

Distinguishing chaos from noise by scale-dependent Lyapunov exponentJ. B. Gao,^{1,*} J. Hu,¹ W. W. Tung,² and Y. H. Cao³¹*Department of Electrical and Computer Engineering, University of Florida, Gainesville, Florida 32611, USA*²*Department of Earth and Atmospheric Sciences, Purdue University, West Lafayette, Indiana 47907, USA*³*BioSieve, 1026 Springfield Drive, Campbell, California 95008, USA*

(Received 27 March 2006; revised manuscript received 31 July 2006; published 13 December 2006)

Time series from complex systems with interacting nonlinear and stochastic subsystems and hierarchical regulations are often multiscaled. In devising measures characterizing such complex time series, it is most desirable to incorporate explicitly the concept of scale in the measures. While excellent scale-dependent measures such as ϵ entropy and the finite size Lyapunov exponent (FSLE) have been proposed, simple algorithms have not been developed to reliably compute them from short noisy time series. To promote widespread application of these concepts, we propose an efficient algorithm to compute a variant of the FSLE, the scale-dependent Lyapunov exponent (SDLE). We show that with our algorithm, the SDLE can be accurately computed from short noisy time series and readily classify various types of motions, including truly low-dimensional chaos, noisy chaos, noise-induced chaos, random $1/f^\alpha$ and α -stable Levy processes, stochastic oscillations, and complex motions with chaotic behavior on small scales but diffusive behavior on large scales. To our knowledge, no other measures are able to accurately characterize all these different types of motions. Based on the distinctive behaviors of the SDLE for different types of motions, we propose a scheme to distinguish chaos from noise.

DOI: [10.1103/PhysRevE.74.066204](https://doi.org/10.1103/PhysRevE.74.066204)

PACS number(s): 05.45.Tp, 05.40.Fb, 05.45.Ac, 89.75.-k

I. INTRODUCTION

Complex systems are usually comprised of multiple subsystems that exhibit both highly nonlinear deterministic, as well as, stochastic characteristics, and are subject to hierarchical regulations. The irregularities in the price of a stock in a financial market, for example, are highly influenced by multilayered decisions made by various policy makers. The Internet, as another example, has been designed in a fundamentally decentralized fashion and consists of a complex web of servers and routers that cannot be effectively controlled or analyzed by traditional tools of queuing theory or control theory, and give rise to highly bursty and multiscale traffic with extremely high variance [1], as well as complex dynamics with both deterministic and stochastic components [2]. Similarly, with our increasing capability to monitor and control biological activities, we have no choice but to deal with signals generated by systems that are by nature heterogeneous, massively distributed, and highly complicated. When devising measures to characterize complex multiscaled signals, it is of fundamental importance to explicitly incorporate the concept of the scale in the measures. Along this line, three measures are most noteworthy. One is the recently introduced multiscale entropy (MSE) [3], which incorporates the concept of scale in the temporal domain, through nonoverlapping running means. This measure is useful for analyzing stochastic systems. Two other measures are ϵ entropy [4] and the finite size Lyapunov exponent (FSLE) [5–7]. They are highly related. Having the concept of scale in phase space, ϵ entropy and FSLE can be applied to both deterministic and stochastic systems. Indeed, theoretically, ϵ entropy has been known for a number of dynamical systems,

including chaos and random fractals. Unfortunately, efficient algorithms have not been proposed to reliably calculate them from short noisy time series. This has severely limited their application to many real world signal processing applications. To fully realize their potential, in this paper, we propose a fast and reliable algorithm to compute a variant of the FSLE, which depends on the scale continuously and therefore, is more appropriate to be called the scale-dependent Lyapunov exponent (SDLE). We shall show that with our computational algorithm, the SDLE is able to characterize all known types of motions by obtaining the most important parameters from a time series of practical length, and therefore, can shed much new light on the problem of distinguishing chaos from noise.

Distinction between chaos and noise is a classic issue arising from life sciences, finance, ecology, physics, fluid mechanics, and geophysics. Albeit tremendous effort has been made to solve this issue [8–11], it remains widely open [6,12–14]. Two major difficulties for solving this issue are (i) chaos can be induced by noise [6,15,16], and (ii) standard Brownian motions may have a deterministic origin [6]. To overcome these two difficulties, we ask two questions: (i) What are the fundamental differences between clean low-dimensional chaos, noisy chaos, and noise-induced chaos? (ii) When a Brownian motion has a deterministic origin, can we determine whether it is from a low-dimensional deterministic system or a high-dimensional deterministic system? The above questions motivate us to propose the following algorithmic scheme to solve the problem of distinguishing chaos from noise. Denote two scales resolvable by the resolution of the data by ϵ_1 and ϵ_2 , where $\epsilon_1 < \epsilon_2$. On this scale range, if the behavior of the data is the same as a chaotic data, then we say the data is chaotic. If the behavior is like a Brownian motion, then we say the data is a Brownian motion, etc. Of course, there may exist scale ranges disjoint with (ϵ_1, ϵ_2) , where the data behave neither like a chaotic

*Electronic address: gao@ece.ufl.edu

motion nor like a Brownian motion, but have features that can define other types of motions. The data on those scales will be classified accordingly. The feasibility of such a scheme critically depends on whether one is able to, by analyzing short noisy time series, (i) classify different types of motions, and (ii) characterize the time series by automatically identifying different scale ranges where different types of motions are manifested. Note that when dealing with complex multiscaled data, the choice of a scale-dependent classification scheme is the most natural, as pointed out by Cencini *et al.* [6]. However, when a Brownian motion is generated by a very high-dimensional deterministic system, we shall simply treat it as a type of stochastic process, so long as it has the defining properties for a Brownian motion.

At this point, we need to make a comment on the concept of resolvable scale. Denote a time series under investigation by $x(1), x(2), \dots, x(n)$. Using time delay embedding [17], one obtains vectors of the form $V_i = [x(i), x(i+L), \dots, x(i+(m-1)L)]$, where the embedding dimension m and the delay time L are chosen according to certain optimization criterion [11,18]. In the reconstructed phase space, the dataset determines two scales, the maximum and the minimum of the distances between two vectors, $\|V_i - V_j\|$, where $i \neq j$. Denote them by ϵ_{\max} and ϵ_{\min} , respectively. This is the resolvable scale range. Analysis of data must be confined within these two scales. Of course, with more data available, the resolvable scale range can be enlarged. Note that one can treat ϵ_{\max} as one unit. This amounts to normalizing the time series $x(1), x(2), \dots, x(n)$ into the unit interval $[0, 1]$ before further analysis.

