Estimation of Lyapunov exponents of dynamical systems using a spatial average

Gary Froyland, Kevin Judd, and Alistair I. Mees
Department of Mathematics, The University of Western Australia, Nedlands WA 6009, Australia
(Received 5 May 1994)

We present an unconventional method of estimating all of the Lyapunov exponents of a dynamical system from either a known map or a set of experimental data. Rather than averaging exponents along a single trajectory, we instead represent each exponent as an integral over all of phase space. The contribution to each exponent, calculated at each point in phase space, is averaged spatially by weighting areas of high density more heavily than areas of low density, according to the invariant measure of the system. Explicit formulas for approximating both the contributions to the exponents and the invariant measure are given, and convergence results stated. The techniques are illustrated in detail for the Hénon system.

PACS number(s): 02.50.Ey, 02.30.−Cj, 02.70.−c, 02.50.Ga

I. LYAPUNOV EXPONENTS OF DETERMINING DYNAMICAL SYSTEMS

Let $f:M\to M$ be a C^2 diffeomorphism of a smooth Riemannian manifold. Lyapunov exponents arise as possible values of the limit

$$\lim_{n\to\infty}\frac{1}{n}\ln\|df^n(x)v\| \tag{1}$$

for various initial points x on some manifold M and vectors v. The above limit, if it exists, represents the long-term local stretching of an orbit beginning at x in the direction v. If x is fixed, and v varied over all of T_xM, only a finite number of values $\lambda^{(1)}(x)\le\lambda^{(2)}(x)\le\cdots\le\lambda^{(d)}(x)$ are possible for a d-dimensional system.

When experimentalists talk about "*the* Lyapunov exponents" of a given dynamical system, they assume that Lebesgue-almost-all initial points in a neighborhood of the "attractor" produce orbits that exhibit the same nontrivial ergodic invariant measure μ; nontrivial in the sense that the support of the measure is more complicated than a union of periodic orbits. Such an assumption is only known to be true for simple systems such as Axiom A diffeomorphisms [1]. For continuous maps, the *existence* of an ergodic invariant measure is assured [2]. As we intend to treat more general (nonuniformly hyperbolic) systems than Axiom A, we must *assume* that there exists an "important" ergodic invariant measure μ. It is the Lyapunov exponent of μ with which we shall concern ourselves. We know from Oseledec's multiplicative ergodic theorem [3], that for μ-almost-all starting points x, Eq. (1) will produce the same d values, known as Lyapunov exponents. For simplicity, we assume that the measure μ has d distinct exponents. Oseledec's theory then also guarantees the pointwise existence of d linearly independent unit vectors $w_x^{(i)}$, $1\le i\le d$, satisfying

$$\lim_{n\to\infty}\frac{1}{n}\ln\|Df^n(x)w_x^{(i)}\|=\lambda^{(i)}\ . \tag{2}$$

To obtain the ith Lyapunov exponent, we must evaluate (2).

A. Calculation of Lyapunov exponents using a time average

The traditional method of computing the largest Lyapunov exponent is to try to directly evaluate the limit in (2) using a randomly chosen x and v; see, for example, [4–6]. This involves rewriting $Df^n(x)$ as $Df(f^{n-1}(x))Df(f^{n-2}(x))\cdots Df(x)$ and then following the orbit of x around for n steps and multiplying all of the Jacobian matrices together in sequence (or alternatively, iteratively applying the matrices to a starting vector). Because the direction associated with the largest exponent stretches more (on average) than the other directions, a randomly chosen vector will nearly always be dragged into line with the correct direction after a number of iterations. This technique, known as time averaging, is fraught with both practical and fundamental theoretical problems. The length of orbit and computing time required for sufficient convergence of the computed exponent to a certain value may be very large. The multiplication of this large number of matrices (either directly, or indirectly by keeping track of a sequence of vectors) may also be numerically inaccurate.

Even if we have infinite precision computations, there is no guarantee that the chosen starting point for the long orbit will produce a "typical" orbit of the system; it may, for instance, be a very long periodic orbit. Alternatively, the orbit may be transient for a long time and not have settled down to "typical" behavior by the end of the computed time series. For example, Gambaudo and Tresser [7] describe a map which behaves "chaotically" for around 600 000 or 1 500 000 iterations (depending on the computer used) before falling into a sink of period 11. If one were to estimate the Lyapunov exponents of such an orbit using a time average, a stable value would be reached after 250 000 iterations and maintained up until the point where the orbit falls into the periodic orbit. Such an example highlights the dangers of estimating Lyapunov exponents, which are descriptions of *limiting* behavior, using a time average of *finite* length. The space average overcomes these difficulties by giving an explicit approximation of the invariant measure from which we obtain the Lyapunov exponents. In this way, we do not select starting points outside the support of the "interest-

ing" measure. Perhaps most importantly, there is little theoretical understanding of what is really going on in the time-averaging procedure. To overcome these potential difficulties, we instead propose to calculate the exponents using a space average rather than a time average.

B. Calculation of Lyapunov exponents via spatial averaging

We propose to average local expansion and contraction spatially, using the invariant measure of the dynamical system. Such a technique is commonly used in one-dimensional systems, though the extension to higher dimensions is not widely known.

By collecting together the vectors $w_x^{(i)}$ over all of phase space, we define a set of d vector fields $\{w_x^{(i)}\}$ which vary measurably with x for each $i=1,\ldots,d$. These vectors $w_x^{(i)}$ are the crucial objects for the spatial averaging technique. They satisfy the important relation

$$Df(x)w_x^{(i)}=a^{(i)}(x)w_{f(x)}^{(i)}\ , \tag{3}$$

where $a^{(i)}(x)$ is a scalar function of x; see [2,8]. The $a^{(i)}(x)$ represent the amount of stretching or shrinking that occurs along the special directions $w_x^{(i)}$. Now,

$$\begin{aligned}
&\lim_{n\to\infty}\frac{1}{n}\ln\|Df^n(x)w_x^{(i)}\| \\
&\quad=\lim_{n\to\infty}\frac{1}{n}\ln\|Df(f^{n-1}(x))\cdots Df(x)w_x^{(i)}\| \\
&\quad=\lim_{n\to\infty}\frac{1}{n}\ln\|a^{(i)}(f^{n-1}(x))\cdots a^{(i)}(x)w_{f^n(x)}^{(i)}\| \\
&\quad=\lim_{n\to\infty}\frac{1}{n}\ln\prod_{j=0}^{n-1}|a^{(i)}(f^j(x))| \qquad (4)\\
&\quad=\lim_{n\to\infty}\frac{1}{n}\sum_{j=0}^{n-1}\ln|a^{(i)}(f^j(x))| \\
&\quad=\lim_{n\to\infty}\frac{1}{n}\sum_{j=0}^{n-1}\ln\|Df(f^j(x))w_{f^j(x)}^{(i)}\| \\
&\quad=\int\ln\|Df(x)w_x^{(i)}\|d\mu(x) \qquad (5)
\end{aligned}$$

using transformation (3), and the Birkhoff theorem for the final equality. The spatial average takes a contribution to the ith exponent from each point x in phase space (the amount of stretching in the appropriate direction $w_x^{(i)}$) and weights this contribution according to the invariant measure μ. The invariant measure we want to use is the natural measure of the system which weights areas of high density more heavily than areas of low density.

In simple cases like fixed points and periodic orbits, it is easy to compute the $w_x^{(i)}$, using transformation (3) and Eq. (4).

Example I.1. Let x be a fixed point, and arrange the eigenvalues of $Df(x)$ in ascending order according to their moduli.

