



ELSEVIER

Physica D 115 (1998) 1–18

PHYSICA D

Effective scaling regime for computing the correlation dimension from chaotic time series

Ying-Cheng Lai^{a,b,*}, David Lerner^{a,c,1}^a Department of Mathematics, The University of Kansas, Lawrence, KS 66045, USA^b Department of Physics and Astronomy, Kansas Institute for Theoretical and Computational Science, The University of Kansas, Lawrence, KS 66045, USA^c Kansas Institute for Theoretical and Computational Science, The University of Kansas, Lawrence, KS 66045, USA

Received 1 April 1997; received in revised form 17 September 1997; accepted 17 September 1997

Communicated by Y. Kuramoto

Abstract

In the analysis of chaotic time series, a standard technique is to reconstruct an image of the original dynamical system using delay coordinates. If the original dynamical system has an attractor, then the correlation dimension D_2 of its image in the reconstruction can be estimated using the Grassberger–Procaccia algorithm. The quality of the reconstruction can be probed by measuring the length of the linear scaling region used in this estimation. In this paper we show that the quality is constrained by both the embedding dimension m and, more importantly, by the delay time τ . For a given embedding dimension and a finite time series, there exists a maximum allowed delay time beyond which the size of the scaling region is no longer reliably discernible. We derive an upper bound for this maximum delay time. Numerical experiments on several model chaotic time series support the theoretical argument. They also clearly indicate the different roles played by the embedding dimension and the delay time in the reconstruction. As the embedding dimension is increased, it is necessary to reduce the delay time substantially to guarantee a reliable estimate of D_2 . Our results imply that it is the delay time itself, rather than the total observation time $(m - 1)\tau$, which plays the most critical role in the determination of the correlation dimension. Copyright © 1998 Elsevier Science B.V.

1. Introduction

In many physical, chemical, engineering and biological experiments, the observations consist of one or more time series. When the underlying dynamics is chaotic, such a time series appears random, and it is important to be able to correctly identify its deter-

ministic origin. One way to distinguish a chaotic time series from the output of a stochastic process is to establish the existence of a finite number of effective degrees of freedom. Estimating the dimensionality of the underlying chaotic process is thus quite important. In this regard, the correlation dimension D_2 , one of the most often considered dimensions characterizing the multifractal property of the underlying chaotic set, has been extensively investigated since Grassberger and Procaccia (GP) [1] first proposed a computationally efficient algorithm to estimate it from chaotic data. Specifically, one first reconstructs

* Corresponding author. Address: Department of Mathematics, The University of Kansas, Lawrence, KS 66045, USA. Fax: +1 913 864 5262; e-mail: lai@poincare.maths.ukans.edu.

¹ E-mail: lerner@poincare.math.ukans.edu.

an equivalent dynamical system using the method of delay-coordinates embedding [2,3]. One then estimates the correlation integral $C(\epsilon)$ for a range of values of the distance ϵ and plots $\log C(\epsilon)$ against $\log \epsilon$. By extracting the slope of this plot in an apparently linear scaling region at a series of increasing embedding dimensions, one obtains an estimate of the correlation dimension. (See Section 2 for precise definitions of delay-coordinate embedding, $C(\epsilon)$, and the fractal dimension spectrum.) We mention, however, that if one is only interested in distinguishing between deterministic chaos and random noise, alternative methods exist which do not require the computation of the correlation dimension [4,5].

There are two critical parameters involved in the construction of delay-coordinate embedding and the subsequent application of the GP algorithm: (1) the embedding dimension m , and (2) the time delay τ . Regarding the embedding dimension, there are general and rigorous results due to Takens [3] and Sauer et al. [6] (see Section 2.1), but if the task is just to compute the correlation dimension, Ding et al. [7] pointed out that the correct value of the correlation dimension can be obtained with m being the smallest integer larger than D_2 provided that the data are infinitely long and noiseless, although for finite and noisy time series, m needs to be larger than D_2 . For the choice of the proper delay time, however, one usually relies on empirical criteria. For a scalar chaotic time series $x(t)$, the basic idea is to choose τ such that the coordinates $x(t)$ and $x(t + \tau)$ are somewhat independent but not completely uncorrelated so that they can be treated as independent coordinates in some reconstructed phase space. If τ is too small, $x(t)$ and $x(t + \tau)$ are too correlated to serve as independent variables in the reconstructed vector space. If τ is too large, then for a chaotic time series, $x(t)$ and $x(t + \tau)$ are uncorrelated as completely random variables and, hence, the underlying deterministic dynamics may be lost. A very elaborate procedure based on the information content of the chaotic attractor was developed by Fraser and Swinney [8]. A straightforward and computationally simple criterion was suggested by Theiler [9]. Using this criterion, one computes the autocorrelation

function $\Psi(\tau)$ from the time series $x(t)$, $\Psi(\tau) = \langle [x(t + \tau) - \bar{x}][x(t) - \bar{x}] \rangle / \langle [x(t) - \bar{x}]^2 \rangle$, where \bar{x} is the averaged value of $x(t)$, and the average $\langle \cdot \rangle$ is with respect to time t . Theiler suggested to select τ such that $\Psi(\tau) \approx 1/e$. Methods based on the correlation integral were developed by Liebert and Schuster [10] and by Kemper and Fowler [11]. Buzug and Pfister [12] also considered obtaining optimal delay time and embedding dimension by analyzing both global and local behaviors of chaotic attractors. More recently, Rosenstein et al. [13] suggested a computationally efficient procedure based on reconstruction expansion for choosing the proper delay time. It should be noted that these empirical procedures all yield similar choice of proper delay times which work quite well in practical applications.

In this paper we show that the effective linear scaling region that can be used reliably for computing D_2 is very sensitive to the choice of time delay used in the reconstruction of the phase space. We develop inequalities relating the maximum allowed delay time to quantities such as the embedding dimension, the length of the time series, and some dynamical invariants of the underlying chaotic process. Our results show that, for a given embedding dimension m , there exists a maximum allowed delay time τ above which the computation of D_2 is no longer reliable. Our main results are Eq. (12), which can be used to estimate the size of the linear scaling region, and Eq. (13), which gives an upper bound for the allowed delay time. An important implication is that the delay time plays a critical role in the determination of the correlation dimension. As the embedding dimension is increased, it is necessary to reduce the delay time substantially to guarantee a reliable estimate of the correlation dimension. The quality of the delay-coordinate reconstruction, as measured by the length of the linear scaling region, deteriorates more rapidly with increasing τ (for fixed m) than it does for increasing m (and fixed τ). Thus, in computing D_2 , the roles played by the delay time and the embedding dimension are apparently not equally influential. This is contrary to some speculation in the literature that the combination $(m - 1)\tau \approx m\tau$, the so-called window size, is the most critical parameter in chaotic data analysis.

A brief account of part of the present work has been published in [14].

The paper is organized as follows: In Section 2, we review the fundamentals of the delay-coordinate embedding technique and the GP algorithm. In Section 3, we present numerical results with the Hénon map which indicate that the length of the linear scaling region in the computation of the correlation dimension changes dramatically as the delay time changes. In Section 4, we derive our main results, Eqs. (12) and (13). In Section 5, we check the validity of our theoretical results by using time series produced from both discrete maps and continuous flows. In Section 6, we discuss the effects of noise, the geometry of large time delays, and the role of τ , m , and the window size $m\tau$ with respect to the length of the linear scaling region. Concluding remarks are in Section 7.

2. Review of the Grassberger–Procaccia algorithm

2.1. Re-embedding a scalar time series

Suppose $\mathbf{z}(t)$ is a trajectory of the d -dimensional dynamical system

$$\frac{d\mathbf{z}}{dt} = \mathbf{f}(\mathbf{z}) \quad \text{with } \mathbf{z}(t_0) = \mathbf{z}_0. \quad (1)$$

Let $\Phi_t: \mathbb{R}^d \rightarrow \mathbb{R}^d$ be the flow of this system at time t , and let $g: \mathbb{R}^d \rightarrow \mathbb{R}$ be an observable. If the measuring process samples g every T units of time beginning at $t = t_0$, the observations corresponding to the trajectory constitute a scalar time series $\{x_n\}$ with

$$\begin{aligned} x_n &= x(t_0 + nT) = g(\mathbf{z}_n) \\ &= g(\mathbf{z}(t_0 + nT)) = g(\Phi_{nT}(\mathbf{z}_0)). \end{aligned} \quad (2)$$

Let us now fix T and simply write \mathbf{F} for Φ_T . This also allows us to consider dynamical systems described by discrete maps such as those arising on the Poincaré surface of section of continuous flows.

The problem, given the scalar time series, is to recover as much information as possible about the original system without explicit knowledge of the

functions g and \mathbf{F} . The first step is to reconstruct a dynamical system which is equivalent to the original. Generically, this can be done using delay coordinates, as was suggested by Packard et al. [2] and was shown rigorously by Takens [3]. Specifically, given the setup described above, for any positive integers m and τ , we can define a map $\mathbf{H}: \mathbb{R}^d \rightarrow \mathbb{R}^m$ by

$$\mathbf{x} = \mathbf{H}(\mathbf{z}) = (g(\mathbf{z}), g(\mathbf{F}^\tau(\mathbf{z})), \dots, g(\mathbf{F}^{(m-1)\tau}(\mathbf{z}))). \quad (3)$$

In terms of the scalar time series, this reads

$$\mathbf{x}_n = \mathbf{H}(\mathbf{z}_n) = (x_n, x_{n+\tau}, \dots, x_{n+(m-1)\tau}). \quad (4)$$

In what follows, we shall refer to τ as the delay time. For smooth dynamical systems, the true time lag is τT . The fundamental theorem by Takens [3] states that for almost all T and observables g , if the motion takes place in a compact region of phase space, and if $m \geq 2d + 1$, where d is the dimension of the manifold in which the attractor lies, then \mathbf{H} is one-to-one. Thus generically, \mathbf{H} embeds the original dynamical system in \mathbb{R}^m . So, if g and \mathbf{F} are differentiable, so is \mathbf{H} . This means that any differential or topological invariants of the original dynamical system can also be computed from the reconstructed system in \mathbb{R}^m .

Since one is typically interested in the geometry of an attractor whose dimension is less than d , and the computational effort needed for dimension calculations increases with the embedding dimension, considerable effort has been directed to minimizing the embedding dimension required. Recently, Sauer et al. [6] proved that an embedding of the attractor can be obtained if $m \geq 2D_0 + 1$, where D_0 is the box-counting dimension of the attractor. Another important result in this direction is due to Ding et al. [7]. They showed that if a strict embedding is not required (i.e., if lower-dimensional self-intersections of the attractor are permitted), as in the case of computing the correlation dimension D_2 , then in the limiting case of noise-free, infinite amount of data, D_2 can be obtained with the GP algorithm if m is any integer larger than D_2 . From a practical point of view, these theoretical results are somewhat tempered by finite sample sizes, sampling

rates which are far from ideal, numerical errors, and noise.

