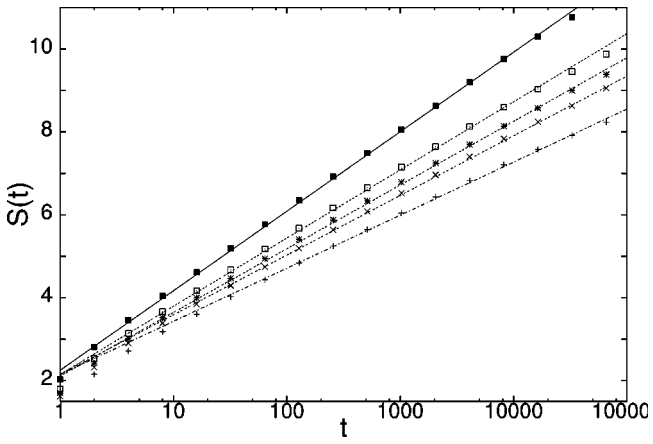

FIG. 3. DEA of the five Lévy flight time series of Sec. VI B. The straight lines of this linear-log representation are fitting functions with the form $f_S(t)=K_S+\delta\ln(t)$. From the top to the bottom we have (1) $\delta=0.833$, $K_S=2.25$; (2) $\delta=0.714$, $K_S=2.15$; (3) $\delta=0.667$, $K_S=2.11$; (4) $\delta=0.625$, $K_S=2.15$; and (5) $\delta=0.556$, $K_S=2.15$.

Eq. (20). However, the Lévy walk is expected to result in a given H, predicted by Eq. (29).

For the illustration purposes of this paper, we realize five sequences to generate Lévy flight and five sequences to generate Lévy walk. This is done by assigning to the distribution of Eq. (21) the following values for μ: (1) $\mu=2.8$, (2) $\mu=2.6$, (3) $\mu=2.5$, (4) $\mu=2.4$, and (5) $\mu=2.2$. In Table I we have reported for reader's convenience the values of δ and H, which according to the theory of Sec. III correspond to each of the five values of μ used for the numerical check. The theoretical prescriptions used are Eq. (22) for δ and Eq. (31) for H.

Figure 3 shows the DEA at work on the time series generating Lévy flight. The straight lines are fitting functions of the form $f_S(t)=K_S+\delta\ln(t)$. As in Fig. 2, these fitting functions become straight lines due to the adoption of a linear-log

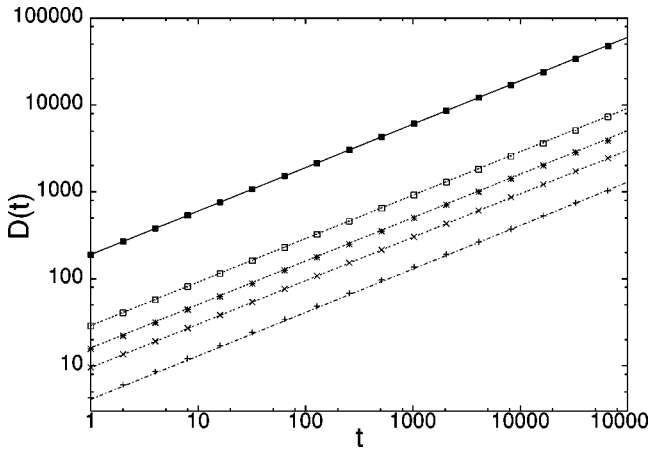

FIG. 4. SDA of the five Lévy flight time series of Sec. VI B. The straight lines of this log-log representation are fitting functions with the form $f_D(t)=K_D\,t^H$. From the top to the bottom we have (1) $H=0.5$, $K_D=190$; (2) $H=0.5$, $K_D=29$; (3) $H=0.5$, $K_D=16$; (4) $H=0.5$, $K_D=9.5$; and (5) $H=0.5$, $K_D=4.1$.