We make another comment on embedding of a stochastic process. Since a stochastic process is infinite dimensional, how high should the embedding dimension be chosen? It turns out that often the major advantage of embedding a self-affine stochastic process to a phase space is to transform it to a self-similar process. When this is the case, specific values of m and L are not important, so long as $m > 1$. For example, one could simply choose $m=2, L=1$. One exception is stochastic oscillations. This will be made clearer when we discuss a concrete example, the stochastic Van de Pol oscillator.

II. SCALE-DEPENDENT LYAPUNOV EXPONENT

The SDLE we shall compute in this paper is a variant of FSLE [5]. The algorithm for calculating the FSLE is very similar to the Wolf *et al.* algorithm [9]. It computes the average r -fold time by monitoring the divergence between a reference trajectory and a perturbed trajectory. To do so, it needs to define ‘‘nearest neighbors,’’ as well as needs to perform, from time to time, a renormalization when the distance between the reference and the perturbed trajectory becomes too large. Such a procedure requires very long time series, and therefore, is not practical. To facilitate derivation of a fast algorithm that works on short data, as well as to ease discussion of continuous but nondifferentiable stochastic processes, we cast the definition of the SDLE as follows.

Consider an ensemble of trajectories. Denote the initial separation between two nearby trajectories by ϵ_0 , and their

average separation at time t and $t + \Delta t$ by ϵ_t and $\epsilon_{t+\Delta t}$, respectively. Being defined in an average sense, ϵ_t and $\epsilon_{t+\Delta t}$ can be readily computed from any processes, even if they are non-differentiable. Next we examine the relation between ϵ_t and $\epsilon_{t+\Delta t}$, where Δt is small. When $\Delta t \rightarrow 0$, we have,

$$\epsilon_{t+\Delta t} = \epsilon_t e^{\lambda(\epsilon_t)\Delta t}, \quad (1)$$

where $\lambda(\epsilon_t)$ is the SDLE. It is given by

$$\lambda(\epsilon_t) = \frac{\ln \epsilon_{t+\Delta t} - \ln \epsilon_t}{\Delta t}. \quad (2)$$

Given a time series data, the smallest Δt possible is the sampling time τ .

The definition of the SDLE suggests a simple ensemble average based scheme to compute it. A straightforward way would be to find all the pairs of vectors in the phase space with their distance approximately ϵ , and then calculate their average distance after a time Δt . The first half of this description amounts to introducing a shell (indexed as k),

$$\epsilon_k \leq \|V_i - V_j\| \leq \epsilon_k + \Delta\epsilon_k. \quad (3)$$

where V_i, V_j are reconstructed vectors, ϵ_k (the radius of the shell) and $\Delta\epsilon_k$ (the width of the shell) are arbitrarily chosen small distances. Such a shell may be considered as a differential element that would facilitate computation of conditional probability. To expedite computation, it is advantageous to introduce a sequence of shells, $k=1, 2, 3, \dots$. Note that this computational procedure is similar to that for computing the so-called time-dependent exponent (TDE) curves [11,18].

With all these shells, we can then monitor the evolution of all of the pairs of vectors (V_i, V_j) within a shell and take average. When each shell is very thin, by assuming that the order of averaging and taking logarithm in Eq. (2) can be interchanged, we have

$$\lambda(\epsilon) = \frac{\langle \ln \|V_{i+t+\Delta t} - V_{j+t+\Delta t}\| - \ln \|V_{i+t} - V_{j+t}\| \rangle}{\Delta t}, \quad (4)$$

where t and Δt are integers in unit of the sampling time, and the angle brackets denote average within a shell. Note that contributions to the SDLE at a specific scale from different shells can be combined, with the weight for each shell being determined by the number of the pairs of vectors (V_i, V_j) in that shell. In the following, to see better how each shell characterizes the dynamics of the data on different scales, we shall plot the $\lambda(\epsilon)$ curves for different shells separately.

In the above formulation, it is implicitly assumed that the initial separation, $\|V_i - V_j\|$, aligns with the most unstable direction instantly. For high-dimensional systems, this is not true, especially when the growth rate is nonuniform and/or the eigenvectors of the Jacobian are non-normal. Fortunately, the problem is not as serious as one might be concerned, since our shells are not infinitesimal. When computing the TDE [11,18], we have found that when difficulties arise, it is often sufficient to introduce an additional condition,

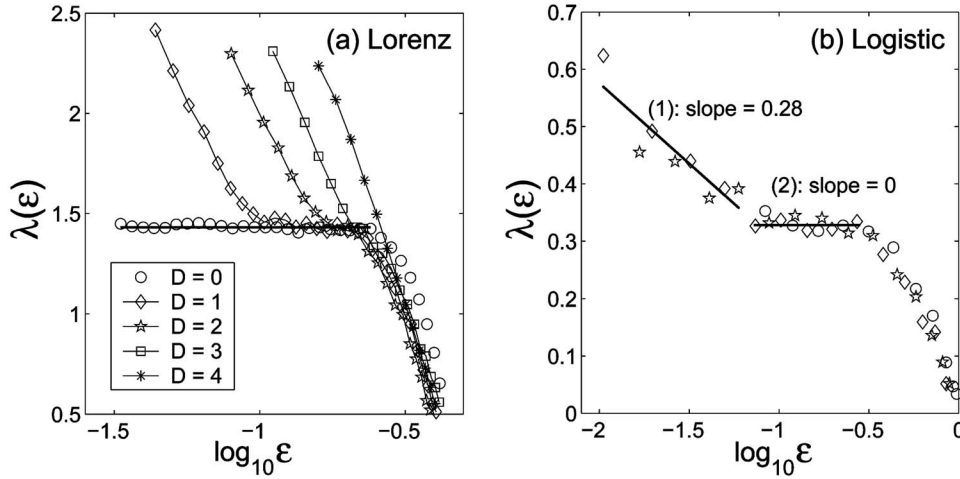


FIG. 1. $\lambda(\epsilon)$ curves for (a) the clean and the noisy Lorenz system, and (b) the noise-induced chaos in the logistic map [curves designated by circles, pentagrams, and diamonds correspond to shells $(2^{-i-1}, 2^{-i})$, $i=9, 14, 19$].

$$|j - i| \geq (m - 1)L, \quad (5)$$

when finding pairs of vectors within each shell. Such a scheme also works well when computing the SDLE. This means that, after taking a time comparable to the embedding window $(m-1)L$, it would be safe to assume that the initial separation has evolved to the most unstable direction of the motion.

Before proceeding on, we wish to emphasize the major difference between our algorithm and the standard method for calculating the FSLE. As we have pointed out, to compute the FSLE, two trajectories, one as reference, another as perturbed, must be defined. This requires huge amounts of data. Our algorithm avoids this by employing two critical operations to fully utilize information about the time evolution of the data: (i) The reference and the perturbed trajectories are replaced by time evolution of all pairs of vectors satisfying the inequality (5) and falling within a shell, and (ii) introduction of a sequence of shells ensures that the number of pairs of vectors within the shells is large while the ensemble average within each shell is well defined. In the following, we shall illustrate the effectiveness of our algorithm by examining various types of complex motions.