2.2. The correlation dimension D_2

In many dissipative dynamical systems, the asymptotic set is often a chaotic attractor. The dimensions of this attractor, which are often fractional, are topological invariants. An accurate estimate of the dimension of the attractor is a primary step towards understanding the underlying dynamical system.

An often computed dimension in chaotic time series analysis is the correlation dimension D_2 which is one of the infinite number of dimensions in the dimension spectrum that characterizes the multifractal structure of the chaotic attractor [15,16]. Grassberger and Procaccia [1] showed in 1983 that D_2 can be evaluated using the correlation integral $C(\epsilon)$, which is defined to be the probability that a pair of points chosen randomly with respect to the natural measure is separated by a distance less than ϵ on the attractor. For a trajectory of length N in the embedding space \mathbb{R}^m , the correlation integral can be approximated by the sum:

$$C_N(\epsilon) = \frac{2}{N(N-1)} \sum_{j=1}^N \sum_{i=j+1}^N \Theta(\epsilon - |\mathbf{x}_i - \mathbf{x}_j|), \quad (5)$$

where $\Theta(\cdot)$ is the Heaviside function given by $\Theta(x) = 1$ for $x \geq 0$ and 0 otherwise, and the norm is defined by

$$|\mathbf{x}| = \max\{|x_i|: 1 \leq i \leq m\}.$$

For N large, we have $C_N(\epsilon) \approx C(\epsilon)$. Grassberger and Procaccia argued that the correlation dimension D_2 is given by

$$D_2 = \lim_{\epsilon \rightarrow 0} \lim_{N \rightarrow \infty} \frac{\log C_N(\epsilon)}{\log \epsilon}. \quad (6)$$

In practice, for a time series of finite length, the sum in Eq. (5) also depends on the delay time τ and the embedding dimension m . Thus, for the rest of the paper we shall denote the correlation sum by $C_N(\epsilon, \tau, m)$. Due to such dependencies, the correlation dimension D_2 is usually estimated by examining the slope of

the linear portion of the plot of $\log C_N(\epsilon, \tau, m)$ versus $\log \epsilon$ for a series of increasing values of m . For $m < D_2$, the dimension of the reconstructed phase space is not high enough to resolve the structure of the attractor and, hence, in this case the slope approximates the embedding dimension. As m increases, the resolution of the attractor in the reconstructed phase space improves. Typically, the slope in the plot of $\log C_N(\epsilon, \tau, m)$ versus $\log \epsilon$ increases with m until it reaches a plateau; its value at the plateau is then taken as the estimate of D_2 [1,7,17,18]. For an infinite and noiseless time series, the value of m at which this plateau begins satisfies $m = \text{Ceil}(D_2)$, where $\text{Ceil}(D_2)$ is the smallest integer greater than or equal to D_2 [7]. On the other hand, short data sets and observational noise can cause the plateau onset to occur at a value of m which is considerably larger than $\text{Ceil}(D_2)$. Even so, the embedding dimension at which the plateau is reached still provides a reasonably sharp upper bound for the true correlation dimension D_2 . Systematic examination of the dependencies of the length of the linear scaling region on fundamental parameters τ , m , and $m\tau$ is the main goal of this paper.

3. Numerical results for the Hénon map

In this section, we present extensive numerical results on the behavior of the correlation sum $C_N(\epsilon, \tau, m)$ for the Hénon map [19],

$$\begin{aligned} x_{n+1} &= a - x_n^2 + by_n, \\ y_{n+1} &= x_n, \end{aligned} \quad (7)$$

with $a = 1.4$ and $b = 0.3$. For this map, $D_2 \approx 1.195$, so for a sufficiently long time series, the correct value of D_2 can be extracted using an embedding dimension $m = 2$ [7]. To generate a time series, we choose a random initial condition, iterate the map 2000 times to get rid of the initial transient, and then record the next 28000 values of x_n . To compare results at various embedding dimensions, we normalize the resulting time series via $\hat{x}_n = (x_n - x_{\min}) / (x_{\max} - x_{\min})$ so that $0 \leq \hat{x}_n \leq 1$, $\forall n$. With the norm $|\mathbf{x} - \mathbf{y}| = \max(|x_i - y_i|: 1 \leq i \leq m)$, for vectors \mathbf{x} and \mathbf{y} in the reconstructed phase space, the maximum distance

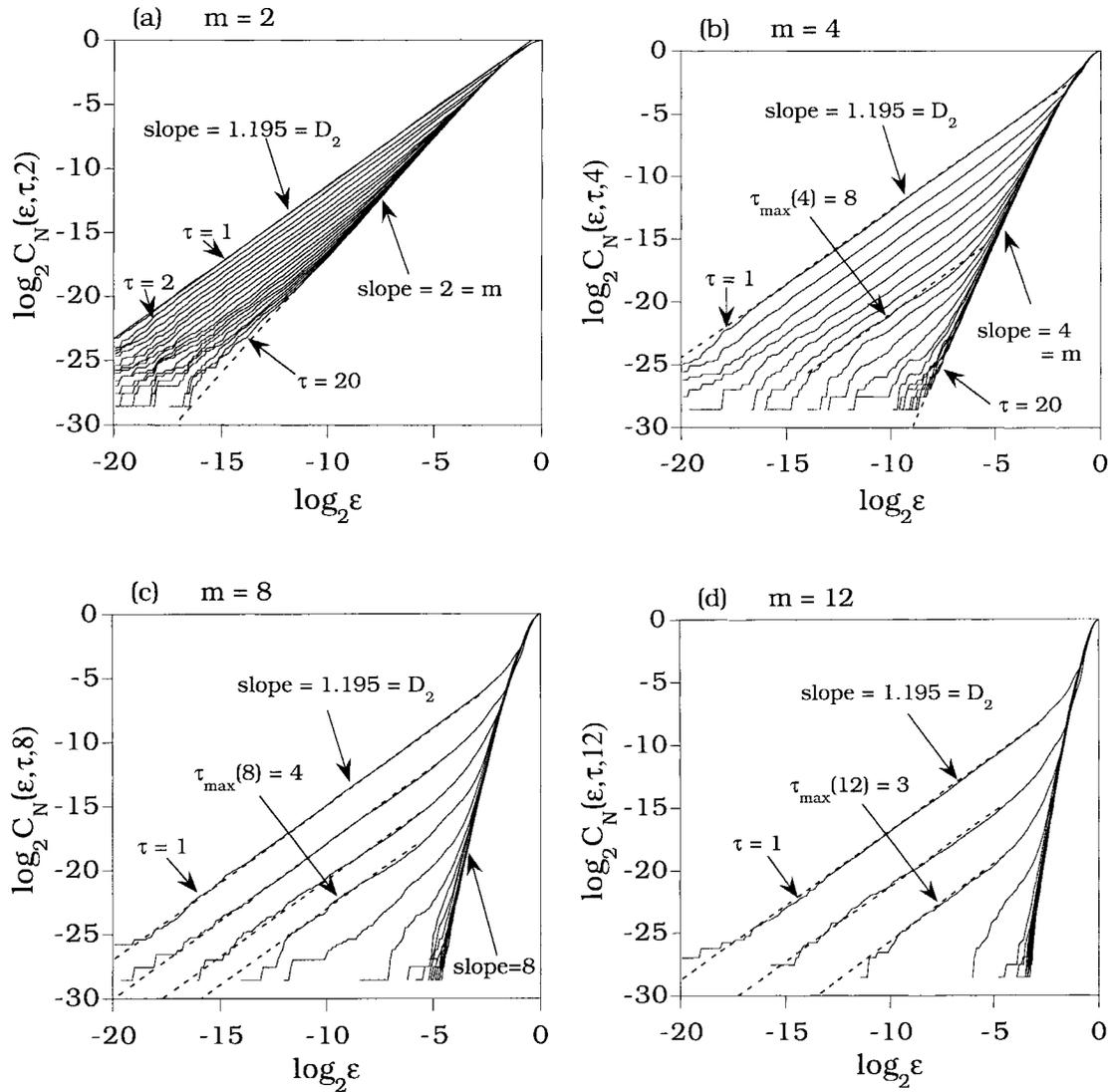


Fig. 1. For a time series of length $N = 28\,000$ from the Hénon map at $a = 1.4$ and $b = 0.3$, $\log_2 C_N(\epsilon, \tau, m)$ versus $\log_2 \epsilon$ for $m = 2$ (a), 4 (b), 8 (c), and 12 (d). In each figure, there are 20 plots corresponding to values of the delay time τ ranging from 1 to 20.

between any pair of points is 1. The correlation sum $C_N(\epsilon, \tau, m)$ is evaluated at 200 values of ϵ for $\log_2 \epsilon \in [-20, 0]$ and for embedding dimensions ranging from $m = 2$ to $m = 12$. For each embedding dimension, the computation is performed with a delay time ranging from $\tau = 1$ to $\tau = 20$.

Figs. 1(a)–(d) show $\log_2 C_N(\epsilon, \tau, m)$ versus $\log_2 \epsilon$ for $m = 2, 4, 8, 12$, respectively, where in each figure, the curves correspond to $\tau = 1$ to $\tau = 20$.

Figs. 2(a)–(d) show $\log_2 C_N(\epsilon, \tau, m)$ versus $\log_2 \epsilon$ for $\tau = 1, 2, 4$ and 8 , respectively, where the curves in each figure correspond to embedding dimensions from $m = 2$ to $m = 12$. From these figures, we observe the following:

1. For the range of embedding dimensions studied, the plot of the correlation integral at $\tau = 1$ yields the largest linear scaling region. The slope of the fitted straight line gives the correct value of D_2 .

