Chaos or noise: Difficulties of a distinction

M. Cencini,[1,2] M. Falcioni,[1] E. Olbrich,[2] H. Kantz,[2] and A. Vulpiani[1]
[1]*Dipartimento di Fisica, Università di Roma ''La Sapienza'' and INFM, Unità di Roma, Piazzale Aldo Moro 2, I-00185 Roma, Italy*
[2]*Max-Planck-Institut für Physik komplexer Systeme, Nöthnitzer Strasse 38, D-01187 Dresden, Germany*
(Received 31 January 2000)

In experiments, the dynamical behavior of systems is reflected in time series. Due to the finiteness of the observational data set, it is not possible to reconstruct the invariant measure up to an arbitrarily fine resolution and an arbitrarily high embedding dimension. These restrictions limit our ability to distinguish between signals generated by different systems, such as regular, chaotic, or stochastic ones, when analyzed from a time series point of view. We propose to classify the signal behavior, without referring to any specific model, as stochastic or deterministic on a certain scale of the resolution ϵ, according to the dependence of the (ϵ,τ) entropy, $h(\epsilon,\tau)$, and the finite size Lyapunov exponent $\lambda(\epsilon)$ on ϵ.

PACS number(s): 05.45.Tp

I. INTRODUCTION

It is a long debated question if and by what means we can distinguish whether an observed irregular signal is deterministically chaotic or stochastic [1–6]. If the signal was obtained by iterating a certain model on a computer, we can give a definite answer, because we know the law which generated the signal.

In the case of time series recorded from experimental measurements, we are in a totally different situation. Indeed, in most cases, there is no unique model of the ''system'' which produced the data. Moreover, we will see that knowing the character of the model might not be an adequate answer to the question of the character of the signal. For example, data of Brownian motion can be modeled by a deterministic regular process as well as by a deterministic chaotic or stochastic process, as we will show in Sec. IV.

In principle, if we were able to determine the maximum Lyapunov exponent (λ) or the Kolmogorov-Sinai entropy per unit time (h_{KS}) of a data sequence, we would know, with no uncertainty, whether the sequence is generated by a deterministic law (in which case $\lambda, h_{KS} < \infty$) or by a stochastic law (in which case $\lambda, h_{KS} \to \infty$).

In spite of their conceptual relevance, there are evident practical problems with such quantities that are defined as infinite time averages taken in the limit of arbitrary fine resolution, since, typically, we have access only to a finite (and often very limited) range of scales. In order to cope with these limitations, in this paper we make use of the ''finite size Lyapunov exponent'' (FSLE) [7], a variant of the maximum Lyapunov exponent, and the (ϵ,τ) entropy per unit time [8–10], a generalization of the Kolmogorov-Sinai entropy per unit time. Basically, while for evaluating λ and h_{KS} one has to detect the properties of a system with infinite resolution, for determining the FSLE, $\lambda(\epsilon)$, or the (ϵ,τ) entropy per unit time, $h(\epsilon,\tau)$, the investigation on the system is performed at a finite scale ϵ, i.e., with a finite resolution. $\lambda(\epsilon)$ gives us the average exponential rate of the divergence between close (on a scale ϵ) trajectories of a system, and $h(\epsilon,\tau)$ is the average rate of information needed for prediction. If properly defined,

$$h(\epsilon,\tau) \underset{\epsilon,\tau\to 0}{\to} h_{KS} \quad \text{and} \quad \lambda(\epsilon) \underset{\epsilon\to 0}{\to} \lambda,$$

if $\lambda \geqslant 0$. Thus if we have the possibility of determining the behavior of $\lambda(\epsilon)$ or $h(\epsilon,\tau)$ for arbitrarily small scales, as pointed out above, we could answer the original question about the character (deterministic or stochastic) of the law that generated the recorded signal.

However, the limits of infinite time and resolution, besides being unattainable when dealing with experimental data, may also result to be physically uninteresting. As a matter of fact, it is now clear that the maximum Lyapunov exponent and the Kolmogorov-Sinai entropy are not completely satisfactory for a proper characterization of the many faces of complexity and predictability of nontrivial systems, such as, for instance, intermittent systems [11] or systems with many degrees of freedom [7,12]. For example, in the case of the maximum Lyapunov exponent, one has to consider infinitesimal perturbations, i.e., infinitesimally close trajectories or infinite resolution, respectively. In systems with many degrees of freedom (e.g. turbulence), an infinitesimal perturbation means, from a physical point of view, that the differences $\delta x_k = x'_k - x_k$ of the components, x'_k and x_k of the initially close state vectors $\mathbf{x}'$ and $\mathbf{x}$, have to be much smaller than the typical values $\tilde{x}_k$ of the variables x_k. If the $\tilde{x}_k$'s take very different values, then the concept of infinitesimal perturbation becomes physically unimportant, in the event one is interested only in the evolution of the components with the largest typical values [7,13] (e.g., the large scales in a turbulent motion).

Taking into account all the limitations mentioned above, in particular the practical impossibility to reach an arbitrarily fine resolution, we propose a different point of view on the distinction between chaos and noise: it neither relies on a particular model for a given data set nor ignores the fact that the character of a signal may depend on the resolution of the observation. Indeed $h(\epsilon,\tau)$ [or equivalently $\lambda(\epsilon)$] usually displays different behaviors as the range of scales is varied. According to these different behaviors, as will become clear through the paper, one can define a notion of deterministic and stochastic behaviors, respectively, on a certain range of scales.

In Sec. II we recall the definitions of the (ϵ,τ) entropy and the finite size Lyapunov exponent. In Sec. III we discuss how one can consistently classify the stochastic or chaotic

character of a signal by using information theoretic concepts such as the (ϵ,τ) entropy or the redundancy and compare our approach with previous attempts. In Sec. IV we discuss some examples showing that systems at opposite ends in the realm of complexity can give similar results when analyzed from a time series point of view. Section V is devoted to a critical discussion of some recent, intriguing and (sometimes) controversial results on data analysis of ''microscopic'' chaos, in particular we comment on the point of view to be adopted in interpreting the result of Sec. IV. In Sec. VI the reader finds some remarks on nontrivial behaviors of high-dimensional systems. Section VII summarizes and concludes the paper.

II. TWO CONCEPTS FOR A RESOLUTION-DEPENDENT TIME SERIES ANALYSIS

A. (ϵ,τ) entropy and redundancy

In this section we recall the definition of the (ϵ,τ) entropy discussing its numerical computation, and possible technical problems, as well as its properties. We start with a continuous (in time) variable $\mathbf{x}(t)\in \mathrm{IR}^d$, which represents the state of a d-dimensional system. We discretize the time by introducing a time interval τ and we consider the new variable

$$\mathbf{X}^{(m)}(t)=(\mathbf{x}(t),\mathbf{x}(t+\tau),\ldots,\mathbf{x}(t+m\tau-\tau)). \tag{1}$$

Of course $\mathbf{X}^{(m)}(t)\in \mathrm{IR}^{md}$, and it corresponds to the discretized trajectory in a time interval $T=m\tau$.

Usually, in data analysis, the space where the state vectors of the system live is not known. Mostly, only a scalar variable $u(t)$ can be measured. In these cases one considers vectors $\mathbf{X}^{(m)}(t)=(u(t),u(t+\tau),\ldots,u(t+m\tau-\tau))$, that live in IR^m and allow a reconstruction of the original phase space, known as delay embedding in the literature [14,15]. It can be viewed as a special case of Eq. (1).

We now introduce a partition of the phase space IR^d, using cells of length ϵ in each of the d directions. Since the region where a bounded motion evolves contains a finite number of cells, each $\mathbf{X}^{(m)}(t)$ can be coded into a word of length m, out of a finite alphabet:

$$\mathbf{X}^{(m)}(t)\rightarrow W^m(\epsilon,t)=(i(\epsilon,t),i(\epsilon,t+\tau),\ldots,i(\epsilon,t+m\tau-\tau)), \tag{2}$$

where $i(\epsilon,t+j\tau)$ labels the cell in IR^d containing $\mathbf{x}(t+j\tau)$. From the time evolution of $\mathbf{X}^{(m)}(t)$ one obtains, under the hypothesis of stationarity, the probabilities $P(W^m(\epsilon))$ of the admissible words $\{W^m(\epsilon)\}$. We can now introduce the (ϵ,τ) entropy per unit time, $h(\epsilon,\tau)$ [9],

$$h_m(\epsilon,\tau)=\frac{1}{\tau}[H_{m+1}(\epsilon,\tau)-H_m(\epsilon,\tau)], \tag{3}$$

$$h(\epsilon,\tau)=\lim_{m\rightarrow\infty} h_m(\epsilon,\tau)=\frac{1}{\tau}\lim_{m\rightarrow\infty}\frac{1}{m}H_m(\epsilon,\tau), \tag{4}$$

where H_m is the block entropy of block length m:

$$H_m(\epsilon,\tau)=-\sum_{\{W^m(\epsilon)\}} P(W^m(\epsilon))\ln P(W^m(\epsilon)). \tag{5}$$

For the sake of simplicity, we ignored the dependence on the details of the partition. For a more rigorous definition one has to take into account all partitions with elements of a size smaller than ϵ, and then define $h(\epsilon,\tau)$ by the infimum over all these partitions (see, e.g., Ref. [10]). In numerical calculations we circumvent this difficulty by using coverings instead of partitions (see below).