III. CLASSIFICATION OF COMPLEX MOTIONS

To understand the SDLE as well as appreciate its power, we apply it to classify seven major types of complex motions.

A. Chaos, noisy chaos, and noise-induced chaos

Obviously, for truly low-dimensional chaos, $\lambda(\epsilon)$ equals the largest positive Lyapunov exponent, and hence, must be independent of ϵ over a wide range of scales. For noisy chaos, we expect $\lambda(\epsilon)$ to depend on small ϵ . To illustrate both features, we consider the chaotic Lorenz system with stochastic forcing:

$$dx/dt = -16(x - y) + D\eta_1(t),$$

$$dy/dt = -xz + 45.92x - y + D\eta_2(t),$$

$$dz/dt = xy - 4z + D\eta_3(t), \quad (6)$$

where $D\eta_i(t)$, $i=1, 2, 3$ are independent Gaussian noise forcing terms with mean 0 and variance D^2 . When $D=0$, the system is clean. Figure 1(a) shows five curves, for the cases of $D=0, 1, 2, 3, 4$. The computations are done with 10 000 points and $m=4$, $L=2$. We observe a few interesting features: (i) For the clean chaotic signal, $\lambda(\epsilon)$ slightly fluctuates around a constant (which numerically equals the largest positive Lyapunov exponent) when ϵ is smaller than a threshold value which is determined by the size of the chaotic attractor. The reason for the small fluctuations in $\lambda(\epsilon)$ is that the divergence rate varies from one region of the attractor to another. (ii) When there is stochastic forcing, $\lambda(\epsilon)$ is no longer a constant when ϵ is small, but increases as $-\gamma \ln \epsilon$ when the scale ϵ is decreased. The coefficient γ does not seem to depend on the strength of the noise. This feature suggests that entropy generation is infinite when the scale ϵ approaches to zero. Note that the relation of $\lambda(\epsilon) \sim -\gamma \ln \epsilon$ has also been observed for the FSLE and the ϵ entropy. In fact, such a relation can be readily proven for the ϵ entropy [4]. (iii) When the noise is increased, the part of the curve with $\lambda(\epsilon) \sim -\gamma \ln \epsilon$ shifts to the right. In fact, little chaotic signature can be identified when D is increased beyond 3. When noise is not too strong, this feature can be readily used to quantify the strength of noise.

Next we consider noise-induced chaos. To illustrate the idea, we follow Ref. [15] and study the noisy logistic map

$$x_{n+1} = \mu x_n (1 - x_n) + P_n, \quad 0 < x_n < 1, \quad (7)$$

where μ is the bifurcation parameter, and P_n is a Gaussian random variable with zero mean and standard deviation σ . In Ref. [15], we reported that at $\mu=3.74$ and $\sigma=0.002$, noise-induced chaos occurs, and thought that it may be difficult to distinguish noise-induced chaos from clean chaos. In Fig. 1(b), we have plotted the $\lambda(\epsilon)$ for this particular noise-induced chaos. The computation was done with $m=4$, $L=1$ and 10 000 points. We observe that Fig. 1(b) is very similar to the curves of noisy chaos plotted in Fig. 1(a). Hence, noise-induced chaos is similar to noisy chaos, but different from clean chaos.

At this point, it is worth making two comments: (i) On very small scales, the effect of measurement noise is similar to that of dynamic noise. (ii) The $\lambda(\epsilon)$ curves shown in Fig. 1 are based on a fairly small shell. The curves computed based on larger shells collapse on the right-hand side of the curves shown in Fig. 1. Because of this, for chaotic systems, one or a few small shells would be sufficient. If one wishes to know the behavior of λ on ever smaller scales, one must use longer and longer time series.

Finally, we consider the Mackey-Glass delay differential system [19],

$$dx/dt = \frac{ax(t+\Gamma)}{1+x(t+\Gamma)^c} - bx(t).$$

When $a=0.2$, $b=0.1$, $c=10$, $\Gamma=30$, it has two positive Lyapunov exponents, with the largest Lyapunov exponent close to 0.007 [11]. Having two positive Lyapunov exponents while the value of the largest Lyapunov exponent of the system is not much greater than 0, one might be concerned that it may be difficult to compute the SDLE of the system. This is not the case. In fact, this system can be analyzed as straightforwardly as other dynamical systems including the Henon map and the Rossler system. An example of the $\lambda(\epsilon)$ curve is shown in Fig. 2, where we have followed Ref. [11] and used $m=5$, $L=1$, and 5000 points sampled with a time interval of 6. Clearly, we observe a well defined plateau, with its value close to 0.007. This example illustrates that when computing the SDLE, one does not need to be very concerned about nonuniform growth rate in high-dimensional systems.

B. $1/f^\alpha$ processes

$1/f^\alpha$ noise is a form of temporal or spatial fluctuation characterized by a power-law decaying power spectral density. It has been observed in many areas of science and engineering (see Refs. [20,21] and references therein). Two important prototypical models for such processes are the fractional Brownian motion (fBm) process [22] and the on/off intermittency with power-law distributed on and off periods [21]. For convenience, we introduce the Hurst parameter $0 < H < 1$ through a simple equation,

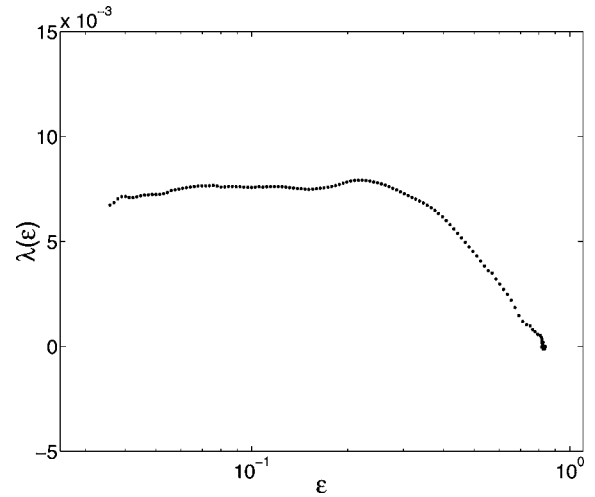


FIG. 2. The $\lambda(\epsilon)$ curve for the Mackey-Glass system. The computation was done with $m=5$, $L=1$, and 5000 points sampled with a time interval of 6.

$$\alpha = 2H + 1. \tag{8}$$

Depending on whether H is smaller than, equal to, or larger than $1/2$, the process is said to have antipersistent correlation, short-range correlation, and persistent long-range correlation [21]. Note that $D=1/H$ is the fractal dimension of such processes, while the famous Kolmogorov's $5/3$ law for the energy spectrum of fully developed turbulence [23] corresponds to $H=1/3$.