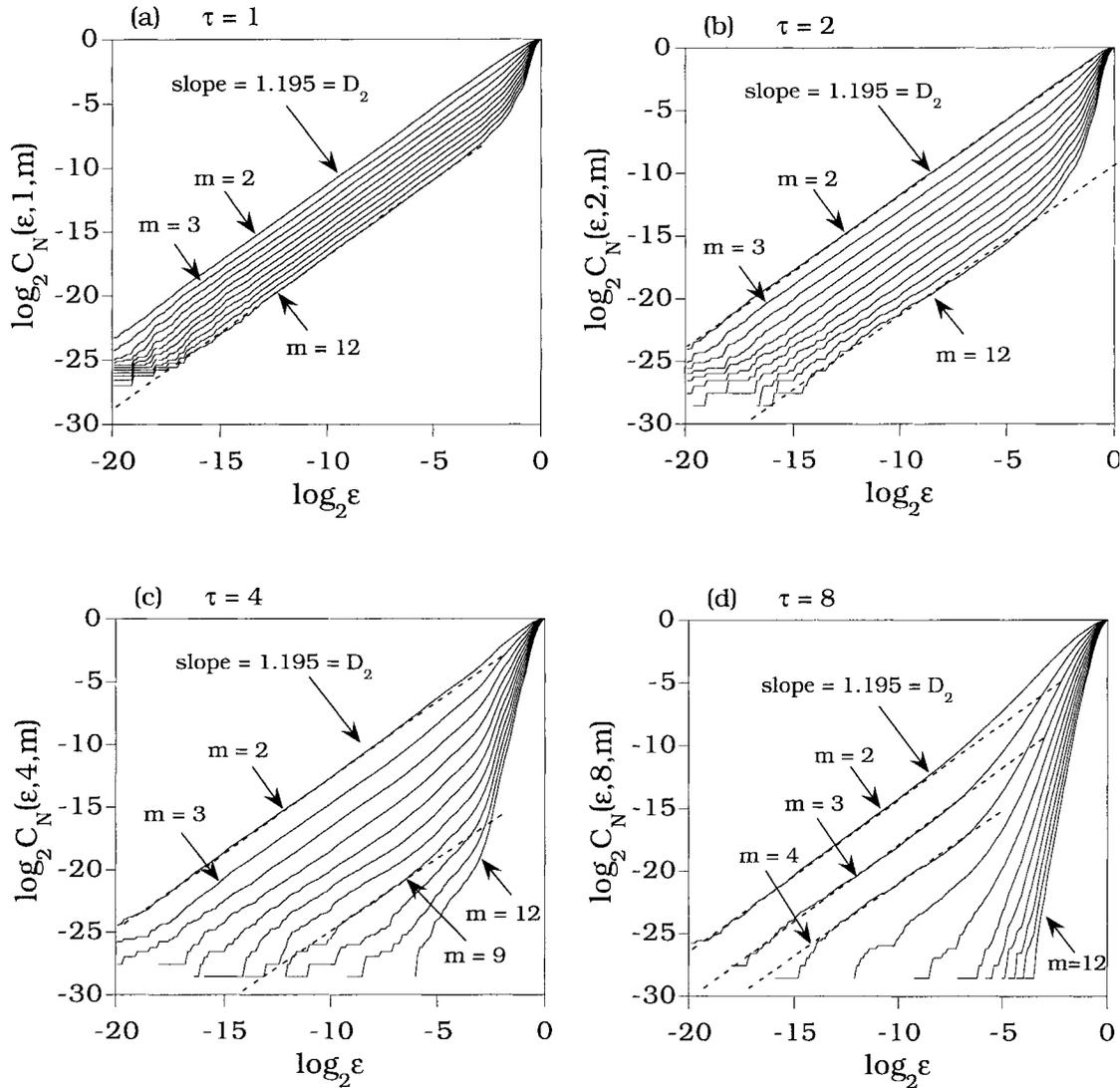


Fig. 2. For the same time series as in Fig. 1, $\log_2 C_N(\epsilon, \tau, m)$ versus $\log_2 \epsilon$ for $\tau = 1$ (a), 2 (b), 4 (c) and 8 (d). In each figure, there are 11 plots corresponding to embedding dimensions ranging from 2 to 12.

2. For a fixed embedding dimension, the linear scaling region decreases with increasing values of τ (Figs. 1(a)–(d)). For instance, when $m = 8$, there is no good scaling region if $\tau \geq 4$. This indicates that if τ is too large, the GP algorithm no longer yields a valid estimation of the correlation dimension D_2 .
3. For small values of τ , while there exists a linear scaling region leading to the correct value of D_2 ,

as ϵ increases there is a crossover from the good scaling region to a straight line whose slope is equal to the embedding dimension.

4. For a fixed delay time, the linear scaling region decreases as the embedding dimension m is increased (Figs. 2(a)–(d)). However, there is no apparent crossover. This can be best seen in Fig. 2(a), where $\tau = 1$, and the sizable linear scaling region appears to persist as m is increased from 1 to 12.

These observations suggest the following:

1. The range of the linear scaling region from which D_2 can be correctly extracted depends on both τ and m , more sensitively on τ , indicating that the roles of τ and m in determining the linear scaling region are different. Thus, utilizing the window size $w \equiv (m - 1)\tau$ to characterize the quality of the delay-coordinate embedding analysis is not sufficient.
2. As the embedding dimension m is increased, the delay time τ should be decreased to ensure a sizable linear scaling region.

In the sequel, we present an analysis to account for the above numerical observations.

4. Estimating the size of the linear scaling region

Our task is to establish conservative upper and lower bounds for the linear scaling region in terms of $\log \epsilon$. This will then permit us to fix an upper limit for the delay time τ at fixed embedding dimension m and fixed number of data points N .

Inside the scaling region, the plots of $\log C_N(\epsilon, \tau, m)$ against $\log \epsilon$ give a family of parallel lines with slope D_2 ; the vertical intercept varies with the precise values of m and τ . For simple dynamical systems which can be solved analytically, this can be seen from the explicit expression of $C_N(\epsilon, \tau, m)$. For instance, for the tent map ($D_2 = 1$), for ϵ near 0 we find (see Appendix A)

$$\log_2 C(\epsilon, \tau, m) = \log_2 \epsilon - m\tau + \tau + 1. \quad (8)$$

For more general dynamical systems, Grassberger and Procaccia [20] derived the following relation:

$$\log_2 C_N(\epsilon, \tau, m) \approx D_2 \log_2 \epsilon - m\tau T K_2 \log_2 e, \quad (9)$$

where K_2 is the order-2 entropy. Equivalently, from Eq. (9), if J is the number of distinct pairs of points on the attractor which are less than ϵ units apart, then in the scaling region, we have

$$J \approx \frac{1}{2} N^2 \epsilon^{D_2} e^{-m\tau T K_2}.$$

Imposing the modest requirement that J be > 1 throughout the linear scaling region, we obtain a

constraint which, when solved for ϵ , establishes the lower bound of the region as

$$\frac{1 - 2 \log_2 N + \tau m T K_2 \log_2 e}{D_2} < \log_2 \epsilon_{\min}(\tau, m). \quad (10)$$

Note that $\epsilon_{\min}(\tau, m)$ is an increasing function of both its arguments, and the values of τ and m affect $\epsilon_{\min}(\tau, m)$ equally.

Next, we consider $\epsilon_{\max}(\tau, m)$, which determines the upper end point of the linear scaling region. For large values of τ , the different delay coordinates are essentially uncorrelated. Consequently, at large distances ϵ , the behavior of $C_N(\epsilon, \tau, m)$ is similar to that obtained from a sequence of random vectors in the m -dimensional embedding space. Thus, we expect the family of curves to have a straight line with slope m as its envelope. This behavior is indeed observed in numerical computations (Figs. 1(a)–(d)). The line given by Eq. (9) intersects the line of slope m through the origin at a value of ϵ which satisfies

$$D_2 \log_2 \epsilon - m\tau T K_2 \log_2 e = m \log_2 \epsilon.$$

Taking this for the upper bound of the scaling region gives

$$\log_2 \epsilon_{\max} < \frac{-m\tau T K_2 \log_2 e}{m - D_2}. \quad (11)$$

Using Eqs. (10) and (11), and letting Δ represent the size of the linear scaling region, we find

$$\begin{aligned} \Delta &= \log_2 \epsilon_{\max}(\tau, m) - \log_2 \epsilon_{\min}(\tau, m) \\ &< -\frac{m K_2 \tau T \log_2 e}{m - D_2} \\ &\quad + \frac{2 \log_2 N - 1 - m\tau T K_2 \log_2 e}{D_2}. \end{aligned} \quad (12)$$

Note that from Eq. (12), the effects of m and τ on $\epsilon_{\max}(\tau, m)$ are apparently different due to the appearance of m in the denominator. Thus, although $\epsilon_{\min}(\tau, m)$ depends equally on m and τ , the size of the linear scaling region changes in a distinctly different manner as m or τ changes. This observation gives a qualitative explanation to Figs. 1 and 2. As a practical matter, in order to estimate D_2 , the linear scaling region must span at least an order of magnitude in ϵ . In

terms of the base-2 logarithmic scale, we shall require that $\Delta \geq 4$. This is a simple, easily checked test of the quality of the reconstruction using delay coordinates. Now we can solve Eq. (12) for τ to obtain an upper bound for the delay time:

$$\tau_{\max} < \frac{(2 \log_2 N - 1 - 4D_2)(m - D_2)}{m^2 K_2 T \log_2 e}. \quad (13)$$

Although our main interest lies in Eqs. (12) and (13), we can also solve (12) to obtain an upper bound for the embedding dimension m :

$$m_{\max} < \frac{\beta + \sqrt{\beta^2 + 16\alpha D_2}}{2\alpha}, \quad (14)$$

where

$$\alpha = K_2 \tau T \log_2(e),$$

and

$$\beta = 2 \log_2(N) - 1 + (\alpha - 4)D_2. \quad (15)$$

5. Numerical confirmation of the length of the linear scaling region and the upper bound for the time delay

To check the validity of Eqs. (12) and (13), we examine low-dimensional discrete maps and flows for which the quantities D_2 and K_2 can be obtained via other avenues. In particular, the correlation dimension D_2 can be calculated by using a box-counting procedure [15]. For dynamical systems described by one-dimensional or two-dimensional maps, the order-2 entropy is the positive Lyapunov exponent [20].