A concept which is complementary to the ϵ entropy is the ϵ redundancy (see, e.g., Ref. [16]), which measures the amount of uncertainty on future observations which can be removed by the knowledge of the past, namely,

$$r_m(\epsilon,\tau)=\frac{1}{\tau}[H_1(\epsilon,\tau)-(H_{m+1}(\epsilon,\tau)-H_m(\epsilon,\tau,))],$$

where $H_1(\epsilon)$ estimates the uncertainty of the single outcome of the measurement, i.e., neglecting possible correlations in the signal. Alternatively, we can write the redundancy in the form

$$r_m(\epsilon,\tau)=\frac{1}{\tau}H_1(\epsilon,\tau)-h_m(\epsilon,\tau), \tag{6}$$

which emphasizes the complementarity between the redundancy and entropy. If the data are totally independent, one has $H_m(\epsilon,\tau)=mH_1(\epsilon)$ and, therefore, $r_m(\epsilon,\tau)=0$. On the opposite side, in the case of a periodic signal the redundancy is maximal $r_m(\epsilon,\tau)=H_1(\epsilon,\tau)/\tau$.

The Kolmogorov-Sinai (KS) entropy h_{KS} is obtained by taking the limit $\epsilon,\tau\rightarrow 0$:

$$h_{KS}=\lim_{\tau\rightarrow 0}\lim_{\epsilon\rightarrow 0} h(\epsilon,\tau). \tag{7}$$

The KS entropy is a dynamical invariant, i.e., it is independent of the employed state representation (1), while this is not the case for the ϵ entropy [Eq. (4)]. To simplify the notation we drop the τ dependence in the following, apart from cases in which the τ dependency is explicitly considered as in Sec. IV.

In a genuine deterministic chaotic system one has $0<h_{KS}<\infty$ ($h_{KS}=0$ for a regular motion), while for a random process $h_{KS}=\infty$. The entropies $H_m(\epsilon)$ were above introduced using a partition and the usual Shannon entropy; however, it is possible to arrive at the same notion starting from other entropylike quantities, which are more suitable for numerical investigations. Following Cohen and Procaccia [17], one can estimate $H_m(\epsilon)$ as follows. Given a signal composed of N successive records and the embedding variable $\mathbf{X}^{(m)}$, let us introduce the quantities

$$n_j^{(m)}=\frac{1}{N-m}\sum_{i\neq j}\Theta(\epsilon-|\mathbf{X}^{(m)}(i\tau)-\mathbf{X}^{(m)}(j\tau)|); \tag{8}$$

then the block entropy $H_m(\epsilon)$ is given by

$$H_m^{(1)}(\epsilon)=-\frac{1}{(N-m+1)}\sum_j \ln n_j^{(m)}(\epsilon). \tag{9}$$

In practice $n_j^{(m)}(\epsilon)$ is an approximation of $P(W^m(\epsilon))$. From the numerical point of view the even more suited quantities are the correlation entropies [18,19]

$$H_m^{(2)}(\epsilon) = -\ln\left(\frac{1}{N-m+1}\sum_j n_j^{(m)}(\epsilon)\right) \leqslant H_m^{(1)}(\epsilon), \quad (10)$$

where one approximates the Shannon entropy by the Renyi entropy of order $q=2$.

In the determination of h_{KS} by data analysis, one has to consider some subtle points (see Ref. [20] for a detailed discussion). Let us just make some remarks about the general problems in the computation of the Kolmogorov-Sinai entropy from a time series of a deterministic system. The first point is the value of the embedding dimension m. Let us assume that the information dimension of the attractor of the deterministic system is D. In order to be able to observe a finite entropy, m has to be larger than D, since the behavior of the entropies in the limit $\epsilon \to 0$ is

$$h_m(\epsilon) = \text{const} + O(\epsilon) \geqslant h_{KS}, \quad (11)$$

provided $m > D$ [21]. The second relevant point is the fact that the saturation, i.e. the regime where the entropy $h_m(\epsilon)$ does not depend on the length scale ϵ, can be observed only on length scales smaller than some ϵ_u. Thus it is possible to distinguish a deterministic signal from a random one only for $\epsilon < \epsilon_u$. Due to the finiteness of the data set there is a lower scale ϵ_l below which no information can be extracted from the data. Taking into account the number of points of the series, N, it is possible to give the following relation between the embedding dimension, the KS entropy, the information dimension, and the saturation range ϵ_u/ϵ_l [22]:

$$\frac{\epsilon_u}{\epsilon_l} \leqslant (N e^{-m\tau h_{KS}})^{1/D}, \quad (12)$$

where ϵ_u and ϵ_l are the upper and lower bounds of the interval of scales at which the deterministic character of a deterministic signal shows up. Note that this relation does not determine ϵ_u. For more details, see Ref. [22]. If m is not large enough and/or ϵ is not small enough, one can obtain misleading results; e.g., see Sec. V.

The ϵ entropy $h(\epsilon,\tau)$ is also well defined for stochastic processes. Its dependence on ϵ can give some insight into the underlying stochastic process [10]. In the case of finite τ, it is possible to define a saturation range; below some length scale $\epsilon_u(\tau)$, we have

$$h_m(\epsilon) = \text{const} - \ln \epsilon + O(\epsilon). \quad (13)$$

However, the limit $\tau \to 0$ will lead to $\epsilon_u \to 0$; thus the saturation will disappear. As shown in Ref. [10], for some stochastic processes it is possible to give an explicit expression of $h(\epsilon,\tau)$ in this limit. For instance, in the case of a stationary Gaussian process with spectrum $S(\omega) \propto \omega^{-2}$, one has [8]

$$\lim_{\tau \to 0} h(\epsilon,\tau) \sim \frac{1}{\epsilon^2}, \quad (14)$$

the same scaling behavior is also expected for Brownian motion [10]. It can be recovered by looking at $h(\epsilon,\tau)$ in a certain (ϵ,τ) region. See Ref. [10] for a detailed derivation of Eq. (14). We have to stress that the behavior predicted by Eq. (14) may be difficult to observed experimentally due to problems related to the choice of τ (see Sec. IV).

B. Finite-size Lyapunov exponent

The finite size Lyapunov exponent was originally introduced in the context of the predictability problem in fully developed turbulence [7]. Such an indicator, as will become clear below, is for some aspects the dynamical systems counterpart of the ϵ entropy.

The basic idea of the FSLE is to define a growth rate for different sizes of the distance between a reference and a perturbed trajectory. In the following we discuss how the FSLE can be computed, by assuming to know the evolution equations. The generalization to data analysis is obtained following the usual ideas of ''embedology'' [15]. First, one has to define a norm to measure the distance $\epsilon(t) = \|\delta \mathbf{x}(t)\|$ between a reference and a perturbed trajectory. In finite-dimensional systems the maximum Lyapunov exponent is independent of the used norm. However, when one considers finite perturbations there could be a dependence on the norm (as for infinite-dimensional systems). Having defined the norm, one has to introduce a series of thresholds starting from a very small one ϵ_0, e.g., $\epsilon_n = r^n \epsilon_0$ $(n=1,\ldots,P)$, and to measure the ''doubling time'' $(T_r(\epsilon_n))$ at different thresholds. $T_r(\epsilon_n)$ is the time a perturbation of size ϵ_n takes to grow up to the next threshold, ϵ_{n+1}. The threshold rate r should not be taken too large, in order to avoid the error to grow through different scales. On the other hand, r cannot be too close to 1, because otherwise the doubling time would be of the order of the time step in the integration (sampling time in data analysis) affecting the statistics. Typically, one uses $r=2$ or $r=\sqrt{2}$. For simplicity T_r is called is ''doubling time'' even if $r \neq 2$.

