Plateau Onset for Correlation Dimension: When Does it Occur?

Mingzhou Ding,[(1)] Celso Grebogi,[(2),(3)] Edward Ott,[(2),(4)] Tim Sauer,[(5)] and James A. Yorke[(3)]

[(1)]*Center for Complex Systems and Department of Mathematics, Florida Atlantic University, Boca Raton, Florida 33431*

[(2)]*Laboratory for Plasma Research, University of Maryland, College Park, Maryland 20742*

[(3)]*Department of Mathematics and Institute for Physical Science and Technology, University of Maryland, College Park, Maryland 20742*

[(4)]*Departments of Physics and of Electrical Engineering, University of Maryland, College Park, Maryland 20742*

[(5)]*Department of Mathematical Sciences, George Mason University, Fairfax, Virginia 22030*

(Received 22 February 1993)

Chaotic experimental systems are often investigated using delay coordinates. Estimated values of the correlation dimension in delay coordinate space typically increase with the number of delays and eventually reach a plateau (on which the dimension estimate is relatively constant) whose value is commonly taken as an estimate of the correlation dimension D_2 of the underlying chaotic attractor. We report a rigorous result which implies that, for long enough data sets, the plateau begins when the number of delay coordinates first exceeds D_2. Numerical experiments are presented. We also discuss how lack of sufficient data can produce results that seem to be inconsistent with the theoretical prediction.

PACS numbers: 05.45.+b

The estimation of the correlation dimension [1] of a presumed chaotic time series has been widely used by scientists to assess the nature of a variety of experimental as well as model systems, ranging from simple circuits to chemical reactions to the human brain. It is also known that many factors, such as noise and a lack of data, can hinder the successful application of the dimension extraction algorithm. In this paper, we address two issues related to the understanding of these difficulties, namely, what happens in an ideal situation (i.e., long data string with low noise) and what could be expected when the data set is small. In particular, we focus on the character of the dependence of the estimated correlation dimension on the dimension of the delay coordinate reconstruction space.

Consider an n-dimensional dynamical system that exhibits a chaotic attractor. A correlation integral $C(\epsilon)$ [1] is defined to be the probability that a pair of points chosen randomly on the attractor with respect to the natural measure ρ is separated by a distance less than ϵ on the attractor. The correlation dimension D_2 [1] of the attractor is then defined as $D_2=\lim_{\epsilon\to 0}\log C(\epsilon)/\log\epsilon$. Assume that we measure and record a trajectory of finite duration L on the attractor at N equally spaced discrete times, $\{\mathbf{x}_i\}_{i=1}^N$, where $\mathbf{x}_i \in \mathbf{R}^n$. The correlation integral $C(\epsilon)$ is then approximated by

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

where $\Theta(x)=1$ for $x>0$ and $\Theta(x)=0$ for $x\le 0$. In the limit $L,N\to\infty$, $C(N,\epsilon)\to C(\epsilon)$.

The dynamical information of a chaotic experimental system is often contained in a time series, $\{y_i=y(t_i)\}_{i=1}^N$, obtained by measuring a single scalar function $y=h(\mathbf{x})$ where $\mathbf{x}\in\mathbf{R}^n$ is the original phase space variable. From $\{y_i\}_{i=1}^N$ one reconstructs an m-dimensional vector $\mathbf{y}_i$ using the delay coordinates [2,3]

$$\mathbf{y}_i=\{y(t_i),y(t_i-T),\ldots,y(t_i-(m-1)T)\}\ , \qquad (2)$$

where $T>0$ is the delay time and m is the dimension of the reconstruction space. We call the mapping from $\{\mathbf{x}_i\}$ in $\mathbf{R}^n$ to $\{\mathbf{y}_i\}$ in $\mathbf{R}^m$ the "delay coordinate map." Results in Ref. [4] show that, for typical $T>0$ and $m>2D_0$, this delay coordinate map is one to one. Here D_0 is the box-counting dimension of the original chaotic attractor.

Our main focus is to estimate correlation dimension from a time series using delay coordinates [Eq. (2)]. As a point of departure for subsequent discussions, we first report a theorem [5,6] which shows that, for estimating the correlation dimension, $m\ge D_2$ suffices. We emphasize that this result holds true irrespective of whether the delay coordinate map is one to one or not. This is contrary to the commonly accepted notion that an embedding (one to one and differentiable) is needed for dimension estimation, leading to the false surmise that m needs to be at least $2D_2+1$ to guarantee a correct dimension estimation (see [7] for further discussion).