5.1. Maps

For the Hénon map at $a = 1.4$ and $b = 0.3$, $D_2 \approx 1.195$. Since there is just one positive Lyapunov exponent, we have $K_2 = \lambda_1 \approx 0.42$. For a time series of $N = 28\,000$ points, we obtain ($T = 1$),

$$\tau_{\max}(m) < 39.22 \left(\frac{m - D_2}{m^2} \right). \quad (16)$$

Selective results are shown in Figs. 1(a)–(d) and 2(a)–(d). Clearly, for large values of τ , delay coordinates

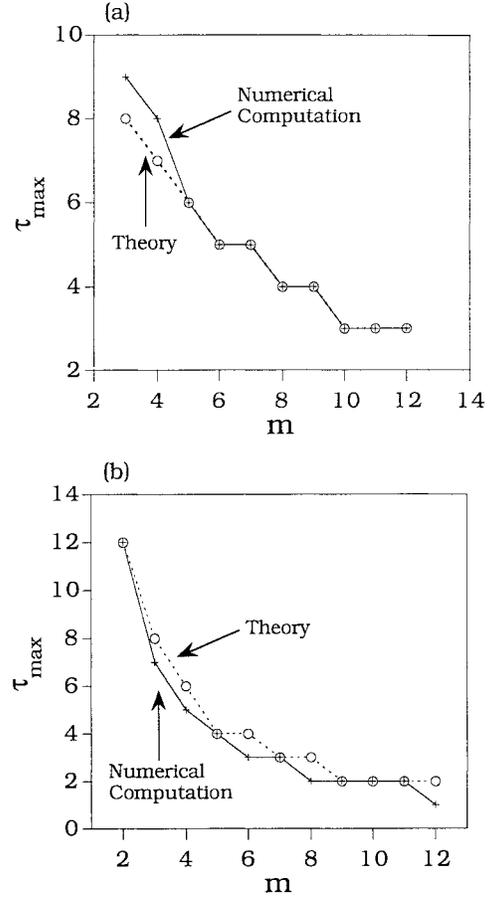


Fig. 3. Theoretical and numerical results for the upper bound of the delay time at various embedding dimensions (a) for the Hénon map, and (b) for the tent map.

are essentially independent of each other and, hence, we see that the curves $\log_2 C_N(\epsilon, \tau, m)$ versus $\log_2 \epsilon$ have a slope m for large values of ϵ . The curves follow this envelope for $\epsilon > \epsilon_{\max}(\tau, m)$ and then level off to obey the asymptotic law $\log_2 C_N(\epsilon, \tau, m) \sim D_2 \log_2 \epsilon - K_2 \tau T m \log_2 e$. For $m = 4$, we have $\tau_{\max}(4) \approx 7$ from Eq. (16), while numerical computation gives $\tau_{\max}(4) \approx 8$ (Fig. 1(b)). For $m = 8$, we have $\tau_{\max}(4) \approx 4$ from Eq. (16), and numerics gives $\tau_{\max}(4) \approx 4$ (Fig. 1(c)). Fig. 3(a) shows both the theoretical prediction for $\tau_{\max}(m)$ and the corresponding numerical results for embedding dimension ranging from $m = 3$ to $m = 12$. The agreement between Eq. (16) and the numerics is good.

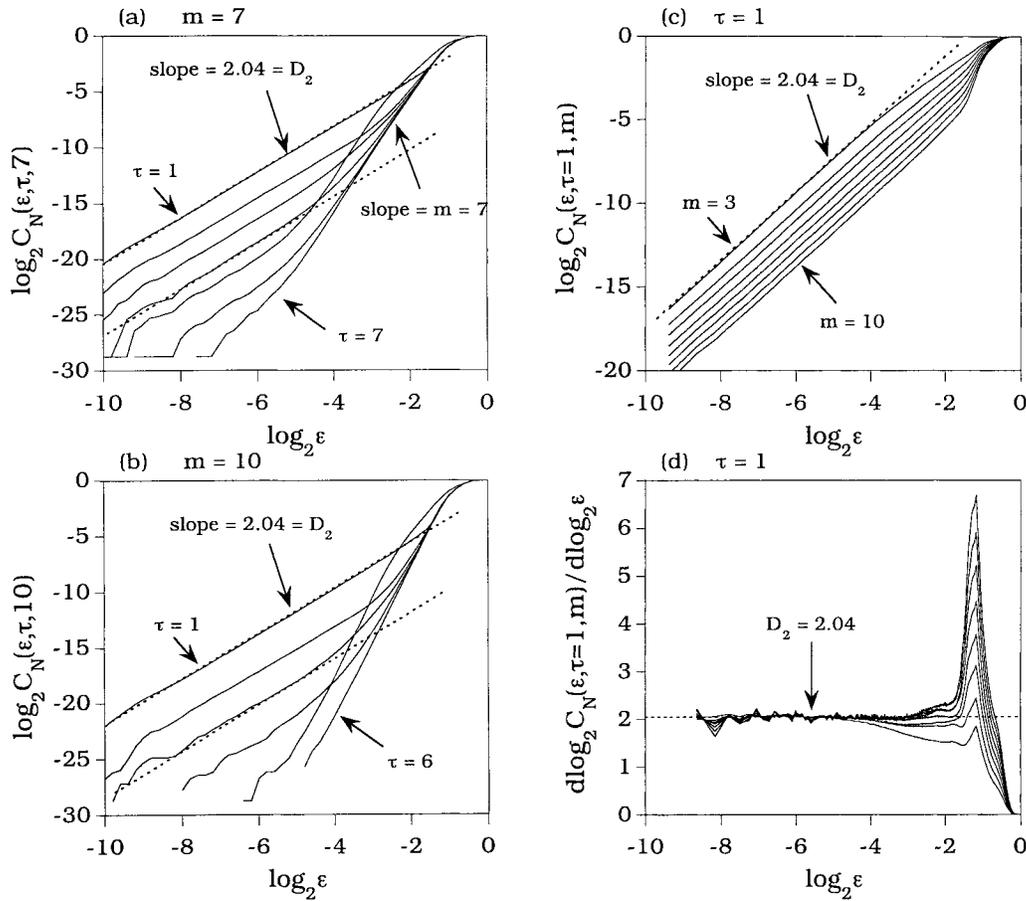


Fig. 4. For the x -measurement of the Lorenz equation (17), $\log_2 C_N(\epsilon, \tau, m)$ versus $\log_2 \epsilon$ for (a) $m = 7$ (fixed), $\tau = 1, \dots, 7$, (b) $m = 10$ (fixed), $\tau = 1, \dots, 6$, and (c) $\tau = 1$ (fixed), $m = 3, \dots, 10$. (d) Pointwise estimates of D_2 : $d \log_2 C_N(\epsilon, \tau, m) / d \log_2 \epsilon$ versus $\log_2 \epsilon$ for $\tau = 1$ (fixed), $m = 3, \dots, 10$.

Results with the tent map are summarized in Fig. 3(b), where the theoretical and numerical estimates of $\tau_{\max}(m)$ are shown for embedding dimensions ranging from $m = 2$ to $m = 12$. Again, a good agreement is obtained.

5.2. Flows

As an example of a continuous dynamical system, we consider the Lorenz equation [21]:

$$\begin{aligned} \frac{dx}{dt} &= -10(x - y), \\ \frac{dy}{dt} &= x(28 - z) - y, \end{aligned} \quad (17)$$

$$\frac{dz}{dt} = -\frac{8}{3}z + xy.$$

After discarding an initial transient of 5000 points, an orbit of 30 000 points on the attractor of (17) was computed numerically. The individual points on the orbit are separated in time by $T = 0.3$ (the time required for the autocorrelation function of the x -coordinate of the solution to (17) to decay to $1/e$ of its original value). Integration between times t and $t + T$ was done with a standard adaptive RK5 routine.

Figs. 4(a) and (b) show the plots of $\log_2 C_N(\epsilon, \tau, m)$ versus $\log_2 \epsilon$ for $m = 7$ ($\tau = 1, \dots, 7$) and $m = 10$ ($\tau = 1, \dots, 6$), respectively. We see that for fixed m , the size of the linear scaling regime decreases

as τ increases. Figs. 4(c) and (d) show for fixed $\tau = 1$, the plots of $\log_2 C_N(\epsilon, \tau, m)$ versus $\log_2 \epsilon$ and $d \log_2 C_N(\epsilon, \tau, m)/d\epsilon$ versus $\log_2 \epsilon$, respectively, for $m = 3, \dots, 10$. As in the case of maps, the length of the linear scaling region is apparently more sensitive to increase in τ than to increase in m (Figs. 4(a) and (b) versus Fig. 4(c)).

To test the goodness of our theoretical prediction, it is necessary to determine the relevant quantities in Eq. (13). The approximately constant value of $d \log_2 C_N(\epsilon, \tau, m)/d\epsilon$ in the sizable plateau region in Fig. 4(d) gives $D_2 \approx 2.04$, while we have $K_2 \approx 1.1$ estimated from $-1/\tau(d \ln C/dm)$ as suggested in [20]. Since $T = 0.3$, the actual time delay is 0.3τ , for some integer τ . In this case, Eq. (13) gives the results $\tau_{\max} \leq 4$ when $m = 7$ and $\tau_{\max} \leq 3$ when $m \geq 9$, once again in good agreement with the experimental results shown in Figs. 4(a) and (b).

Similar results were obtained for the Rössler system and for the forced damped pendulum.

6. Noise, geometrical interpretation of large time delays, window size, and the “knee” behavior

6.1. Effect of noise

When the time series is contaminated by noise, structures on the attractor with scales less than the noise amplitude ϵ_{noise} can no longer be resolved; for $\epsilon \leq \epsilon_{\text{noise}}$, slopes in the plots $\log_2 C_N(\epsilon, \tau, m)$ versus $\log_2 \epsilon$ keep increasing as the embedding dimension m is increased. The correct slopes (D_2) can be extracted only for $\epsilon > \epsilon_{\text{noise}}$. Figs. 5(a) and (b) illustrate this phenomenon for a noise contaminated time series of $N = 28\,000$ points from the Hénon map, which shows plots of $\log_2 C_N(\epsilon, \tau, m)$ versus $\log_2 \epsilon$ for $\tau = 1$ to 20 at $m = 3$ and $m = 6$, respectively. The time series was generated by iterating the Hénon map with an additive noise term $2^{-8}\sigma_n$, where σ_n is a random variable uniformly distributed in $[0, 1]$. After normalizing the time series to the unit interval, the noise amplitude becomes approximately $2^{-8}/4 = 2^{-10}$. Clearly, there is an abrupt change in each of the plots at $\log_2 \epsilon \approx -10$ for all τ values shown. The slopes of the plot for

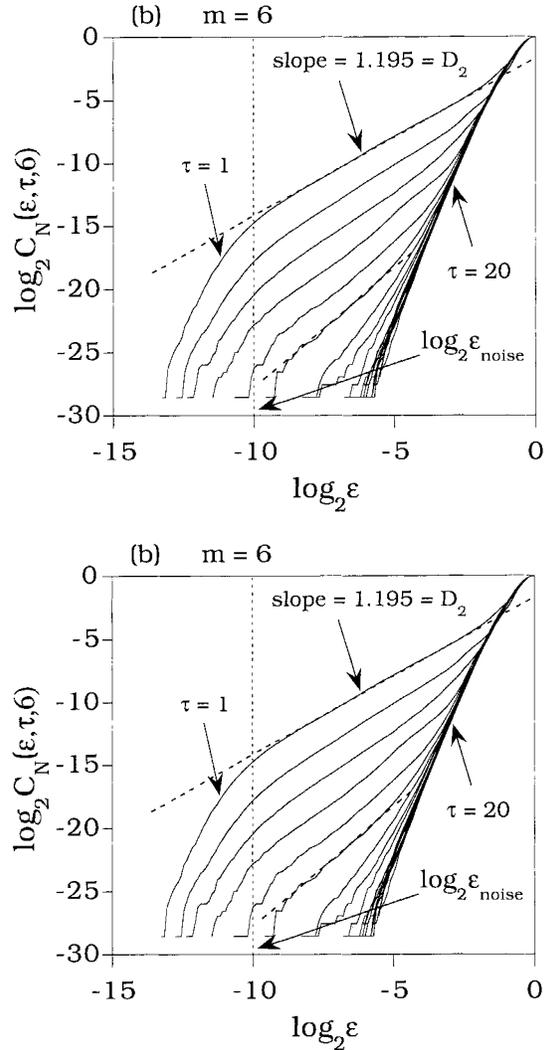


Fig. 5. For a noise contaminated time series of $N = 28\,000$ points from the Hénon map, the plots of $\log_2 C_N(\epsilon, \tau, m)$ versus $\log_2 \epsilon$ for $\tau = 1$ to 20 at $m = 3$. The noisy time series is generated by iterating the Hénon map with an additive noise term $2^{-8}\sigma_n$, where σ_n is a random number uniformly distributed in the $[0, 1]$. After normalizing, the noise amplitude is approximately 2^{-10} . The effect of noise induces a crossover from random to deterministic behavior at $\log_2 \epsilon \approx -10$.