The doubling times $T_r(\epsilon_n)$ are obtained by following the evolution of the distance $\|\delta \mathbf{x}(t)\|$ from its initial value $\epsilon_{min} \leqslant \epsilon_0$ up to the largest threshold ϵ_P. Knowing the evolution equations, this is obtained by integrating the two trajectories of the system starting at an initial distance ϵ_{min}. In general, one must choose $\epsilon_{min} \leqslant \epsilon_0$, in order to allow the direction of the initial perturbation to align with the most unstable direction in the phase space. Moreover, one must pay attention to keep $\epsilon_P < \epsilon_{saturation}$, so that all the thresholds can be attained ($\epsilon_{saturation}$ is the typical distance of two uncorrelated trajectories).

The evolution of the error from the initial value ϵ_{min} to the largest threshold ϵ_P carries out a single error-doubling experiment. At this point the model trajectory is rescaled at a distance ϵ_{min} with respect to the true one, and another experiment begins. After $\mathcal{N}$ error-doubling experiments, we can estimate the expectation value of some quantity A as

$$\langle A \rangle_e = \frac{1}{\mathcal{N}} \sum_{i=1}^{\mathcal{N}} A_i. \quad (15)$$

This is not the same as taking the time average, because each error-doubling experiment may require a different time than the others. For the doubling time we have

$$\lambda(\epsilon_n) = \frac{1}{\langle T_r(\epsilon_n) \rangle_e} \ln r; \quad (16)$$

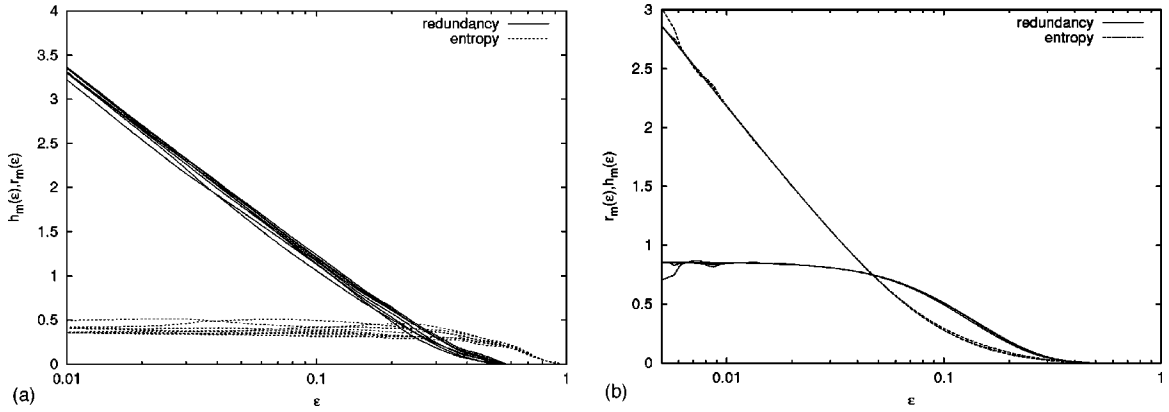

FIG. 1. (a) $h_m(\epsilon)$ (dashed line) and $r_m(\epsilon)$ (solid line) for the Henon map with the standard parameters ($a=1.4$ and $b=0.3$), with $m=2,\ldots,9$, and (b) The same for an AR(1) process, with $m=1,\ldots,5$ and fixed τ.

for details, see Ref. [7]. The method described above assumes that the distance between the two trajectories is continuous in time. This is not the case for maps or for a discrete sampling in time, thus the method has to be slightly modified. In this case $T_r(\epsilon_n)$ is defined as the minimum time at which $\epsilon(T_r) \geqslant r\epsilon_n$, and now we have [7]

$$\lambda(\epsilon_n) = \frac{1}{\langle T_r(\epsilon_n)\rangle_e} \left\langle \ln\left(\frac{\epsilon(T_r)}{\epsilon_n}\right)\right\rangle_e . \tag{17}$$

It is worth noting that the computation of the FSLE is not more expensive than the computation of the Lyapunov exponent by the standard algorithm. One simply has to integrate two copies of the system and this can also be done for very complex simulations.

One can expect that in systems with only one positive Lyapunov exponent, one has $\lambda(\epsilon) \simeq h(\epsilon)$; see Ref. [7] for details. Additionally it was shown in Ref. [23] how it is possible to use the FSLE for characterizing the predictability (also from the data analysis point of view) of systems containing a slow component and a fast one. Let us comment on some advantages of the FSLE with respect to the (ϵ,τ) entropy. For the FSLE it is not necessary to introduce an ϵ partition; most importantly, at variance with the (ϵ,τ) entropy, the algorithmic procedure automatically finds the ''proper time,'' so that it is not necessary to decide on the right sampling time and to test the convergence at varying the words block size N. This point will be discussed in Sec. IV C.

III. CLASSIFICATION BY ϵ DEPENDENCE

In Sec. II we discussed the ϵ entropy and the FSLE as tools to characterize dynamical processes. Let us re-examine the question of distinguishing chaos and noise posed in Sec. I. Equations (11) and (13) allow us to make rigorous statements about the behavior of the entropy in the limit $\epsilon \to 0$. Then the behavior of the redundancy can be determined by using the relation to the entropy, given by Eq. (6), if we take into account that $H_1(\epsilon) \propto -\ln\epsilon$ for continuously valued nonperiodic process. Both the behavior of the entropy and the redundancy are summarized in the following table [6]:

deterministic ($m>D$)		stochastic	
$r_m(\epsilon)\to\infty$		$h_m(\epsilon)\to\infty$	
chaotic	nonchaotic	white noise	colored noise
$\lim_{m\to\infty} h_m(\epsilon)>0$	$\lim_{m\to\infty} h_m(\epsilon)=0$	$r_m(\epsilon)=0$	$r_m(\epsilon)>0$

The behavior of the FSLE in the limit $\epsilon \to 0$ is similar to that of the ϵ entropy, $h(\epsilon)$. It is worth noting that the FSLE defined through the doubling times (see Sec. II B) is also zero if $\lambda<0$.

In all practical situations we have only a finite amount of data. Let us assume we have embedded the time series in an m-dimensional space, e.g., by time delay embedding. Then to this set of points one can relate an empirical measure μ^*:

$$\mu^*(\mathbf{X}^{(m)}) = \frac{1}{N}\sum_{i=1}^{N} \delta(\mathbf{X}^{(m)} - \mathbf{X}^{(m)}(i\tau)). \tag{18}$$

This empirical measure μ^* approximates the true measure only on length scales larger than a finite length scale ϵ_l. This means that we cannot perform the limit $\epsilon \to 0$. Of course, on a finite scale ϵ_l, both entropy and redundancy are always finite; therefore, we are unable to decide which will reach infinity for $\epsilon \to 0$. But we can define stochastic and deterministic behaviors of a time series at the length scale dependence of the entropy and redundancy. Figure 1 shows the typical behavior of the entropy $h_m(\epsilon)$ and redundancy $r_m(\epsilon)$ in the case of a deterministic model (a two-dimensional chaotic map) and a stochastic model (autoregressive model of first order, AR(1)).

For a time series long enough, a ''typical'' system can show a saturation range for both the entropy and the redundancy. For decreasing length scales ϵ with $\epsilon < \epsilon_u$, one observes the following behaviors:

Deterministic	Stochastic
$r_m(\epsilon) \propto -\ln\epsilon$	$h_m(\epsilon) \propto -\ln\epsilon$
$h_m(\epsilon) \approx$ const	$r_m(\epsilon) \approx$ const

In addition, as far as stochastic behaviors are concerned, the ϵ entropy can exhibit power laws on large scales, e.g., in the case of the diffusion (14) (see Sec. IV and Ref. [10] for further details).

If on some range of length scales either the entropy $h_m(\epsilon)$ or redundancy $r_m(\epsilon)$ is a constant, we call the signal deterministic or stochastic on these length scales, respectively. Thus we have a practical tool to classify the character of a signal as deterministic or stochastic without referring to a model, and we are no longer obliged to answer the metaphysical question of whether the system which produced the data was a deterministic or stochastic one.