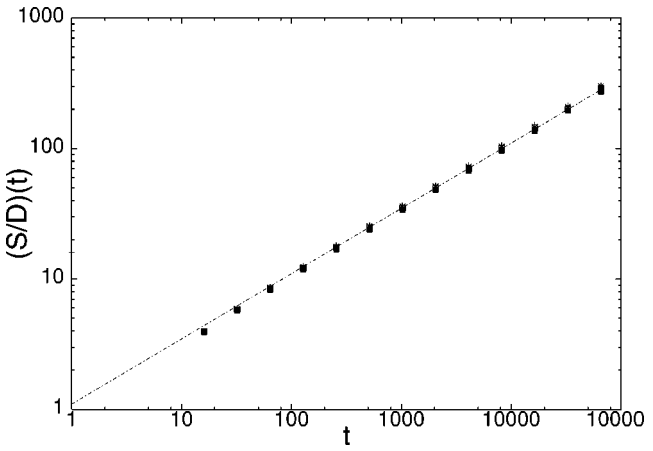

FIG. 5. RRA of the five Lévy flight time series of Sec. VI B. All the five cases fitted by the straight line of this log-log representation are fitting functions with the form $f_D(t)=K_Dt^H$, with $K_D=1.1$ and $H=0.5$.

representation. The values of the parameters δ coincide with the theoretical prediction of Table I. Figures 4 and 5 illustrate the results of the SDA and RRA, respectively, applied to the same five time series of Fig. 3. For these figures we adopt a log-log representation, and consequently fitting functions with the form $f_D(t)=K_Dt^H$ that become straight lines in this representation. Both figures yield for H a value independent of μ. This value is $H=0.5$ in all cases. According to the traditional wisdom, this would suggest the wrong conclusion that we are in the presence of ordinary Brownian motion. We are not, and the DEA is warning us that this would be a wrong conclusion. The reason for this misleading result is that these techniques are determined by both the finite value of the variance, due to statistical limitation, and the memoryless nature of the sequence $\{r_i\}$. The smaller the parameter μ, the smaller the variance, as shown by Fig. 4. The RRA

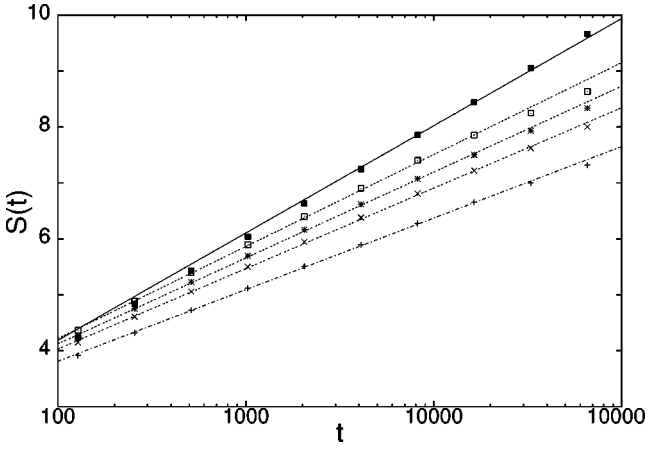

FIG. 6. DEA of the five Lévy walk time series of Sec. VI B. The straight lines of this linear-log representation are fitting functions with the form $f_S(t)=K_S+\delta\ln(t)$. From the top to the bottom we have (1) $\delta=0.833$, $K_S=0.35$; (2) $\delta=0.714$, $K_S=0.93$; (3) $\delta=0.667$, $K_S=1.05$; (4) $\delta=0.625$, $K_S=1.15$; and (5) $\delta=0.556$, $K_S=1.25$.