It is well known that the variance of such stochastic processes increases with t as t^{2H} . Translating into the average distance between nearby trajectories, we can write

$$\epsilon_t = \epsilon_0 t^H. \tag{9}$$

Using Eq. (2), we readily find $\lambda(\epsilon_t) \sim H/t$. Expressing t by ϵ_t , we finally find

$$\lambda(\epsilon_t) \sim H \epsilon_t^{-1/H}. \tag{10}$$

Note that the same functional relation can be derived for ϵ entropy [4]. However, using ϵ entropy, it is not easy to estimate H through analysis of short time series. In contrast, Eq. (10) can be conveniently used to estimate H . To illustrate this, we first study fBm processes $B_H(t)$, which are Gaussian

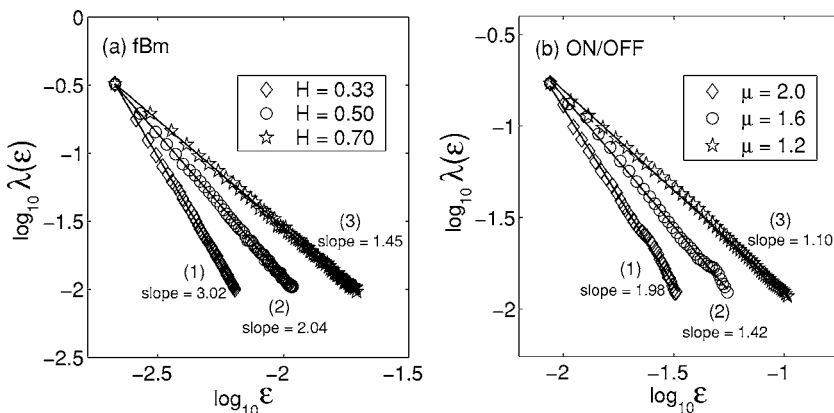


FIG. 3. $\lambda(\epsilon)$ curves for (a) fBm processes with $H=1/3$, 0.5, and 0.7, and (b) on/off models with $\mu=1.2$, 1.6, and 2.0.

process with mean 0, stationary increments, variance

$$E[(B_H(t))^2] = t^{2H} \quad (11)$$

and covariance

$$E[B_H(s)B_H(t)] = \frac{1}{2}(s^{2H} + t^{2H} - |s-t|^{2H}). \quad (12)$$

In Fig. 3(a), we have shown three curves for $H=0.33, 0.5$, and 0.7 , where the calculation is done with 2^{14} points and $m=2, L=1$. We observe that the estimated $1/H$ clearly match those used in simulating these processes.

Next, we study on/off models with power-law distributed on and off periods,

$$P[X \geq x] = \left(\frac{b}{x}\right)^\mu, \quad x \geq b > 0, \quad 0 < \mu \leq 2. \quad (13)$$

When $\mu < 2$, such processes have infinite variance. It is well known that [21]

$$H = (3 - \mu)/2. \quad (14)$$

In Fig. 3(b), we have shown three curves for $\mu=1.2, 1.6, 2$, which corresponds to $H=0.9, 0.7$, and 0.5 , respectively. Again, the estimated H accurately reflect the values of μ used in the simulations.

Note that as far as estimating the H parameter from fBm and on/off processes is concerned, our algorithm is as accurate as any other methods, including the detrended fluctuation analysis (DFA). [24]

Before proceeding on, a comment should be made on the meaning of ϵ in the power-law distributed on/off model. As we have pointed out, when $\mu < 2$, on/off processes have infinite variance. Therefore, to effectively characterize on/off processes, the absolute value of ϵ also has to be large. So long as ϵ is small relative to the largest scale resolved by the dataset, our algorithm should work. This comment is also pertinent to the Levy flights to be studied next. In fact, the example of a Levy flight to be discussed not only has infinite variance, but also has infinite mean. Interestingly, the defining parameter for such a Levy flight can still be readily estimated by our method.

C. Levy flights

We now consider Levy processes, one of the prototypical random fractal models that have found numerous applications [25,26].

There are two types of Levy processes [21]. One is Levy flights, which are random processes consisting of many independent steps, each step being characterized by a stable law, and consuming a unit time regardless of its length. The other is Levy walkers, where each step takes time proportional to its length. A Levy walker can be viewed as sampled from a Levy flight with a uniform speed. The increment process of a Levy walker, obtained by differencing the Levy walker, is very similar to an on/off train with power-law distributed on and off periods. Therefore, in the following, we shall not be further concerned about it. We shall focus on Levy flights. Note that Levy flights, having independent steps, are memoryless processes characterized by $H=1/2$,

irrespective of the value of the exponent α characterizing the stable laws [21]. In other words, methods such as DFA cannot be used to estimate the α parameter.

In order to derive an analytic expression for $\lambda(\epsilon)$ to estimate α , we define Levy flights more precisely. A (standard) symmetric α -stable Levy flight $\{L_\alpha(t), t \geq 0\}$ is a stochastic process [27] which is almost surely zero at $t=0$, has independent increments, and $L_\alpha(t) - L_\alpha(s)$ follows an α -stable distribution with characteristic function $e^{-(t-s)|u|^\alpha}$, where $0 \leq s < t < \infty$. A random variable Y is called (strictly) stable if the distribution for $\sum_{i=1}^n Y_i$ is the same as that for $n^{1/\alpha}Y$,

$$\sum_{i=1}^n Y_i \stackrel{d}{=} n^{1/\alpha}Y, \quad (15)$$

where Y_1, Y_2, \dots are independent random variables, each having the same distribution as Y . From Eq. (15), one then finds that $n \text{Var} Y = n^{2/\alpha} \text{Var} Y$. For the distribution to be valid, $0 < \alpha \leq 2$. When $\alpha=2$, the distribution is Gaussian, and hence, the corresponding Levy flight is just the standard Brownian motion. When $0 < \alpha < 2$, the distribution is heavy tailed, $P[X \geq x] \sim x^{-\alpha}, x \rightarrow \infty$, and has infinite variance. Furthermore, when $0 < \alpha \leq 1$, the mean is also infinite.