$\log_2 \epsilon > -10$ are still approximately the value of D_2 , in so far as τ is small enough.

The presence of noise can change the effective scaling region in the computation of D_2 . For given τ and m , if $\epsilon_{\text{noise}} \leq \epsilon_{\min}(\tau, m)$, the scaling region, and hence Eqs. (12) and (13) are unaffected. If, however, $\epsilon_{\text{noise}} >$

$\epsilon_{\min}(\tau, m)$, then the effective scaling region is given by

$$\Delta \lesssim \frac{-m\tau T K_2 \log_2 e}{m - D_2} - \log_2 \epsilon_{\text{noise}}. \quad (18)$$

Requiring $\Delta \gtrsim 4$, we obtain

$$\tau_{\max}(m) \lesssim \frac{[\log_2(1/\epsilon_{\text{noise}}) - 4](m - D_2)}{m K_2 T \log_2 e}. \quad (19)$$

Clearly, if $\epsilon_{\text{noise}} > \epsilon_{\max}(\tau, m)$, it is no longer possible to compute D_2 . In fact, a stronger condition holds: if

$$\log_2(\epsilon_{\text{noise}}) \gtrsim - \left[4 + \frac{m K_2 T \log_2 e}{(m - D_2)} \right]$$

(corresponding to $\tau_{\max} < 1$), then D_2 can no longer be reliably computed using the correlation integral.

6.2. Some geometrical considerations

The correlation integral is a spatial average; distances between pairs of points are computed without regard to their temporal ordering. It is known [22] that, in the case of flows, too short a delay time results in an underestimate of the correlation dimension due to the fact that pairs of points which are close in time dominate the integral for small values of ϵ . For this reason, it is important to choose τ greater than some τ_{\min} at which the temporal correlations become unimportant [8,23–25].

Here we wish to point out that there is a straightforward geometrical explanation for the rapid disappearance of the good scaling region for large τ as well. Consider the time-delay embeddings of the time series from a one-dimensional map $x_{n+1} = f(x_n)$ that possesses a chaotic attractor, for instance, so that $D_2 = 1$. Using values of $\tau \geq 1$ is equivalent to examining the τ th iterations of the map. Consider, say, $m = 2$. For large values of τ , the plot of $f^\tau(x_n)$ versus x_n is effectively two-dimensional due to the folds of the map [26]. Thus, when τ is large, one obtains the erroneous result $D_2 = 2$. This type of behavior is also observable for embedding continuous flows. Generically, successive increments of the delay time τ correspond geometrically to successive iterations of a fixed map σ which depends only on the initial time delay.

In other words, the lack of correlation observed between the variables x_n and $x_{n+k\tau}$ is due not only to the fact that the data correspond to points that are far apart on the original attractor. It is also, and primarily, due to the fact that the extrinsic geometry of the embedding map itself is a function of τ , and becomes considerably more complicated with increasing τ .

The above behavior can be summarized more formally by factoring the map \mathbf{H} of Eq. (3) through an intermediate copy of the original manifold on which the dynamical system lives: Let F be a differentiable dynamical system on \mathbb{R}^d or more generally on the d -dimensional manifold M . Let $g: M \rightarrow \mathbb{R}$ be an observable, and suppose we embed M into \mathbb{R}^m in the manner previously discussed. Define

$$\begin{aligned} \sigma : \underbrace{M \times M \times \dots \times M}_{m \text{ times}} &\longrightarrow M^m \quad \text{by} \\ \sigma(\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_m) &= (\mathbf{z}_1, F(\mathbf{z}_2), \dots, F^{m-1}(\mathbf{z}_m)). \end{aligned} \quad (20)$$

Then $\sigma^\tau(\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_m) = (\mathbf{z}_1, F^\tau(\mathbf{z}_2), \dots, F^{\tau(m-1)}(\mathbf{z}_m))$. We can embed M onto the diagonal of M^m by sending $\mathbf{z} \in M \hookrightarrow \iota(\mathbf{z}) = (\mathbf{z}, \mathbf{z}, \dots, \mathbf{z})$. Note that $\sigma \circ \iota$ (and hence $\sigma^\tau \circ \iota$) is a diffeomorphism even if F is not. The observable g induces a map $\hat{g}: M^m \rightarrow \mathbb{R}^m$ via

$$\hat{g}(\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_m) = (g(\mathbf{z}_1), g(\mathbf{z}_2), \dots, g(\mathbf{z}_m)).$$

We can now express the map \mathbf{H} of Eq. (3) as the composition of the three separate maps

$$\begin{aligned} \mathbf{H}_\tau(\mathbf{z}) &= (\hat{g} \circ \sigma^\tau \circ \iota)(\mathbf{z}) \\ &= (g(\mathbf{z}), g(F^\tau(\mathbf{z})), \dots, g(F^{(m-1)\tau}(\mathbf{z}))). \end{aligned} \quad (21)$$

Recall that our scalar time series is assumed to arise from a sequence $\mathbf{z}_k = F^k(\mathbf{z}_0) \in M$, so the re-embedded trajectories are given by the sequences

$$\{\mathbf{H}_\tau(\mathbf{z}_k); k = 0, 1, 2, \dots\} \subset \mathbb{R}^m.$$

We thus see that using the delay time τ corresponds precisely to computing τ iterations of the map σ for each of these points. The maps ι and \hat{g} do not change with τ , so all the observed computational effects result from this iterative process. Although we do not

give a rigorous proof here, it is clear that the effect of these compositions is to increase the complexity of the embedding. For a fixed distance ϵ , increasing τ increases the probability that any two points lie on different folds of the attractor. It is this geometric effect which renders the computation of D_2 problematic at large τ . In particular, as τ increases, the value of ϵ_{\max} must also decrease to be less than half the average distance between the folds. This is quite different from the effect of increasing the embedding dimension while leaving τ fixed. Indeed, as is clear from Figs. 2(a)–(d) and 4(c), the linear scaling region near $\log_2 \epsilon_{\max}$ is only slightly affected when m is increased while τ is held fixed at small values.

6.3. The relation between τ and window size w

For continuous dynamical systems, the quantity $w = (m - 1)\tau T$ is called the window size. Because it represents the total time interval over which the original trajectory must be observed to construct each delay vector, it is often considered to be a more fundamental quantity than either m or τ separately [22,23,25].

From the standpoint of accurately determining the correlation dimension however, it is not clear that the quality of the reconstructed dynamics will depend in a useful way on the window size. In fact, the scaling region in Eq. (12) can be written as

$$\Delta \lesssim \frac{2 \log_2(N) - 1 - w K_2 \log_2(\epsilon)}{D_2} - \frac{w K_2 \log_2(\epsilon)}{m - D_2}, \quad (22)$$

where $w \approx m\tau T$ has been used. For fixed window size, this has the form

$$\Delta \lesssim \frac{A}{D_2} - \frac{B}{m - D_2},$$

where A and B are positive constants. This implies that in order to increase Δ , one should increase the embedding dimension m . Since w is fixed, this means that the delay time τ should be decreased. This expectation is borne out by Fig. 6, in which log–log plots of correlation integrals versus distance are shown for the Lorenz x -coordinate, using a fixed window size

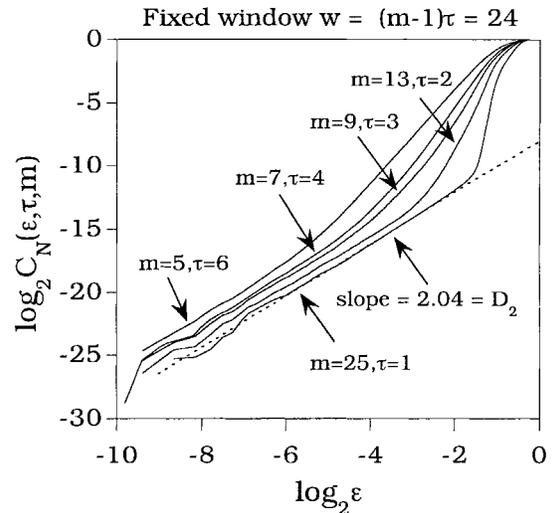


Fig. 6. Correlation integrals for the x -measurement of the Lorenz flow Eq. (17) at fixed window $w = (m - 1)\tau = 24$ but different combinations of τ and m . Clearly, for fixed w , larger m and smaller τ yield larger linear scaling region.

of $w = 24$ for different admissible values of m and τ (five combinations of m and τ). The largest linear scaling region clearly occurs for $m = 25$ and $\tau = 1$, while for $m = 5$ and $\tau = 6$, there is no apparent linear scaling region. These considerations lend some support to the assertion that one should always choose τ as close as possible to τ_{\min} (e.g., as the *first* minimum or zero of the autocorrelation function [23] or the first minimum of the mutual information [8]). Although our reasoning is somewhat different, this is in agreement with the conclusions of Gibson et. al. [25].