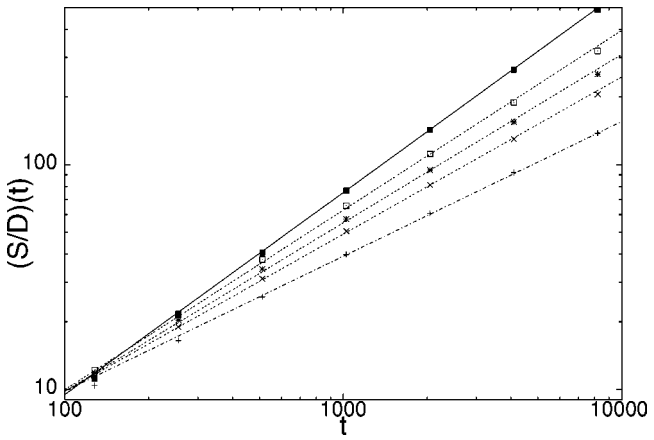

FIG. 7. RRA of the five Lévy walk time series of Sec. VI B. The straight lines of this log-log representation are fitting functions with the form $f_D(t)=K_D t^H$. From the top to the bottom we have (1) $H=0.9$, $K_D=0.15$; (2) $H=0.8$, $K_D=0.25$; (3) $H=0.75$, $K_D=0.75$; (4) $H=0.7$, $K_D=0.39$; and (5) $H=0.6$, $K_D=0.62$.

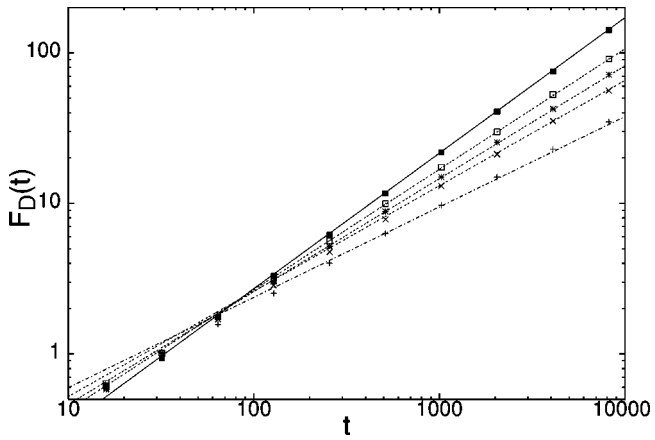

FIG. 9. DFA of the five Lévy walk time series of Sec. VI B. The straight lines of this log-log representation are fitting functions with the form $f_D(t)=K_D\, t^H$. From the curve with highest slope to that with lowest slope we have (1) $H=0.9$, $K_D=0.043$; (2) $H=0.8$, $K_D=0.067$; (3) $H=0.75$, $K_D=0.082$; (4) $H=0.7$, $K_D=0.104$; and (5) $H=0.6$, $K_D=0.15$.

eliminates this spreading, due to the fact that it normalizes the data by dividing by the standard deviation.

Figures 6–10 refer to the time series generating Lévy walk. Figure 6 illustrates the result of the DEA. As in Figs. 2 and 3, the straight lines are fitting functions of the form $f_S(t)=K_S+\delta\ln(t)$, made linear by the adoption of a linear-log representation, and, again the parameters δ coincide with the theoretical prediction of Table I. Figures 7, 8, and 9 illustrate the results of RRA, SDA, and DFA, respectively. For all these figures we adopt the log-log representation, and, consequently we change into straight lines the fitting functions with the form $f_D(t)=K_D\, t^H$. Finally, Fig. 10 shows the results of SWA. The SWA is made upon the integral of the signal and in the ordinate the square root of $S_W(\tau)$ is plotted. In this way, we can adopt fitting functions with the form $f_W(\tau)=K_W\tau^H$ for Fig. 7. In all four cases the parameter H corresponds to the theoretical value of H of Table I. Still more important than this, is the fact that for the same μ all four techniques yield the same value of H, as the fitting curves show, thereby supporting our conviction that they are different forms of the same technique of analysis, and that this technique of analysis is reliable only in the FBM case. On the other hand, we notice that the values of δ and the values of H reported in Table I fit the condition of Eq. (30), and this is a strong evidence that the statistics generated by the time series is Lévy statistics. This means that the disagreement between the scaling exponent δ detected by the DEA and the exponent H detected by the variance techniques of analysis can be used for the important purpose of defining the nature of statistics generated by strange kinetics.