The symmetric α -stable Levy flight is $1/\alpha$ self-similar. That is, for $c > 0$, the processes $\{L_\alpha(ct), t \geq 0\}$ and $\{c^{1/\alpha}L_\alpha(t), t \geq 0\}$ have the same finite-dimensional distributions. By this argument as well as Eq. (15), it is clear that on average, the length of the motion, $\Delta L(\Delta t)$, in a time span of Δt , is given by the following scaling:

$$\Delta L(\Delta t) \propto \Delta t^{1/\alpha}. \quad (16)$$

Comparing to $1/f^{2H+1}$ processes, we identify that $1/\alpha$ plays the role of H . Therefore, the scaling for the SDLE is

$$\lambda(\epsilon_t) \sim \frac{1}{\alpha} \epsilon_t^{-\alpha}. \quad (17)$$

We have simulated a number of Levy flights with different α . One realization for the flights with $\alpha=1$ is shown in Fig. 4(a). The computed $\lambda(\epsilon_t)$ curves (based on 2^{15} points and $m=2, L=1$) are shown in Fig. 4(b). In fact, the $\lambda(\epsilon_t)$ curves from a number of different shells are plotted together. We observe that the slope for the *envelope* is -1 , consistent with the value of α chosen in simulating the flights.

To understand why the envelope of the $\lambda(\epsilon_t)$ gives a better estimate of α , we resort to the two small boxes, denoted as A and B , in Fig. 4(a). When those two small boxes are enlarged, they show similar pattern as that of Fig. 4(a). Obviously, the scales involved in A and B are different. Those different scales are captured by shells of different sizes. Since the mean as well as the variance of the flights are both infinite, while the number of points in each small box is not large, $\lambda(\epsilon_t)$ from each shell cannot have a long scaling region. However, when $\lambda(\epsilon_t)$ from different shells are plotted together, they form an envelope with a long scaling region. With this argument, it is clear that when computing $\lambda(\epsilon_t)$ for a Levy flight, a number of shells is more advantageous than a single shell, especially when the time series is not too long.

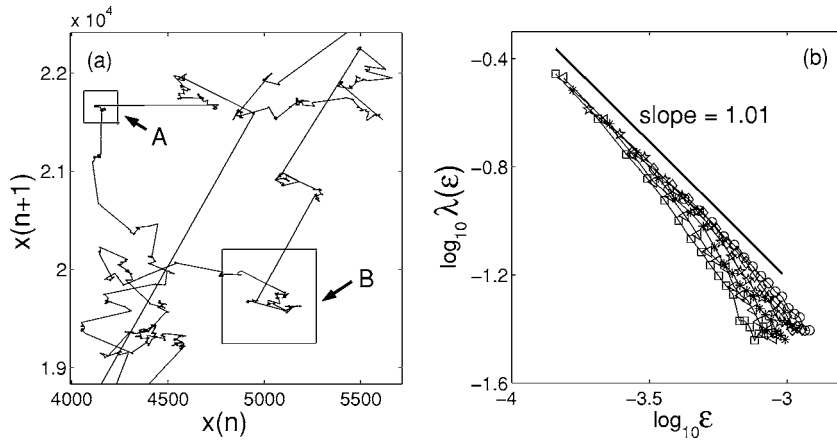


FIG. 4. (a) 2D Levy flights with $\alpha=1.0$, (b) $\lambda(\epsilon)$ curves for the Levy process [curves designated by circles, diamonds, pentagrams, asterisks, left triangles, and squares correspond to shells $(2^{-i-1}, 2^{-i})$, $i=24, 25, \dots, 29$]. The estimated slope 1.01 is close to the given α , which is 1.0.

D. Stochastic oscillations

Stochastic oscillation is an important type of complex motion that has been observed in many different disciplines of science and engineering. A stochastic oscillator, having structures due to oscillatory motions but not having closed orbits in the phase space, may be interpreted as chaos. Which functional form of $\lambda(\epsilon_i)$ characterizes such motions? To gain insights into this problem, we study a stochastically driven Van de Pol's oscillator

$$dx/dt = y + D_1 \eta_1(t),$$

$$dy/dt = -(x^2 - 1)y - x + D_2 \eta_2(t), \quad (18)$$

where $\langle \eta_i(t) \rangle = 0$, $\langle \eta_i(t) \eta_j(t') \rangle = \delta_{ij} \delta(t-t')$, $i, j=1, 2$, and the parameters D_i , $i=1, 2$ characterize the strength of noise. We have calculated the SDLE for many different choices of the noise level. There are two types of behaviors for $\lambda(\epsilon_i)$, quite independent of the noise level. One functional form for $\lambda(\epsilon_i)$ is $\lambda(\epsilon) \sim -\ln \epsilon$, observed when the embedding dimension m and delay time L are both small. Another type of behavior is $\lambda(\epsilon) \sim \epsilon^{-1/H}$, where $H \approx 1/2$, observed when $(m-1)L$ is comparable to one-half of the period of the oscillation. An example of the latter is shown in Fig. 5.

The relation $\lambda(\epsilon) \sim -\ln \epsilon$ for small embedding window suggests that locally, the dynamics of a stochastic oscillator is just like other noisy dynamics, such as those shown in Fig. 1. The relation $\lambda(\epsilon) \sim \epsilon^{-2}$ for large embedding window suggests that the motion is like a Brownian motion. This reflects the variation of amplitude of the oscillation on longer time scales, as studied in Ref. [28].

Note that stochastic oscillators may be characterized by more general relations $\lambda(\epsilon) \sim \epsilon^{-1/H}$, with $H \neq 1/2$ [29], especially in the case of noise-induced chaos, H is usually larger than 1 when noise level is appropriate [15].

Finally, we emphasize that relations such as $\lambda(\epsilon) \sim \epsilon^{-1/H}$ for stochastic oscillators cannot be characterized by many other methods, such as the DFA.

E. Complex motions with multiple scaling behaviors

Some dynamical systems may exhibit multiple scaling behaviors, such as chaotic behavior on small scales but diffu-

sive behavior on large scales. To see how our algorithm can characterize such systems, we follow Cencini *et al.* [6] and study the following map:

$$x_{n+1} = [x_n] + F(x_n - [x_n]) + \sigma \eta_t \quad (19)$$

where $[x_n]$ denotes the integer part of x_n , η_t is a noise uniformly distributed in the interval $[-1, 1]$, σ is a parameter quantifying the strength of noise, and $F(y)$ is given by

$$F(y) = \begin{cases} (2 + \Delta)y & \text{if } y \in [0, 1/2), \\ (2 + \Delta)y - (1 + \Delta) & \text{if } y \in [1/2, 1]. \end{cases} \quad (20)$$

The map $F(y)$ is shown in Fig. 6 as the dashed lines. It gives a chaotic dynamics with a positive Lyapunov exponent $\ln(2 + \Delta)$. On the other hand, the term $[x_n]$ introduces a random walk on integer grids.