(i) If the ith eigenvalue is real and of unit multiplicity, then the logarithm of the modulus of this eigenvalue is the ith Lyapunov exponent and $w_x^{(i)}$ is the corresponding (normalized) eigenvector.

(ii) If the eigenvalues i up to $i+j-1$ are real and equal (multiplicity j), then rather than a vector, we have a j-dimensional subspace $W_x^{(i)}$ associated with the ith exponent, which is equal to the logarithm of the eigenvalues. The subspace $W_x^{(i)}$ is spanned by the j eigenvectors corresponding to eigenvalues i up to $i+j-1$.

(iii) If the eigenvalues i up to $i+j-1$ are complex and of equal modulus, then we again have a j-dimensional subspace $W_x^{(i)}$ associated with the ith exponent. The ith exponent is equal to the logarithm of the modulus of the eigenvalues.

Example I.2. If x is a periodic point of period p, then x is a fixed point of f^p. For simplicity, assume that $Df^p(x)$ has all its eigenvalues real and distinct. The $w_x^{(i)}$ are then normalized eigenvectors of $Df^p(x)$. By writing $Df^p(x)$ in the form $Df(f^{p-1}(x))Df(f^{p-2}(x))\cdots Df(x)$ and cyclically permuting the order of the matrices, the $w_{f^j(x)}^{(i)}$ may be obtained for $j=1,..,p-1$. From (4), it is readily seen that the corresponding exponents are given by the logarithms of the product of the $a^{(i)}(f^j(x))$, that is, $\lambda^{(i)}=\ln[\prod_{j=0}^{p-1}|a^{(i)}(f^j(x))|]^{1/p}$.

As a more complicated example, the vector fields $\{w_x^{(i)}\}$ and $\{w_x^{(2)}\}$ have been approximated and plotted for the Hénon map (strictly Hénon attractor) in Fig. 1. The Hénon system has two distinct exponents, one positive and the other negative. The positive exponent generates the vector field $\{w_x^{(2)}\}$ and $\{w_x^{(1)}\}$ is generated by the negative exponent. Note that the $w^{(2)}$ appear to be tangent to the attractor, which contains the unstable manifold of the fixed point.

To compute the vector corresponding to the largest exponent for the point x, we iterate x backwards a few times (three or four) and choose a random vector. We then apply the Jacobian matrices in sequence to "push forward" the random vector to the point x. Clearly, this method is not guaranteed to work for nonuniformly hyperbolic systems such as Hénon as you may hit a con-

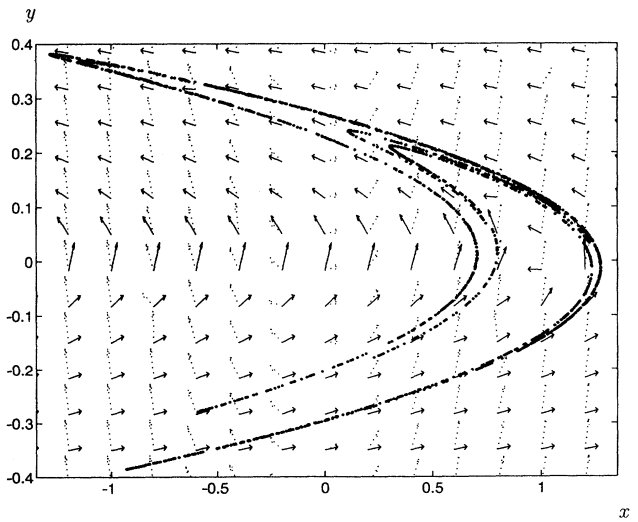

FIG. 1. Approximation of the vector fields for the Hénon map; $\{w_x^{(1)}\}$ (dotted) and $\{w_x^{(2)}\}$ (solid).

tracting region where there should be expansion, but "most" of the time the computed vector looks roughly correct. To compute the smallest exponent, reverse the procedure, iterating forward and then "pulling back" a random vector with Jacobian matrices. A brief discussion of the "folding up" of vectors is given in Wolf *et al.* [5].

The existence of the vector fields $\{w_x^{(i)}\}$ addresses the problem identified in the widely referenced paper [5]. "The apparent multivaluedness of the most rapidly growing direction in some regions of the attractor shows that this direction is not simply a function of position on the attractor." We also remark that the authors of [5] appear to suggest that what we are proposing to do is not possible. "Attempts to estimate exponents by averaging local contraction and expansion rates of phase space are likely to fail at the point where these contributions to the exponents are combined... We conclude that exponent calculation by averaging local divergence estimates is a dangerous procedure."

C. Invariant measures of dynamical systems

The main practical difficulty with the spatial averaging technique is that is requires an estimate of the "physical" or "natural" invariant measure of the dynamical system. Such notions have precise mathematical definitions only in very simple and well understood systems, such as expanding or unimodal maps in one dimension and Axiom A diffeomorphisms in higher dimensions. For Axiom A systems, it is known that there is a unique "physical" measure, known as the Sinai-Bowen-Ruelle (SBR) measure, which is exhibited by orbits of Lebesgue-almost-all starting points in a neighborhood of the attractor. In the nonuniformly hyperbolic case, very little is known about the existence of "physical" measures. For example, it has only recently been shown [9] that the Hénon map (with parameters $b>0$ fixed and sufficiently small, and $a\in\Delta_b$, a set of positive measure) admits an SBR measure. However, nothing can yet be said about "physical" measures for the Hénon map with the usual parameters, $a=1.4, b=0.3$. So the situation is very difficult and to say anything precisely, strong assumptions have to be laid down.

For general dynamical systems, the natural measure is usually taken by experimentalists to mean the limiting distribution of long orbits that show up on a computer for nearly all initial points. This description is far from satisfactory as it depends on machine properties that vary from computer to computer. Two problems with computer modeling are round-off errors and the fact that the computer operates on a finite, discrete state space. The finite state space means that *all* orbits are eventually periodic; a property that may be in stark contrast to the real system. The round-off error (combined with sensitive dependence on initial conditions) means that computer orbits may be completely different from true orbits of the system. The shadowing lemma is often cited as justification of computer modeling, but even in cases where it may be applied, there is no guarantee that the shadowing orbit exhibits the "important" measure. Minimal progress has been made on relating convergence of computer orbits (under increasing computer precision) to "important" invariant measures of the true system. Most of the work has centered on one-dimensional systems, where absolutely continuous invariant measures are considered to be the "important" measures. In the case of maps of the unit interval which admit a unique absolutely continuous invariant measure, Góra and Boyarsky [10] have shown that histograms produced by long computer orbits approach the histogram given by the invariant measure as the computer precision increases. This result depends on the existence of long, nonperiodic orbits whose length grows at a rate proportional to the cardinality of the set of computer distinguishable numbers, an assumption which appears to be false even in simple cases [11]. Diamond, Kloeden, and Pokrovskii [12] have given a characterization of invariant measures exhibited by the "circle-doubling" map under various symmetric random perturbations of binary arithmetic, showing that the binary arithmetic leaves the Lebesgue measure invariant in the infinite precision limit only for special perturbation values.

We do not concern ourselves with these problems in this paper. Our method of approximating the invariant measure, discussed in the following section, is to compute invariant measures of systems which are subjected to small random perturbations. The magnitude of our perturbations will be much greater than the machine epsilon of our computer so that we ignore perturbation effects due to roundoff. As the perturbations go to zero, we extract a limiting measure which is invariant for the unperturbed system. We argue that invariant measures that arise as limits of invariant measures of perturbed systems have physical significance as they are in some sense robust with respect to small perturbations.

