Calculating topological entropy for transient chaos with an application to communicating with chaos

Joeri Jacobs and Edward Ott*
Institute for Plasma Research and Department of Physics, University of Maryland, College Park, Maryland 20742

Brian R. Hunt†
Institute for Physical Science and Technology and Department of Mathematics, University of Maryland, College Park, Maryland 20742

(Received 26 November 1997)

Recent work on communicating with chaos provides a practical motivation for being able to determine numerically the topological entropy for chaotic invariant sets. In this paper we discuss numerical methods for evaluating topological entropy. To assess the accuracy and convergence of the methods, we test them in situations where the topological entropy is known independently. We also discuss the entropy of invariant chaotic saddles formed by those points in a given attractor that never visit some forbidden ''gap'' region. Such gaps have been proposed as a means of providing noise immunity in schemes for communication with chaos, and we discuss the dependence of the topological entropy on the size of the gap. [S1063-651X(98)00806-X]

PACS number(s): 05.45.+b

I. INTRODUCTION

One way of quantitatively characterizing chaotic motion is via Lyapunov exponents. Lyapunov exponents provide a way of quantifying the exponential divergence of the orbits of two nearby initial conditions. The definition of Lyapunov exponents leads rather straightforwardly to efficient numerical algorithms for their computation [1,2].

A second way of quantifying the complexity of a dynamical system is the entropy, of which two kinds are most commonly discussed. One kind, the metric entropy, weights possible motions according to their probability of being observed, and therefore tells something about the typical behavior of the system. The topological entropy [3], on the other hand, measures complexity by looking at all possible motions without regard to their likelihood. Topological entropy is therefore more difficult to compute, and less attention has been paid to methods for accurately computing it.

In the context of control of chaos for communication [4], which we briefly describe below, any of the possible motions can be selected by manipulating the system with small perturbations. Therefore the relative likelihood of these motions in the uncontrolled system is unimportant and the topological entropy is the measure of complexity that is most relevant in this case.

Topological entropy can be thought of as follows (an exact definition will be given later). One can associate a symbolic dynamics with a chaotic dynamical system by partitioning the state space into r regions W_i, where the index i runs from 1 to r. Given an initial condition $\boldsymbol{x}$ in state space, one constructs a symbolic sequence associated with the orbit of $\boldsymbol{x}$ by recording the consecutive indices i_j of the regions the orbit visits on iterate j for $j=0,1,2,\ldots$. We call such a sequence of M consecutive symbols a word of length M. If we look at all initial conditions of the dynamical system, and record the possible sequences of regions they pass through on their first M iterates, then we know all the words of length M that are allowed by the dynamical system. In general, this number will be smaller than r^M. A sequence of M symbols $i_1 i_2, \ldots, i_M$ that is not realized by any of the points of the state space we call ''grammatically forbidden.'' The number of allowed words of length M typically grows exponentially with M,

$$W(M) \sim \exp(MH\{W_i\}), \tag{1}$$

where we denote the ''entropy'' associated with the partition $\{W_i\}$ by $H\{W_i\}$, and the number of allowed words of length M by $W(M)$. The topological entropy H_{top} is the supremum of $H\{W_i\}$ over all possible partitions,

$$H_{\mathrm{top}} = \sup_{\{W_i\}} H\{W_i\}. \tag{2}$$

Therefore, if we are given the first $M-1$ symbols of a word of length M, the topological entropy tells us on average how many different possibilities there are for completing the word with the last symbol when we use an optimal partition (assuming one exists). Thus we can think of the entropy as the average amount of information carried by each symbol.

By using small perturbations, it has been shown that one can cause the symbolic dynamics of a chaotic system to track a prescribed symbol sequence, thus allowing one to encode any desired message in the signal from a chaotic oscillator [4]. This can be done with arbitrarily small perturbations provided the prescribed symbol sequence is consistent with the symbolic dynamics of the unperturbed dynamical system (i.e., the signal to be encoded by the orbit of the chaotic system does not contain any sequences that are grammatically forbidden by the system). Considering the topological entropy to be the average amount of information that can be stored in one symbol, it is thus a measure of the information transmission rate that can be achieved with the scheme of

*Also at Department of Electrical Engineering and Institute for Systems Research, University of Maryland, College Park, MD 20742.

†FAX: (301) 314-9363. Electronic address: bhunt@ipst.umd.edu

Ref. [4] for ''communicating with chaos.'' Alternatively, given a dynamical system and a rule to construct symbolic dynamics (i.e., a partition used to form ''letters'' as the orbit visits the various partition elements), knowledge of the exact topological entropy would allow us to conclude if the partition is well chosen, in the sense of allowing the optimal transmission rate. That is, we can compare H_{top} to the exponential growth of the number of allowed words with a given word length. If the latter is significantly smaller than H_{top}, then better throughput could be achieved by choosing a different partition.

However, unlike for Lyapunov exponents, the numerical computation of the topological entropy according to its definition is difficult, because it involves computing a supremum over all possible partitions. For certain systems, this supremum is realized by a finite partition called the generating partition. Such partitions have been constructed for several systems [5], but are in general difficult to obtain.

Newhouse [6] has related the topological entropy to the maximal growth rate of any volume element with dimension varying from one to the dimension of the state space. He uses this relation for two-dimensional maps to obtain a bound on the entropy by computing the exponential growth rate of the length of a line segment.

For certain systems (Axiom A systems, see Ref. [7] for a definition), the topological entropy is also given by the exponential growth rate of the number of periodic points as a function of the period.

For chaotic scattering in two-dimensional Hamiltonian flows, Kovács and Tél [8–10] associate a topological entropy with an invariant chaotic saddle. They consider a compact region, the scattering region, that contains the invariant chaotic saddle, and from which almost every initial condition eventually leaves. They then look at a line segment that intersects the stable manifold of the invariant chaotic saddle, and count the number of connected intervals on this line segment that do not leave the scattering region for n iterates. The exponential growth of this number with n is then used to estimate the topological entropy. Chen *et al.* [11] introduce an efficient algorithm that is similar to this one, but uses the inverse of the map rather than the forward mapping. This relies on the fact that the topological entropy for the forward and backward mappings is the same. Furthermore, it has the advantage that it can be used for area contracting mappings such as the dissipative Hénon map as well, since under the inverse mapping almost every initial condition moves away from the attractor.

In this paper we introduce and assess the accuracy and convergence of different algorithms for calculating topological entropy for transient chaotic sets (chaotic saddles) of two-dimensional maps. Transient chaotic sets are the sets of the most interest from the point of view of communicating with chaos, since, even when the orbit is on an attractor, we are generally interested in orbits that always avoid certain regions of the attractor. (As discussed in Sec. IV, this results from considerations having to do with making the signal more immune to noise contamination.) Thus the orbit of the information bearing signal is contained in a smaller chaotic set embedded within the full chaotic attractor, and this smaller set is a transiently chaotic set (small perturbations from it yield orbits that wander over the full chaotic attractor).

In Sec. II we deduce a relation giving the topological entropy in terms of the average stretching factor and the mean decay time. The numerical algorithms are presented in Sec. III. In Sec. IV we make explicit the conditions under which the numerical algorithms yield the correct results. We numerically verify the performance of the different methods on one- and two-dimensional systems in Sec. V. Finally, in Sec. VI, we discuss the dependence of the topological entropy on a ''noise gap'' parameter introduced to model situations of interest for communication with chaos.

II. RELATION WITH STRETCHING AND TRANSIENT TIME

Under certain circumstances it can be shown that the supremum in Eq. (2) is achieved by a partition into a finite number of components $\{W_i\}$, called a generating partition. For such a generating partition, one can obtain a transition matrix $[a_{ij}]$, where $a_{ij}=1$ if points in W_i map into W_j and zero otherwise. The topological entropy then equals the logarithm of the largest eigenvalue of this matrix $[a_{ij}]$.

Let M be a continuous map. It can be shown that the topological entropies of $\boldsymbol{M}$ and $\boldsymbol{M}^{-1}$ are the same. Also, H_{top} is the same for the map $\boldsymbol{M}$ and for any map derived from $\boldsymbol{M}$ by a continuous, invertible change of state space variables (a map topologically conjugate to $\boldsymbol{M}$).