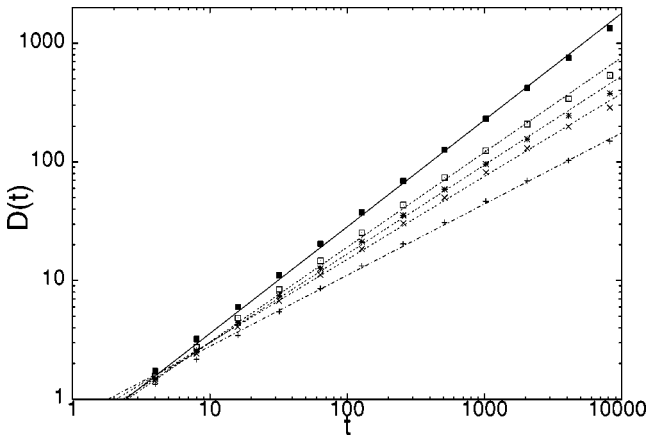

FIG. 8. SDA of the five Lévy walk time series of Sec. VI B. The straight lines of this log-log representation are fitting functions with the form $f_D(t)=K_D t^H$. From the curve with highest slope to that with lowest slope we have (1) $H=0.9$, $K_D=0.45$; (2) $H=0.8$, $K_D=0.48$; (3) $H=0.75$, $K_D=0.53$; (4) $H=0.7$, $K_D=0.6$; and (5) $H=0.6$, $K_D=0.7$.

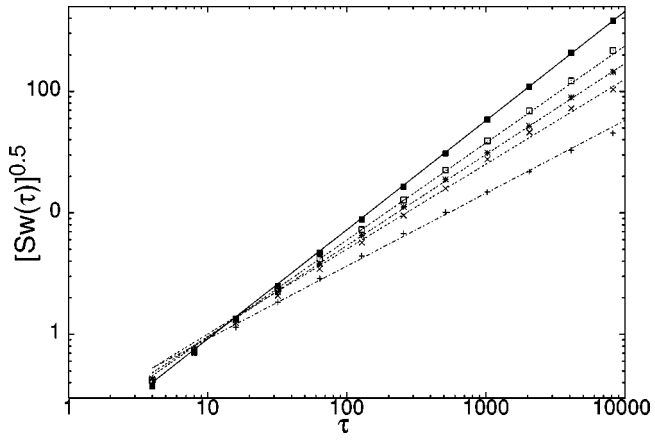

FIG. 10. SWA of the five Lévy walk time series of Sec. VI B. The straight lines of this log-log representation are fitting functions with the form $f_W(\tau)=K_W\tau^H$. From the curve with highest slope to that with lowest slope we have (1) $H=0.9$, $K_W=0.115$; (2) $H=0.8$, $K_W=0.15$; (3) $H=0.75$, $K_W=0.17$; (4) $H=0.7$, $K_W=0.2$; and (5) $H=0.6$, $K_W=0.23$. The wavelet spectral density is calculated using the maximum overlap discrete wavelet transform [8] with the Daubechies $H4$ discrete wavelet.

VII. SIGNIFICANCE OF THE RESULTS OBTAINED

This paper affords the compelling evidence that the DEA is the only method leading in all conditions to the detection of the correct scaling exponent δ. In the case of a sequence of random numbers, which according to the GCLT should result in an anomalous scaling, the popular Hurst method would lead to the wrong conclusion that the process observed is equivalent to the ordinary Brownian motion. All the traditional methods would lead to quite correct conclusions only in the case of Gaussian statistics, a condition that does not mean, of course, ordinary Brownian diffusion, as made evident by the FBM theory of Mandelbrot. It is also evident that these traditional methods ought not to be abandoned, even if they have to be used with caution. The results of Sec. VI B prove that the departure of δ from H is a clear indication of the occurrence of Lévy statistics. More generally, the departure of the traditional methods from the DEA finding might be used to shed light on statistics as well as on dynamics. Paraphrasing the title of a recent paper [27], ''Do strange kinetics imply unusual thermodynamics?'', we can say that one of the basic problems concerning complex systems is that of establishing if anomalous diffusion (strange kinetics) is compatible or not with ordinary Gaussian distribution (orinary thermodynamics). In statistical mechanics, thermodynamics is used to establish the most plausible form of equilibrium distribution, thereby implying that the transition from an out-of-equilibrium initial condition to the final equilibrium condition is thought of as a transition from dynamics to thermodynamics. We consider the scaling regime as a form of equilibrium, and consequently as a thermodynamic equilibrium. If we look at the results of this paper from within this perspective, we can conclude that FBM is an example of strange kinetics compatible with ordinary thermodynamics. We can thus conclude that the joint use of DEA and techniques of analysis based on variance can assess when strange kinetics force the complex system to depart from ordinary thermodynamics.

