State-space reconstruction using averaged scalar products of the dynamical system flow vectors

R. Huerta, C. Santa Cruz, J. R. Dorronsoro, and Vicente Lopez

Instituto de Ingenieria del Conocimiento, Universidad Autonoma de Madrid, Cantoblanco, C-XVI, 28049, Madrid, Spain

(Received 10 September 1993)

The averaged scalar product, P, of the dynamical system flow vectors evaluated along the trajectory is shown to be a very simple, efficient, and useful quantity for the purpose of selecting adequate embedding dimensions and time delays. The effectiveness of the method is shown in different examples, such as ordinary differential equations, delayed differential equations, and experimental time series.

PACS number(s): 05.45.+b, 02.40.−k

I. INTRODUCTION

In the last few years ideas from nonlinear dynamics have lead to improvement in the study of experimental time series. Often, the time evolution of complex dynamical systems can be approximated with low-dimensional nonlinear models. Those models can yield a better understanding of the system or can be the starting point for forecasting procedures. In this paper we focus our attention on deterministic dynamical systems from which only a univariate time series is available. This is a common situation which occurs in many situations in natural and social sciences.

A deterministic state-space reconstruction for the system is an embedding of the measured time series. We can say that an embedding is a smooth map from the original time series to a multidimensional space such that points and tangent vectors maintain their individuality. State space is normally used to represent a multidimensional space of the original time series constructed with any mapping. If the original dynamical system is deterministic, the embedding produces a particular state-space that is also deterministic. The auxiliary variables usually chosen to create state spaces are time delays of the measured variable [1–3]. Given the time series $x(t)$, the state-space reconstruction is given by $\mathbf{x}(t) = (x(t), x(t-T), ..., x(t-(d-1)T))$, where T is the time delay and d is the dimension. This state-space reconstruction may be embedded in a d dimension if $d > 2m$ as proposed by Takens [2], with m being the dimension of the manifold on which the original dynamical system evolves. Sauer *et al.* [4] showed that under certain conditions the state space may be embedded in $d > m$. An estimate of m can be obtained by means of the correlation dimension [5,6]; for a rigorous result see [7]. Therefore, each state-space reconstruction is characterized by the values $\{T, d\}$.

Proposed methods [8–11] to approximate the embedding with state spaces pay attention to state-space points and they do not take into account the flow vectors. Casdagli's method [12] comes directly from the forecast problem. In this work we are going to consider flow vectors rather than state-space points, as was proposed by Kaplan and Glass [13] for discrimination between random and deterministic time series. We base our approach on this idea.

The basic idea is very simple: an appropriate embedding should avoid self-crossings of the reconstruction space trajectories. An obvious measure of such a crossing at a given point is the average inner product of the flow vectors in a neighborhood of that point. It is obvious that crossing trajectories will have crossing flows; thus, since at self-crossing points normalized flow vectors will have a scalar product different from one, it is reasonable to use scalar products between normalized flow vectors to quantify the degree of self-crossing. In order to get a global quantification of the degree of self-crossing along a complete reconstructed trajectory, we average the above inner products along that trajectory. A state space where the global average is extremely close to one will be a good approximation to the deterministic state space. Lower values of this global average will indicate sizeable self-crossing and, therefore, poorer state-space reconstructions. An important point is that this measure of the average scalar product is performed after a transformation by singular value decomposition, and so the measure is made in similar conditions for any state-space reconstruction.

A search of maximal P has been performed in the $\{T, d\}$ space for a time series extracted from Rossler and Mackey-Glass. Very similar results have been obtained for the Lorentz attractor and the Henon map, but these examples will not be exposed here. Also, an experimental time series proposed by the Santa Fe Institute has been studied.

II. QUANTIFICATION OF THE SELF-CROSSING DEGREE IN A TRAJECTORY

In order to estimate the average scalar product along the trajectory, we associate to a generic point $\mathbf{x}_i = \mathbf{x}(i\Delta)$, where Δ is the sampled interval, of the reconstructed state space a set of neighboring points $O(\mathbf{x}_i)$, that belong to a closed ball of dimension d and radius ϵ. That is,

$$O(\mathbf{x}_i) = \{\mathbf{x_j} : \; \|\mathbf{x}_i - \mathbf{x}_j\| \leq \epsilon, \; j = 1, ..., N, \; j \neq i\},$$

where $\|...\|$ is the Euclidean norm and we call μ_i the number of points in $O(\mathbf{x_i})$.