Consider an n-dimensional map $G:\mathbf{R}^n\to\mathbf{R}^n$. Let A be an attractor of G in $\mathbf{R}^n$ with a natural probability measure ρ. For a function $h:\mathbf{R}^n\to\mathbf{R}$, define a delay coordinate map $F_h:\mathbf{R}^n\to\mathbf{R}^m$ as

$$F_h(\mathbf{x})=[h(\mathbf{x}),h(G^{-1}(x)),\ldots,h(G^{-(m-1)}(\mathbf{x}))]\ .$$

The projected image of the attractor A under F_h has an induced natural probability measure $F_h(\rho)$ in $\mathbf{R}^m$. Furthermore, assume that G has only a finite number of periodic points of period less than or equal to m in A. The following result then applies.

Theorem.—If $D_2(\rho)\le m$, then for almost every h, $D_2(F_h(\rho))=D_2(\rho)$.

The theorem says that the correlation dimension is preserved under the delay coordinate map with $m\ge D_2(\rho)$. Similar results hold for flows generated by ordinary differential equations. "Almost every" in the statement is understood in the sense of prevalence defined in Ref. [4]; roughly speaking, we can regard this "almost every" as meaning that the functions h that do not give

0031-9007/93/70(25)/3872(4)$06.00

the stated result are very scarce and are not expected to occur in practice. The above dimension preservation result also holds for almost all general projection maps meeting the condition in the theorem. To illustrate, consider a closed curve with a uniform measure in $\mathbf{R}^3$. The dimension of this curve is 1. The projected image of this curve onto the plane still has a dimension of 1 but is generally self-intersecting. Thus the map is not one to one but preserves dimension information. One can further project the image to $\mathbf{R}^1$ and obtain an interval, which again has a dimension of 1 but bears little resemblance to the original curve in $\mathbf{R}^3$.

In applications D_2 is commonly extracted from a time series as follows (see Refs. [8-11] for reviews). First, an m-dimensional trajectory is constructed using Eq. (2). Then, the correlation integral $C_m(N,\epsilon)$ is computed according to Eq. (1), where m indicates the dimensionality of the reconstruction space. From the curve $\log C_m(N,\epsilon)$ vs $\log\epsilon$ one then locates a linear scaling region for small ϵ and estimates the slope of the curve over the linear region. This slope, denoted $\bar{D}_2^{(m)}$, is then taken as an estimate of the correlation dimension $D_2^{(m)}$ of the projection of the attractor to the m-dimensional reconstruction space. If these estimates $\bar{D}_2^{(m)}$, plotted as a function of m, appear to reach a plateau for a range of large enough m values, then we denote the plateaued value $\bar{D}_2$ and take $\bar{D}_2$ an estimate of the true correlation dimension D_2 for the system. From the theorem it is clear that the onset of this plateau should ideally start at $m=\text{Ceil}(D_2)$, where $\text{Ceil}(D_2)$, standing for ceiling of D_2, denotes the smallest integer greater than or equal to D_2.

Our original interest in the current problem was motivated by published reports (see Refs. [12-17] for a sample) where $\bar{D}_2^{(m)}$ plateaus at m values that are considerably greater than $\bar{D}_2$. A particular concern is that, when this happens, what does it imply regarding the correctness of the assertion that $\bar{D}_2$ is an estimate of the true correlation dimension D_2 of the underlying chaotic process? In an attempt to answer this question we have obtained new results on the systematic behavior of the correlation integrals. Based on these behaviors we are able to explain how factors such as a lack of sufficient data can produce results, resembling those seen in the experimental reports cited above, which seem to be inconsistent with the theorem. Furthermore, we find that even in cases where the plateau onset of $\bar{D}_2^{(m)}$ occurs at m values considerably greater than $\text{Ceil}(\bar{D}_2)$, there are situations where the plateaued $\bar{D}_2$ is a good estimate of the true correlation dimension D_2. See Refs. [18-25] for other relevant works addressing the issue of short data sets and noise.