We now obtain a relation between the topological entropy of a chaotic saddle, the stretching properties of the map, and the transient time, i.e., the time it takes a typical point to move away from the saddle. The methods that we introduce will be based on this relation. We consider a nonattracting chaotic ergodic invariant set of a two-dimensional invertible smooth map. We enclose the invariant set in a region (the ''restraining region'') from which almost every initial condition eventually leaves, and make the simplifying assumption that the invariant set is hyperbolic, so that we can conceptualize the region as a rectangle whose edges are parallel to the directions of stretching and compression. By suitable normalization, we can take the rectangle to be a square whose edges are of unit length.

We now identify the initial conditions in the square that stay inside the square for at least n iterates. As in Kovács and Tél [8–10], we note that these initial conditions will be given by $B(n)$ bands that lie along the stable manifold segments, where

$$B(n)\sim\exp(H_{\text{top}}n). \tag{3}$$

There will be a transient time associated with the chaotic invariant set as follows. Suppose we sprinkle a very large number $N(0)$ of initial conditions uniformly in the square. Let $N(n)$ denote the number of those points that have not left the square by iterate n. This number will decay exponentially with time scale τ:

$$N(n)\sim N(0)\exp(-n/\tau). \tag{4}$$

We define the stable manifold natural measure of a set A in the square as follows. Denote by $N_s(A,n)$ the number of

the $N(n)$ initial conditions whose trajectories remain in the square for n iterates that lie in A. Then

$$\mu_s(A)=\lim_{n\to\infty}\lim_{N(0)\to\infty}\frac{N_s(A,n)}{N(n)}. \tag{5}$$

We then label the bands that remain for n iterates by $i=1,\ldots,B(n)$. The natural measure within one of the $B(n)$ bands is proportional to its width (because the initial points were uniformly sprinkled in the square). Points in the same band closely follow each other for n iterates [2]. Thus for $n\gg 1$, the $(1/n)$th power of the magnitude of the largest eigenvalue (the ''stretching factor'') of $\boldsymbol{DM}^n(\boldsymbol{x})$ [where $\boldsymbol{DM}^n(\boldsymbol{x})$ is the tangent map] is approximately the same for all initial points $\boldsymbol{x}$ in a band i. We denote the stretching factor for the ith band of the $B(n)$ bands obtained after n iterates by $\lambda_i^{(n)}$. We also note that the width of this band is roughly $1/\lambda_i^{(n)}$. This is because, as the band is iterated, it is stretched in the direction parallel to its width (by the factor $\lambda_i^{(n)}$), until it just fits (in the stretching direction) in the 1×1 square. Hence the natural measure of a band is inversely proportional to $\lambda_i^{(n)}$.

Therefore computing the natural measure-weighted stretching factor of the points remaining for n iterates yields

$$\langle\lambda\rangle=\lim_{n\to\infty}\lim_{N(0)\to\infty}\left(\left[\sum_{i=1}^{i=B(n)}\frac{1}{\lambda_i^{(n)}}\lambda_i^{(n)}\right]\Big/\left[\sum_{i=1}^{i=B(n)}\frac{1}{\lambda_i^{(n)}}\right]\right)^{1/n}$$
$$=\lim_{n\to\infty}\lim_{N(0)\to\infty}\left(\frac{B(n)}{N(n)/N(0)}\right)^{1/n}. \tag{6}$$

Notice that we first average the n-iterate stretching factors and then take the nth root in order to get a per-iterate quantity. Taking the logarithm on both sides, and using Eqs. (3) and (4), we get

$$H_{\rm top}=\ln(\langle\lambda\rangle)-\frac{1}{\tau}. \tag{7}$$

In a numerical experiment, we can compute this average by sprinkling a very large number $N(0)$ of points in the square, keeping track of the ones $j=1,\ldots,N(n)$ that remain for n iterates ($n\gg 1$) and computing the stretching factor $\lambda_j^{(n)}$ of these points in n iterates. Then,

$$\langle\lambda\rangle\cong\left[\frac{\sum_{j=1}^{N(n)}\lambda_j^{(n)}}{N(n)}\right]^{1/n}=\left[\frac{\sum(n)}{N(n)}\right]^{1/n}, \tag{8}$$

where $\Sigma(n)=\Sigma_{j=1}^{N(n)}\lambda_j^{(n)}$.

To illustrate the application of Eq. (7), consider the one-dimensional model of Fig. 1, where we consider the invariant set of points that never leave the interval $[0,1]$. For this system, one can show that the natural measure of the invariant set in $[0,\alpha]$ is $\alpha/(\alpha+\beta)$, whereas the natural measure in $[\alpha+a,\alpha+a+\beta]$ is $\beta/(\alpha+\beta)$. Therefore

$$\langle\lambda\rangle=\alpha^{-1}\frac{\alpha}{\alpha+\beta}+\beta^{-1}\frac{\beta}{\alpha+\beta}=\frac{2}{\alpha+\beta}. \tag{9}$$

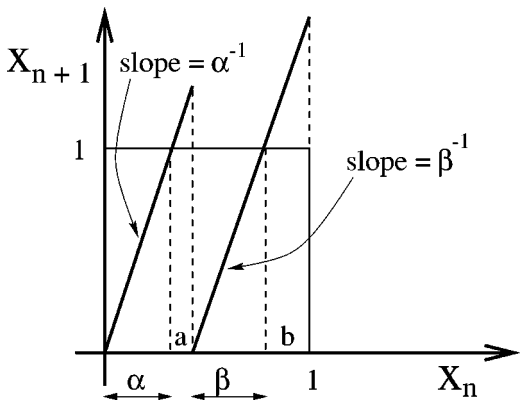

FIG. 1. Piecewise linear one-dimensional model.

Upon each iterate, a fraction $a+b$ falls outside the interval $[0,1]$ so that

$$\exp(-n/\tau)=[1-(a+b)]^n=(\alpha+\beta)^n, \tag{10}$$

or

$$\frac{1}{\tau}=\ln\left(\frac{1}{\alpha+\beta}\right). \tag{11}$$

Therefore Eq. (7) yields

$$H_{\rm top}=\ln\left(\frac{2}{\alpha+\beta}\right)-\ln\left(\frac{1}{\alpha+\beta}\right)=\ln 2. \tag{12}$$

(This is the correct value because the number of intervals remaining in the square after n iterates is 2^n.)

III. NUMERICAL METHODS FOR ESTIMATING $H_{\rm top}$

The methods that we introduce are based on Eqs. (7) and (8). First of all, suppose that we sprinkle a very large number $N(0)$ of points in a region that encloses a nonattracting chaotic invariant set. We keep track of the $N(n)$ points that do not leave that region after n iterates, and compute $\langle\lambda\rangle$ using Eq. (8). Taking the logarithm on both sides of Eq. (8), we get

$$\ln(\langle\lambda\rangle)\cong\frac{1}{n}\ln[\Sigma(n)]+\frac{1}{n}\ln\left(\frac{1}{N(n)}\right). \tag{13}$$

Using Eqs. (4) and (7), we obtain

$$\ln[\Sigma(n)]\sim n\left[\ln(\langle\lambda\rangle)-\frac{1}{\tau}\right]\sim nH_{\rm top}, \tag{14}$$

which says that, if we plot the quantity $\ln[\Sigma(n)]$ versus n, the topological entropy will be given by the asymptotic slope of this graph. We refer to this as method 1. We note that this method, when applied to an attractor (in which case $\tau=\infty$), becomes similar to the method introduced by Newhouse [6]. For the situations we consider in this paper (smooth, two-dimensional maps that do not expand areas), Newhouse proves that the topological entropy is given by the exponential growth rate of the arc length of the forward iteration of a line segment γ. Newhouse uses this rigorous result as a basis for numerical computations of topological entropy. To numerically estimate the growth of the length of γ under suc-

cessive iterations, Newhouse and Pignataro [6] consider N evenly spaced points on γ and approximate the length of the nth iterate of γ by averaging the stretching of a differential line segment at these N points over n iterates. For the case of an attractor, Newhouse's method is therefore slightly different from our method 1 in that he averages stretching rates of a number of initial points on a line segment, whereas we average over points all across the region that contains the invariant set [12]. The similarity of our method 1 applied to attractors ($\tau=\infty$) with Newhouse's method and the fact that his rigorous results apply generally (i.e., do not assume hyperbolicity) lead us to believe that method 1 will work for nonhyperbolic systems, despite the fact that our reasoning in its derivation assumed hyperbolicity.