II. PRACTICAL ESTIMATION OF LYAPUNOV EXPONENTS OF DETERMINISTIC DYNAMICAL SYSTEMS

At the outset, we wish to make it clear that our proposed space-averaging technique is far from superseding the time average as the experimentalist's preferred method of computing Lyapunov exponents. Rather, it presents an alternative to the traditional method; one which provides greater theoretical and, it is hoped, intuitive understanding of what Lyapunov exponents are and how they are calculated.

In practice, there are two positions from which one wants to estimate Lyapunov exponents. Either the map is known, or it is unknown and only a set of data generated by the map is available. The difficulty with directly applying the spatial averaging method to the deterministic system is that one does not have an easily accessible estimate of the invariant measure. By representing the reconstructed deterministic system as a random dynamical system governed by a Markov chain, we obtain an approximation of the invariant measure, namely, the invariant density of the chain. For practical purposes, we need a finite state Markov chain for calculations to be done by a computer. To define the Lyapunov exponents of a Mar-

kov chain, we require a matrix to be associated with each state, which means a finite number of Jacobian matrices. In other words, our map, reconstructed or not, should be piecewise linear. The invariant density we obtain will assign a weight to each portion of linearity. We produce a sequence of increasingly accurate piecewise linear approximations-reconstructions of our map, which induces a sequence of Markov chains. In the case of an unknown map, the sequence of reconstructions may be enumerated, for example, as more data points become available. We require that the sequence of Jacobian matrices of our piecewise linear maps converge uniformly to the Jacobian matrices of the true map. We also want the Markov chains to be small random perturbations [13] of the true map. The hope is that the computed exponents of the Markov chains approach the Lyapunov exponents of the *true* deterministic map. Young [14] has pointed out that in general this is not true, i.e., there exist maps which, when varied continuously in the C^1 topology, have Lyapunov exponents that vary discontinuously. In other words, by approximating a map, we do not necessarily approximate its Lyapunov exponents. If we impose a strong condition on the map f, amounting to each vector field in the Oseledec splitting $\{w_x^{(i)}: i=1,\ldots,d\}$ varying continuously with x, then our hope is substantiated in the sense that diffeomorphisms near f have similar Lyapunov exponents.

In the case of two-dimensional systems with exponents of opposite sign, uniform hyperbolicity is enough, as this gives us a continuous splitting. So, for two-dimensional, uniformly hyperbolic systems, convergence of the exponents is assured but in higher dimensions further restrictions apply.

This lack of continuity is not a failure of the space-averaging method; it is a general fact of Lyapunov exponents. Lyapunov exponents are commonly estimated in the physics literature using time averages of approximate maps under the (false) implicit assumption that approximate maps possess similar exponents. However, computer experiments seem to somehow hide such pathological behavior, and for numerical estimation of Lyapunov exponents, the authors will turn one blind eye and not discourage spatially averaging exponents of approximate maps.

In our example we estimate exponents from a finite set of data. We use the method of triangulation (Mees [15]) to provide us with an approximation of the true system. Initially, the convex cover of the data points, denoted by $\mathcal{C}$, is triangulated using d simplices; see Fig. 2.

On each simplex, the map is approximated linearly, so we obtain a continuous piecewise linear approximation to the system. From the reconstructed map, the derivative may also be computed, and will be constant on each simplex.

Approximating the invariant measure

Having split the space up into m disjoint simplices, we look for an approximation of the invariant measure which gives a weight to the centroid of each simplex. The image of the vertices of each simplex is known, and

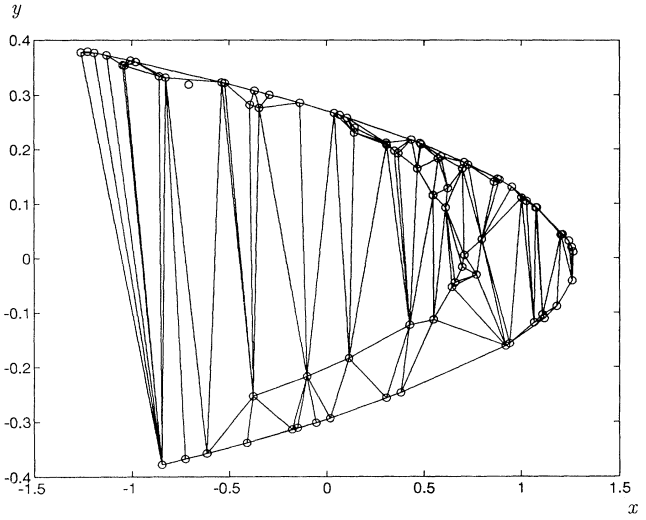

FIG. 2. Triangulation of an 80 point orbit.

since the map is affine on each simplex, the image of any simplex will lie in the convex cover of the image of its defining vertices. The transition probability from simplex (i) to simplex (j) represents the probability that the image of a point in simplex (i) lies in simplex (j), and is denoted by P_{ij}. Appealing to the linearity of the map on simplices, a suitable approximation of transition probabilities is obtained by setting

$$P_{ij}=\frac{l((\text{image of } S_i)\cap S_j)}{l(\text{image of } S_i)} ,$$

where S_i denotes simplex (i) and l is the natural Riemannian volume. From this definition, it is readily seen that

$$\sum_{j=1}^{m} P_{ij}=1 \quad \text{for } i=1,\ldots,m$$

and

$$P_{ij}\geq 0 \quad \text{for } i=1,\ldots,m ,$$

so that P is a stochastic matrix. The matrix P represents the evolution of our system from a probabilistic point of view. Instead of a starting point for an orbit, we now have an initial distribution on our m states, $\pi^{(0)}=(\pi_1^{(0)},\pi_2^{(0)},\ldots,\pi_m^{(0)})$, and subsequent distributions are given by

$$\pi^{(i)}=\pi^{(i-1)}P=\pi^{(0)}P^i .$$

The invariant distribution we seek will satisfy $\pi=\pi P$, and so is just a left eigenvector of P with unit eigenvalue. Associated work has been done by Murao and co-workers in Ref. [16] where invariant measures are estimated on tori using a related technique. Figure 3 shows the approximation to the invariant measure for the map reconstructed from the 80 data point orbit.

The Markov process governed by P_{ij} can be thought of as a deterministic mapping on the finite space $\{x_1,\ldots,x_m\}$ of centroids of simplices with random noise added after each computation. After computing

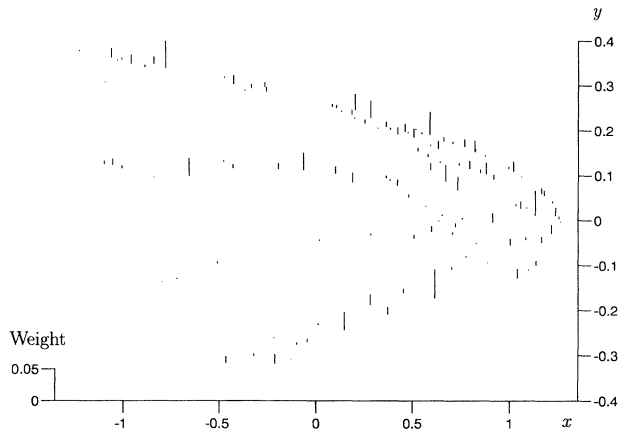

FIG. 3. Approximation of the invariant measure of the Hénon map using an 80 point orbit. The measure is a finite combination of δ-function measures shown as spikes on the two-dimensional plot. The base of the spike represents the position of the δ-function measure, and the height represents the relative weight given to the centroid.

the true image of x_i, the result is randomly perturbed so that the resulting value is one of the points $\{x_j\}$ with $P_{ij}\neq 0$. Because of the way P_{ij} is defined, as the distance between points decreases, the "radius" of perturbation also lessens. It may be shown [17] that the Markov chain defined with transition matrix P_{ij} as above is a small random perturbation of the true deterministic system. In the next section we see how to compute the Lyapunov exponents of the induced Markov chain.