To study the numerical aspects of dimension estimation we use chaotic time series generated by the Mackey-Glass equation [26] $dy(t)/dt=ay(t-\tau)/\{1+[y(t-\tau)]^c\}-by(t)$, where we fix $a=0.2$, $b=0.1$, $c=10.0$, and $\tau=100.0$, and take as the initial condition $y(t)=0.5$ for $t\in[-\tau,0]$. The numerical integration of this equation is done by the following iterative scheme [1]:

$$y(t+\delta t)=\frac{2-b\delta t}{2+b\delta t}y(t)+\frac{\delta t}{2+b\delta t}\left\{\frac{ay(t-\tau)}{1+[y(t-\tau)]^{10}}+\frac{ay(t-\tau+\delta t)}{1+[y(t-\tau+\delta t)]^{10}}\right\}, \tag{3}$$

where δt is the integration step size. We choose $\delta t=0.1$. Equation (3) is then a 1000-dimensional map, which, aside from being an approximation to the original equation, is itself a dynamical system. The time series, generated with a sampling time $t_s=10.0$, are normalized to the unit interval so that the reconstructed attractor lies in the unit hypercube in the reconstruction space. The norm we use to calculate distances in Eq. (1) is the max-norm in which the distance between two points is the largest of all the component differences. To reconstruct the attractor, we follow Eq. (2) and take the delay time to be $T=20.0$ The dimension of the reconstruction space is varied from $m=2$ to $m=25$.

The first time series, used to illustrate the theorem, consists of 50000 points. For each reconstructed attractor at a given m we calculate the correlation integral $C_m(N,\epsilon)$ according to Eq. (1). In Fig. 1 we display $\log_2 C_m(N,\epsilon)$ vs $\log_2\epsilon$ for $m=2$-8, 11, 15, 19, 23. For each m we identify a scaling region for small ϵ and fit a straight line through the region. The open circles in Fig. 2 show the values of $\bar{D}_2^{(m)}$ so estimated as a function of m. For $m\leq 7$, $\bar{D}_2^{(m)}\approx m$. For $m\geq 8$, $\bar{D}_2^{(m)}$ plateaus at $\bar{D}_2$ which has a value of about 7.1. Identifying $\bar{D}_2$ with the true correlation dimension D_2 of the underlying attractor, this result is consistent with the prediction by the theorem that the onset of the plateau occurs at $m=\text{Ceil}(D_2)$.

The second time series, used to illustrate the effect

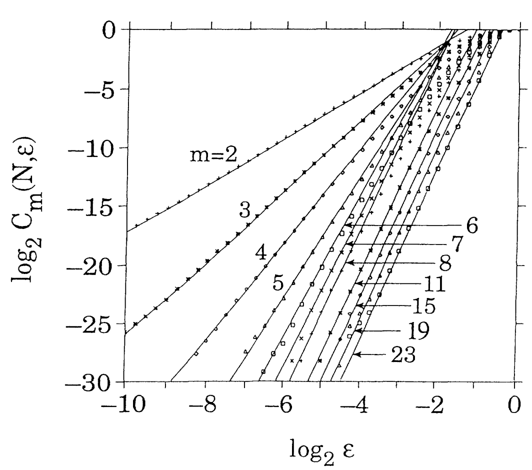

FIG. 1. Log-log plots of the correlation integrals $C_m(N,\epsilon)$ for the data set of 50000 points generated by Eq. (3).

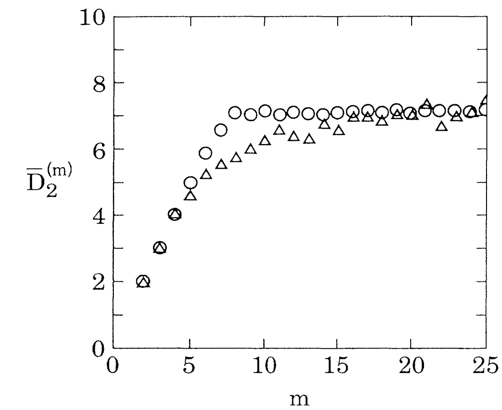

FIG. 2. $\bar{D}_2^{(m)}$ vs m plotted as open circles for the long data set ($N=50000$ and Fig. 1), $\bar{D}_2^{(m)}$ vs m plotted as triangles for the short data set ($N=2000$ and Fig. 3).