There are two main problems when implementing this idea. On the one hand, the radius of the d-dimensional sphere, ϵ, over which the average is performed. For those cases where a degree number of points is available, a local approximation to the flow around the studied point would be obtained for a very small ϵ. However, the amount of data available is normally limited, and there are practical constraints to the reduction of the radius size. For a given time series, the study could be carried out by fixing either the radius size or the number of points belonging to the set $O(\mathbf{x}_i)$. We have decided to fix the radius ϵ for every state-space reconstruction of a given time series. A rough method of estimating a useful ϵ value will be given below considering a uniform distribution of points.

On the other hand, there is a difficulty for those time delays where state-space points collapse close to the identity line. For instance, when $T \to 0$ and for a fixed ϵ, we may have too many points in $O(\mathbf{x}_i)$. In order to avoid this problem, we rotate and scale state-space reconstructions using singular value descomposition [14]. In the transformed coordinates the average number of points in $O(\mathbf{x}_i)$ is similar for any state-space reconstruction, $\{T, d\}$, because the attractor has the same variance in any axis. Therefore, state-space points are rotated using the principal component axis (PCA: $\mathbf{x} \to \mathbf{y}$) and each coordinate is scaled ($z_i = y_i/\sqrt{\lambda_i}$, where λ_i is the variance) to the standard deviation of the new axis.

With these new points, $\mathbf{z}_j$, the normalized flow vectors are calculated as

$$\mathbf{f}(\mathbf{z}_j) = \frac{\mathbf{z}_{j+1} - \mathbf{z}_j}{\|\mathbf{z}_{j+1} - \mathbf{z}_j\|},$$

and the mean flux vector value, V_i, inside the set $O(\mathbf{z}_i)$, is given by

$$\mathbf{V}_i = \frac{1}{\mu_i} \sum_{\mathbf{z}_j \in O(z_i)} \mathbf{f}(\mathbf{z}_j). \tag{2.1}$$

Finally, the scalar product is averaged along the trajectory using the expression

$$P = \frac{1}{N} \sum_{i=1}^{N} \mathbf{f}(\mathbf{z}_i) \mathbf{V}_i. \tag{2.2}$$

We can intuitively describe the above set of calculations as a travel along the trajectory with a tube of radius, ϵ, averaging the scalar product with all the flows inside this tube.

As previously mentioned, the estimate of P would be more accurate for low values of ϵ, provided there is a large enough number of points in $O(\mathbf{z_i})$. It is useful to give a starting value of ϵ. This value can be adjusted in order to perform more accurate measurements. For a time series having a uniform distribution of points covering the state space, we can calculate the value of ϵ needed to obtain a number of points, n_0, falling inside the closed ball of radius ϵ (see the Appendix),

$$\epsilon = \sqrt{\frac{2+d}{d}} \left(\frac{n_0}{N}\right)^{\frac{1}{d}},$$

where N is the total number of points available in the time series. Although the distribution of points on an attractor is far from uniform, we assume uniformly here to get a rough estimate.

The value of the averaged scalar product, P, is a function of T and d for the set of possible state spaces drawn from the original time series. The procedure that we have found to be more accurate to determine the optimum $\{T, d\}$ values consists of maximizing P as a function of d for a fixed T value. It is expected that P will reach a plateau or an asymptotic value when approaching a large enough dimension. In a second step the optimum T value is calculated by maximizing P for the d value previously determined.

A brute force implementation of the averaged scalar products along trajectories will imply operations of the order of N^2. This is due to the fact that the distance from a point $\mathbf{z}_i$ to the $N-1$ remaining points have to be determined. However, the computational cost is reduced if the calculation of these distances, required to determine the set $O(\mathbf{z}_i)$, is reduced to determine the distance between $\mathbf{z}_i$ and those values $\mathbf{z}_j$, for which the first component satisfies $z_j^0 - \epsilon \leq z_i^0 \leq z_j^0 + \epsilon$, where ϵ is the radius of the sphere. To do that, the first component of $\mathbf{z}_j$ is sorted in increasing order and only those points which satisfy the above relationship are used to evaluate the distances. In this way the number of operations is of $n(\epsilon)N$ where $n(\epsilon)$ is the number of points contained in the above inequality.