Another way to exploit Eq. (7) is to consider the distribution of finite time Lyapunov exponents. Given an initial condition $\boldsymbol{x}$ that stays in the restraining region for n iterates, we define the finite time Lyapunov exponent as

$$h(\boldsymbol{x},n)=\frac{1}{n}\ln[\lambda^{(n)}(\boldsymbol{x})], \tag{15}$$

where $\lambda^{(n)}(\boldsymbol{x})$ denotes the stretching factor for the n times iterated orbit starting at $\boldsymbol{x}$. We then compute this quantity for all uniformly sprinkled initial conditions that stay in the restraining region, and obtain the distribution $P(h,n)$ of finite time Lyapunov exponents. For large n, the distribution is asymptotically of the form [13]

$$P(h,n)\sim(2\pi n)^{-1/2}\exp[-nG(h)], \tag{16}$$

where the minimum value of the function G is zero and occurs at $h=\overline{h}$, where $\overline{h}$ is the asymptotic Lyapunov exponent. Note that as n is increased, the distribution $P(h,n)$ becomes more and more peaked about $h=\overline{h}$ and approaches a delta function, $\delta(h-\overline{h})$, as $n\to\infty$.

In terms of this distribution, we can express the average stretching factor as

$$\langle\lambda\rangle=\left[\int P(h,n)\exp(nh)dh\right]^{1/n}$$
$$\sim\left[\int \exp[n(h-G(h)]dh\right]^{1/n}. \tag{17}$$

For large n, the dominant contribution to this integral will come from h values near the point h_1, where

$$\left.\frac{dG(h)}{dh}\right|_{h_1}=1. \tag{18}$$

Therefore

$$\ln(\langle\lambda\rangle)=\frac{1}{n}[nh_1-nG(h_1)]=h_1-G(h_1). \tag{19}$$

Another way to compute the topological entropy is therefore to find the G function and determine the point where its slope is one. Independently, one can determine the decay time τ from a fit of the slope of a plot of the logarithm of the remaining number of points versus number of iterates. With these ingredients, one can estimate the topological entropy using Eqs. (7), (18), and (19). We refer to this as method 2. Its numerical implementation will be discussed further.

Finally, we will compare the results for the topological entropy obtained by these methods with the method of Kovács and Tél [8–10], which we refer to as method 3. This method consists of estimating the exponential growth, as a function of n, of the number of connected intervals on a line across the restraining region such that points in these intervals do not leave the restraining region for n iterates. [Chen *et al.* [11] apply this method to chaotic attractors by noting that the entropy for the forward and inverse mapping is the same. Thus they apply the same method for a chaotic attractor (under the inverse mapping) by enclosing it in a region from which almost every initial condition eventually leaves under the inverse map.]

IV. ESTIMATION OF THE REQUIRED SAMPLING SIZE FOR THE DIFFERENT METHODS

The numerical methods that we described in the preceding section relied on sprinkling a large number of points in the region that contains the chaotic saddle (methods 1 and 2), or a large number of points on a line segment that stretches across this region (method 3). We now wish to obtain an estimate of how large this number of sampling points needs to be, in order for these methods to yield reliable results. Equivalently, given a number of points sprinkled, we want to find the number of iterates n over which the scalings predicted by Eqs. (14) and (3) hold.

A. Methods 1 and 2

Consider again the distribution of finite time Lyapunov exponents, Eq. (16). According to this equation, when sprinkling a large number of points and computing stretching factors for those points that stay in the region for n iterates, the logarithm of the per-iterate stretching will be distributed around $\overline{h}$. Moreover, the distribution will be peaked more and more around this value as n increases. For the computation of $\langle\lambda\rangle$, however, values of h near h_1 (the value of h for which the slope of the G function is one) will form the dominant contribution, as explained below Eq. (17). Therefore we want to be sure that in the numerical method we have a significant number of points that will yield h values near h_1. The number of h values that are close to h_1 will be proportional to $P(h_1,n)$, so if we denote by $N(n)$ the number of the initially sprinkled points that stay in the region for n iterates, we want

$$N(n)\exp[-nG(h_1)]\gg 1, \tag{20}$$

or, using $N(0)\sim N(n)\exp(n/\tau)$ and Eqs. (7) and (19),

$$N(0)\gg\exp[n(h_1-H_{\text{top}})]. \tag{21}$$

B. Method 3

For the method based on Eq. (3) [7], we have to find all the intervals on a line segment that stay for n iterates. When we have $N(0)$ points on such a line segment, we can compare the spacing between the points $[\sim 1/N(0)]$ with the

typical size of one of the $B(n)$ intervals. Consider the median interval size, which is the interval size such that 50% of all intervals are smaller than this value. Then, as we increase the number of iterates, the median interval size eventually becomes smaller than $1/N(0)$. Past this point, less than half of the connected intervals that stay for n iterates will be found, and the found fraction decreases exponentially with further increase. The numerical implementation of method 3 based on Eq. (3) then fails past this point.

To estimate the median interval size, consider the $P(h,n)$. This is the probability density of finding a particular stretching factor $\exp(hn)$. This probability is proportional to the number of intervals with this stretching factor [and therefore of length $\approx\exp(-nh)$] and the length of these intervals. Therefore the fraction of all intervals with length $\exp(-nh)$, which we denote by $\tilde{P}(h,n)$, is proportional to

$$\tilde{P}(h,n)\sim P(h,n)\exp(nh)\sim\exp\{-n[G(h)-h]\}. \quad (22)$$

The maximum of this distribution will occur at $h=h_1$ (where $dG/dh|_{h_1}=1$). For large n, we can expand the exponent in Eq. (22) about $h=h_1$ to obtain a Gaussian $\tilde{P}\sim\exp[-nG''(h_1)(h-h_1)^2]$ for h near h_1. Thus, for n large, h_1 is approximately the median h and therefore the approximate median interval size becomes $\exp(-nh_1)$. The criterion for the required number of points for method 3 then becomes $\exp(-nh_1)\gg 1/N(0)$, or

$$N(0)\gg\exp(nh_1). \quad (23)$$

Comparing Eq. (23) with Eq. (21) shows that the number of points required for method 3 is always larger than for method 1. This is another advantage of method 1.

C. One-dimensional map example

Consider again the one-dimensional map model of Fig. 1. The number of intervals that stay for n iterates is 2^n and there are $(n!)/[m!(n-m)!]$ of length $\alpha^m\beta^{n-m}$. The number $(n!)/[m!(n-m)!]$ is the binomial coefficient. For fixed large n the binomial coefficient is strongly peaked and symmetric about $(m/n)=1/2$, and the width of the peak in m/n decreases to zero with increasing n as $n^{-1/2}$. The median interval size is given by $\alpha^{n/2}\beta^{n/2}=(\alpha\beta)^{n/2}$, and therefore we have that $h_1=\ln(\alpha\beta)^{-1/2}$.

According to Eq. (21), we see that, for method 1, the scaling breaks down for a lower number of iterates if h_1 differs substantially from the topological entropy H_{top}. Therefore to most easily test the scaling for the model system of Fig. 1 we choose α and β so as to make h_1-H_{top} substantial. We apply method 1 for $\alpha=0.1$ and $\beta=0.8$ with the initial number of points sprinkled equal to 6×10^9, 6×10^8, 6×10^7, and 6×10^6. With these values for α and β $h_1=\ln(\alpha\beta)^{-1/2}=1.26$ as compared to $H_{\text{top}}=\ln(2)=0.693$. The resulting scalings are shown in Figs. 2(a)–2(d). The dashed line in each figure represents a line with slope ln2 (the topological entropy). Solving for n in Eq. (21) to obtain an upper bound for the scaling range, we obtain