Moreover, from this point of view, we are now able to give the notion of noisy chaos a clear meaning: chaotic scaling on large scales and stochastic scaling on small scales. We also have the notion of chaotic noise, namely, stochastic scaling on large scales and deterministic scaling on small scales. These notions will become clear with the examples in the following sections.

The method presented for distinguishing between chaos and noise is a refinement and generalization of one of the first discussed methods of approaching this problem: estimating the correlation dimension, and taking a finite value as a sign for the deterministic nature of the signal [24]. The main criticism of this approach is based on the work of Osborne and Provenzale [2], who claimed that stochastic systems with a power-law spectrum will produce time series which exhibit a finite correlation dimension. A detailed discussion of this problem is beyond the scope of this paper, but a main step toward clarifying the problem was taken by Theiler [25]. First, he noted that the discussed signals were nonstationary and highly correlated with correlation times of the order of the length of the time series. From a conservative point of view one has to stop at this point in any attempt to calculate dimensions or entropies. If one proceeds nonetheless, Theiler showed that the result will depend on the number of data points and the length scale. If one has a sufficient number of data points, for this kind of signals one will also encounter the embedding dimension which leads to the typical behavior given in Eq. (13) for the entropy. Moreover, if one uses a typical time delay embedding like Eq. (1) in contrast to Refs. [2] and [25], the result depends strongly on the chosen delay time.

We are aware that there are many other attempts to distinguish chaos from noise discussed in the literature. They are based on the difference in the predictability using prediction algorithms rather than the estimating the entropy [3,4], or they relate determinism to the smoothness of the signal [5,26]. All these methods have in common that one has to choose a certain length scale ϵ and a particular embedding dimension m. Thus they also could, in principle, shed light on the interesting crossover scenarios we are going to describe in Sec. IV.

IV. DIFFICULTIES IN THE DISTINCTION BETWEEN CHAOS AND NOISE: EXAMPLES

In this section we analyze in a detailed way some examples which illustrate how subtle the transition from large scale behavior to small scale behavior can be; and thus the difficulties arising in distinguishing, from data analysis alone, a genuine deterministic chaotic system from one with intrinsic randomness.

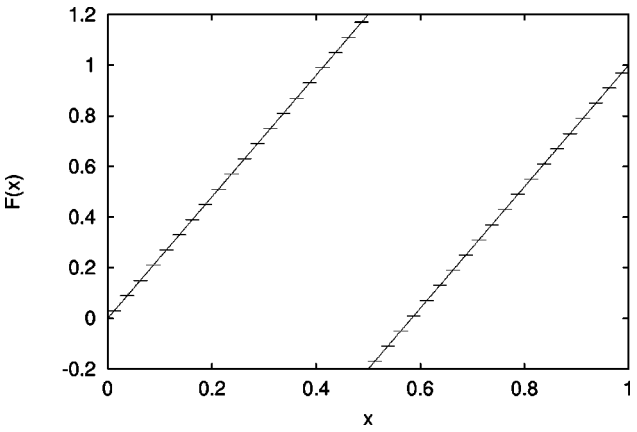

FIG. 2. The map $F(x)$ (20) for $\Delta=0.4$ is shown, superimposed with the approximating (regular) map $G(x)$ [Eq. (27)] obtained by using 40 intervals of slope 0.

A. Diffusive regime

We first discuss problems, due to the finite resolution, which one can have in analyzing experimental data. We consider a map which generates a diffusive behavior on the large scales [27],

$$x_{t+1}=[x_t]+F(x_t-[x_t]), \tag{19}$$

where $[x_t]$ indicates the integer part of x_t, and $F(y)$ is given by

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

The maximum Lyapunov exponent λ can be obtained immediately: $\lambda=\ln|F'|$, with $F'=dF/dy=2+\Delta$. Therefore one expects the following scenario for $h(\epsilon)$ [and for $\lambda(\epsilon)$]:

$$h(\epsilon)\approx\lambda \quad \text{for } \epsilon<1, \tag{21}$$

$$h(\epsilon)\propto\frac{D}{\epsilon^2} \quad \text{for } \epsilon>1, \tag{22}$$

where D is the diffusion coefficient, i.e.,

$$\langle(x_t-x_0)^2\rangle\approx 2\ D\ t \quad \text{for large } t. \tag{23}$$

Figures 3 and 4 show $\lambda(\epsilon)$ and $h(\epsilon)$, respectively. Let us briefly comment on a technical aspect. The numerical computation of $\lambda(\epsilon)$ does not present any particular difficulties; on the other hand, the results for $h(\epsilon)$ depend on the employed sampling time τ. This can be appreciated by looking at Fig. 25b of the review by Gaspard and Wang [10], where the power law behavior [Eq. (22)] in the diffusive region is obtained only if one considers the envelope of $h_m(\epsilon,\tau)$ evaluated for different values of τ; while looking at a single τ, one has a rather inconclusive result. This is due to the fact that, at variance with the FSLE, when computing $h(\epsilon,\tau)$ one has to consider very large m, in order to obtain a good convergence for $H_m(\epsilon)-H_{m-1}(\epsilon)$.

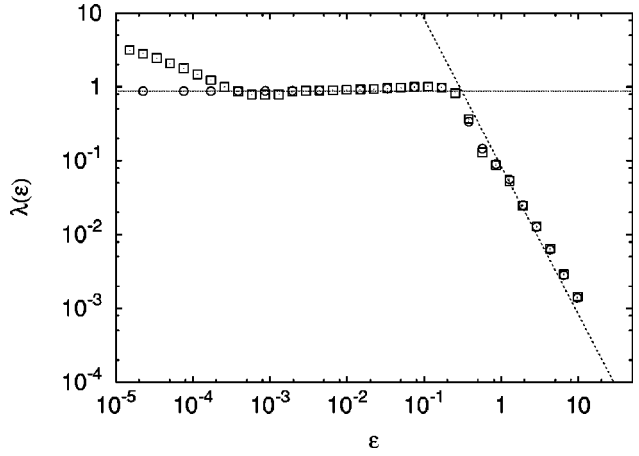

FIG. 3. $\lambda(\epsilon)$ vs ϵ obtained with the map $F(y)$ [Eq. (20)] with $\Delta=0.4$ ($\bigcirc$), and with the noisy (regular) map ($\square$) [Eq. (27)] with 10000 intervals of slope 0.9 with $\sigma=10^{-4}$. The straight lines indicates the Lyapunov exponent $\lambda=\ln 2.4$ and the diffusive behavior $\lambda(\epsilon)\sim\epsilon^{-2}$.

Because of the diffusive behavior, a simple dimensional argument shows that, by sampling the system every elementary time step, a good convergence holds for $m\geqslant\epsilon^2/D$. Thus, for $\epsilon=10$ and typical values of the diffusion coefficient $D\simeq 10^{-1}$, one has to consider an enormous block size. A possible way out of this computational difficulty may be the following. We call $l\,(=1)$ the length of the interval $[0,1]$ where $F(y)$ is defined; if we adopt a coarse-grained description on a scale $\epsilon=l$, i.e. we follow the evolution of the integer part of x_t, the dynamical system (19) is well described by means of a random walk [27], with a given probability p_0 that in a time step $s\,(=1)$ the integer part of x does not change: $[x_{t+s}]=[x_t]$, and probabilities $p_\pm$ that $[x_{t+s}]=[x_t]\pm 1$. A diffusive behavior means that the probability for changing $[x_t]$ by $\pm k$ in n elementary time steps is given by [28]

$$P(k,n)\simeq\frac{e^{-(k^2/2\alpha n)(l^2/s)}}{\sqrt{2\pi\alpha n}}\sqrt{\frac{l^2}{s}}, \tag{24}$$

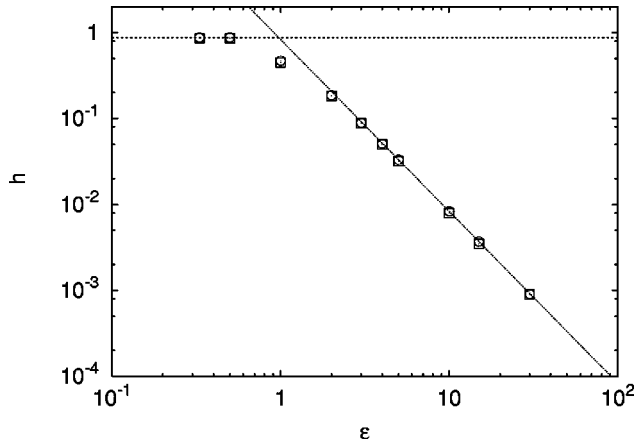

FIG. 4. (ϵ,τ) entropy for noisy ($\square$) and chaotic ($\bigcirc$) maps with the same parameters as in Fig. 3; the encoding method is explained in the text. The straight lines indicates the KS entropy $h_{KS}=\lambda=\ln 2.4$ and the diffusive behavior $h(\epsilon)\sim\epsilon^{-2}$. The region $\epsilon<\sigma$ has not been explored, due to computational costs.