III. LYAPUNOV EXPONENTS OF RANDOM DYNAMICAL SYSTEMS

Lyapunov exponents may be defined for random systems analogously to the deterministic case. For a deterministic dynamical system (M,f), each point $x\in M$ has a matrix associated with it, namely, $Df(x)$, the Jacobian matrix of f evaluated at the point x. For a random dynamical system, we again have matrices associated with each point in our domain, and as the evolution of the system moves randomly about from point to point, the matrices corresponding to the points in the orbit are multiplied together in sequence. Thus we may talk about the Lyapunov exponents of this random product of matrices. For our purposes, we are only interested in Lyapunov exponents of Markov chains, as this is the simplest type of random system that can be used to model a deterministic system. We will be looking at random products of matrices with Markovian dependence.

A. Calculation of Lyapunov exponents of a Markov chain via spatial averaging

Random dynamical systems are a generalization of deterministic dynamical systems in the following sense. Let the state space be some smooth manifold M, and let the set Ω contain all the possible histories of the random system as it moves from state to state. In the deterministic case, once an initial point $x\in M$ is specified, the only possible history is

$$(\ldots,f^{-i}(x),\ldots,f^{-1}(x),x,f(x),\ldots,f^{i}(x),\ldots)\ .$$

This is in contrast to the random system, which for a given starting point may have many different histories. In the case of the deterministic system, we were able to spatially average the exponents over the state (points in M) by measuring the stretching caused by the Jacobian matrix of that state (Df evaluated at the point in M) along a single direction specified by the state. For random systems, each state has many histories which pass through it, and so we must average the stretching induced by a state's matrix over a "smear" of directions, to compensate for the increased complexity. To represent these "smears" we consider $(d-1)$-dimensional real projective space, denoted $\mathbb{R}P^{d-1}$. Elements of this space are equivalence classes of vectors in $\mathbb{R}^d$, where two vectors are considered equivalent if they are real scalar multiples of one another. $\mathbb{R}P^{d-1}$ can be thought of as the "top" half of S^{d-1} embedded in $\mathbb{R}^d$, with opposing points on the "diameter" being identified as equal. In fact, we always adopt this point of view, with the statement $v\in\mathbb{R}P^{d-1}$ meaning v is a unit vector in $\mathbb{R}^d$. When we speak of the action of a $d\times d$ matrix on an element of $\mathbb{R}P^{d-1}$, we mean matrix multiplication of the unit vector representing the point in S^{d-1}. In order to take in contributions to the stretching from the matrix for state x, we associate with the state a probability measure μ_x on $\mathcal{B}(\mathbb{R}P^{d-1})$, the Borel subsets of $\mathbb{R}P^{d-1}$.

1. *The primary method of determining the largest Lyapunov exponent*

As in the case of the deterministic system, we obtain a contribution to the Lyapunov exponents from each state x and weight this contribution according to the invariant measure. In what we have called the primary method, we average the stretching induced by the original Jacobian matrices to calculate the largest exponent. The next section describes the use of the *inverse* Jacobian matrices to calculate the largest exponent. This is known as the secondary method. The roles of the original and inverse Jacobian matrices are exchanged for the primary and secondary methods of calculating the smallest Lyapunov exponent.

Define random variables $A_n:\Omega\to GL(d,\mathbb{R})$ by $A_n(\omega)=Df(x_n)$ where

$$\omega=(\ldots,x_{-i},\ldots,x_{-1},x_0,x_1,\ldots,x_i,\ldots),\omega\in\Omega,x_i\in M\ .$$

Let

$$\begin{aligned}S_n(\omega)&=A_n(\omega)A_{n-1}(\omega)\cdots A_0(\omega)\\&=Df(x_n)Df(x_{n-1})\cdots Df(x_0)\ .\end{aligned}$$

For some $v'\in\mathbb{R}P^{d-1}$ (see [18]), the largest Lyapunov exponent λ is given by

$$\lambda = \lim_{n\to\infty} \frac{1}{n} E(\ln\|S_n(\omega)v'\|)$$
$$= \lim_{n\to\infty} \frac{1}{n} \int_\Omega \ln\|S_n(\omega)v'\| dP_\pi(\omega) ,$$

where P_π is the measure on Ω invariant under the shift map. We may express this integral as an integral over the product of the spaces M and $\mathbb{R}P^{d-1}$ using a suitable product measure μ [17,18], so that

$$\lambda = \int_{M\times\mathbb{R}P^{d-1}} \ln\|Df(x)v\| d\mu(x,v)$$
$$= \int_M \int_{\mathbb{R}P^{d-1}} \ln\|Df(x)v\| d\mu_x(v) d\pi(x) , \quad (6)$$

by disintegration of the measure μ with respect to π; see [19], p. 4. Here, π is the invariant measure of the Markov system satisfying $\int_M P(x,\Gamma) d\pi(x) = \pi(\Gamma)$ for all $x\in M, \Gamma\in\mathcal{B}(M)$. The family of measures $\{\mu_x\}$ satisfy the invariance conditions

$$\mu_y(A) = \int_M \mu_x([Df(x)]^{-1}A) P^*(y,dx) , \quad (7)$$

for all $y\in M$ and $A\in\mathcal{B}(\mathbb{R}P^{d-1})$, where P^* is the transition function representing the reverse Markov process. Thus the largest exponent will be given by (6) using a family of measures $\{\mu_x\}$ on $\mathbb{R}P^{d-1}$ such that (7) holds.

Note that (6) is a generalization of the spatial average formula for deterministic systems (5). In (6) we first average the stretching induced by state x's matrix over a smear of directions given by the probability measure μ_x, before weighting this contribution according to the invariant measure π.

Since we have a finite state Markov chain, the integral over M may be replaced by a finite sum. We also wish to reduce the second integral over $\mathbb{R}P^{d-1}$ to a finite sum using a suitable approximation. Denote our finite number of states by $x_1, \ldots, x_m$. Associated with each of these states is a matrix $Df(x_k)$, $k=1,\ldots,m$. For want of a better method, we approximate measures on $\mathbb{R}P^{d-1}$ using histograms. To avoid confusion, we denote the approximation to μ_x by ν_x. For simplicity, we only describe the case where $d=2$, so that $\mathbb{R}P^1$ is just one side of S^1, with the ends being identified as the same point. Split the semicircle up into N equal segments s_j, $j=1,\ldots,N$, and consider the action of a fixed matrix $g\in GL(2,\mathbb{R})$ on $\mathbb{R}P^1$. Let ν be the measure on this discretized space, with the density on the jth segment s_j being uniform. Figure 4 shows a density function of an absolutely continuous measure on $\mathbb{R}P^1$ overlaid with a possible histogram approximation on 100 segments.

Now

$$\nu(g\cdot s_j) = \sum_i \frac{l((g\cdot s_j)\cap s_i)}{l(s_i)} \nu(s_i)$$
$$= \sum_i (\text{fraction of } s_i \text{ overlapped by } g\cdot s_j)\nu(s_i) , \quad (8)$$

by uniformity of the density on s_i. The matrix t_{ij} (equal to the fraction of s_i overlapped by $g\cdot s_j$) is stochastic and represents the action of the matrix g on projective space. We may fully describe the approximate measure ν by its measure of each segment s_j. Set $\nu(s_j)=\nu_j$, where the ν_j are real numbers and $\sum_j \nu_j = 1$. Defining $\nu' = \nu\circ g$, the density of ν' is given by $\nu'_j = \sum_i \nu_i t_{ij}$.