$$n\ll\frac{N(0)}{h_1-H_{\text{top}}}. \quad (24)$$

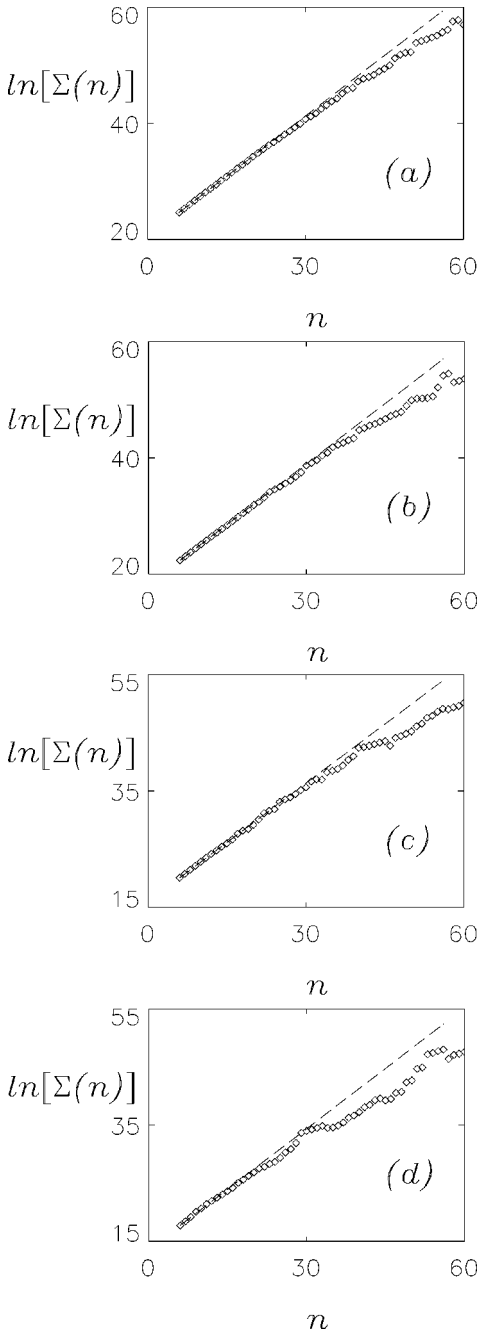

FIG. 2. Limitations to the predicted scaling for method 1. The method is applied with the number of points sprinkled equal to (a) 6×10^9, (b) 6×10^9, (c) 6×10^7, and (d) 6×10^6. The respective upper boundaries for the scaling regime according to Eq. (21) are approximately 40, 35, 31, and 27.

With the numbers used this gives as the upper limit for the scaling ranges of Figs. 2(a)–2(d), respectively, 40, 35, 31, and 27. These numbers can be seen to agree with the range in which the predicted scaling holds.

For method 3, we also check whether breakdown of the scaling occurs where predicted by Eq. (23). With $\alpha=0.4$ and $\beta=0.1$, the criterion is $N(0)>(0.2)^{-n}$. Therefore with $N(0)=65\times10^6$ we expect bad scaling from $n\sim12$ onwards. This number agrees with the onset of curvature in Fig. 3.

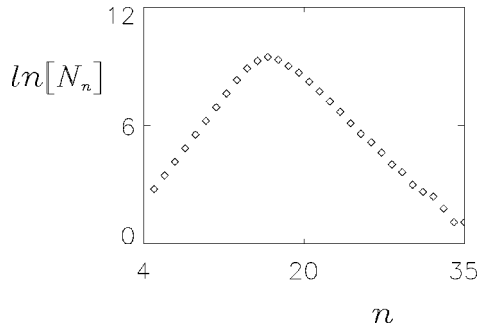

FIG. 3. Scaling of the number of connected intervals N_n that stay for at least n iterates as counted with a linear grid. The slope is fitted from $n=6$ to $n=12$.

D. More easily computable approximate criteria

Since the numerical determination of h_1 in principle involves computation of $G(h)$ which is not straightforward, we wish to express the criteria (21) and (23) in terms of quantities that are easier to compute. Therefore we assume the G function to be approximately quadratic about its minimum value $h=\overline{h}$, and we will need this approximation to hold in $\overline{h}\leqslant h\leqslant h_1$,

$$G(h)\cong\frac{(h-\overline{h})^2}{(2\sigma^2)},$$

$$\sigma=[2G''(\overline{h})]^{-1/2}, \tag{25}$$

so that $P(h,n)$ is Gaussian,

$$P(h,n)\sim\frac{1}{\sqrt{(2\pi)/n}\sigma}\exp\left(-n\frac{(h-\overline{h})^2}{2\sigma^2}\right). \tag{26}$$

With this notation, Eq. (18) yields

$$h_1=\overline{h}+\sigma^2. \tag{27}$$

Consider the schematic illustration of the G function in Fig. 4. The quadratic approximation yields $h_1-G(h_1)=\overline{h}+(h_1-\overline{h})/2=\overline{h}+\sigma^2/2$. Also, from Eqs. (7) and (19) we have that $h_1-G(h_1)=H_{\text{top}}+(1/\tau)$. Therefore the criterion for method 1 becomes

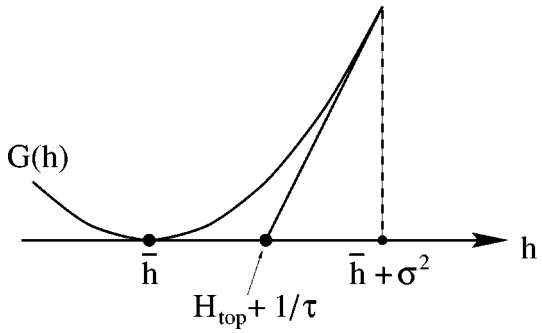

FIG. 4. Relation between G function, $\overline{h}$, h_1, and σ^2 when the G function is given by its lowest order approximation, Eq. (25).

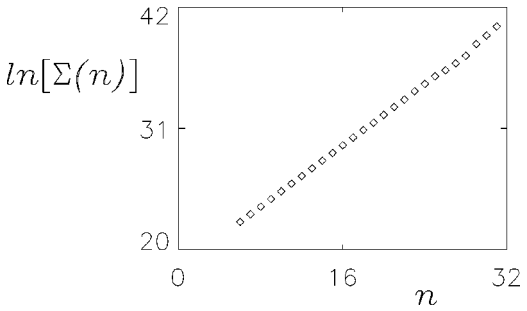

FIG. 5. $\ln[\Sigma(n)]$ versus n for the piecewise linear model of Fig. 1.

$$N(0)\gg\exp(n/\tau)\exp[n(H_{\text{top}}+1/\tau-\overline{h})]$$
$$=\exp[n(H_{\text{top}}-\overline{h}+2/\tau)]. \tag{28}$$

The criterion for method 3 becomes

$$N(0)\gg\exp[n(2H_{\text{top}}-\overline{h}+2/\tau)]. \tag{29}$$

The quantities appearing in Eqs. (28) and (29) are readily available from the computations. The topological entropy is, of course, already computed; $\overline{h}$ is approximated by

$$\langle\ln\lambda\rangle=\frac{1}{n}\frac{\sum_j\ln\lambda_j^{(n)}}{N(n)}, \tag{30}$$

and $1/\tau$ is available from the slope of $\ln N(n)$ versus n. The quadratic approximation does not hold well for the map of Fig. 1 when $|\alpha-\beta|$ is substantial, as in our example in Sec. IV C. We find, however, that it does hold well for our other examples in this paper (the Ikeda and Hénon maps).

In the rest of this paper, we have checked that all of our estimations of the entropy are consistent with the limits on their scaling range.

V. NUMERICAL EXPERIMENTS

A. One-dimensional model system

To experimentally test the validity of these methods, we first wish to apply them to a system where we know the topological entropy exactly, namely, the model system of Fig. 1. The topological entropy for this system is ln2.

Since we wish to compare the correctness and efficiency of the different methods, we do the following. For method 1, we randomly sprinkle 65×10^6 points in the interval $[0,1]$ and compute the stretching factor of each of these points until they leave. In this way, we obtain $\Sigma(n)$ for n up to approximately 30. One can see in Fig. 5 that the predicted exponential scaling, Eq. (14), holds very well over this number of iterates. The slope is estimated with a least squares fit. We then repeat this procedure 50 times (the randomly sprinkled points are different each time) and find the average and root mean square deviation. The results are summarized in Table I.

For method 2, we again use about 65×10^6 points and repeat the procedure 50 times. Results for the same quantities as for method 1 are listed in Table I. As can be seen in Fig. 6, the distribution of Lyapunov exponents $P(h,n)$ does

TABLE I. Comparison of the estimated topological entropies for the one-dimensional model system of Fig. 1 obtained using the three different methods. For each case, the average and root mean square deviation of 50 independent attempts is listed. Note that the correct value of the topological entropy is ln(2)=0.693 147.