It turns out this system is very easy to analyze. When $\Delta = 0.4$, with only 5000 points and $m=2$, $L=1$, we can resolve both the chaotic behavior on very small scales, and the nor-

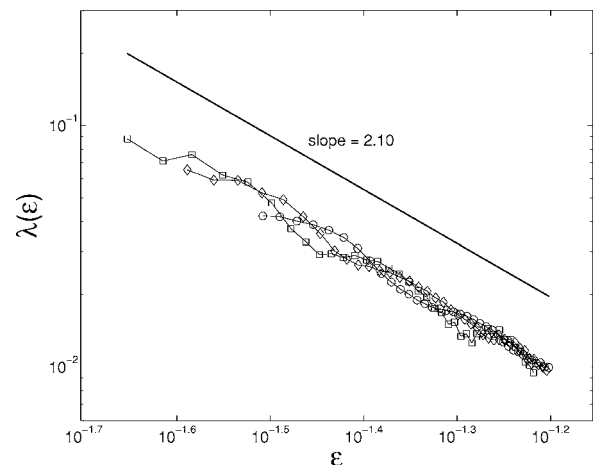


FIG. 5. $\lambda(\epsilon)$ curves for the $y(t)$ data of the stochastic Van de Pol's oscillator with $D_1=D_2=0.02$ [curves designated by circles, diamonds, and squares correspond to shells $(2^{-i-1}, 2^{-i})$, $i=11, 12, 13$]. The estimated slope 2.10 yields an H close to 0.5. The data was sampled with a time interval of 0.2. The period of the oscillation is about 30 sample points. The embedding parameters are $m=8$, $L=2$, therefore, the embedding window $(m-1)L$ is about one-half of the period.

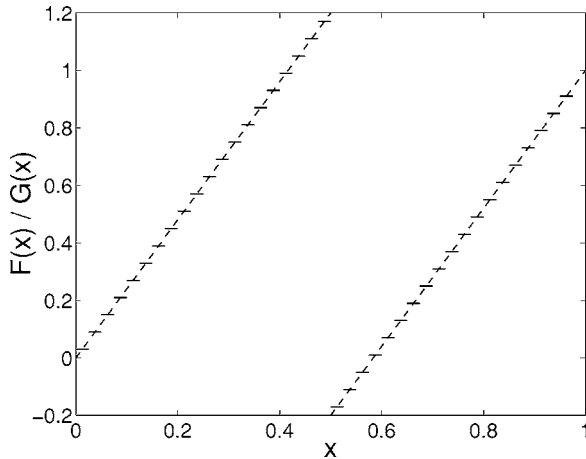


FIG. 6. The function $F(x)$ (20) for $\Delta=0.4$ is shown as the dashed lines. The function $G(x)$ (21) is an approximation of $F(x)$, obtained using 40 intervals of slope 0. In the case of noise-induced chaos discussed in the paper, $G(x)$ is obtained from $F(x)$ using 10^4 intervals of slope 0.9.

mal diffusive behavior (with slope -2) on large scales. See Fig. 7(a).

We now ask a question: Given a small dataset, which type of behavior, the chaotic or the diffusive, is resolved first? To answer this, we have tried to compute the SDLE with only 500 points. The result is shown in Fig. 7(b). It is interesting to observe that the chaotic behavior can be well resolved by only a few hundred points. However, the diffusive behavior needs much more data to resolve. Intuitively, this makes sense, since the diffusive behavior amounts to a Brownian motion on the integer grids and is of much higher dimension than the small scale chaotic behavior. Therefore, more data are needed to resolve it. As will be shown in the next section, this important result will allow us to determine whether a deterministic Brownian motion is of low or high dimension.

We have also studied the noisy map. The resulting SDLE for $\sigma=0.001$ is shown in Fig. 7(a), as squares. We have again used 5000 points. While the behavior of the SDLE suggests noisy dynamics, with 5000 points, we are not able to well resolve the relation $\lambda(\epsilon) \sim -\ln \epsilon$. This indicates that for the noisy map, on very small scales, the dimension is very high.

Map (19) can be modified to give rise to an interesting system with noise-induced chaos. This can be done by re-

placing the function $F(y)$ in map (19) by $G(y)$ to obtain the following map [6]:

$$x_{t+1} = [x_t] + G(x_t - [x_t]) + \sigma \eta_t, \quad (21)$$

where η_t is a noise uniformly distributed in the interval $[-1, 1]$, σ is a parameter quantifying the strength of noise, and $G(y)$ is a piecewise linear function which approximates $F(y)$ of Eq. (20). An example of $G(y)$ is shown in Fig. 6. In our numerical simulations, we have followed Cencini *et al.* [6] and used 10^4 intervals of slope 0.9 to obtain $G(y)$. With such a choice of $G(y)$, in the absence of noise, the time evolution described by the map (21) is nonchaotic, since the largest Lyapunov exponent $\ln(0.9)$ is negative. With appropriate noise level (e.g., $\sigma=10^{-4}$ or 10^{-3}), the SDLE for the system becomes indistinguishable to the noisy SDLE shown in Fig. 7 for the map (19). Having a diffusive regime on large scales, this is a more complicated noise-induced chaos than the one we have found from the logistic map.

Before proceeding on, we make a comment on the computation of the SDLE from deterministic systems with negative largest Lyapunov exponents, such as the map (21) without noise. A transient-free time series from such systems is a constant time series. Therefore, there is no need to compute the SDLE or other metrics. When the time series contains transients, if the time series is sampled with high enough sampling frequency, then the SDLE is negative. In the case of simple exponential decay to a fixed point, such as expressible as $e^{-\lambda t}$, where $\lambda > 0$, one can readily prove that the SDLE is $-\lambda$. Since such systems are not complex, we shall not be further concerned about them.

IV. DISTINGUISHING CHAOS FROM NOISE

In the introduction, we have argued that the key to distinguish chaos from noise is to identify different scale ranges where different types of motions are manifested. Now that we have found distinctive behaviors of the SDLE for seven important types of motions that have been known so far, we are ready to answer how chaos can be distinguished from noise. Since our purpose is for practical applications, we shall assume that our dataset is not only finite, but of small or medium size.

We first ask a question. When all three distinctive behaviors of the SDLE, (i) $\lambda(\epsilon) \sim -\ln \epsilon$, (ii) $\lambda(\epsilon) \sim \text{constant}$, and

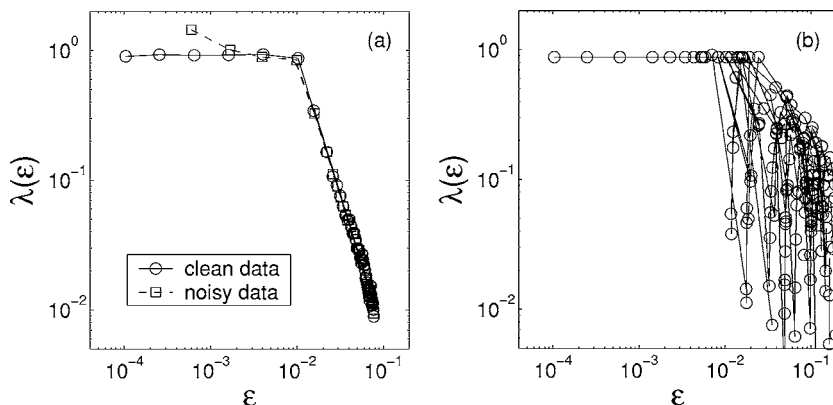


FIG. 7. $\lambda(\epsilon)$ for the model described by Eq. (19). (a) 5000 points were used; for the noisy case, $\sigma=0.001$. (b) 500 points were used.