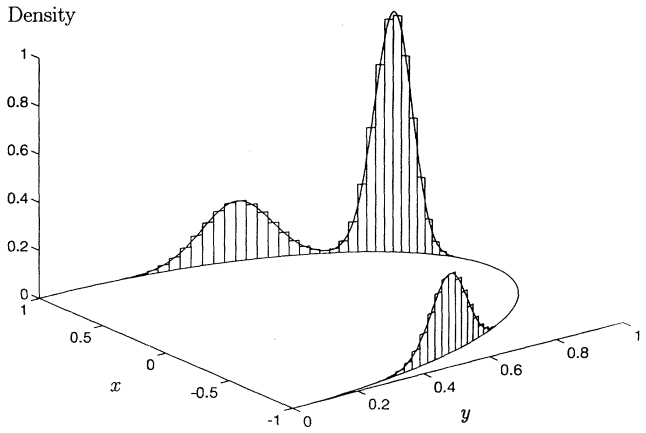

FIG. 4. Approximation of a measure on $\mathbb{R}P^1$ using 100 bins.

We consider the direct method of finding the largest exponent, and approximate the family of measures satisfying (7). Denote by $t^{(k)}$ the stochastic matrix approximating the action of $Df(x_k)^{-1}$ on projective space. For clarity, we abbreviate the states $x_1,\ldots,x_m$ as the integers $1,\ldots,m$, and denote the transition matrix for the reversed Markov system by $\hat{P}$. In our approximate setup, (7) becomes

$$\nu_j(s_l) := \nu_{j,l} = \sum_{i=1}^m \nu_i (Df(i)^{-1}s_l)\hat{P}_{ji}$$
$$= \sum_{i=1}^m \sum_{k=1}^N t_{kl}^{(i)} \nu_i(s_k)\hat{P}_{ji}$$
$$= \sum_{i=1}^m \hat{P}_{ji} \sum_{k=1}^N \nu_{i,k} t_{kl}^{(i)} \quad (9)$$

for all $j=1,\ldots,m$ and $l=1,\ldots,N$. Note that these are all strict equalities. For an arbitrary measurable set $A\subset\mathbb{R}P^{d-1}$, we proceed as follows.

$$\nu_j(A) = \sum_{l=1}^N \nu_j(A\cap s_l)$$
$$= \sum_{l=1}^N \frac{l(A\cap s_l)}{l(s_l)} \nu_j(s_l)$$
$$= \sum_{l=1}^N \frac{l(A\cap s_l)}{l(s_l)} \sum_{i=1}^m \hat{P}_{ji}\nu_i(Df(i)^{-1}s_l)$$
$$= \sum_{i=1}^m \hat{P}_{ji} \sum_{l=1}^N \frac{l(A\cap s_l)}{l(s_l)} \nu_i(Df(i)^{-1}s_l)$$
$$\approx \sum_{i=1}^m \hat{P}_{ji}\nu_i(Df(i)^{-1}A) .$$

Our approximation is exact for sets that are unions of full individual bins s_l, but is inexact for sets covering parts of bins. The set of $m\times N$ linear equations (9) may be given a matrix representation as

$$(\nu_{1,1},\ldots,\nu_{1,N},\nu_{2,1},\ldots,\nu_{2,N},\ldots,\nu_{m-1},\ldots,\nu_{m,N})$$
$$=(\nu_{1,1},\ldots,\nu_{1,N},\nu_{2,1},\ldots,\nu_{2,N},\ldots,\nu_{m,1},\ldots,\nu_{m,N})\begin{bmatrix}\hat{P}_{11}t^{(1)} & \hat{P}_{21}t^{(1)} & \cdots & \hat{P}_{m1}t^{(1)}\\ \hat{P}_{12}t^{(2)} & \hat{P}_{22}t^{(2)} & \cdots & \hat{P}_{m2}t^{(2)}\\ \vdots & \vdots & \ddots & \vdots\\ \hat{P}_{1m}t^{(m)} & \hat{P}_{2m}t^{(m)} & \cdots & \hat{P}_{mm}t^{(m)}\end{bmatrix},$$

where each entry displayed in the matrix is an $N\times N$ block. The required vector of length mN is an eigenvector of unit eigenvalue. All components of this eigenvector must be non-negative, as each component represents the measure of a slice of projective space. Another important restriction that must be placed on this vector is that $\sum_l \nu_{j,l}=1$ for all $j=1,\ldots,m$. We require this because ν_j is a probability measure for each j and hence $\sum_l \nu_j(s_l)=1$.

The following proposition [17] shows that the $mN\times mN$ matrix $\mathcal{M}$ has a left eigenvector of eigenvalue one, satisfying the above requirements.

Proposition III.1. The matrix $\mathcal{M}$ has a non-negative left eigenvector x of eigenvalue 1 satisfying

$$\sum_l x_{(j-1)N+l}=1 \quad \text{for each } j=1,\ldots,m\ .$$

However, as $\mathcal{M}$ is an $mN\times mN$ matrix, memory constraints could become a problem for computer routines that calculate eigenvalues. It would be nice if the eigenvalue one were the largest eigenvalue in modulus as we could then use the power method to approximate the eigvenvector. This would only require the storage of $\mathcal{M}$ itself, which is typically sparse, and the vector of length mN. Fortunately, it may be shown [17] that the moduli of all eigenvalues of $\mathcal{M}$ are bounded by unity. By choosing a starting vector with all positive entries, say all ones, we can be certain that the power method will provide us with an eigenvector that satisfies (9). This is because 1 is the dominant positive eigenvalue, and other eigenvalues of modulus one cannot have non-negative eigenvectors. Now that we know how to find a solution to (9), we would like to see [17] that it does in fact represent a good approximation of a measure satisfying (7).

Proposition III.2. Let $\nu_j^{(N)}$ denote the approximate measure on $\mathbb{R}P^1$ for state j obtained using N equal "bins" in projective space. If the solutions of (9) approach a continuous density as the number of bins increases, then weak limits of the sequence $\{\nu_j^{(N)}\}$ satisfy (7).

2. *The secondary method of determining the largest Lyapunov exponent*

So far, we have been using the matrices $Df(x)$ to measure the local stretching. We can repeat [17] the whole above procedure, this time using the inverse matrices $[Df(x)]^{-1}$ to find the largest exponent. The appropriate condition for the family of measured $\{\mu_x^2\}$ is

$$\mu_y^2(A)=\int_X \mu_x^2([Df(y)]^{-1}A)P^*(y,dx) \tag{10}$$

for all $y\in M$. We then have that

$$\lambda=-\int_M\int_{\mathbb{R}P^{d-1}} \ln\|[Df(x)]^{-1}v\|d\mu_x^2(v)d\pi(x)\ . \tag{11}$$

We now have two different families of measures that will provide us with the largest Lyapunov exponent of our Markov system.