[1] H.E. Stanley, L.A.N. Amaral, P. Gopikrishnan, P.Ch. Ivanov, T.H. Keitt, and V. Plerou, Physica A **281**, 60 (2000).

[2] N. Goldenfeld, *Lectures on Phase Transitions and the Renormalization Group* (Perseus, Reading, MA, 1985).

[3] C.-K. Peng, S.V. Buldyrev, S. Havlin, M. Simons, H.E. Stanley, and A.L. Goldberger, Phys. Rev. E **49**, 1685 (1994).

[4] A.I. Khinchin, *Mathematical Foundations of Statistical Mechanics* (Dover Publications, Inc., New York, 1949).

[5] N. Scafetta, P. Hamilton, and P. Grigolini, Fractals **9**, 193 (2001).

[6] P. Grigolini, L. Palatella, and G. Raffaelli, Fractals **9**, 439 (2001).

[7] P. Grigolini, D. Leddon, and N. Scafetta, Phys. Rev. E **65**, 046203 (2002).

[8] N. Scafetta, V. Latora, and P. Grigolini, Phys. Rev. E (to be published).

[9] N. Scafetta, V. Latora, and P. Grigolini, Phys. Lett. A (to be published).

[10] P. Allegrini, M. Barbi, P. Grigolini, and B.J. West, Phys. Rev. E **52**, 5281 (1995).

[11] H.E. Hurst, R.P. Black, and Y.M. Simaika, *LongTerm Storage: An Experimental Study* (Constable, London, 1965).

[12] D.B. Percival and A.T. Walden, *Wavelet Methods for Time Series Analysis* (Cambridge University Press, Cambridge, England, 2000).

[13] B.B. Mandelbrot, *The Fractal Geometry of Nature* (Freeman, New York, 1983).

[14] R. Kubo, M. Toda, and N. Hashitsume, *Statistical Physics II* (Springer-Verlag, Berlin, 1991).

[15] M. Annunziato, P. Grigolini, and J. Riccardi, Phys. Rev. E **61**, 4801 (2000).

[16] G. Trefan, E. Floriani, B.J. West, and P. Grigolini, Phys. Rev. E **50**, 2564 (1994).

[17] J. Klafter, M.F. Shlesinger, and G. Zumofen, Phys. Today **49**(2), 33 (1996).

[18] M.F. Shlesinger, Nature (London) **411**, 641 (2001).

[19] B.V. Gnedenko and A.N. Kolmogorov, *Limit Distributions for Sum of Independent Random Variables* (Addison-Wesley, Reading, MA, 1954).

[20] T. Geisel, J. Nierwetberg, and A. Zacherl, Phys. Rev. Lett. **54**, 616 (1985).

[21] M.F. Shlesinger and J. Klafter, Phys. Rev. Lett. **54**, 2551 (1985).

[22] M.F. Shlesinger, B.J. West, and J. Klafter, Phys. Rev. Lett. **58**, 1100 (1987).

[23] T. Geisel and S. Thomae, Phys. Rev. Lett. **52**, 1936 (1984).

[24] K. Hu, P.Ch. Ivanov, Z. Chen, P. Carpena, and H.E. Stanley, Phys. Rev. E **64**, 011114 (2001).

[25] J. Feders, *Fractals* (Plenum Publishers, New York, 1988).

[26] H.A. Makse, S. Havlin, M. Schwartz, and H.E. Stanley, Phys. Rev. E **53**, 5445 (1996).

[27] I.M. Sokolov, J. Klafter, and A. Blumen, Phys. Rev. E **64**, 021107 (2001).