(iii) $\lambda(\epsilon) \sim \epsilon^{-1/H}$, coexist, is it typically true that the scales for them to occur are in an ascending order? The answer is yes, as can be readily understood by the following argument. Behavior (i) amounts to stochastic forcing. Its effect is to kick the dynamics to larger scales. For the chaotic motion to be resolvable, its effect must be limited to scales that are smaller than the scales showing chaotic motion. To understand why diffusive behavior (iii) must occur on scales larger than those showing chaotic motions, it suffices to note that diffusive motion is a nonstationary process. It needs a huge or even an unbounded region of phase space to play out. It is unimaginable to have chaotic motion beyond those scales still resolvable by a finite data. Therefore, the scales for the three behaviors to occur must be in an increasing order.

The above discussion suggests that if a Brownian motion is generated by a deterministic system, then the dimension of the system can be readily determined. This is because for deterministic low-dimensional systems such as the map (19) to generate large scale diffusive motions, locally its dynamics must be unstable. Low-dimensional local unstable dynamics amounts to a low-dimensional chaotic motion. Therefore, we must have a constant $\lambda(\epsilon)$ on small scales. In other words, coexistence of $\lambda(\epsilon) \sim \text{constant}$ and $\lambda(\epsilon) \sim \epsilon^{-2}$ would indicate that the Brownian motion is from a low-dimensional deterministic system. We emphasize that $\lambda(\epsilon) \sim \text{constant}$ should occur on the scales smaller than those for diffusion to occur, and that the behavior of $\lambda(\epsilon) \sim -\ln \epsilon$ should not occur, since the system is low-dimensionally deterministic.

We are now ready to answer to what extent chaos can be distinguished from noise. First, in order to say that a time series under study has a chaotic scaling regime, $\lambda(\epsilon)$ must show a plateau (i.e., almost constant) for a scale range (ϵ_1, ϵ_2) . Since many arbitrary functions locally could be treated as a constant, if $\lambda(\epsilon)$ is constant only for ϵ_2 very close to ϵ_1 , then it should not be interpreted as a plateau. In other words, ϵ_2 must be considerably larger than ϵ_1 . Unfortunately, it is not easy to state exactly how much larger ϵ_2 should be than ϵ_1 , since this depends on the dimension of the data. We recommend that $\log_{10}(\epsilon_2/\epsilon_1)$ must be at least around 1/2.

After knowing existence of a chaotic scaling regime in the data, it is natural to ask if the dynamics is deterministically chaotic or not. This can be readily inferred by checking the behavior of $\lambda(\epsilon)$ on scales smaller than ϵ_1 . If on those scales, $\lambda(\epsilon)$ is above the plateau, then we have good reason to suspect that the data is noisy. If on those scales, $\lambda(\epsilon) \sim -\ln \epsilon$, then we can be sure that it is noisy chaos or noise-induced chaos. However, if all is known is the time series, then we will not be able to further distinguish between noisy chaos and noise-induced chaos.

In most situations, when the data is found to be noisy, it is important to find out which type of noise process the data is. The power-law behavior of the SDLE is especially interesting for this purpose. For example, we can combine analysis using the SDLE with spectral analysis and distributional analysis, to determine whether the data is a type of $1/f^\alpha$ process, or Levy processes, or stochastic oscillations. More precisely, a power-law behavior of the SDLE together with sharp spectral peaks revealed by Fourier analysis would indicate a stochastic oscillation, a power-law behavior of the

SDLE together with a Gaussian-type distribution would indicate a fBm-like data. If instead, heavy-tailed distributions are observed, then the data is of on/off intermittency type or similar to Levy processes. While many excellent methods for estimating the Hurst parameter have been proposed (see a recent comparative study [21]), the SDLE offers another effective approach. In fact, as we have pointed out in Sec. III, the SDLE provides better characterization of Levy flights and stochastic oscillations, since most other methods would only give $H=0.5$ for Levy flights, irrespective of the values of the defining parameter α , and would fail to characterize the power-law behavior of the SDLE for stochastic oscillations.

V. CONCLUSIONS

Motivated by lack of efficient computational approaches to simultaneously characterize complex signals on a wide range of scales, we have proposed a algorithm to compute the SDLE from short noisy time series. We have shown that with our algorithm, the SDLE can readily characterize seven important types of complex motions, including truly low-dimensional chaos, noisy chaos, noise-induced chaos, random $1/f^\alpha$ processes, random α -stable Levy processes, stochastic oscillations, and motions with multiple scaling behaviors such as chaotic behavior on small scales and diffusive behavior on large scales. In particular, the power-law relation in the SDLE offers another excellent method for estimating the defining parameters for major models of stochastic processes, including $1/f^\alpha$ processes and Levy processes. Classification of different types of motions based on the SDLE of short time series has helped us understand how much chaos can be distinguished from noise, and enabled us to propose a simple scheme to solve this problem.

Since ϵ entropy and the SDLE (as well as the FSLE) are equivalent [6], our work has shed considerable light on whether microscopic chaos exists or not [14] as well as why the K_2 entropies for $1/f^\alpha$ processes converge to zero when the embedding dimension increases [13]. In light of our work, the power-law decaying ϵ entropy (with slope -2) reported in Ref. [14] simply indicates Brownian motions instead of chaos on the scales examined. To understand why the K_2 entropies for $1/f^\alpha$ processes converge to zero when the embedding dimension increases, it suffices to note that after one chooses a norm for the reconstructed phase space, the scale increases with the embedding dimension m (in the case of the Euclidean norm, the scaling is $m^{1/2}$). Following the power-law behavior of $\lambda(\epsilon) \sim \epsilon^{-1/H}$, if we start from a power-law decaying K_2 entropy, $K_2(\epsilon) \sim \epsilon^{-1/H}$, then in the case of the Euclidean norm, we can express it in terms of the embedding dimension,

$$K_2(m) \sim m^{-1/2H},$$

therefore, it converges rapidly to 0 when m is increased.

We have found that for noisy chaotic dynamics, on very small scales, the SDLE scales with ϵ as $\lambda(\epsilon) \sim -\gamma \ln \epsilon$. It is of considerable interest to determine the coefficient γ in terms of other quantities such as the dimension. This will be one of our future tasks.