All of the theory in the preceding section applies to the approximations for the remaining three families of measures, $\{\nu^2\},\{\nu^3\},\{\nu^4\}$. To approximate the family of measures $\{\nu^2\}$, we apply the same approximation to Eq. (10).

$$\nu_j^2(s_l)=\nu_{j,l}^2=\sum_{i=1}^m \nu_i^2(Df(j)^{-1}s_l)\hat{P}_{ji}$$
$$=\sum_{i=1}^m\sum_{k=1}^N t_{kl}^{(j)}\nu_i^2(s_k)\hat{P}_{ji}$$
$$=\sum_{i=1}^m \hat{P}_{ji}\sum_{k=1}^N \nu_{i,k}^2 t_{kl}^{(j)}$$

for all $j=1,\ldots,m$ and $l=1,\ldots,N$. This set of linear equations has matrix representation

$$(\nu_{1,2}^2,\ldots,\nu_{1,N}^2,\nu_{2,1}^2,\ldots,\nu_{2,N}^2,\ldots,\nu_{m,1}^2,\ldots,\nu_{m,N}^2)$$
$$=(\nu_{1,1}^2,\ldots,\nu_{1,N}^2,\nu_{2,1}^2,\ldots,\nu_{2,N}^2,\ldots,\nu_{m,1}^2,\ldots,\nu_{m,N}^2)\begin{bmatrix}\hat{P}_{11}t^{(1)} & \hat{P}_{21}t^{(2)} & \cdots & \hat{P}_{m1}t^{(m)}\\ \hat{P}_{12}t^{(1)} & \hat{P}_{22}t^{(2)} & \cdots & \hat{P}_{m2}t^{(m)}\\ \vdots & \vdots & \ddots & \vdots\\ \hat{P}_{1m}t^{(1)} & \hat{P}_{2m}t^{(2)} & \cdots & \hat{P}_{mm}t^{(m)}\end{bmatrix}.$$

3. *The primary method of determining the smallest Lyapunov exponent*

By running our system backwards we may obtain the smallest Lyapunov exponent, as the smallest exponent of the forward system is equal to the negative of the largest exponent of the reversed system. Recall that this is also the case for deterministic systems. The corresponding invariance condition for the family $\{\mu_x^3\}$ is

$$\mu_x^3(A)=\int_M \mu_y^3(Df(y)A)P(x,dy) \tag{12}$$

for all $x\in M$. We then have

$$\lambda^R=\int_M \int_{\mathbb{R}P^{d-1}} \ln\|[Df(x)]^{-1}v\|d\mu_x^3(v)d\pi(x)\ . \tag{13}$$

For the remaining two families $\{\nu^3\}$ and $\{\nu^4\}$, we need to approximate the action of $Df(x)$ rather than $Df(x)^{-1}$, and we denote the approximation of the action of $Df(k)$ by $t^{(k)'}$. To approximate the family of measures $\{\nu^3\}$, we use Eq. (12).

$$\begin{aligned}\nu_i^3(s_l)=\nu_{i,l}^3&=\sum_{j=1}^m \nu_j^3(Df(j)s_l)P_{ij}\\&=\sum_{j=1}^m\sum_{k=1}^N t_{kl}^{(j)'}\nu_j^3(s_k)P_{ij}\\&=\sum_{j=1}^m P_{ij}\sum_{k=1}^N \nu_{j,k}^3 t_{kl}^{(j)'}\end{aligned}$$

for all $i=1,,\ldots,m$ and $l=1,\ldots,N$. The matrix we seek a left unit eigvenvector of is

$$\begin{bmatrix} P_{11}t^{(1)'} & P_{21}t^{(1)'} & \cdots & P_{m1}t^{(1)'}\\ P_{12}t^{(2)'} & P_{22}t^{(2)'} & \cdots & P_{m2}t^{(2)'}\\ \vdots & \vdots & \ddots & \vdots\\ P_{1m}t^{(m)'} & P_{2m}t^{(m)'} & \cdots & P_{mm}t^{(m)'}\end{bmatrix}.$$

4. *The secondary method of determining the smallest Lyapunov exponent*

We may again apply our technique to the reversed system. The invariance formula for the family $\{\mu_x^4\}$ reads

$$\mu_x^4(A)=\int_M \mu_y(Df(x)A)P(x,dy) \tag{14}$$

for all $x\in M$, and

$$\lambda^R=-\int_M \int_{\mathbb{R}P^{d-1}} \ln\|Df(x)v\|d\mu_x^4(v)d\pi(x)\ . \tag{15}$$

As in the case of the largest exponent, we have two different families of measures that provide us with the smallest Lyapunov exponent of our Markov system.

The family of measures $\{\nu^4\}$ approximate Eq. (14).

$$\begin{aligned}\nu_i^4(s_l)=\nu_{i,l}^4&=\sum_{j=1}^m \nu_j^4(Df(i)s_l)P_{ij}\\&=\sum_{j=1}^m\sum_{k=1}^N t_{kl}^{(i)'}\nu_j^4(s_k)P_{ij}\\&=\sum_{j=1}^m P_{ij}\sum_{k=1}^N \nu_{j,k}^4 t_{kl}^{(i)'}\end{aligned}$$

for all $i=1,\ldots,m$, and $l=1,\ldots,N$. The matrix we seek a left unit eigenvector of is

$$\begin{bmatrix} P_{11}t^{(1)'} & P_{21}t^{(2)'} & \cdots & P_{m1}t^{(m)'}\\ P_{12}t^{(1)'} & P_{22}t^{(2)'} & \cdots & P_{m2}t^{(m)'}\\ \vdots & \vdots & \ddots & \vdots\\ P_{1m}t^{(1)'} & P_{2m}t^{(2)'} & \cdots & P_{mm}t^{(m)'}\end{bmatrix}.$$

B. How to obtain all the Lyapunov exponents

To obtain all of the exponents, the above theory may be repeated using the pth exterior powers $\wedge^p\mathbb{R}^d$ of $\mathbb{R}^d$. The matrices $Df(x)$ are replaced by the induced linear maps $Df^{\wedge p}:\wedge^p\mathbb{R}^d\to\wedge^p\mathbb{R}^d$. The top Lyapunov exponent of this new system is equal to $\lambda^{(d)}+\lambda^{(d-1)}+\cdots+\lambda^{(d-p+1)}$, where $\lambda^{(d)},\ldots,\lambda^{(d-p+1)}$ are the p largest exponents of the original system. By varying p we may obtain the sum of the p largest exponents for $p=1,\ldots,d$, and hence all of the separate exponents. From a computational point of view it may be difficult to resolve the negative exponents. This may be overcome by measuring all of the positive exponents, then reversing the system and measuring all of the (now) positive exponents, which were the originally negative ones.

IV. THE RELATIONSHIP BETWEEN THE LYAPUNOV EXPONENTS OF THE RANDOM AND DETERMINISTIC SYSTEMS

As the diameters of the triangles in our triangulated map go to zero by virtue of there being more data points made available, three important properties of the sequence of Markov chains converge to properties possessed by the true deterministic map. First, the Jacobian matrices of the piecewise linear approximation converge uniformly to the Jacobian matrices of the true map. Secondly, the evolution of the Markov chains becomes increasingly similar to that of the true map. Finally, the invariant measures of the sequence of Markov chains approach an invariant measure of the true map. Under the assumption that the true map satisfies the conditions outlined in Sec. II, and that the limiting invariant measure is an ergodic "physical" measure of the system, it may be shown (according to [14] and results of [17]) that the exponents of the sequence of Markov chains approach the Lyapunov exponents of the *true* map. Such a result is also true of any sequence of piecewise linear approximations which are small random perturbations of the true system and whose Jacobian matrices uniformly converge to the Jacobian matrices of the true map.

There is a similar convergence result for the case where one has a finite set of data or a fixed piecewise linear approximation. For the purpose of description, we consider a fixed triangulation in two dimensions obtained from a finite set of points. We may decrease the random perturbation of the system by breaking up the triangles into four identical triangles, each similar to the original. These new triangles will have a diameter half that of the

original. An upper bound for the perturbation of the triangulated map is half the diameter of the largest triangle plus half the diameter of the largest image triangle. Triangles whose image is large or which are themselves large should be targeted for the splitting up procedure. As the maximum diameter of the triangles goes to zero, the exponents of the Markov system approach the Lyapunov exponents of the *triangulated* map, under the above assumption on the map and the limiting measure. We cannot expect to find the Lyapunov exponents of the true map as no new information about the Jacobian matrices is obtained through our splitting up technique; for this we require more data points. Similar splitting procedures may be performed in higher dimensions. The same principle may also be applied to general piecewise linear approximations, not arising from triangulations.