6.4. “Knee” due to large time delay

It is known that time series either from stochastic process or from deterministic chaotic process may exhibit anomalous structures in the correlation integral plots [27]. One example is the occurrence of a “knee” where the plot of the correlation integral on a logarithmic scale exhibits distinct linear regions with different slopes. Theiler [27] showed that in certain cases, this anomalous behavior is caused by a high degree of correlation among nearby data points in the time series, and he proposed to ignore the contribution to the correlation integral from nearby points in the embedding

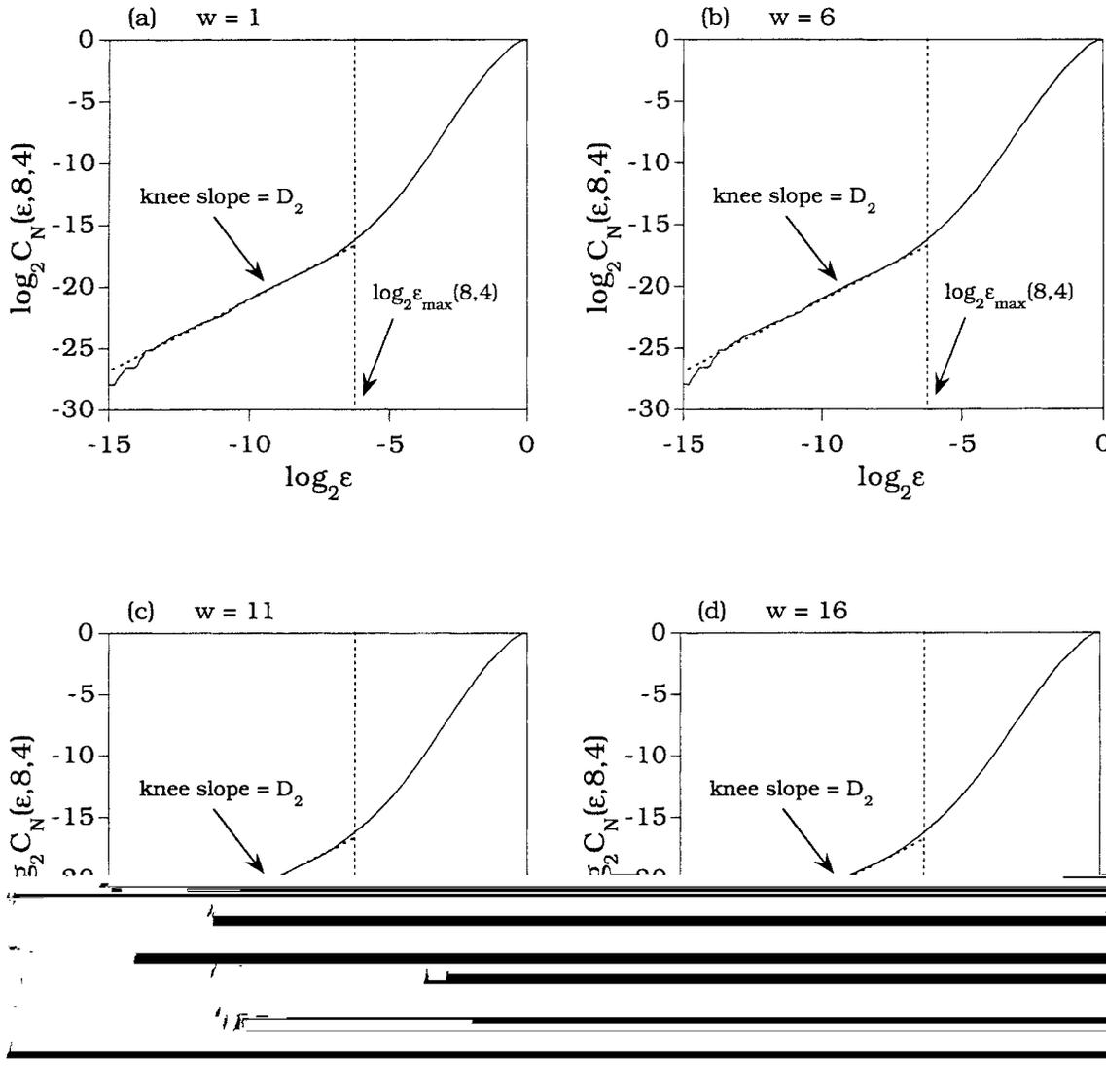


Fig. 7. For a time series length $N = 40000$ from the Hénon map, $\log_2 C_N(\epsilon, \tau = 8, m = 4, W)$ versus $\log_2 \epsilon$ for $W = 1$ (a), 6 (b), 11 (c) and 16 (d). The relatively large delay time $\tau = 8$ used induces the “knee” behavior in the plot (a). Note that using $W > \tau$ in Eq. (23) does not eliminate the observed “knee” behavior.

space to remedy this behavior. Specifically, Theiler suggested the following modified correlation integral in the GP algorithm:

$$C_N(\epsilon, \tau, m, W) = \frac{2}{N^2} \sum_{j=W}^N \sum_{i=1}^{N-j} \Theta(\epsilon - |\mathbf{x}_i - \mathbf{x}_j|), \quad (23)$$

where $W > 1$ is an integer. It was demonstrated that if $W > \tau$, then the “knee” behavior in the correlation integral can be eliminated and the resulting integral would yield the correct value of D_2 [27]. Here, we point out that large time delays may also cause the occurrence of “knee” in the correlation integral. One case is illustrated in Fig. 7(a), where

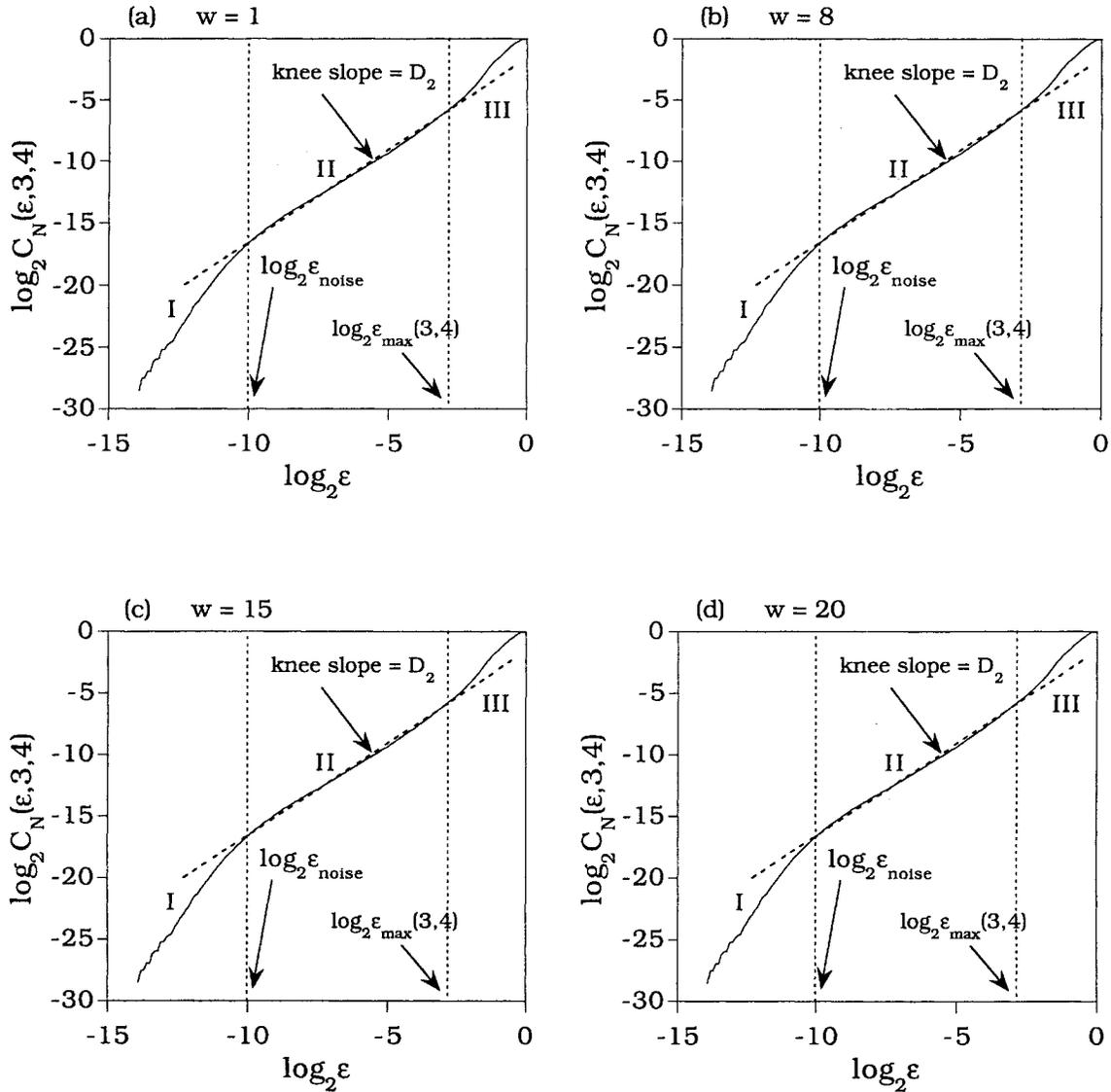


Fig. 8. For the same noise contaminated time series as in Fig. 7, $\log_2 C_N(\epsilon, \tau = 3, m = 4, W)$ versus $\log_2 \epsilon$ for $W = 1$ (a), 8 (b), 15 (c) and 20 (d). There are now three distinct regions in the plot, as described in the text. The slope in region II, the “knee” part, gives a good approximation to the correlation dimension D_2 .

$\log_2 C_N(\epsilon, \tau = 8, m = 4)$ versus $\log_2 \epsilon$ is plotted for a noiseless time series of $N = 40\,000$ points from the Hénon map. The “knee” behavior is not eliminated by using $W > 1$ in Eq. (23), as shown in Figs. 7(b)–(d) for $W = 6, 11,$ and 16 , respectively, where we notice that $W > \tau$ in Figs. 7(c) and (d). The reason is that the occurrence of this type of “knee” is mainly due to the large delay time used, or equivalently, is

due to *loss* of the correlation between components of points in the embedding space, in contrast to the “knee” behavior discussed by Theiler [27] which is due to the excessive correlation between nearby points in the time series. The “knee” behavior due to loss of correlation is particularly severe in the presence of small random noise. Fig. 8(a) shows a plot of $\log_2 C_N(\epsilon, \tau, m)$ versus $\log_2 \epsilon$ for a time series

from the Hénon map, where $N = 28\,000$, $\tau = 3$, $m = 4$, and there is random noise with amplitude 2^{-10} mixed with chaotic data. It is clear that the plot can be divided roughly into three regions: $\log_2 \epsilon < -10$ (I), $-10 \leq \log_2 \epsilon < \log_2 \epsilon_{\max}(\tau, m)$ (II) and $\log_2 \epsilon > \log_2 \epsilon_{\max}(\tau, m)$ (III). Region I represents the influence of random noise, and region III reflects the effect of large delay time in a large ϵ range. Both curves in I and III fail to yield the correct D_2 . Only the curve in region II gives a good estimate of D_2 , as ϵ in region II is large than the noise amplitude, but ϵ is not too large so that the correlation integral is still a correct manifestation of the fractal properties of the chaotic attractor. Figs. 8(b)–(d) show the same correlation integral computed using $W = 8, 15, \text{ and } 20$, respectively. In contrast to the occurrence of “knee” due to autocorrelation of the time series, here the slope in the “knee” part gives a good approximation of the correlation dimension D_2 .