with α a function of p_0. Now we increase the graining, and identify all the x values in a cell of size $\epsilon=L$, with L an integer multiple of l. If we observe the coarse-grained state of the system only every $\tau(>s)$ time steps, and define the variables $\tilde{k}$ and $\tilde{n}$, such that $kl\simeq\tilde{k}L$ and $ns\simeq\tilde{n}\tau$, we can write

$$P(\tilde{k},\tilde{n})\simeq\frac{e^{-(\tilde{k}^2/2\alpha\tilde{n})(L^2/\tau)}}{\sqrt{2\pi\alpha\tilde{n}}}\sqrt{\frac{L^2}{\tau}} \tag{25}$$

for the probability of finding the system $\tilde{k}$ L cells apart after $\tilde{n}$ τ intervals. Thus, by choosing $L^2/\tau=l^2/s$, we expect that the sequences generated by checking the system either on a scale L every τ steps or on a scale l every elementary time step s, have the same statistics, in particular the same entropy, as a signature of a diffusive behavior. Note that if $\lim_{m\to\infty}H_m(\epsilon,\tau=\epsilon^2)/m$ is constant at varying ϵ, as we found numerically, then the ϵ entropy per unit time $\lim_{m\to\infty}H_m(\epsilon,\tau)/m\tau$ goes like $1/\epsilon^2$.

Since the equality $L^2/\tau=l^2/s$ is assured by the choices $\epsilon=L=\gamma l$ and $\tau=\gamma^2 s$, one can expect that, for a diffusion process, the following scaling relation holds: $\lim_{m\to\infty}H_m(\epsilon,\tau)/m=\lim_{m\to\infty}H_m(\gamma\epsilon,\gamma^2\tau)/m$, with γ an arbitrary scaling parameter. This scaling relation allows us to see why the power law behavior (14) is expected to be valid generally for the Brownian motion. Indeed, if we choose $\gamma=1/\epsilon$ we have $H_m(\epsilon,\tau)=H_m(1,\tau/\epsilon^2)$ and, finally, taking the limit $\tau\to 0$, the ϵ entropy is given by

$$\begin{aligned} h(\epsilon)&=\lim_{\tau\to 0}\frac{[H_{m+1}(1,\tau/\epsilon^2)-H_m(1,\tau/\epsilon^2)]}{\tau}\\ &\simeq\lim_{\tau\to 0}\frac{\text{const}\times\tau/\epsilon^2+O(\tau^2)}{\tau}\propto\frac{1}{\epsilon^2}, \end{aligned} \tag{26}$$

which is Eq. (14). Note that the first equality in Eq. (26) was obtained by a Taylor expansion around $\tau=0$, and by noting that $h(1,0)=0$ otherwise, the entropy for unit time will be infinite at finite ϵ, which is impossible.

B. Finite resolution effects

We now consider a stochastic system, namely, a map with dynamical noise

$$x_{t+1}=[x_t]+G(x_t-[x_t])+\sigma\eta_t, \tag{27}$$

where $G(y)$ is shown in Fig. 2, and η_t is a noise with uniform distribution in the interval $[-1,1]$, and no correlation in time. As can be seen from Fig. 2, the new map $G(y)$ is a piecewise linear map which approximates the map $F(y)$. When $dG/dy<1$, as is the case we consider, map (27), in the absence of noise, gives a nonchaotic time evolution.

Now one can compare the chaotic case, i.e., Eq. (19) with the approximated map (27) with noise. For example let us start with the computation of the finite size Lyapunov exponent for the two cases. Of course from a data analysis point of view we have to compute the FSLE by reconstructing the dynamics by embedding. However, in this example we are interested only in discussing the resolution effects. There-

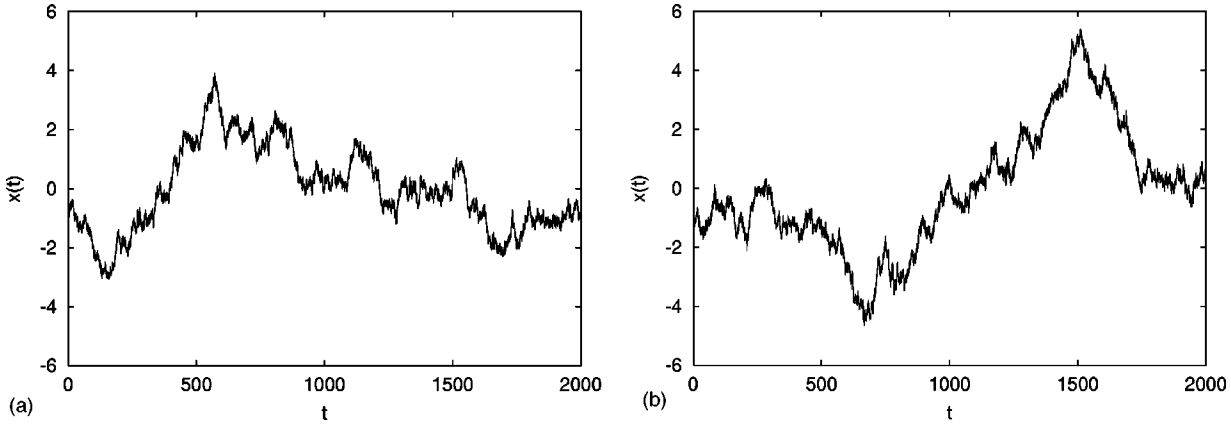

FIG. 5. (a) Time record obtained from Eq. (28) with the frequencies chosen as discussed in the text with $M=10^4$, $C=0.005$ and $\mu_0/\mu=10^{-6}$, the numerically computed diffusion constant is $D\approx 0.007$. The length of the data set is 10^5, and the data are sampled with $\Delta t=0.02$. (b) Time record obtained from an artificial Brownian motion [Eq. (30)], with the same value of the diffusion constant as in (a).

fore, we compute the FSLE directly by integrating the evolution equations for two (initially) very close trajectories, in the case of noisy maps, using two different realizations of the noise. In Fig. 3 we show $\lambda(\epsilon)$ versus ϵ both for the chaotic [Eq. (19)] and the noisy [Eq. (27)] maps. As one can see, the two curves are practically indistinguishable in the region $\epsilon > \sigma$. The differences appear only at very small scales $\epsilon<\sigma$, where one has a $\lambda(\epsilon)$ which grow with ϵ for the noisy case, remaining at the same value for the chaotic deterministic chase.

Both the FSLE and the (ϵ,τ)-entropy analysis (see Figs. 3 and 4) show that we can distinguish three different regimes, observing the dynamics of Eq. (27) on different length scales. On the large length scales $\epsilon>1$ we observe diffusive behavior in both models. At length scales $\sigma<\epsilon<1$, both models show a chaotic deterministic behavior, because the entropy and the FSLE are independent of ϵ and larger than zero. Finally on the smallest length scales $\epsilon<\sigma$ we see stochastic behavior for system (27), while system (19) still shows a chaotic behavior.

C. Effects of finite block length

In Sec. IV B we discussed the difficulties arising in classifying a signal as chaotic or stochastic because of the impossibility of reaching an arbitrary fine resolution. Here we investigate the reasons which make it difficult to distinguish a stochastic behavior from a deterministic nonchaotic one. In particular, we show that a nonchaotic deterministic system may produce a signal practically indistinguishable from a stochastic one, provided its phase space dimension is large enough.

The simplest way to generate a nonchaotic (regular) signal having statistical properties similar to a stochastic one is by considering the Fourier expansion of a random signal [2]. One can consider the signal

$$x(t)=\sum_{i=1}^{M} X_{0i}\sin(\Omega_i t+\phi_i), \tag{28}$$

where the frequencies are equispaced discrete frequencies, i.e., $\Omega_i=\Omega_0+i\Delta\Omega$, the phases ϕ_i are random variables uniformly distributed in $[0,2\pi]$, and the coefficient X_{0i} are chosen in such a way to have a definite power spectrum, e.g., a power law spectrum, which is a common characteristic of many natural signals. Of course Eq. (28) can be considered as the Fourier expansion of a stochastic signal only if one consider a set of $2M$ points such that $M\Delta\Omega=\pi/\Delta t$, where Δt is the sampling time [2]. Time series like Eq. (28) have been used to claim that suitable stochastic signals may display a finite correlation dimension [2,29]; see the discussion in Sec. III.