V. EXAMPLES AND RESULTS

We use the much studied Hénon map to illustrate the performance of the techniques described in this paper. The spatial averaging technique was applied to reconstructions of the Hénon map obtained from time series of varying lengths. Recall that the accuracy of both the reconstruction and the estimates of the exponents depends on the diameters of the triangles rather than the number of data points. The results are shown in Table I.

Accepted values are $\lambda^{(2)} \approx 0.418, \lambda^{(1)} \approx -1.62$.

The estimate of $\lambda^{(1)}$ has been calculated indirectly, by subtracting the estimate of $\lambda^{(2)}$ from the estimate of $\lambda^{(2)}+\lambda^{(1)}$. As the number of data points increases, the estimates can be seen to improve for both exponents. We conjecture that the estimates of the larger exponent are much better than the estimates of the smaller exponent because all data points lie near the unstable manifold of the fixed point. Thus we are obtaining information about the stretching in preference to information about contraction. To expect better estimates of the smaller exponent we would require some transient data to contain the contraction information. To compute the time average of the exponents, we assumed the Hénon map to be known and multiplied together the Jacobian matrices of the finite orbits of length 24, 44, 80, and 201. This product was then applied to unit vectors in uniformly distributed directions and the results averaged over the starting vectors. As the determinant of the Jacobian matrices of the Hénon map is constant (ln3), $\lambda^{(1)}$ may be obtained by subtracting $\lambda^{(2)}$ from ln3. These values are included as a guide to the relative accuracy of various techniques. Of course, once we have our reconstructed map, we could perform the standard time average on arbitrarily long orbits. However, this defeats our purpose, as one is once again dogged by the problems alluded to in Sec. I A.

The maximum perturbation is given by

$$\text{(maximum perturbation)} = \max\{\text{radius of triangle}_i + \text{radius of image of triangle}_j : \text{triangle}_i \cap \text{image of triangle}_j \neq \phi\} .$$

This value is an upper bound for any perturbation of the triangulated map. The mean perturbation is a weighted average of the perturbations,

$$\text{(mean perturbation)} = \sum_{i=1}^{m} \pi_i \sum_{j=1}^{m} P_{ij} d_{ij} ,$$

where d_{ij} is the distance between the centroids of $f(S_i)$ and S_j.

Because of the shape of the Hénon attractor and the action of the map, these perturbations are quite large for the triangulations obtained from time series. For example, triangles that are mapped to the "top" and "bottom" edges of the attractor are often stretched into long, thin triangles with large diameters. The large gap on the left hand side of the attractor also creates large triangles, although the triangles spanning this gap are not as large as they look because of the difference in scaling in the x and y directions. We need to split up these triangles to reduce the perturbation. For our examples of the splitting procedure, we chose a suitable set of data points for which the original perturbation was relatively small, rather than a data set arising from a time series; see Fig. 5.

This choice is merely to reduce computing time and data storage. The splitting principle works equally well on any set of data; it just takes longer in some cases. The Hénon map was reconstructed from a set of 87 data points, and the splitting procedure applied twice. The results of our algorithms are shown in Table II.

The sum of the exponents is exactly $\ln 0.3 \approx -1.2040$.

For this example, the estimates of both exponents were directly calculated, so that the sum of the two estimates could be compared with the estimate of the sum of the exponents. Comparing the results of Tables I and II, we

TABLE I. Estimates of Lyapunov exponents for the Hénon map using data from time series.

Number of data points	Number of states	Maximum perturbation	Mean perturbation	Estimate of $\lambda^{(2)}$	Estimate of $\lambda^{(1)}$	Estimate of $\lambda^{(2)}+\lambda^{(1)}$	"Time average" estimate of $\lambda^{(1)}$
24	29	0.6585	0.1265	0.09	−0.86	−0.7702	0.5221
44	60	0.6630	0.1657	0.27	−1.01	−0.7325	0.4412
80	120	0.4728	0.0853	0.39	−1.25	−0.8630	0.4013
80	123	0.5829	0.0845	0.42	−1.35	−0.9250	0.4080
201	336	0.4411	0.0521	0.42	−1.37	−0.9496	0.3981

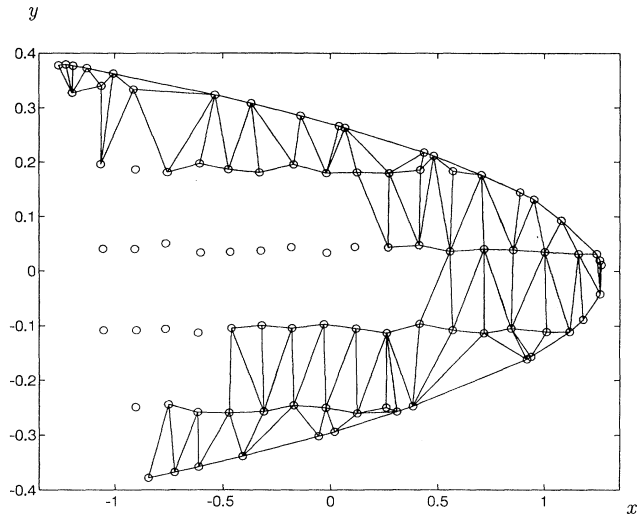

FIG. 5. Triangulation of 87 data points with triangles not intersecting the image triangulation removed.

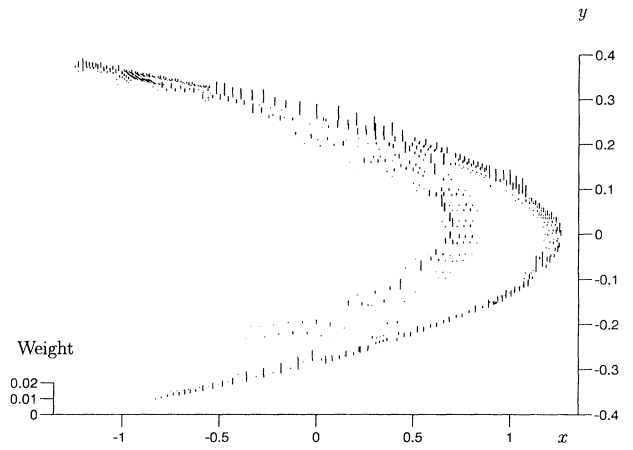

FIG. 6. Singular approximation of the invariant measure using 708 states. The measure is a convex combination of 708 δ-function measures, which are shown as spikes in this two-dimensional plot. The base of the spike represents the position of the support of the measure and the height represents the relative weight given to the centroid.

see that the estimated sum of the exponents is much better for the less perturbed map and is improving as the perturbations decrease. This effect is not due entirely to the perturbations; the Jacobian matrices obtained from the reconstruction also play a part and depend on the position of the data points. Note that it may not be necessary to split up the entire convex cover of the data points. After computing the transition matrix, it may be easily seen that there may be some triangles which have empty intersection with the image of the triangulation. These triangles may be immediately neglected and need not be split up. After each splitting, this process may be repeated to rid the triangulation of unnecessary triangles. In our example, the original triangulation contained 139 triangles. Computation of the transition matrix revealed that 86 triangles intersected the image triangulation, and that all of these were nontransient. The 86 triangles were then split up into 482 triangles to reach a prescribed level of perturbation. Of these, 339 intersected the image, and only 314 were nontransient. The 339 triangles were split again to produce a lower, predetermined level of perturbation.