	Average	rms
Method 1	0.69315	0.00015
Method 2	0.68463	0.01500
Method 3	0.69313	0.00003

indeed become narrower as n is increased, whereas the G function remains approximately unaltered, as predicted. In order to find the point where the slope of the G function is equal to one, we fit a cubic polynomial through the histogram data with a least squares fitting method, in a range where the estimates of the G function for different n agree well with one another.

Finally, for method 3, we wish to obtain an estimation of the entropy for a similar computational effort as for the other methods. Therefore we lay down 65×10^6 evenly spaced points along the interval [0,1], and obtain the lifetime of each point (i.e., the time it takes to leave the interval [0,1]). For each n, we then go through this list and count the number of intervals of consecutive points all having lifetimes larger than n. In this way, our estimation for the number of connected intervals with lifetime at least n is a lower bound, and we expect to miss a lot of intervals with lifetime n when a typical such interval is of the order of magnitude of the grid size or less. This can be seen in Fig. 3, where we see that the scaling starts to be very bad from around $n=14$ onwards. Therefore for this method we do a fitting of the slopes for $n\leqslant12$, and for consistency choose the same scaling range for method 1. Here also, the procedure is repeated 50 times with the grid being shifted between different realizations by a random fraction of the grid size. The average and rms deviation for method 3 are also shown in Table I.

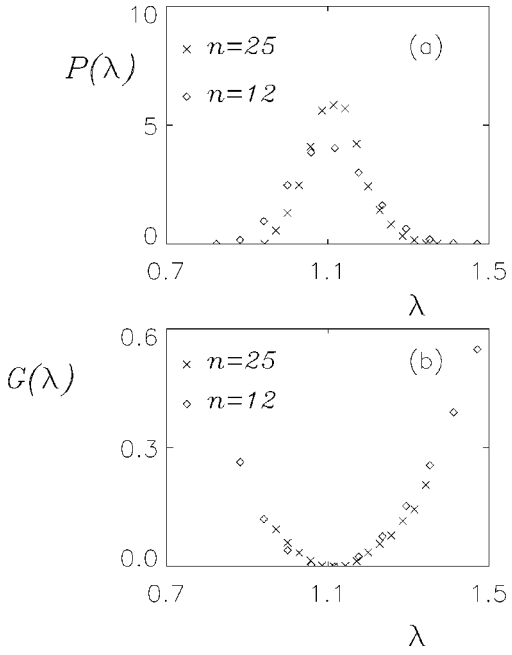

FIG. 6. (a) Distribution of the finite time Lyapunov exponents for the one-dimensional model system for $n=12$ and $n=25$ and (b) G functions obtained from these distributions.

From this table, we can conclude that both method 1 and method 3 give very accurate results for the estimation of the topological entropies with small deviations from the exact value for each estimation. Method 2 is in this respect significantly less accurate, and the fluctuations between different estimations are substantial. We remark that the root mean square deviation for method 3 is artificially small for this model (compared to the two-dimensional models that we will discuss further). As a matter of fact, the 50 different estimations only gave rise to two possible outcomes, one of which is correct (up to six digits), with the other one being slightly below the correct value. In fact, method 3 necessarily gives a lower bound for the topological entropy. Estimation with method 3 is, however, very dependent on the scaling range (fitting the slope with $n\leqslant15$ would change the results greatly), and this suggests that method 1 may be more reliable in general.

B. Ikeda map

To perform a numerical test on a two-dimensional system, we consider the Ikeda map [14] for the complex number $z_n=x_n+iy_n$ where x_n and y_n are real,

$$z_{n+1}=a+bz_n\exp\left(i\kappa-\frac{i\eta}{1+|z_n|^2}\right), \tag{31}$$

with parameter values $a=1.0027$, $b=0.9$, $\kappa=0.4$, and $\eta=6.0$ [15]. These parameter values give rise to a chaotic attractor. However, we can find an invariant subset of this attractor that never leaves the rectangle [16]

$$0.1<x<1.1, \quad -1.0<y<0.8.$$

This invariant set is shown in Fig. 7. We can again associate a lifetime with each point in this rectangle by counting how many iterates it takes to leave the rectangle. In this way we can associate a transient time τ with this set. We then perform the same procedure as for the one-dimensional system: estimating the topological entropy 50 times independently and computing the average and rms deviation. The results are summarized in Table II. For each estimation 180×10^6 points were used. The 50 line segments for method 3 were obtained by connecting two randomly chosen points on the right and left boundary of the rectangle. From Table II, we see that the results of methods 1 and 3 are consistent with each other, with method 2 slightly deviating. The fact that method 1 has the smallest fluctuations, and that its predicted scaling holds over a very wide range leads us to believe that method 1 is the most reliable. Note also that the numerical value that it yields is consistently slightly above the value of method 3, which gives a lower bound. Further evidence for this conclusion comes from a separate estimation with the transition matrix method, described at the beginning of Sec. II. Labeling the five regions as shown in Fig. 7 by the numbers 1 to 5, and examining orbits that stay in these regions we obtain the following 5×5 transition matrix:

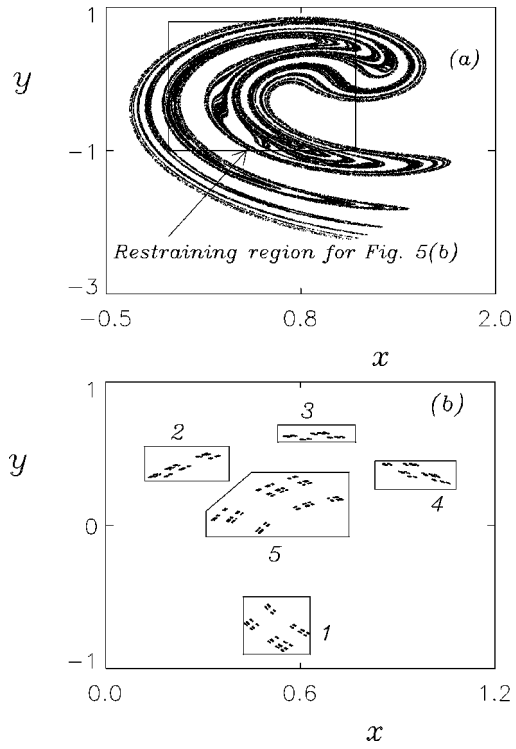

FIG. 7. (a) Ikeda attractor and (b) invariant set that is formed by points that never get mapped outside of the rectangle under forward or inverse mapping.

$$\boldsymbol{A}=\begin{bmatrix} 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix}. \tag{32}$$

The logarithm of the largest magnitude eigenvalue of $\boldsymbol{A}$ is 0.413 807 6, which is consistent with the results of method 1 and method 3. (Note, however, that there is no guarantee that the partition shown in Fig. 7 is a generating partition. Thus the logarithm of this eigenvalue is not necessarily H_{top} though it should, at least, be a good estimate for it.)

C. Hénon map

Finally, we illustrate how the three methods work for attractors. We consider the attractor of the Hénon map,

$$(x,y)\rightarrow(a-x^2+by,x), \tag{33}$$

TABLE II. Comparison of the estimated topological entropies for the invariant set contained in the Ikeda attractor (Fig. 7). Again, the listed values are the average and root mean square deviation for 50 independent estimations. The eigenvalue of the transition matrix for the partition in Fig. 7 is 0.413 807 6.

	Average	rms
Method 1	0.41405	0.00030
Method 2	0.40443	0.00200
Method 3	0.41381	0.00150

TABLE III. Comparison of the estimated topological entropies for the attractor of the Hénon map. Again, the listed values are the average and root mean square deviation for 50 independent estimations. The previously estimated [5] value for the topological entropy is 0.4650 ± 0.0002.

	Average	rms
Method 1	0.46493	0.00003
Method 2	0.46621	0.00300
Method 3	0.44924	0.03000

with the ''classical'' values $a=1.4, b=0.3$. We assume that Eq. (7) still holds with $\tau=\infty$, and that therefore the methods we have discussed still yield a reliable estimation of the topological entropy. For method 3 we must use the inverse mapping, since we are working with an attractor. For methods 1 and 2 we use the forward mapping. We compare our results with the estimation obtained from the very intensive computation of Grassberger, Kantz, and Moenig [5], who constructed a partition meant to give the topological entropy. They conclude the topological entropy to be 0.4650 ± 0.0002. We use 9×10^8 points for each estimation in each of the three methods. The random lines for method 3 are selected by connecting random points on the lines $x=\pm3$ between $y=-1.5$ and $y=1.5$. The results are summarized in Table III. Again, we see that method 1 has very small fluctuations and that the scaling holds over a wide range. Moreover, the quantitative result is consistent with the most precise previous value of Grassberger, Kantz, and Moenig. Method 2 yields values that are consistent with this estimation as well, but the fluctuations are significantly larger. The uncertainty with method 3 is particularly large for this test case. We attribute this to the typical lifetime under the inverse mapping being very short. Thus we conclude that for this attractor case method 1 again seems to be most reliable.