Here we adopt a slightly different point of view. Signal (28) can also be considered as the displacement of a harmonic oscillator linearly coupled to other harmonic oscillators. Indeed, it has been well known for a long time that a large ensemble of harmonic oscillators can originate stochasticlike behaviors. In particular, we refer to Ref. [30], where it was proved that an impurity of mass μ linearly coupled to a one-dimensional equal mass, μ_0, chain of M oscillators coupled by a nearest-neighbor harmonic interaction, in the limit of $\mu\gg\mu_0$ and of infinite oscillators ($M\to\infty$), undergoes a Brownian motion. Our observable is practically given by the sum of harmonic oscillations as in Eq. (28), where the frequencies Ω_i were derived in the limit $\mu_0/\mu\ll 1$ by Cukier and Mazur [30]. The phases ϕ_i are chosen as uniformly distributed random variables in $[0,2\pi]$ and the amplitudes X_{0i} are chosen as

$$X_{0i}=C\Omega_i^{-1}, \tag{29}$$

where the C is an arbitrary constant and the Ω dependence is just to obtain a diffusivelike behavior. Note that for a signal of length $2M$ the random phases and the X_{0i}'s represent a initial condition of the M oscillators, because their phase space is $2M$ dimensional.

In Fig. 5(a) we show an output of signal (28); for comparison, in Fig. 5(b) we also show an artificial continuous time Brownian motion obtained by integrating the equation

$$\frac{dx(t)}{dt}=\xi(t), \tag{30}$$

where $\xi(t)$ is a Gaussian white noise, produced by a random number generator [the variance of the process is chosen as to mimic that obtained by Eq. (28)]. Because the random num-

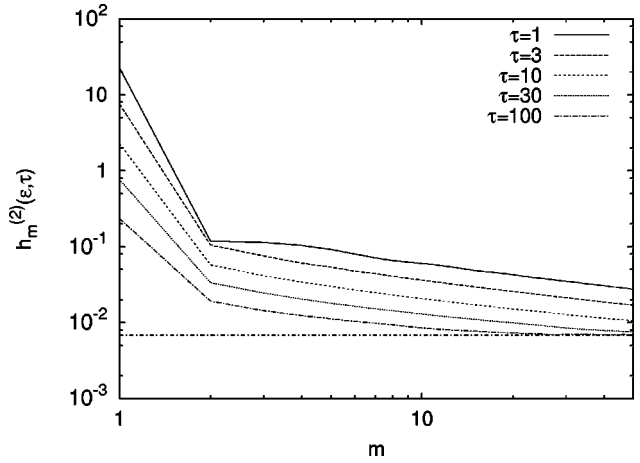

FIG. 6. Dependence of the embedding dimension of the ϵ entropy calculated with the Grassberger-Procaccia algorithm using $\epsilon = 1.85$, from the time series shown in Fig. 5(a). The horizontal line only indicates a possible numerically evaluated value for the saturation of the entropy.

ber generator uses a high entropic one-dimensional deterministic map, this is an example of a high entropic low-dimensional system, which produces a stochastic behavior. It is possible to see that the two signals appear to be very similar.

Now it is important to stress that if $M<\infty$ the signals obtained according to Eq. (28) cannot develop a true Brownian motion, especially if one is interested in long time series. Indeed for a long enough record one should be able to recognize the regularities in the trajectory of $x(t)$. However, even if the time record is long enough, in order to give a definite answer about the value of the entropy one also requires very large embedding dimensions. The basic fact is that a deterministic behavior can be observed only if the embedding dimension m is larger than the dimension of the manifold where the motion takes place, which is M for M harmonic oscillators. This means that although the entropy h_{KS} is zero, the conditional entropies $h_m(\epsilon,\tau)=(H_{m+1}(\epsilon) - H_m(\epsilon))/\tau$ for finite m are nonzero, and may even be slowly decreasing for $m>M$. Moreover, one can encounter some quasiconvergences with respect to m for $m<M$, if τ is large enough; i.e., the entropy can seem to be independent of m, e.g. see Fig. 6.

In Figs. 7(a) and 7(b) we show the ϵ entropy, calculated by the Grassberger-Procaccia method [18]. The deterministic signal [Eq. (28)] and the stochastic one [Eq. (30)] indeed produce very similar results. Note that we calculated $h_m^{(2)}(\epsilon,\tau)$ instead of $h_m^{(1)}(\epsilon,\tau)$, because it is used more often due to better statistics in most cases. However, in Fig. 8 one can see that, on relevant length scales, both entropies are almost equal.

As in Refs. [10,31,32], we considered different time delays τ in computing the ϵ entropy because of the problems discussed in Sec. III. The power law behavior ϵ^{-2} for the ϵ entropy is finally obtained only as an envelope of different computations with different delay times. The results for the FSLE calculated from the time series are shown in Fig. 9. Both the ϵ entropy and the FSLE display $1/\epsilon^2$ behavior, which denotes that the signals can be classified as Brownian motion [10]. It is worth noting that the FSLE computed from the time record is not too sensitive on the choice of the delay time τ and the embedding dimension m. From this simple example is it easy to understand that the impossibility of reaching high enough embedding dimensions severely limits our ability to make definite statements about the ''true'' character of the system which generated a given time series, as well as the already analyzed problem of the lack of resolution.

V. SOME REMARKS ON A RECENT DEBATE ABOUT ''MICROSCOPIC'' CHAOS

The issue of the detection of ''microscopic'' chaos by data analysis has recently received some attention [32,33] after a work by Gaspard *et al.* [31]. Gaspard *et al.*, from an entropic analysis of an ingenious experiment on the position of a Brownian particle in a liquid, claim to give an empirical evidence for microscopic chaos, i.e., they claimed to give evidence that the diffusive behavior observed for a Brownian particle is the consequence of chaos on the molecular scale. Their work can be briefly summarized as follows: from a long ($\approx 1.5\times 10^5$ data) record on the position of a Brownian particle they computed the ϵ entropy with the Cohen-Procaccia [17] method, described in Sec. II, from which they obtained

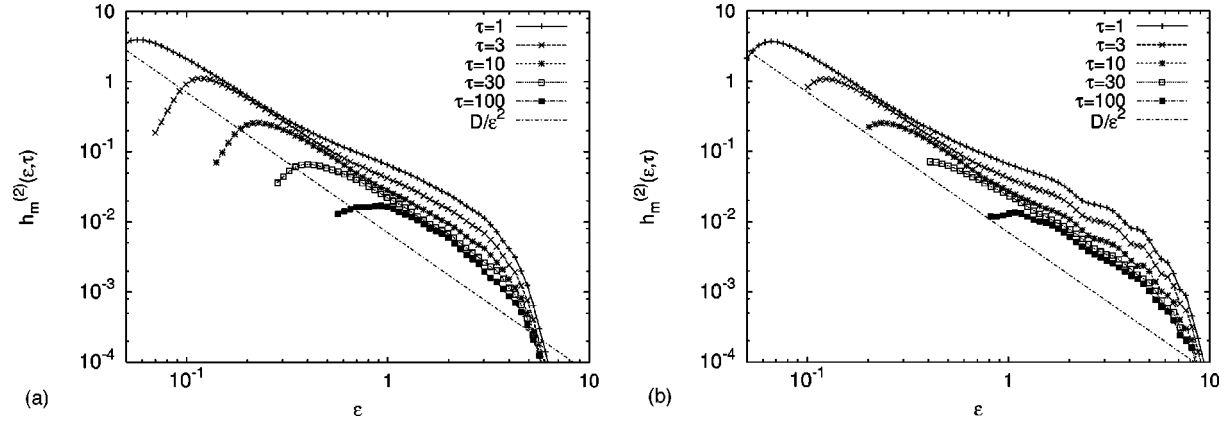

FIG. 7. ϵ entropy calculated with the Grassberger-Procaccia algorithm using using 10^5 points from the time series shown in Fig. 5. We show the results for an embedding dimension $m=50$. The two straight lines show the D/ϵ^2 behavior.