III. EXAMPLES

The method proposed in this work has been tested in three different types of times series. The first one is obtained from a set of three ordinary differential equations, the second one is obtained from a delay differential equation, and the third one is an experimental series.

A. Rössler attractor

The set of nonlinear ordinary differential equations (ODE's) [15]:

$$\begin{aligned} \dot{x} &= -z - y, \\ \dot{y} &= x + ay, \\ \dot{z} &= b + z(x - c), \end{aligned} \tag{3.1}$$

with parameter values $a = 0.15$, $b = 0.20$, and $c = 10.0$, have been used to generate $x(t)$. These ODE's have been integrated with a Bulirsch-Stoer method with a fixed time step of $\pi/100$. Given a Poincaré section the average number of steps needed to cross it since the last crossing is 193.3 steps, we call this the average period of the attractor. This quantity is noted T_0 and it is used as a time unit in what follows. A value of $\epsilon = 0.1$ was used and 10 000 sampled points were used. Initial conditions

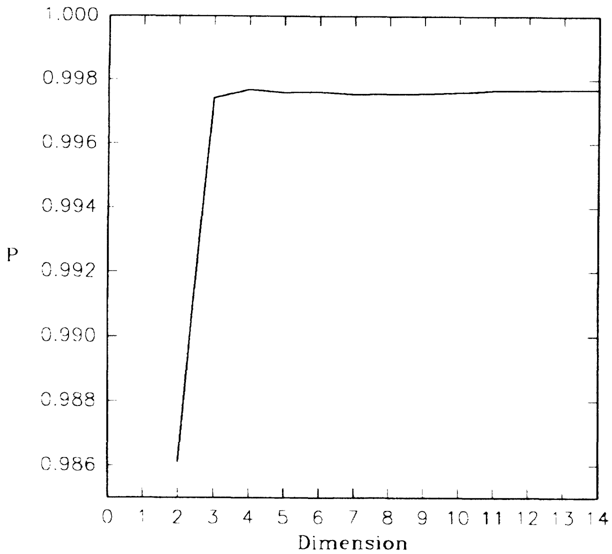

FIG. 1. Estimation of the dimension for Rössler attractor, P as function of the dimension for $\epsilon = 0.1$.

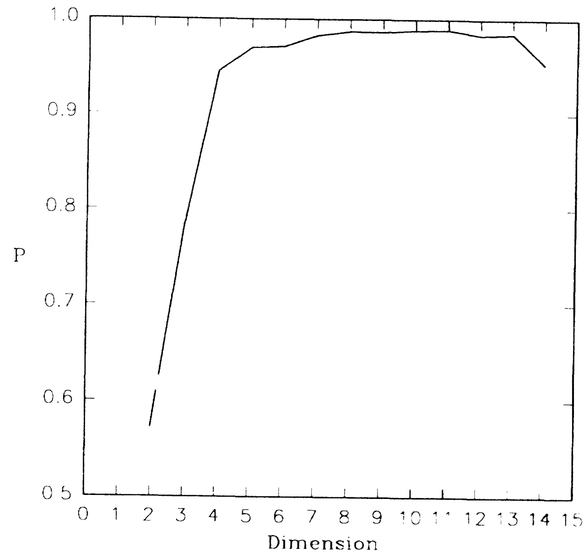

FIG. 3. Estimation of the dimension for Mackey-Glass system, P as function of d for $\epsilon = 0.1$.

are not relevant since the system is ergodic.

In Fig. 1 we present a plot of the average scalar product P as a function of the dimension d for $T = 4$. It is observed that for $d > 2$, the average scalar product reaches a plateau. Therefore, we take $d = 3$ as the embedding dimension. You can see from the ordinary differential equations that the Rössler attractor is embedded in dimension three and its fractal dimension obtained from the Lyapunov exponents is around 2.01. Therefore there is a good agreement with what is expected.

In Fig. 2 we plot P as a function of T for $d = 2, 3, 4$. It is seen that low values of $T \in (0, 0.45)$, are those for which P is maximum. Regions with low P appear around time delays commensurable with the mean period of the attractor.