VI. VARIATION OF THE TOPOLOGICAL ENTROPY AS A FUNCTION OF ''NOISE GAP'' WIDTH

In this section we discuss examples indicating the utility of topological entropy computations in a practical application. In particular, when trying to encode a message in the signal from a chaotic oscillator for the purpose of forming a communication signal [4], one must avoid confusion about which partition element the orbit is in. For illustration, consider the chaotic attractor for the standard Hénon map as an encoding system. One can make a partition by dividing the attractor into two parts using $y=0$ as the dividing line. However, noise added to the signal upon transmission may make it impossible to attach the correct symbol to points in the boundary region $y\simeq0$. One way to counter this is to use the small control (already employed to control the symbol sequence that forms the signal) so that orbits never fall in the ''noise gap'' region $-\epsilon<y<\epsilon$ (where ϵ is chosen larger than the noise level). This technique was used in the proof-of-principle laboratory demonstration of communication with chaos in Ref. [17]. Because the control is small, we cannot expect it to create new orbits, topologically distinct from those already existing on the attractor. Therefore, with imposition of our noise gap, we are concerned with points on

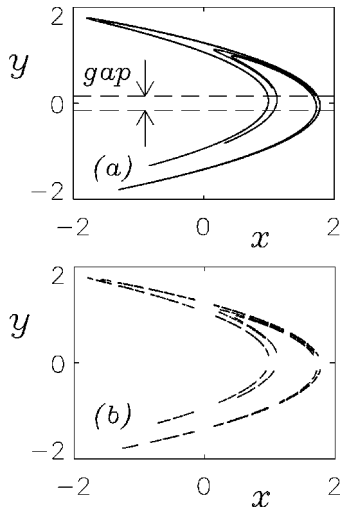

FIG. 8. (a) Hénon attractor for $a=1.4$ and $b=0.3$ with noise gap and (b) resulting chaotic saddle made up of points on the attractor that never map into the noise gap.

the original uncontrolled attractor whose orbits never visit the noise gap. This scenario is depicted in Fig. 8. In Fig. 8(a) we show the attractor for the Hénon map, Eq. (33), with parameters $a=1.4$ and $b=0.3$, and the noise gap, i.e., the region $-\epsilon<y<\epsilon$. The resulting invariant chaotic saddle of points that never maps into this noise gap is shown in Fig. 8(b).

Since points in the invariant set may be eliminated but not added with increasing ϵ, the topological entropy is nonincreasing with ϵ. Thus there is an inherent trade-off between information capacity (H_{top}) and noise immunity. To assess this tradeoff, a first step (which we consider here) is to evaluate H_{top} as a function of the noise gap size [18].

We first investigate the dependence of the topological entropy on this noise gap for a simple one-dimensional system, $x\rightarrow 2x \mod 1$. For symbolic encoding associated with a partition of $[0,1]$ we write down a zero if an orbit lands to the left of 0.5 and a one otherwise. The noise gap is set from 0.5 to $0.5+\epsilon$. We can then compute the topological entropy of the Cantor set of points that never fall in $[0.5,0.5+\epsilon]$ as follows. Since the stretching factor is 2 for every orbit, the topological entropy $[H_{\text{top}}=\ln(2)-1/\tau]$ is determined if we can compute the decay time τ. (Our method 1 reduces to this computation when the stretching is constant on each iterate.) We therefore sprinkle a very large number of points in the interval $[0,1]$ and compute how long it takes for each of these points to fall in the gap $[0.5,0.5+\epsilon]$. From a fit of the slope of the logarithm of the remaining points, $\ln[N(t)]$, versus t, we can obtain $1/\tau$. As can be seen from Fig. 9, the resulting graph of the topological entropy versus noise gap exhibits many intervals of constant entropy. Such intervals of constant entropy for saddle sets were also found in Ref. [11] for the Hénon map as a function of the parameter a for $a>1.6$ and in [18] for noise gaps.

This phenomenon can be understood as follows. Consider the binary expansion of points in the interval $[0,1]$. If $x=\Sigma_i \sigma_i (1/2)^i$, where $\sigma_i=0$ or 1, then we associate with x the (infinite) symbol sequence,

$$\sigma(x)=\sigma_1\sigma_2\sigma_3\ldots\ . \tag{34}$$

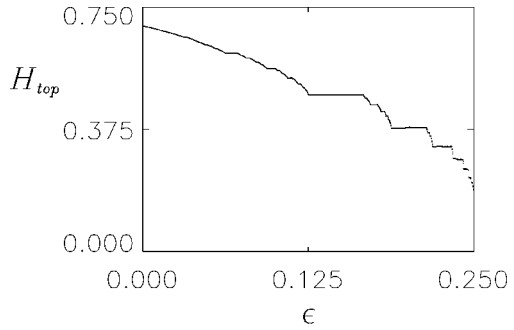

FIG. 9. Topological entropy versus noise gap for $x\rightarrow 2x \mod 1$.

The map $M(x)=2x \mod 1$ then has the property that

$$\sigma[M(x)]=\sigma_2\sigma_3\sigma_4\ldots, \tag{35}$$

i.e., applying the map to the binary representation of any point is equivalent to removing the first digit and shifting the remaining digits one place to the left. Now consider the set that we obtain by setting the gap from 1/2 to 1/2+1/8. Note that $\sigma(1/2)=1$ and $\sigma(1/2+1/8)=101$ (by convention we omit the infinite string of zeros that follow these numbers). Removing the interval $[1/2,1/2+1/8[$ means that all points whose binary representation starts with 100 are not a part of the invariant set. Removing all the points that eventually map into $[1/2,1/2+1/8[$ means that all the points that have $\sigma_i\sigma_{i+1}\sigma_{i+2}=100$ for some i are not part of the invariant set. In particular, we remove all the points whose binary representation starts with 10100; that is, the interval $[1/2+1/8,1/2+1/8+1/32[$. Therefore enlarging the noise gap from $[1/2,1/2+1/8]$ to $[1/2,1/2+1/8+1/32[$ will not alter the invariant set since no new points are excluded through this enlargement of the gap: points that map into $[1/2+1/8,1/2+1/8+1/32[$ will be mapped into $[1/2,1/2+1/8]$ two iterates later, and therefore are excluded anyway. This explains the plateau in Fig. 9 starting at 0.125. One can see that for similar reasons, moving the noise gap from 1/2+1/8 to the point with binary representation 1010101 . . . will result in the same set and will not affect the entropy. Therefore we expect the length of the largest constant entropy interval to be

$$1/32+1/128+1/512+\cdots=1/2^5\sum_i 1/2^{2i}=1/24, \tag{36}$$

which is indeed the value that can be read from the graph. The other constant entropy intervals can be explained in a similar way. As a matter of fact, the graph represents a complete devil's staircase, i.e., it is a continuous, monotonically decreasing function that is constant almost everywhere. In particular, the set of points that are not included in an interval of constant entropy form a Cantor set of zero Lebesgue measure.

We now investigate how we can use the concept of fractal dimension to characterize the Cantor set of ϵ values that are not included in a constant entropy interval. Call this set A. We now consider the dimension of the intersection of A with an interval $[0.5+\epsilon,0.5+\epsilon+\delta]$. For fixed ϵ consider the set

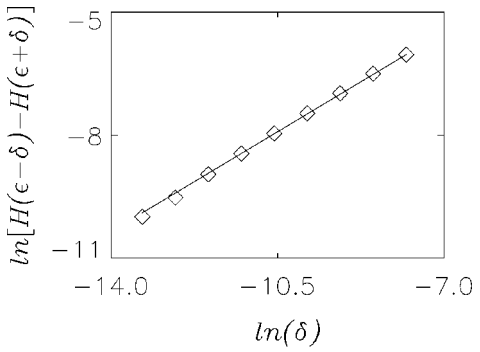

FIG. 10. Measurement of the Hölder exponent of the function $H_{\text{top}}(\epsilon)$ at $\epsilon=0.166\,67$. The solid line represents a line of slope $H_{\text{top}}(0.166\,67)/\ln(2)=0.6936$ (the prediction for the Hölder exponent).