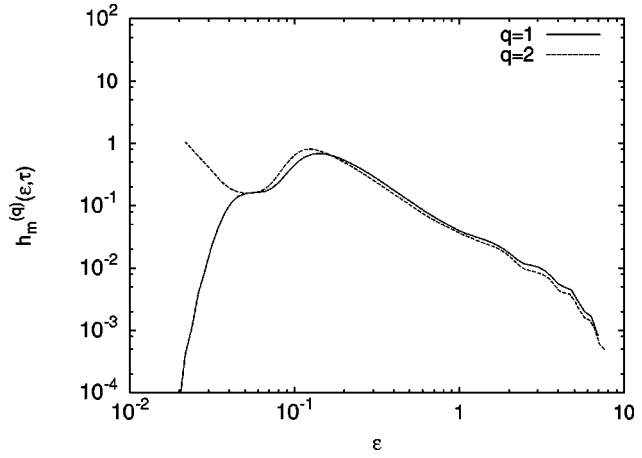

FIG. 8. ϵ entropy estimated by the Cohen-Procaccia algorithm ($q=1$) compared to the entropies calculated by the Grassberger-Procaccia algorithm ($q=2$) for $m=20$ and $\tau=0.2$. Let us stress that the behavior below $\epsilon=0.2$ is essentially a finite sample effect.

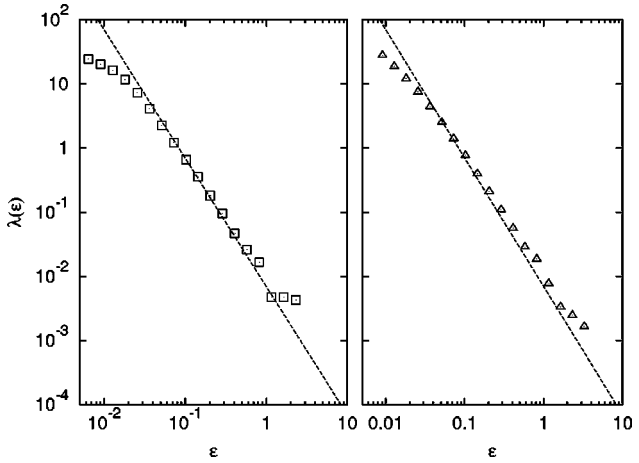

FIG. 9. Finite size Lyapunov exponent from an artificial Brownian motion (right), and for the time series (28) (left). Embedding dimension $m=20$. We consider only neighbors which are distant in time more than 1000 sampling times (Theiler windows). The straight lines show the behavior D/ϵ^2, where $D=0.007$ is the diffusion constant. The analysis was performed on 10^5 points. Different computations that change τ and m have led to the same result.

$$h(\epsilon)\sim\frac{D}{\epsilon^2}, \tag{31}$$

where D is the diffusion coefficient. Then they *assumed* that the system is deterministic and therefore, because of the inequality $h(\epsilon>0)\leqslant h_{KS}$, they concluded that the system is chaotic. However, from the results presented in the previous sections we can understand that their result does not give direct evidence that the system is deterministic and chaotic. Indeed, power law (31) can be produced in different ways.

(1) A genuine chaotic system with diffusive behavior [like map (20) of Sec. IV A).

(2) A nonchaotic system with some noise, like map (27).

(3) A genuine Brownian system, like Eq. (30).

(4) A deterministic linear nonchaotic system with many degrees of freedom [like Eq. (28)].

(5) A ''complicated'' nonchaotic system like the Ehrenfest wind-tree model, where a particle diffuses in a plane due to collisions with randomly placed, fixed oriented square scatters (as discussed by Dettman *et al.* [32] in their comment on Gaspard *et al.* [31]).

It seems to us that the very weak points of Gaspard *et al.* are (a) the explicit assumption that the system is deterministic, and (b) the neglect of the limited number of data points and, therefore, of both limitations in the resolution and the block length. Point (a) is crucial; without this assumption (even with an enormous data set) it is not possible to distinguish between cases (1) and (2). One has to say that in cases (4) and (5), at least in principle, it is possible to understand that the systems are ''trivial'' (i.e., not chaotic); however, for this one has to use a very huge number of data. For example Dettman *et al.* estimated that in order to distinguish between cases (1) and (5), using realistic parameters of a typical liquid, the number of data points required has to be at least $\sim 10^{34}$. Let us point out that Gaspard *et al.* [31] used $\sim 1.5\times 10^5$ points.

It seems to us that the distinction between chaos and noise makes sense only in very peculiar cases, e.g. very low-dimensional systems. Nevertheless even in such a case an entropic analysis can be unable to recognize the ''true'' character of the system due to the lack of resolution. This is particularly evident in the comparison between the diffusive map [Eq. (20)] and the noisy map [Eq. (27)], if one only observes at scales $\epsilon>\sigma$. According to the pragmatic proposal of classification discussed in Sec. III, one has that for $\sigma\leqslant\epsilon\leqslant 1$ both systems (20) and (27), in spite of their ''true'' character, can be classified as chaotic, while for $\epsilon\geqslant 1$ both can be considered stochastic.

In this respect, the problem of the lack of resolution is even more severe for high entropic systems. One can roughly estimate the critical ϵ (ϵ_u), below which the saturation can be observed, to be $\epsilon_u\propto\exp(-h_{KS})$. Indeed $\exp(h_{KS})$ estimates the number of symbols, i.e., cells of the partition required to reconstruct the dynamics. Therefore, in the Brownian motion studied by Gaspard *et al.* [31], where the KS entropy is expected to be proportional to the number of molecules present in the fluid, the possible small ϵ_u is pushed on scales far from being reachable with the finest experimental resolution available.

VI. DISCUSSIONS

Here we briefly review two examples, studied in detail in Refs. [34,35], showing that high-dimensional systems can display nontrivial behaviors at varying resolution scales. We believe this discussion can be useful in further clarifying the subtle aspects of the distinction between stochastic and deterministic behaviors in dynamical systems.

Olbrich *et al.* [34] analyzed an open flow system described by a unidirectionally coupled map lattice,

$$x_j(t+1)=(1-\sigma)f(x_{j+1}(t))+\sigma x_j(t), \tag{32}$$

where $j=1,\ldots,N$ denotes the site of a lattice of size N, t the discrete time, and σ the coupling strength. A detailed numerical study (also supported by analytical arguments) of the ϵ entropy $h(\epsilon)$ at different values of ϵ, in the limit of small coupling, gives the following scale-dependent scenario: for

$1 \geqslant \epsilon \geqslant \sigma$ there is a plateau $h(\epsilon)=\lambda_s$ where λ_s is the Lyapunov exponent of the single map $x(t+1)=f(x(t))$. For $\sigma \geqslant \epsilon \geqslant \sigma^2$ another plateau appears at $h(\epsilon)\approx 2\lambda_s$, and so, for $\sigma^{n-1} \geqslant \epsilon \geqslant \sigma^n$, one has $h(\epsilon)\approx n\lambda_s$. Similar results hold for the correlation dimension, which increases step by step as the resolution increases, showing that the high dimensionality of the system becomes evident only as $\epsilon \to 0$.

Therefore, one has strong evidence that the dynamics at different scales is basically ruled by a hierarchy of low-dimensional systems whose ''effective'' dimension $n_{eff}(\epsilon)$ increase as ϵ decreases,

$$n_{eff}(\epsilon) \sim \left[\frac{\ln(1/\epsilon)}{\ln(1/\sigma)} \right], \tag{33}$$

where $[\cdots]$ indicates the integer part. In addition, for a resolution larger than ϵ, it is possible to find a suitable low-dimensional noisy system (depending on ϵ) which is able to mimic $x_1(t)$ given by Eq. (32). It is interesting to note that, looking at $h(\epsilon)$ on an extended range of values of ϵ, for $\epsilon \geqslant \sigma^N$ one observes

$$h(\epsilon) \sim \ln \frac{1}{\epsilon}, \tag{34}$$

i.e., the typical behavior of a stochastic process. Of course for $\epsilon \leqslant \sigma^N$ one has to realize that the system is deterministic and $h(\epsilon)\approx N\lambda_s$. Even if this study mainly concerns small unidirectional coupling, the diffusive and strong coupling cases deserve further analysis; they represent a first step toward an understanding of the issue of data analysis of high-dimensional systems.