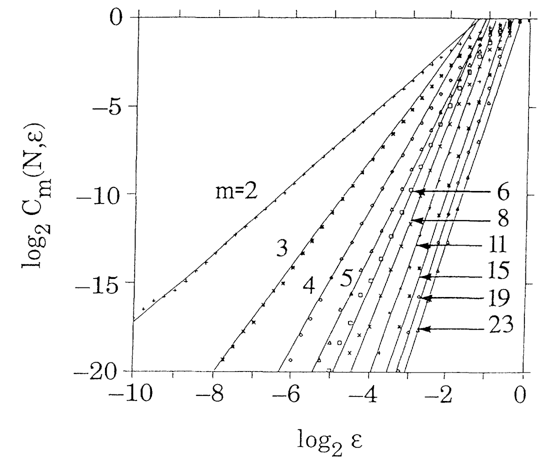

FIG. 3. $\log_2 C_m(N,\epsilon)$ vs $\log_2 \epsilon$ for the data set of 2000 points.

due to a lack of data, consists of 2000 points. The $\log_2 C_m(N,\epsilon)$ vs $\log_2\epsilon$ curves are shown in Fig. 3 for $m=2$–6, 8, 11,15, 19, 23. The values of $\bar{D}_2^{(m)}$ in this case are plotted using triangles as a function of m in Fig. 2. This function attains an approximate plateau which begins at $m=16$ and extends beyond $m=25$. The slope averaged over the plateau is about 7.05 which is consistent with the value of 7.1 obtained using the long data set ($N=50000$) plotted as open circles in Fig. 2. But the D_2 estimates for the short data set fall systematically under that for the long data set for $5\leq m\leq 13$. Thus the plateau does not begin until m is substantially larger than Ceil($\bar{D}_2$). This behavior has also been seen in many experimental studies. In what follows we explain the origin of this apparent inconsistency by exploring the systematic behavior of correlation integrals.

Figure 4 is a schematic diagram of a set of correlation integrals for $m=2$ to $m=13$. A dashed line is fit through the scaling region for each m. For $m\leq 5$, $\bar{D}_2^{(m)}\approx m$. For $m\geq 6$, $\bar{D}_2^{(m)}$ plateaus at $\bar{D}_2\approx 5.7$. This value is an estimate of the true D_2 for the system. This figure exhibits several features that are typical of correlation integrals for chaotic systems. The first feature we note is that the horizontal distance between $\log C_m(N,\epsilon)$ and $\log C_{m+1}(N,\epsilon)$ for $m\geq 6$ in the scaling regions is roughly a constant. This constant is predicted, for large m and small ϵ, to be $\Delta h=\Delta v/D_2$, where $\Delta v=TK_2$ with K_2 the correlation entropy [27] and T the delay time in Eq. (2). Two other significant features exhibited by Fig. 4 are as follows. For $m\leq 9$, $\log_2 C_m(N,\epsilon)$ increases with a gradually diminishing slope; while for $m\geq 11$, after exiting the linear region, the log-log plots in Fig. 4 first increase with a slope that is steeper than that in the linear scaling region and then level off to meet the point (0,0). These two different trends give rise to an uneven distribution in the extent of the scaling regions for different m with the most extended scaling region occurring at $m=10$.

In Refs. [8,28] arguments are presented to show that the trend observed for relatively small m is due to an "edge effect" resulting from the finite extent to the reconstructed attractor. Ding *et al.* [5] show that the steeper slope observed for relatively large m is caused by foldings occurring on the original attractor. This can be illustrated analytically [5] for the tent map [29]. For $m=1$, the correlation integral for the reconstructed tent map attractor is $C_1(\epsilon)=\epsilon(2-\epsilon)$. For $m=2$, $C_2(\epsilon)$ is written as $C_2(\epsilon)=C_1(\epsilon/2)+R(\epsilon)$. The first term arises because a pair of points y_j and y_l in the time series satisfying $|y_j-y_l|<\epsilon/2$ give rise to a pair of two-dimensional points, $\mathbf{y}_{j+1}=\{y_{j+1},y_j\}$ and $\mathbf{y}_{l+1}=\{y_{l+1},y_l\}$, satisfying $|\mathbf{y}_{j+1}-\mathbf{y}_{l+1}|<\epsilon$. The folding of the tent map at $y=\frac{1}{2}$ leads to situations in which $|y_j-y_l|>\epsilon/2$, but $|y_{j+1}-y_{l+1}|<\epsilon$. Thus the folding in the attractor underlies the correction term $R(\epsilon)$ which is calculated [5] to be $R(\epsilon)=\epsilon^2/2$ for $0\leq\epsilon<\frac{2}{3}$ and $R(\epsilon)=3\epsilon-7\epsilon^2/4-1$ for $\frac{2}{3}\leq\epsilon\leq 1$. For $0\leq\epsilon<\frac{2}{3}$, $d\log_2 C_2(\epsilon)/d\log_2\epsilon=1+\epsilon/(2+\epsilon)$. This derivative is 1 when $\epsilon=0$ ($D_2=1$ for the attractor) and increases due to the term $\epsilon/(2+\epsilon)$ whose presence reflects the influence of $R(\epsilon)$ which, in turn, is caused by the folding on the attractor.