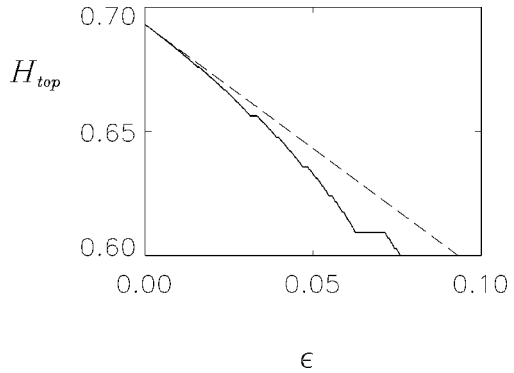

FIG. 11. The topological entropy for $x\to 2x$ mod 1 for small ϵ. The dashed line represents $\ln(2)-\epsilon$, which approximates $H_{\text{top}}(\epsilon)$ for small ϵ.

of points whose orbits under the $2x$ mod 1 map never land in $[0.5,0.5+\epsilon[$. Call this Cantor set B_ϵ. One can show [2] that the dimension of this set is

$$\overline{d}(B_\epsilon)=\frac{\ln(2)-1/\tau}{\ln(2)}=\frac{H_{\text{top}}}{\ln(2)}, \tag{37}$$

where $H_{\text{top}}=\ln(2)-1/\tau$. [Information dimension of the natural measure and capacity dimension are equal for this set since the stretching rate, ln(2), is constant over the set. Similarly, the metric and topological entropies are equal.] Now, consider the noise gap $[0.5,0.5+\epsilon+\delta[$ where δ increases continuously from 0. The entropy can decrease only for values of δ such that $0.5+\epsilon+\delta$ is one of the points in the Cantor set B_ϵ. Therefore we conjecture that the local dimension of the Cantor set A near a value of the noise gap ϵ [with entropy $H_{\text{top}}(\epsilon)$] is the dimension of the Cantor set B_ϵ, namely, $H_{\text{top}}/\ln(2)$. In particular, if we let $d(H_{\text{top}})$ denote the local dimension of A at the value of ϵ with $H_{\text{top}}(\epsilon)=H_{\text{top}}$, then $d(H_{\text{top}})=\overline{d}(B_\epsilon)$, or

$$d(H_{\text{top}})=\frac{H_{\text{top}}}{\overline{h}}, \tag{38}$$

where $\overline{h}=\ln(2)$ is the Lyapunov exponent of the Cantor set B_ϵ.

To test the relation, Eq. (38), numerically, we do the following. Take a noise gap width ϵ such that ϵ is on the Cantor set A and compute the Hölder exponent of the entropy function (H_{top} versus ϵ) at that value. This involves computing the entropy for $\epsilon\pm\delta$ for small and varying δ. The Hölder exponent is then given by the slope of the graph of $\ln[H_{\text{top}}(\epsilon-\delta)-H_{\text{top}}(\epsilon+\delta)]$ versus $\ln(\delta)$. The resulting graph for $\epsilon=0.125$ is shown in Fig. 10 and an exponent of 0.7 is measured in this way. This is in good agreement with a value of $H_{\text{top}}/\ln(2)=0.4814/0.6931=0.6946$. The association of the Hölder exponent with the local dimension of A is made as follows. Consider the interval $[\epsilon-a,\epsilon+a]$ for some small fixed a. Now divide this region into N intervals of equal δ. Let x_i $(i=0,\ldots,N)$ be the points on the interval boundaries; then

$$H_{\text{top}}(\epsilon-a)-H_{\text{top}}(\epsilon+a)=\sum_{i=0}^{N-1}[H_{\text{top}}(x_i)-H_{\text{top}}(x_{i+1})]. \tag{39}$$

The terms on the right-hand side will be zero for all the constant entropy intervals, and if we assume the Hölder exponent at points in A to be nearly constant across $[\epsilon-a,\epsilon+a]$, then the intervals for which the entropy is not constant contribute approximately δ^α (α being the Hölder exponent). If we denote by $N(\delta)$ the number of nonconstant entropy δ width intervals, we then have that

$$N(\delta)\,\delta^\alpha\sim H_{\text{top}}(\epsilon+a)-H_{\text{top}}(\epsilon-a). \tag{40}$$

Since the right-hand side is independent of δ, we have $N(\delta)\sim\delta^{-\alpha}$ implying that the local dimension of A is the Hölder exponent α.

As can be seen from Fig. 9, the intervals become smaller as ϵ approaches zero, and the graph approaches a linear function. This is because the local dimension of the Cantor set of points that are not contained in a constant entropy interval goes to 1 as ϵ goes to zero. The scaling of the entropy versus ϵ for small ϵ goes as [19]

$$H_{\text{top}}(\epsilon)\cong\ln(2)-\epsilon. \tag{41}$$

This follows from Eq. (7) and the fact that the natural measure of the $\epsilon=0$ attractor that is contained in $[1/2,1/2+\epsilon[$ is ϵ. Assuming that the attractor measure is only slightly altered by a small ϵ gap, we estimate that $\exp(1/\tau)\cong 1-\epsilon$, and thus

$$1/\tau\cong\epsilon. \tag{42}$$

One can see from Fig. 11 that the tangent to the entropy versus ϵ graph at $\epsilon=0$ (the dashed line) has a slope of -1 as expected.

Similar structure of the graph of the topological entropy versus the noise gap is observed for other one-dimensional maps. As further examples, Fig. 12 shows such graphs for the tent map,

$$x\to\begin{cases}2x & \text{for } x \text{ in } [0,0.5]\\ 2-2x & \text{for } x \text{ in } [0.5,1],\end{cases} \tag{43}$$

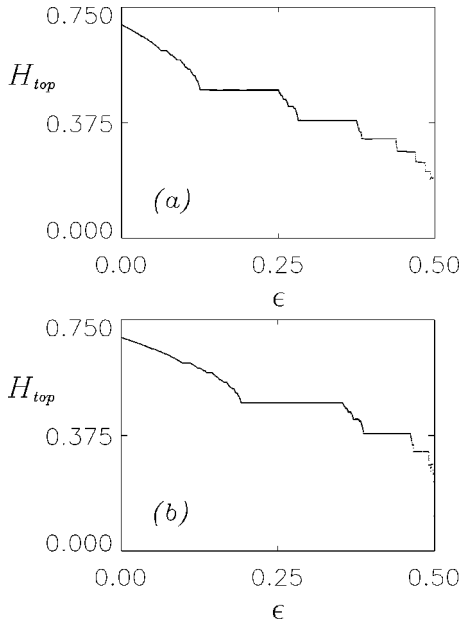

FIG. 12. (a) Topological entropy versus noise gap for the tent map, and (b) topological entropy versus noise gap for the logistic map.

and for the logistic map $x \to 4x(1-x)$. For small ϵ, the entropy scales as $\ln(2)-\epsilon$ for the tent map, and $\ln(2)-(2/\pi)\epsilon$ for the logistic map, for the same reason as explained above. Numerical estimations of the slope of the entropy versus noise gap graph near $\epsilon=0$ confirm these predictions.

In two dimensions, we use method 1 to obtain $H_{\rm top}$ as a function of ϵ for the Hénon map (with $a=1.4$ and $b=0.3$) as illustrated in Fig. 8. The resulting graph is shown in Fig. 13(a). Although the function looks rather smooth for ϵ up to 0.25, we recognize the same structure as for the one-dimensional case. From the enlargement in Fig. 13(b), it can

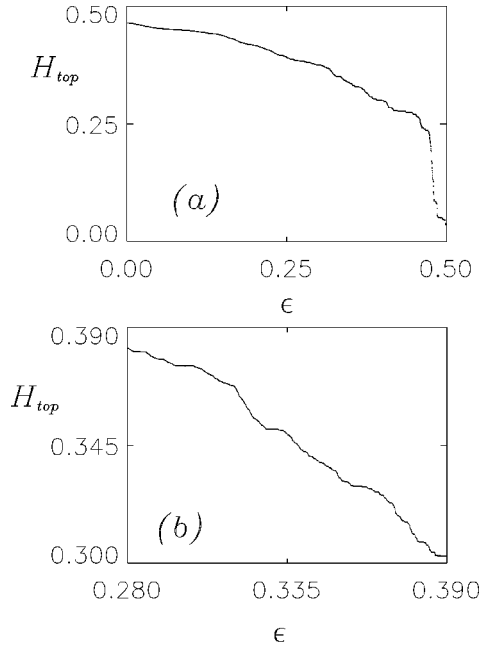

FIG. 13. (a) Topological entropy versus noise gap for the Hénon map ($a=1.4$, $b=0.3$). (b) An enlargement of the same function shows the constant entropy plateaus more clearly.