Let us now briefly discuss the issue of the macroscopic chaos, i.e., the hydrodynamical-like irregular behavior of some global observable, with typical times much longer than the times related to the evolution of the single (microscopic) elements composing a certain system. This interesting kind of dynamical behavior has been studied in some recent works [35,36] for globally coupled maps evolving according to the equation

$$x_i(t+1)=(1-\sigma) f_a(x_i(t)) + \frac{\sigma}{N} \sum_{j=1}^{N} f_a(x_j(t)), \tag{35}$$

where N is the total number of elements, and f_a is a nonlinear function depending on a parameter a. Cencini *et al.* [35] (see also Shibata and Kaneko [36] for a related work) studied the behavior of a global variable (i.e., the center of mass) using the FSLE analysis. Their results can be summarized as follows: (i) at small ϵ ($\ll 1/\sqrt{N}$), one recovers the ''microscopic'' Lyapunov exponent, i.e. $\lambda(\epsilon)\approx\lambda_{micro}$; and (ii) at large ϵ ($\gg 1/\sqrt{N}$) one observe another plateau (corresponding to what we can call the ''macroscopic'' Lyapunov exponent) $\lambda(\epsilon)\approx\lambda_{macro}$, which can be much smaller than the microscopic one. The above results suggest that at a coarse-grained level, i.e., $\epsilon \gg 1/\sqrt{N}$, the system can be described by an ''effective'' hydrodynamical equation (which in some cases can be low-dimensional), while the ''true'' high-dimensional character appears only at a very high resolution, i.e.,

$$\epsilon \leqslant \epsilon_c \approx O\left(\frac{1}{\sqrt{N}} \right).$$

The presence of two plateaus for the FSLE at different length scales is present in generic systems with slow and fast dynamics [23]. The interesting fact is that in systems like Eq. (35) the two temporal scales are generated by the dynamics itself.

Let us stress that the behaviors $h(\epsilon)=\text{const}$ at small ϵ, and $h(\epsilon)$ decreasing for larger ϵ, and not pecularities of the diffusive map [Eq. (20)]. In typical high-dimensional chaotic systems one has $h(\epsilon)=h_{KS}\sim O(N)$ for $\epsilon \leqslant \epsilon_c$ (where N is the number of degrees of freedom, and $\epsilon_c \to 0$ as $N \to \infty$) while for $\epsilon \geqslant \epsilon_c$, $h(\epsilon)$ decreases (often with a power law). From this point of view, the fact that in certain stochastic processes $h(\epsilon)\sim\epsilon^{-\alpha}$ can indeed be extremely useful for modeling such high-dimensional systems. As a relevant example we mention fully developed turbulence, which is a very high-dimensional system whose noninfinitesimal (the so-called inertial range) properties can be successfully mimicked in terms of a suitable stochastic process [37].

VII. CONCLUSIONS

We have shown how an entropic analysis at different resolution scales (in terms of ϵ entropy and the finite size Lyapunov exponent) of a given data record allows us a classification of the stochastic or chaotic character of a signal. In practice, without any reference to a particular model, one can define the notion of deterministic or chaotic behavior of a system on a certain range of scales. In our examples we show that, according to the pragmatic classification proposed in Sec. III, one can consider (on a certain resolution) a system random or deterministic independently of its ''true'' nature. At first glance this can appear disturbing; however, if one adopts a ''nonmetaphysical'' point of view there is an advantage in the freedom of modeling the behavior of the system, at least if one is interested in a certain (noninfinitesimal) coarse-graining property.

ACKNOWLEDGMENTS

We thank R. Hegger and T. Schreiber for useful discussions and suggestions. M.C., M.F. and A.V. were partially supported by INFM (PRA-TURBO) and by the European Network *Intermittency in Turbulent Systems* (Contract No. FMRX-CT98-0175).

[1] C. Nicolis and G. Nicolis, Nature (London) **311**, 529 (1984); P. Grassberger, *ibid.* **323**, 609 (1986); C. Nicolis and G. Nicolis, *ibid.* **326**, 523 (1987); P. Grassberger, *ibid.* **326**, 524 (1987).

[2] A. R. Osborne and A. Provenzale, Physica D **35**, 357 (1989).

[3] G. Sugihara and R. May, Nature (London) **344**, 734 (1990).

[4] M. Casdagli, J. R. Stat. Soc. Ser. B **54**, 303 (1991).

[5] D. T. Kaplan and L. Glass, Phys. Rev. Lett. **68**, 427 (1992).

[6] G. Kubin, in *Workshop on Nonlinear Signal and Image Processing* (IEEE, Halkidiki, Greece, 1995), Vol. 1, pp. 141–145.
[7] E. Aurell, G. Boffetta, A. Crisanti, G. Paladin, and A. Vulpiani, Phys. Rev. Lett. **77**, 1262 (1996); , J. Phys. A **30**, 1 (1997).
[8] A. Kolmogorov, IRE Trans. Inf. Theory **1**, 102 (1956).
[9] C. Shannon and W. Weaver, *The Mathematical Theory of Communication* (University of Illinois Press, 1993).
[10] P. Gaspard and X.-J. Wang, Phys. Rep. **235**, 291 (1993).
[11] R. Benzi, G. Paladin, G. Parisi, and A. Vulpiani, J. Phys. A **18**, 2157 (1985).
[12] P. Grassberger, Phys. Scr. **40**, 346 (1989).
[13] E. Lorenz, Tellus **21**, 3 (1969).
[14] F. Takens, in *Dynamical Systems and Turbulence (Warwick 1980)*, edited by D. A. Rand and L.-S. Young, Lecture Notes in Mathematics Vol. 898 (Springer-Verlag, Berlin, 1980), pp. 366–381.
[15] T. Sauer, J. A. Yorke, and M. Casdagli, J. Stat. Phys. **65**, 579 (1991).
[16] A. M. Fraser, IEEE Trans. Inf. Theory **35**, 245 (1989); D. Prichard and J. Theiler, Physica D **84**, 476 (1995).
[17] A. Cohen and I. Procaccia, Phys. Rev. A **31**, 1872 (1985).
[18] P. Grassberger and I. Procaccia, Phys. Rev. A **28**, 2591 (1983).
[19] F. Takens and E. Verbitski, Nonlinearity **11**, 771 (1998).
[20] H. Kantz and T. Schreiber, *Nonlinear Time Series Analysis* (Cambridge University Press, Cambridge, 1997).
[21] M. Ding, C. Grebogi, E. Ott, T. Sauer, and J. A. Yorke, Physica D **69**, 404 (1992); , Phys. Rev. Lett. **70**, 3872 (1993).
[22] E. Olbrich and H. Kantz, Phys. Lett. A **232**, 63 (1997).
[23] G. Boffetta, A. Crisanti, F. Paparella, A. Provenzale, and A. Vulpiani, Physica D **116**, 301 (1998); G. Boffetta, P. Giuliani, G. Paladin, and A. Vulpiani, J. Atmos. Sci. **55**, 3409 (1998).
[24] P. Grassberger and I. Procaccia, Physica D **9**, 189 (1983).
[25] J. Theiler, Phys. Rev. Lett. **155**, 480 (1991).
[26] D. T. Kaplan and L. Glass, Physica D **64**, 431 (1993).
[27] M. Schell, S. Fraser, and R. Kapral, Phys. Rev. A **26**, 504 (1982).
[28] W. Feller, *An Introduction to Probability Theory and its Applications* (Wiley, New York, 1968).
[29] A. Provenzale, L. Smith, R. Vio, and G. Murante, Physica D **58**, 31 (1992).
[30] P. Mazur and E. Montroll, J. Math. Phys. **1**, 70 (1960); R. I. Cukier and P. Mazur, Physica (Amsterdam) **53**, 157 (1971).
[31] P. Gaspard, M. E. Briggs, M. K. Francis, J. V. Sengers, R. W. Gammon, J. R. Dorfman, and R. V. Calabrese, Nature (London) **394**, 865 (1998).
[32] C. Dettman, E. Cohen, and H. van Beijeren, Nature (London) **401**, 875 (1999).
[33] P. Grassbeger and T. Schreiber, Nature (London) **401**, 875 (1999).
[34] E. Olbrich, R. Hegger, and H. Kantz, Phys. Lett. A **244**, 538 (1998).
[35] M. Cencini, M. Falcioni, D. Vergni, and A. Vulpiani, Physica D **130**, 58 (1999).
[36] T. Shibata and K. Kaneko, Phys. Rev. Lett. **81**, 4116 (1998).
[37] L. Biferale, G. Boffetta, A. Celani, A. Crisanti, and A. Vulpiani, Phys. Rev. E **57**, R6261 (1998).