From Eq. (1), the range of $C_m(N,\epsilon)$ is $\log_2 2/N^2$

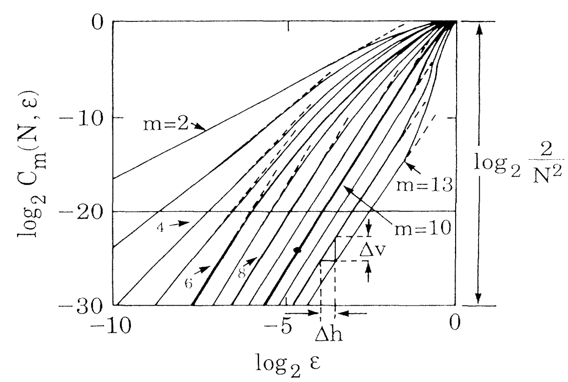

FIG. 4. Schematic diagram of correlation integrals.

$\leq \log_2 C_m(N,\epsilon) \leq 0$. Imagine a time series of $N=2000$ points generated by the system underlying Fig. 4. The plots of $\log_2 C_m(N,\epsilon)$ vs $\log_2\epsilon$ for this data set roughly correspond to the portion of Fig. 4 above the horizontal line drawn at $\log_2 C_m(N,\epsilon) = \log_2[2/(2000)^2] \approx -20$. Since the upper boundary points of the scaling regions for $m=6$ and 7 are under this horizontal line, the correct dimension is not obtained for $m=6$ and 7. In fact, if we fit a straight line to an *apparent* linear region above $\log_2 C_m(N,\epsilon) = -20$ for $m=6$ we obtain a slope which is markedly smaller than the actual dimension. However, since the upper boundary points of the scaling regions for $m \geq 8$ are above the horizontal line, we can still expect to obtain the correct estimate of $D_2 = 5.7$ for $m \geq 8$. Thus the plateau onset is delayed due to a lack of data.

The same consideration applies to the short data set generated by Eq. (3). In particular, imagine that we restrict our attention to the region $\log_2 C_m(N,\epsilon) > -20$ in Fig. 1 and fit a line through an *apparent* linear range for the $m=8$ data in this region. The slope of this straight line is about 5.9, which is roughly the same as that of 5.8 estimated using 2000 points. Thus, by knowing the correlation integrals for a large data set, we can roughly predict the outcome of a dimension measurement based on a smaller subset of this data.

We remark that if one extends the range of m values beyond what is shown in Fig. 2, at large enough m, $\bar{D}_2^{(m)}$ will start to deviate from the plateau behavior and increase monotonically with m. This is caused by the finite length of the data set and can be understood from the systematic behavior of correlation integrals seen in Fig. 4. A lack of sufficient data will not only delay the plateau onset, but also make the deviation from the plateau behavior occur at smaller values of m, thus shortening the plateau length from both sides. This can again be understood with reference to Fig. 4.

This work was supported by the U.S. DOE (Office of Scientific Computing, Office of Energy Research) and the NSF (Divisions of Mathematical and Physical Sciences). M.D.'s research is also supported in part by a grant from the National Institute of Mental Health.

---

[1] P. Grassberger and I. Procaccia, Phys. Rev. Lett. **50**, 346 (1983); Physica (Amsterdam) **9D**, 189 (1983).