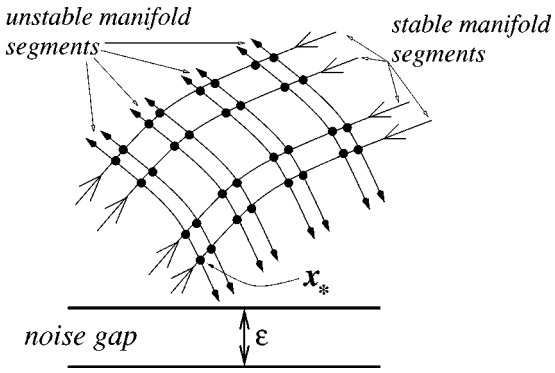

FIG. 14. Schematic illustration of the chaotic saddle that consists of all points on the chaotic attractor that, under the forward or backward mapping, will never map into the noise gap. The chaotic set shown as black dots is the intersection of its stable and unstable manifold.

be seen that the constant entropy plateaus are present, and we believe this graph to be a complete devil's staircase as well. The numerics are not accurate enough to make a numerical estimation of the local dimension of the set of points that are not included in a constant entropy interval. We can, however, give a simple theoretical treatment of this question. To do this, consider the schematic illustration at noise gap ϵ in Fig. 14. The black dots at the intersections of the stable and unstable manifolds schematically represent the chaotic saddle. Again, call this saddle B_ϵ and the Cantor set of ϵ values not included in a constant entropy interval A. This time, when we increase ϵ, the chaotic saddle will get smaller each time ϵ is such that the gap widens to include one of the chaotic saddle points. Referring to Fig. 14, we see that this will occur when the corner point on the chaotic saddle $\boldsymbol{x}_*$ is included in the gap. In addition, as the gap increases through $\boldsymbol{x}_*$, all the points of the saddle that lie on the stable manifold segment through $\boldsymbol{x}_*$ will also be removed from the chaotic saddle. (For a related consideration see [19].) Therefore we can effectively consider the process as successive removal of stable manifold line segments as ϵ increases. Consequently, if we let $d(H_{\rm top})$ denote the local dimension of A at $H_{\rm top}(\epsilon)$, then $d(H_{\rm top})=\overline{d}_s(B_\epsilon)-1$, where $\overline{d}_s(B_\epsilon)$ is the dimension of the stable manifold of the chaotic saddle B_ϵ. The information dimension of the stable manifold is given by [2] $2-1/(\tau\overline{h})$, where $\overline{h}$ is the positive Lyapunov exponent of the chaotic saddle. For the Hénon attractor, numerical computations of the capacity dimension and the information dimension give values that are very nearly the same (the capacity dimension is slightly higher), and the situation is similar with respect to the metric entropy (given by $\overline{h}-1/\tau$) and the topological entropy (with the topological entropy slightly higher). In such situations, a rough estimate is provided by approximately equating the capacity and information dimensions, and the topological and the metric entropies. This gives $d(H_{\rm top})\cong 1-(\overline{h}\tau)^{-1}$, and thus

$$d(H_{\rm top})\cong\frac{H_{\rm top}}{\overline{h}}. \tag{44}$$

This coincides with our previous result [Eq. (38)] for the tent map and the $2x$ mod 1 map (with the important difference

that for the tent and the $2x \bmod 1$ maps we expect exact, rather than approximate, equality due to the uniform stretching by the constant factor 2 in those cases).

VII. CONCLUSION

We have discussed methods to numerically compute the topological entropy for chaotic saddles of two-dimensional maps. The methods indirectly rely on computing a transient time and stretching properties of the map. These are quantities which are much easier to compute than the quantities that occur explicitly in the definition of the topological entropy. Numerical experiments for cases where the entropy is known from independent sources indicate that the methods provide the correct answer. One of the methods (method 1) seems to be particularly efficient. For method 1, in the cases we tested, the fluctuations between different estimations are smaller than for both of the other two methods. In addition, we find that method 1 obeys its predicted scaling over a wider range, and is not dependent on the lifetime τ being long or short.

We also used method 1 to study the dependence of the topological entropy on the noise gap, i.e., the region that is forbidden for the dynamics to enhance noise immunity in communicating with chaos [4,17]. The structure of the entropy versus noise gap function is established to be a complete devil's staircase. Numerical tests provide evidence for our prediction of the dimension of the Cantor set of points that are not included in a constant entropy interval.

ACKNOWLEDGMENTS

This work was supported by the Office of Naval Research (Physics) and by the U.S. Department of Energy. The numerical computations reported in this paper were supported in part by a grant from the W.M. Keck Foundation.

[1] J.-P. Eckmann and D. Ruelle, Rev. Mod. Phys. **57**, 617 (1985).
[2] E. Ott, *Chaos in Dynamical Systems* (Cambridge University Press, Cambridge, England, 1993), and appropriate references therein.
[3] R. C. Adler, A. C. Konheim, and M. H. McAndrew, Trans. Am. Math. Soc. **114**, 309 (1965).
[4] S. Hayes, C. Grebogi, and E. Ott, Phys. Rev. Lett. **70**, 3031 (1993).
[5] P. Grassberger, H. Kantz, and U. Moenig, J. Phys. A **22**, 5217 (1989).
[6] S. Newhouse, in *The Physics of Phase Space*, Lecture Notes in Physics Vol. 278 (Springer, Berlin, 1986), pp. 2–8; Ergod. Theor. Dynam. Syst. **8**, 283 (1988); S. Newhouse and T. Pignataro, J. Stat. Phys. **72**, 1331 (1993).
[7] R. Bowen, Trans. Am. Math. Soc. **154**, 377 (1971); A. B. Katok, Publ. Math. IHES **51**, 137 (1980).
[8] Z. Kovács and T. Tél, Phys. Rev. Lett. **64**, 1617 (1990).
[9] T. Tél, Phys. Rev. A **44**, 1034 (1991).
[10] T. Tél, J. Phys. A **22**, L691 (1989).
[11] Q. Chen, E. Ott, and L. P. Hurd, Phys. Lett. A **156**, 48 (1991).
[12] A method similar to that of Newhouse and Pignataro but applied to a chaotic saddle in a physical situation (motion of tracer particles in a fluid flow) can be found in E. M. Ziemniak, C. Jung, and T. Tél, Physica D **76**, 123 (1994).
[13] R. S. Ellis, *Entropy, Large Deviations and Statistical Mechanics* (Springer-Verlag, New York, 1985).
[14] S. Hammel, C. K. R. T. Jones, and J. Maloney, J. Opt. Soc. Am. B **2**, 552 (1985).
[15] C. Grebogi, E. Ott, and J. Yorke, Phys. Rev. Lett. **57**, 1284 (1986).
[16] J. Jacobs, E. Ott, and C. Grebogi, Physica D **108**, 1 (1997).
[17] S. Hayes, C. Grebogi, E. Ott, and A. Mark, Phys. Rev. Lett. **73**, 1781 (1994).
[18] See also E. Bollt, Y.-C. Lai, and C. Grebogi, Phys. Rev. Lett. **79**, 3787 (1997).
[19] This is consistent with J. Vollmer and W. Breymann, Europhys. Lett. **27**, 23 (1994). This paper considers the parameter dependence of the entropy of a chaotic attractor. As the parameter increases, more orbits are forbidden, pruned. At the onset of pruning, i.e., the critical parameter beyond which orbits in the chaotic attractor are forbidden, it is shown that the decrease of the topological entropy satisfies a power law behavior. The scaling exponent is shown to be the pointwise dimension of the first pruned orbit. For the one-dimensional map we consider here, pruning starts as soon as the size of the gap is strictly positive. Since the pointwise dimension of any point on the original attractor is one, the scaling exponent should be one as well. This is clearly the case in Eq. (41).