We remark at this point that it is very simple indeed to obtain the sum of all of the exponents for our Markov system. For a d-dimensional Markov system with m states,

$$\lambda^{(d)}+\cdots+\lambda^{(1)}=\sum_{i=1}^{m}\pi_i \det[Df(i)]\ , \tag{16}$$

where $Df(i)$ is the matrix associated with state (i). This easily computable quantity may be used to check the values of the individual exponents, which are much more difficult to obtain. In all the cases that we have studied, the sum of the individual exponents has been very close to the value given by (16). From Table II, it seems that the estimated sum of the exponents is converging to a value near that of the true Hénon map. This suggests that the average area contraction has been accurately described by the reconstructed map. The individual exponents, however, appear to be converging to slightly incorrect values for the true Hénon map, indicating differences between the reconstructed piecewise linear map and the true map, in terms of average local expansion and contraction.

Two possible representations of the approximation of the invariant measures of the fixed triangulated map are shown in Figs. 6 and 7. Compare Fig. 6 with Fig. 3; clearly, decreasing the perturbations provides a much more accurate approximation.

The absolutely continuous approximation has been constructed from the singular approximation by spread-

TABLE II. Data from the sequence of splittings of a fixed triangulated map.

Number of data points	Number of nontransient states	Maximum perturbation	Mean perturbation	Estimate of $\lambda^{(2)}$	Estimate of $\lambda^{(1)}$	Estimate of $\lambda^{(2)}+\lambda^{(1)}$
87	86	0.4169	0.0866	0.3652	−1.6660	−1.3001
87	314	0.1647	0.0349	0.3531	−1.5845	−1.2302
87	708	0.0990	0.0191	0.3705	−1.5845	−1.2138

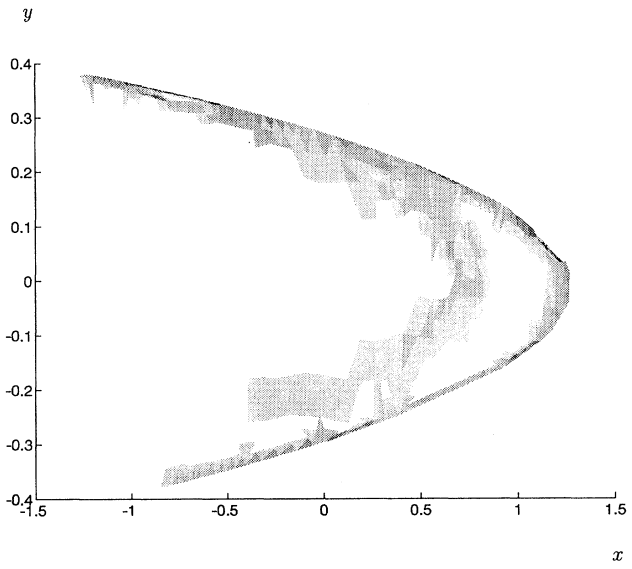

FIG. 7. Absolutely continuous approximation of the invariant measure using 708 states. Darker shades of gray represent regions of higher density.

ing the weight concentrated at the centroid of the triangle uniformly over the whole triangle. Here, darker areas represent higher density. In the context of a finite state Markov chain, the absolutely continuous measure is meaningless, as it has support on an infinite set of points. We show it here merely as an alternative representation of the approximation to the invariant measure, though the two measures are equivalent in the following sense. As previously mentioned, a convergent sequence of the singular measures tends to an invariant measure of the unperturbed system as the radii of the triangles go to zero. It may be easily shown that the sequence of absolutely continuous measures constructed from the sequence of singular measure converges to the same invariant measure. For all of the calculations, however, it is the singular measure that must be used.

VI. DISCUSSION

To our knowledge, Kim and Hsu [20] made the first attempt at practical estimation of Lyapunov exponents using a spatial average. This method had a number of shortcomings, however. It only produced an estimate of the largest exponent and even for this it required that the map be known. It also lacked theoretical justification for the computation of the appropriate directions to measure stretching in and had no results on convergence of the exponents to the true values. Similar work has been done by Boyarsky [21] in one dimension for piecewise monotonic expanding maps and unimodel maps with negative Schwarzian derivative. Such maps admit absolutely continuous invariant measures. Boyarsky approximates these maps with piecewise linear maps to obtain estimates of the absolutely continuous invariant measure and the single Lyapunov exponent. In one dimension, the problem of discontinuity of the single exponent does not arise, and it is true that C^1-close maps have nearby exponents, provided that the approximate map possesses an ergodic invariant measure close to the original.

We again make the point that as our techniques stand, they are not as easy to use, nor as quick or memory efficient as the traditional time average. However, it is hoped that they will form a basis for a methodology of exponent calculation in the future, one which has a firmer mathematical footing than the current accepted practice. The major stumbling block at the moment is a rigorous treatment of characterizing and approximating invariant measures of dynamical systems. The existence of "physical" measures has only been shown in special cases, and little is known about the relationship between the invariant measure of the *real* system and the long cycles produced by the *computer representation* of the system on a finite state computer. We hope to make a small amount of progress in this area, to show that our perturbed measure is indeed a good approximation of the "physical" measure of the system, whatever that may turn out to be.

[1] David Ruelle, Am. J. Math. **98**, 619 (1976).
[2] Peter Walters, *An Introduction to Ergodic Theory,* Graduate Texts in Mathematics Vol. 79 (Springer-Verlag, New York, 1982).
[3] V. I. Oseledec, Trans. Mosc. Math. **19**, 197 (1968).
[4] Giancarlo Benettin, Luigi Galgani, Antonio Giorgilli, and Jean-Marie Strelcyn, Meccanica **15**, 9 (1980).
[5] Alan Wolf, Jack B. Swift, Harry L. Swinney, and John A. Vastano, Physica D **16**, 285 (1985).
[6] Karlheinz Geist, Ulrich Parlitz, and Werner Lauterborn, Prog. Theor. Phys. **83**, 875 (1990).
[7] Jean-Marc Gambaudo and Charles Tresser, Phys. Lett. A **94**, 412 (1983).
[8] David Ruelle, IHES Publ. Math. **50**, 275 (1979).
[9] Michael Benedicks and Lai-Sang Young, Invent. Math. **112**, 541 (1993).
[10] Pawel Góra and Abraham Boyarsky, Comput. Math. Appl. **16**, 321 (1988).
[11] C. Beck and G. Roepstorff, Physica D **25**, 173 (1987).
[12] Phil Diamond, Peter Kloeden, and Alexei Pokrovskii, J. Nonlinear Sci. **4**, 59 (1994).
[13] Yuri Kifer, *Random Perturbations of Dynamical Systems,* Progress in Probability and Statistics, Vol. 16 (Birkhauser, Boston, 1988).
[14] Lai-Sang Young (private communication).
[15] Alistair I. Mees, Int. J. Bifurcation Chaos **1**, 777 (1991).
[16] Alistair I. Mees, Kenji Murao, Kevin Judd, and Gary Froyland (unpublished).
[17] Gary Froyland, Kevin Judd, and Alistair I. Mees (unpublished).

[18] Yves Guivarc'h, in *Probability Measures on Groups VII. Oberwolfach, 1983,* edited by H. Heyer (Springer-Verlag, Berlin, 1984) (French).
[19] *Lyapunov Exponents. Proceedings, Oberwolfach 1990,* edited by Ludwig Arnold, Hans Crauel, and Jean-Pierre Eckmann, Lecture Notes in Mathematics, Vol. 1486 (Springer-Verlag, Berlin, 1991).
[20] Myun C. Kim and C. S. Hsu, J. Stat. Phys. **45**, 49 (1986).
[21] Abraham Boyarsky, J. Stat. Phys. **50**, 213 (1988).