[2] J.-P. Eckmann and D. Ruelle, Rev. Mod. Phys. **57**, 617 (1985).

[3] F. Takens, in *Dynamical Systems and Turbulence*, edited by D. Rand and L. S. Young (Springer-Verlag, Berlin, 1981), p. 366; R. Mañe, *ibid.*, p. 230.

[4] T. Sauer, J. A. Yorke, and M. Casdagli, J. Stat. Phys. **65**, 579 (1991).

[5] M. Ding, C. Grebogi, E. Ott, T. Sauer, and J. A. Yorke, Physica (Amsterdam) D (to be published).

[6] T. Sauer and J. A. Yorke (to be published).

[7] To get a rough idea of how widespread the misconception about the relevance of embedding to dimension estimation is, we conducted a literature search using the Science Citation Index for the years from 1987 to 1992 by looking for papers citing both Ref. [1] and Ref. [3]. We found 183 such papers. We then randomly selected a sample of 22 of these papers for closer examination. The following is what we found. Among the 22 papers there are 15 of them that calculate correlation dimension from time series. Five of these 15 papers make explicit connections between $2D+1$ and dimension estimation. The rest of these papers ignore this issue entirely. Based on this information we estimate that, during the period from 1987 to 1992, there are at least 42 papers (probably many more) where the authors implicitly or explicitly assumed that a one-to-one embedding is needed for dimension calculation. In addition, among the papers we researched for this work, only [8] and [9] imply that $m \geq D_2$ is sufficient for estimating D_2, although no justification is given.

[8] J. Theiler, J. Opt. Soc. Am. A **7**, 1055 (1990).

[9] P. Grassberger, in *Chaos*, edited by A. V. Holden (Manchester Univ. Press, Manchester, 1986).

[10] P. Grassberger, T. Schreiber, and C. Schaffrath, Int. J. Bifurcation Chaos **1**, 521 (1991).

[11] *Dimensions and Entropies in Chaotic Systems*, edited by G. Mayer-Kress (Springer-Verlag, Berlin, 1986).

[12] J. Holzfuss and G. Mayer-Kress, in *Dimensions and Entropies in Chaotic Systems* (Ref. [11]).

[13] A. Branstater and H. L. Swinney, Phys. Rev. A **35**, 2207 (1987).

[14] U. Hübner, W. Klische, N. B. Abraham, and C. O. Weiss, in *Measures of Complexity and Chaos*, edited by N. B. Abraham *et al.* (Plenum, New York, 1990).

[15] K. R. Sreenivasan, in *Dimensions and Entropies in Chaotic Systems* (Ref. [11]).

[16] A. A. Tsonis and J. B. Elsner, Nature (London) **333**, 545 (1988).

[17] D. V. Vassiliadis, A. S. Sharma, T. T. Eastman, and K. Papadopoulos, Geophys. Res. Lett. **17**, 1841 (1990).

[18] A. Ben-Mizrachi, I. Procaccia, and P. Grassberger, Phys. Rev. A **29**, 975 (1984).

[19] J. G. Caputo, in *Measures of Complexity and Chaos*, edited by N. B. Abraham *et al.* (Plenum, New York, 1989).

[20] J. G. Caputo, B. Malraison, and P. Atten, in *Dimensions and Entropies in Chaotic Systems* (Ref. [11]).

[21] J.-P. Eckmann and D. Ruelle, Physica (Amsterdam) **56D**, 185 (1992).

[22] N. A. Gershenfeld, Physica (Amsterdam) **55D**, 155 (1992).

[23] J. W. Havstad and C. L. Ehlers, Phys. Rev. A **39**, 845 (1989).

[24] M. Möller, W. Lange, F. Mischke, N. B. Abraham, and U. Hübner, Phys. Lett. A **138**, 176 (1989).

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

[26] M. C. Mackey and L. Glass, Science **197**, 287 (1977).

[27] P. Grassberger and I. Procaccia, Phys. Rev. A **28**, 2591 (1983); Physica (Amsterdam) **13D**, 34 (1984).

[28] L. A. Smith, Phys. Lett. A **133**, 283 (1988).

[29] The tent map is defined as $y_{n+1} = T(y_n)$, where $T(y) = 2y$ if $0 \leq y < 1/2$ and $T(y) = 2-2y$ if $1/2 \leq y \leq 1$.