7. Conclusion

As the field of chaotic dynamics matures [28], it becomes important to establish and test theoretical relations between various dynamical quantities such as D_2 and λ_1 and those such as m and τ which are to some extent under the control of the observer. A well-known example is the result of Eckmann and Ruelle [29] who derived a quantitative requirement for the amount of the data in order to compute dynamical invariants such as the correlation dimension and the Lyapunov exponent. In this spirit, we have shown the existence of such a relationship for the size of the effective linear scaling region which incorporates an upper bound for the delay time.

In Eqs. (12) and (13) we give what are essentially first-order estimates of both the length of the effective scaling region and the maximal permitted delay time for fixed embedding dimension and length of the time series. These estimates are then shown to be consistent with the data from several low-dimensional model chaotic systems. The quantity Δ in (12) (or modified as in Section 5 in the presence of noise) is a reliable indicator of the quality of the reconstruction. We also

show that the extrinsic geometry of the re-embedded attractor plays a critical role in determining this quality. The increased folding observed with increasing τ is due to the repeated iteration of a fixed, nonlinear map. The good linear scaling region does not depend on the window length alone: for fixed values of w , with m large enough to guarantee an embedding, and τ large enough to avoid the problems of autocorrelation, the length of the scaling region is largest for the smallest admissible value of τ .

Acknowledgements

This work was supported by AFOSR, Air Force Materiel Command, USAF, under Grant no. F49620-96-1-0066, by NSF under Grant no. DMS-962659, and by the General Research Fund from the University of Kansas. We thank Robert Hayden for preparing the figures used in the appendix.

Appendix A. Evaluation of the correlation integral for the tent map

In this appendix, we analytically compute the correlation integral for the tent map for certain values of the embedding dimension m and delay time τ .

For $\tau = 1$ and $m = 1$, the correlation integral, $C(\epsilon, 1, 1)$, is just the probability of two elements in the time series to be within ϵ distance of one another, i.e., $C(\epsilon, 1, 1) = P(|p - q| < \epsilon)$. Since the tent map has a uniform probability distribution, this probability is equal to the area defined by $|p - q| < \epsilon$ in the unit square $0 \leq p \leq 1$ and $0 \leq q \leq 1$. We obtain, $C(\epsilon, 1, 1) = 2\epsilon - \epsilon^2$.

For $\tau = 1$ and $m = 2$, the correlation integral is the probability of two events occurring, i.e. $P(|p - q| < \epsilon \text{ and } |T(p) - T(q)| < \epsilon)$, where $T(p)$ is one iteration of the tent map. Ding et al. [7] explicitly calculated $C(\epsilon, 1, 2)$, the result is

$$\begin{aligned} C(\epsilon, 2, 1) &= C(\frac{1}{2}\epsilon, 1, 1) + R(\epsilon) \\ &= \epsilon - \frac{1}{4}\epsilon^2 + \begin{cases} \frac{1}{2}\epsilon^2, & 0 < \epsilon < \frac{2}{3}, \\ 3\epsilon - \frac{7}{4}\epsilon^2, & \frac{2}{3} < \epsilon < 1. \end{cases} \end{aligned} \quad (\text{A.1})$$

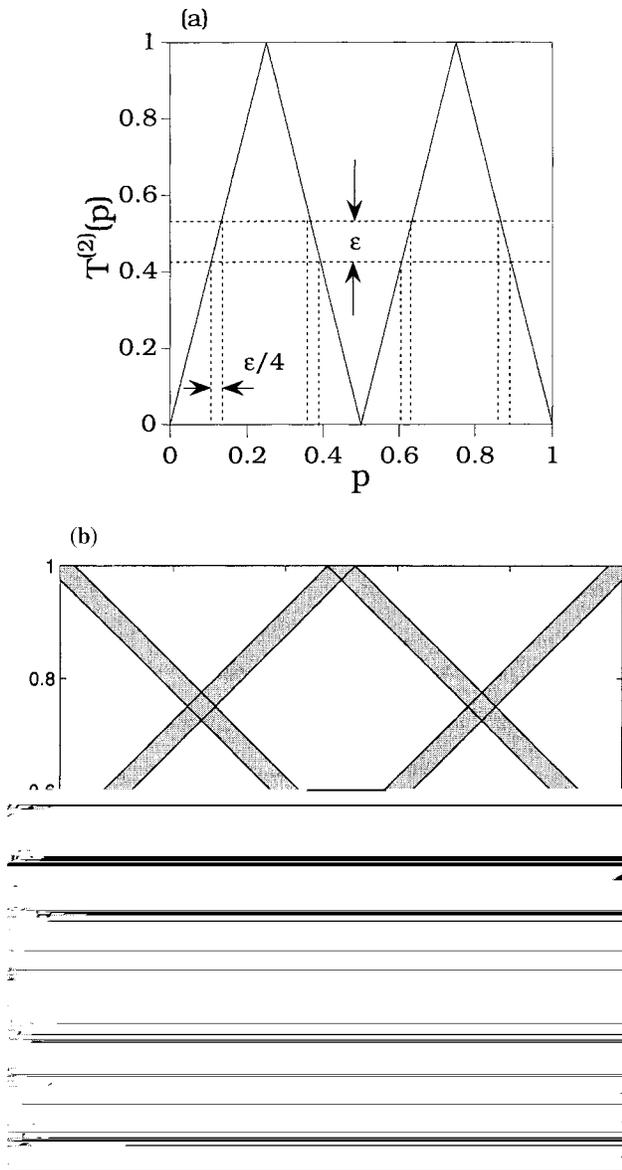


Fig. 9. (a) The second iterate of the tent map $T(x)$, and (b) regions in the unit square of (p, q) where $|T^2(p) - T^2(q)| < \epsilon$.

In Eq. (A.1), $R(\epsilon)$ is the correction term due to folding and is defined to be the probability of $\frac{1}{2}\epsilon < |p - q| < \epsilon$ and $|T(p) - T(q)| < \epsilon$.

For $\tau = 2$ and $m = 2$, the computation becomes more complicated. In this case the correlation integral is the probability that $|p - q| < \epsilon$ and $|T^{(2)}(p) - T^{(2)}(q)| < \epsilon$, where $T^{(2)}(p) = T(T(p))$ is the second

iteration of the tent map. Consider the graph of $T^{(2)}(p)$ versus p , as shown in Fig. 9(a). If the distance between p and q is less than $\frac{1}{4}\epsilon$, then the distance between $T^2(p)$ and $T^2(q)$ is less than ϵ . This is seen by the horizontal lines with width ϵ and the vertical lines with width $\frac{1}{4}\epsilon$ in Fig. 9(a). There are 16 possibilities for this to occur, depending on where p and q are. Each possibility is represented by a leg of one of the 4 X's, with each of horizontal width being equal to $2(\frac{1}{4}\epsilon)$, as shown in Fig. 9(b), where the shaded area represents the probability that $|T^{(2)}(p) - T^{(2)}(q)| < \epsilon$. Note that the probability that $|p - q| < \epsilon$ is the area enclosed by the dashed lines, as shown in Figs. 10(a)–(d). Therefore the correlation integral is equal to the intersection of the shaded areas in Fig. 9(b) with the area enclosed by the two dashed lines. Depending on the value of ϵ , there are four distinct cases, as shown in Figs. 10(a)–(d), respectively. The correlation integral, $C(\epsilon, 2, 2)$, is equal to the sum of the areas of the shaded regions. By calculating the shaded areas in Figs. 10(a)–(d), we obtain

$$C(\epsilon, 2, 2) = \begin{cases} \frac{1}{2}\epsilon + \frac{17}{16}\epsilon^2, & 0 \leq \epsilon < \frac{2}{5}, \\ -\frac{3}{4} + \frac{7}{2}\epsilon - \frac{7}{4}\epsilon^2, & \frac{2}{5} \leq \epsilon < \frac{2}{3}, \\ \frac{3}{2}\epsilon - \frac{7}{16}\epsilon^2, & \frac{2}{3} \leq \epsilon < \frac{4}{5}, \\ -1 + 4\epsilon - 2\epsilon^2, & \frac{4}{5} \leq \epsilon \leq 1. \end{cases} \quad (\text{A.2})$$

When ϵ is small, $C(\epsilon, 2, 2)$ scales as $\frac{1}{2}\epsilon$. This occurs because the intersecting area in Fig. 10(a) is primarily in the main diagonal strip for very small ϵ values. The area in the main diagonal strip is approximately $\frac{1}{2}\epsilon$.

From the above calculations, we see that the correlation integral becomes increasingly more difficult to compute analytically as τ increases. For example, for $\tau = 3$, there exist 64 cases for which the distance from $T^{(3)}(p)$ and $T^{(3)}(q)$ are guaranteed to be less than ϵ if $|p - q| < \frac{1}{8}\epsilon$. Nonetheless, for small ϵ values, the correlation integral can be approximated by the area along the main diagonal, as shown in Fig. 11. This area is $\frac{1}{4}\epsilon - \frac{1}{64}\epsilon^2$ for $\tau = 3$.

For higher embedding dimensions, it is necessary to consider the probability for m events to occur. Geometrically this involves the intersection of m cases.