Throughout the paper, we have assumed that the dataset under consideration is not only finite, but of small or medium size. What would the SDLE behave if the dataset can be extremely large? One possibility is that on scales smaller than where the behavior of $\lambda(\epsilon) \sim -\gamma \ln \epsilon$ is observed,

one could observe another plateau, indicating existence of a high-entropic chaotic state. Nonlinear maps used as random number generators [30] belong to this category. For practical applications, one may not be concerned about this, however.

-
- [1] W. E. Leland, M. S. Taqqu, W. Willinger, and D. V. Wilson, *IEEE/ACM Trans. Netw.* **2**, 1 (1994); J. Beran, R. Sherman, M. S. Taqqu, and W. Willinger, *IEEE Trans. Commun.* **43**, 1566 (1995).
- [2] J. B. Gao, N. S. V. Rao, J. Hu, and J. Ai, *Phys. Rev. Lett.* **94**, 198702 (2005); J. B. Gao and N. S. V. Rao, *IEEE Commun. Lett.* **9**, 4 (2005).
- [3] M. Costa, A. L. Goldberger, and C. K. Peng, *Phys. Rev. E* **71**, 021906 (2005).
- [4] P. Gaspard and X. J. Wang, *Phys. Rep.* **235**, 291 (1993).
- [5] E. Aurell, G. Boffetta, A. Crisanti, G. Paladin, and A. Vulpiani, *J. Phys. A* **30**, 1 (1997).
- [6] M. Cencini, M. Falcioni, E. Olbrich, H. Kantz, and A. Vulpiani, *Phys. Rev. E* **62**, 427 (2000).
- [7] G. Boffetta, M. Cencini, M. Falcioni, and A. Vulpiani, *Phys. Rep.* **356**, 367 (2002).
- [8] P. Grassberger and I. Procaccia, *Phys. Rev. Lett.* **50**, 346 (1983); *Phys. Rev. A* **28**, 2591 (1983).
- [9] A. Wolf, J. B. Swift, H. L. Swinney, and J. A. Vastano, *Physica D* **16**, 285 (1985).
- [10] G. Sugihara and R. M. May, *Nature (London)* **344**, 734 (1990); D. T. Kaplan and L. Glass, *Phys. Rev. Lett.* **68**, 427 (1992); R. Wayland, D. Bromley, D. Pickett, and A. Passamante, *ibid.* **70**, 580 (1993); L. W. Salvino and R. Cawley, *ibid.* **73**, 1091 (1994); C. S. Poon and M. Barahona, *Proc. Natl. Acad. Sci. U.S.A.* **98**, 7107 (2001); J. Hu, W. W. Tung, J. B. Gao, and Y. H. Cao, *Phys. Rev. E* **72**, 056207 (2005).
- [11] J. Gao and Z. Zheng, *Phys. Rev. E* **49**, 3807 (1994); J. B. Gao and Z. M. Zheng, *Europhys. Lett.* **25**, 485 (1994).
- [12] A. R. Osborne and A. Provenzale, *Physica D* **35**, 357 (1989).
- [13] A. Provenzale, A. R. Osborne, and R. Soj, *Physica D* **47**, 361 (1991).
- [14] P. Gaspard, M. E. Briggs, M. K. Francis, J. V. Sengers, R. W. Gammons, J. R. Dorfman, and R. V. Calabrese, *Nature (London)* **394**, 865 (1998); C. P. Dettmann and E. G. D. Cohen, *J. Stat. Phys.* **101**, 775 (2000).
- [15] J. B. Gao, S. K. Hwang, and J. M. Liu, *Phys. Rev. Lett.* **82**, 1132 (1999); J. B. Gao, C. C. Chen, S. K. Hwang, and J. M. Liu, *Int. J. Mod. Phys. B* **13**, 3283 (1999).
- [16] S. K. Hwang, J. B. Gao, and J. M. Liu, *Phys. Rev. E* **61**, 5162 (2000).
- [17] N. H. Packard, J. P. Crutchfield, J. D. Farmer, and R. S. Shaw, *Phys. Rev. Lett.* **45**, 712 (1980); F. Takens, in *Dynamical Systems and Turbulence*, edited by D. A. Rand and L. S. Young, *Lecture Notes in Mathematics*, Vol. 898 (Springer-Verlag, Berlin, 1981), p. 366; T. Sauer, J. A. Yorke, and M. Casdagli, *J. Stat. Phys.* **65**, 579 (1991).
- [18] J. B. Gao and Z. M. Zheng, *Phys. Lett. A* **181**, 153 (1993).
- [19] M. C. Mackey and L. Glass, *Science* **197**, 287 (1977).
- [20] J. B. Gao, Y. H. Cao, and J. M. Lee, *Phys. Lett. A* **314**, 392 (2003).
- [21] J. Gao, J. Hu, W. W. Tung, Y. Cao, N. Sarshar, and V. P. Roychowdhury, *Phys. Rev. E* **73**, 016117 (2006).
- [22] B. B. Mandelbrot, *The Fractal Geometry of Nature* (Freeman, San Francisco, 1982).
- [23] U. Frisch, *Turbulence—The Legacy of A. N. Kolmogorov* (Cambridge University Press, Cambridge, 1995).
- [24] C.-K. Peng, S. V. Buldyrev, S. Havlin, M. Simons, H. E. Stanley, and A. L. Goldberger, *Phys. Rev. E* **49**, 1685 (1994).
- [25] B. B. Mandelbrot, *Fractals and Scaling in Finance* (Springer, New York, 1982).
- [26] T. H. Solomon, E. R. Weeks, and H. L. Swinney, *Phys. Rev. Lett.* **71**, 3975 (1993); S. Stapf, R. Kimmich, and R. O. Seitter, *ibid.* **75**, 2855 (1993); G. M. Viswanathan, V. Afanasyev, S. V. Buldyrev, E. J. Murphy, P. A. Prince, and H. E. Stanley, *Nature (London)* **381**, 413 (1996).
- [27] A. Janicki and A. Weron, *Stat. Sci.* **9**, 109 (1994).
- [28] J. B. Gao and W. W. Tung, *Biol. Cybern.* **86**, 263 (2002); J. B. Gao, *Med. Biol. Eng. Comput.* **52**, 345 (2004).
- [29] J. B. Gao, *Physica D* **106**, 49 (1997); J. B. Gao, S. K. Hwang, and J. M. Liu, *Phys. Rev. A* **59**, 1582 (1999).
- [30] M. Falcioni, L. Palatella, S. Pigolotti, and A. Vulpiani, *Phys. Rev. E* **72**, 016220 (2005); H. Kantz and E. Olbrich, *Physica A* **280**, 34 (2000).