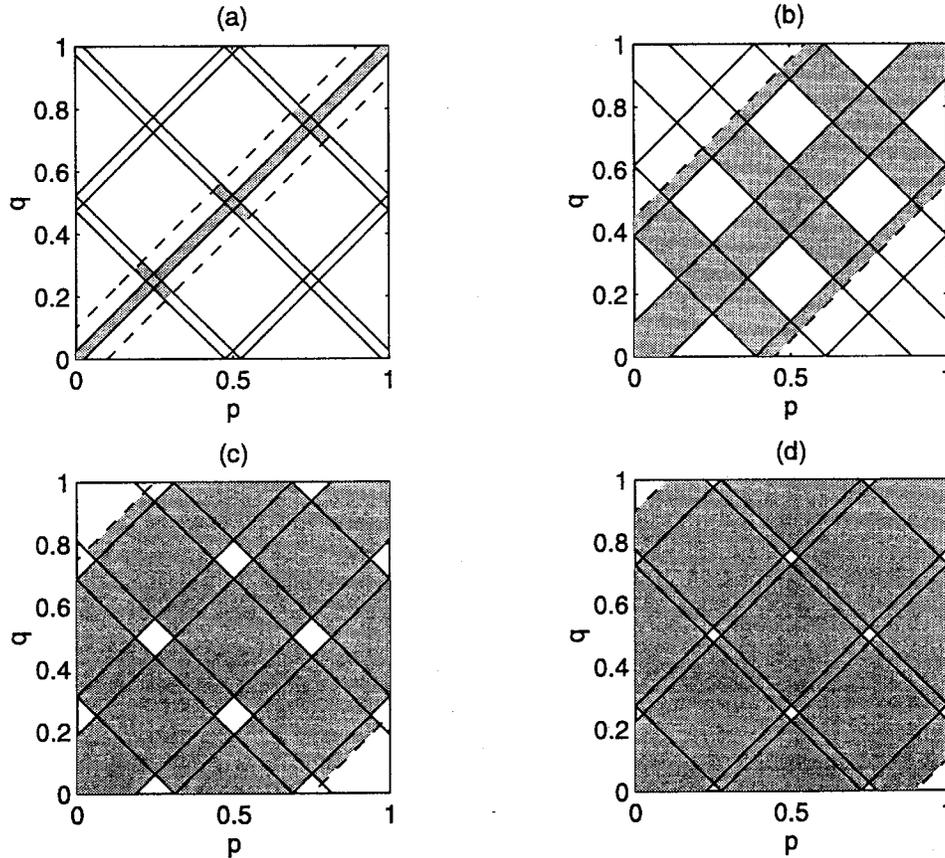


Fig. 10. Four different possibilities in the computation of the correlation integral $C(\epsilon, 2, 2)$ for the tent map: (a) $0 \leq \epsilon < \frac{2}{5}$, (b) $\frac{2}{5} \leq \epsilon < \frac{2}{3}$, (c) $\frac{2}{3} \leq \epsilon < \frac{4}{5}$ and (d) $\frac{4}{5} \leq \epsilon \leq 1$.

For example, when $m = 3$ and $\tau = 1$, the correlation integral is the probability that $|p - q| < \epsilon$, $|T(p) - T(q)| < \epsilon$, and $|T^{(2)}(p) - T^{(2)}(q)| < \epsilon$. Nonetheless, when ϵ is small, the correlation integral scales as $2(\epsilon/2^{(m-1)\tau})$ or $2^{1-(m-1)\tau}\epsilon$. This is due to the geometrical observation that the appropriate area equated to the correlation integral is essentially a strip, whose horizontal width is equal to $2(\epsilon/2^{(m-1)\tau})$, along the diagonal of the unit square. The diagonal strip is related to the window length, $(m - 1)\tau$, because it represents the approximate probability that

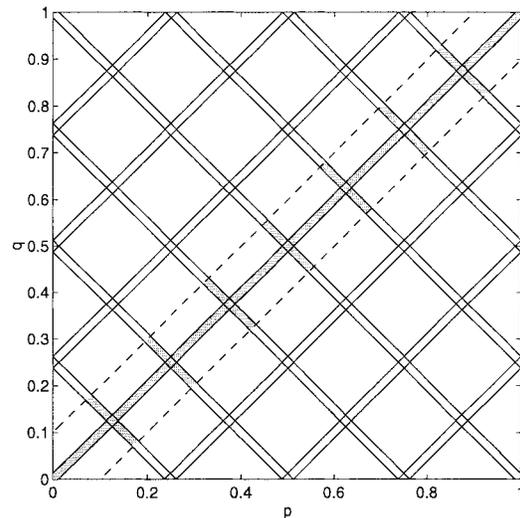


Fig. 11. Areas contributing to $C(\epsilon, 3, 2)$ for $\epsilon \ll 1$ (shaded regions). The diagonal area gives the main contribution to $C(\epsilon, 3, 2)$ which is first-order in ϵ . Other small shaded regions give a contribution which is on the order of ϵ^2 .

$|T^{(m-1)\tau}(p) - T^{(m-1)\tau}(q)| < \epsilon$. Thus, we have

$$C(\epsilon, \tau, m) \approx \frac{2\epsilon}{2^{(m-1)\tau}} \approx \frac{C(\epsilon, 1, 1)}{2^{(m-1)\tau}} \text{ for } \epsilon \ll 1. \quad (\text{A.3})$$

References

- [1] P. Grassberger, I. Procaccia, *Phys. Rev. Lett.* 50 (346) (1983); *Physica D* 9 (1983) 189.
- [2] N.H. Packard, J.P. Crutchfield, J.D. Farmer, R.S. Shaw, *Phys. Rev. Lett.* 45 (1980) 712.
- [3] F. Takens, in: D. Rand, L.S. Young (Eds.), *Dynamical Systems and Turbulence, Lecture Notes in Mathematics*, vol. 898, Springer, Berlin, 1981, p. 366.
- [4] D.T. Kaplan, L. Glass, *Phys. Rev. Lett.* 68 (1992) 427; M.B. Kennel, S. Isabelle, *Phys. Rev. A* 46 (1992) 3111; A.A. Tsonis, J.B. Elsner, *Nature* 358 (1992) 217; D. Kaplan, L. Glass, *Physica D* 64 (1993) 431; R. Wayland, S. Bromley, D. Pickett, A. Passamante, *Phys. Rev. Lett.* 70 (1993) 580; D.T. Kaplan, *Physica D* 73 (1994) 38; L.W. Salvina, R. Cawley, *Phys. Rev. Lett.* 73 (1994) 1091; J. Cao, Z. Zheng, *Phys. Rev. E* 49 (1994) 3907.
- [5] P. Grassberger, *Nature* 323 (1986) 609; J. Kurths, H. Herzel, *Physica D* 25 (1987) 165; J. Theiler, Ph.D. Thesis, Caltech, 1988; A.M. Fraser, *Physica D* 34 (1989) 391; D.T. Kaplan, R.J. Cohen, *Circulation Res.* 67 (1990) 886; S.J. Schiff, T. Chang, *Biol. Cyber.* 67 (1992) 387; J. Theiler, S. Eubank, A. Longtin, B. Galdrikian, J.D. Farmer, *Physica D* 58 (1992) 77; D. Prichard, J. Theiler, *Phys. Rev. Lett.* 73 (1994) 951.
- [6] T. Sauer, J.A. Yorke, M. Casdagli, *J. Stat. Phys.* 65 (1991) 579.
- [7] M. Ding, C. Grebogi, E. Ott, T. Sauer, J.A. Yorke, *Phys. Rev. Lett.* 70 (1993) 3872; *Physica D* 69 (1993) 404.
- [8] A.M. Fraser, H.L. Swinney, *Phys. Rev. A* 33 (1986) 1134; A.M. Fraser, *IEEE Trans. Inform. Theory* 35 (1989) 245.
- [9] J. Theiler, *J. Opt. Soc. Am. A* 7 (1990) 1055; *Phys. Rev. A* 41 (1990) 3038.
- [10] W. Liebert, H.G. Schuster, *Phys. Lett. A* 142 (1989) 107; W. Liebert, K. Pawelzik, H.G. Schuster, *Europhys. Lett.* 14 (1991) 521.
- [11] G. Kember, A.C. Fowler, *Phys. Lett. A* 179 (1993) 72.
- [12] T. Buzug, G. Pfister, *Phys. Rev. A* 45 (1992) 7073.
- [13] M.T. Rosenstein, J.J. Collins, C.J. De Luca, *Physica D* 65 (1993) 117; 73 (1994) 82.
- [14] Y.-C. Lai, D. Lerner, R. Hayden, *Phys. Lett. A* 218 (1996) 30.
- [15] P. Grassberger, *Phys. Lett. A* 97 (1983) 227; H.G.E. Hentschel, I. Procaccia, *Physica D* 8 (1983) 435; P. Grassberger, *Phys. Lett. A* 107 (1985) 101.
- [16] T.C. Halsey, M.J. Jensen, L.P. Kadanoff, I. Procaccia, B.I. Shraiman, *Phys. Rev. A* 33 (1986) 1141; C. Grebogi, E. Ott, J.A. Yorke, *Phys. Rev. A* 37 (1988) 1711.
- [17] G. Mayer-Kress (Ed.), *Dimensions and Entropies in Chaotic Systems*, Springer, Berlin, 1986; M.A.H. Nerenberg, C. Essex, *Phys. Rev. A* 42 (1990) 7065; J.C. Schouten, F. Takens, C.M. van den Bleek, *Phys. Rev. E* 50 (1994) 1851; K. Judd, *Physica D* 71 (1994) 421.
- [18] A. Ben-Mizrachi, I. Procaccia, P. Grassberger, *Phys. Rev. A* 29 (1984) 975; P. Grassberger, in: A.V. Holden *Chaos* (Ed.), Manchester University Press, Manchester, 1986; P. Grassberger, T. Schreiber, C. Schaffrath, *Int. J. Bifurc. Chaos* 1 (1991) 521.
- [19] M. Hénon, *Comm. Math. Phys.* 50 (1976) 69.
- [20] P. Grassberger, I. Procaccia, *Physica D* 13 (1984) 34.
- [21] E.N. Lorenz, *J. Atmos. Sci.* 20 (1963) 130.
- [22] A.M. Albano, J. Muench, C. Schwartz, A.I. Mees, P.E. Rapp, *Phys. Rev. A* 38 (1988) 3017; P.E. Rapp, A.M. Albano, T.I. Schmah, L.A. Farwell, *Phys. Rev. E* 47 (1993) 2289; D. Kugiumtzis, *Physica D* 95 (1996) 13.
- [23] M. Casdagli, S. Eubank, J.D. Farmer, J. Gibson, *Physica D* 51 (1991) 52.
- [24] A.R. Osborne, A. Provenzale, *Physica D* 35 (1989) 357; J.W. Havstad, C.L. Ehlers, *Phys. Rev. A* 39 (1989) 845.
- [25] J.F. Gibson, J.D. Farmer, M. Casdagli, S. Eubank, *Physica D* 57 (1992) 1.
- [26] H.-O. Peitgen, H. Jürgens, A. Saupe, *Chaos and Fractals, New Frontiers of Science*, chapter 12, Springer, New York, 1992.
- [27] J. Theiler, *Phys. Rev. A* 34 (1986) 2427.
- [28] H.D.I. Abarbanel, R. Brown, J.J. Sidorowich, L.S. Tsimring, *Rev. Mod. Phys.* 65 (1993) 1331; H.D.I. Abarbanel, *Analysis of Observed Chaotic Data*, Springer, New York, 1996.
- [29] J.-P. Eckmann, D. Ruelle, *Physica D* 56 (1992) 185.