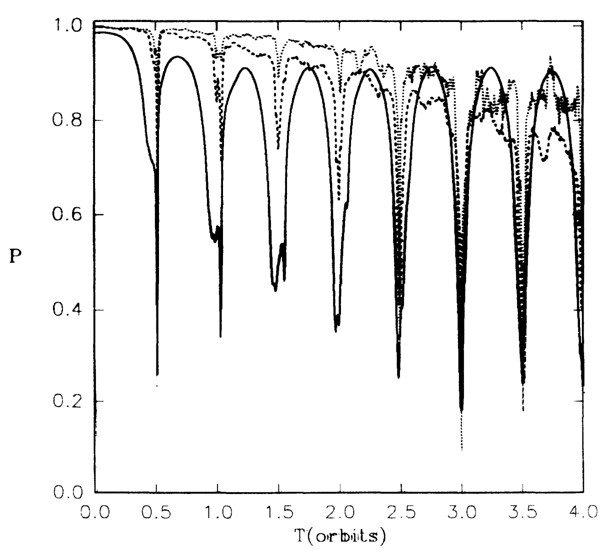

FIG. 2. Variation of P as function of T for Rössler attractor. The solid line is for $d = 2$, the dashed line is for $d = 3$, and the dotted line is for $d = 4$.

B. The Mackey-Glass system

Consider the delay differential equation

$$\dot{x} = -\gamma x(t) + \beta x(t-\tau) \frac{\theta^n}{\theta^n + x(t-\tau)^n}$$

which is a well-known model for the white blood cell population in humans [16]. We have integrated the equation, with parameter values $\gamma = 0.1$, $\beta = 0.2$, $\theta = 1$, $n = 10$, and $\tau = 30$, with a Bulirsch-Stoer method. The trajectory is sampled with 10 000 points and the average scalar product is calculated with $\epsilon = 0.1$.

The estimated dimension, m, for this system is around 3.6. We can see in Fig. 3 that $d = 4$ is beginning to reach

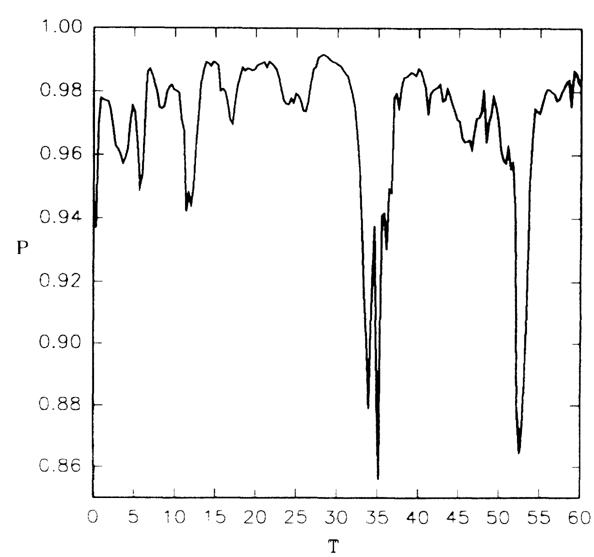

FIG. 4. Variation of P as function of T for Mackey-Glass system, for $d = 5$.

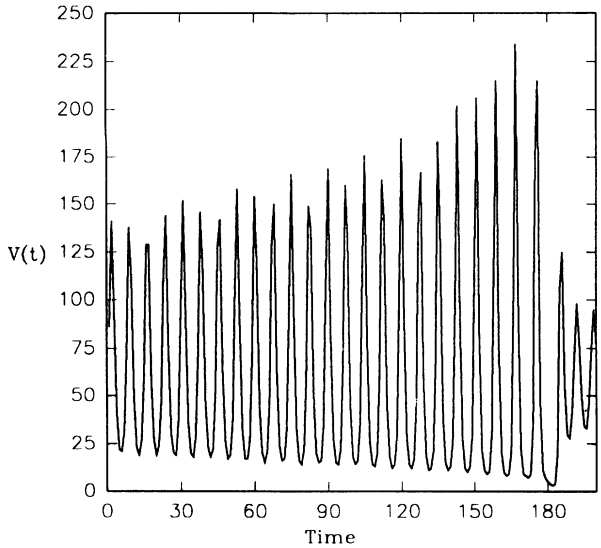

FIG. 5. The time series proposed by the Santa Fe Institute, which was measured in a physics laboratory [17].

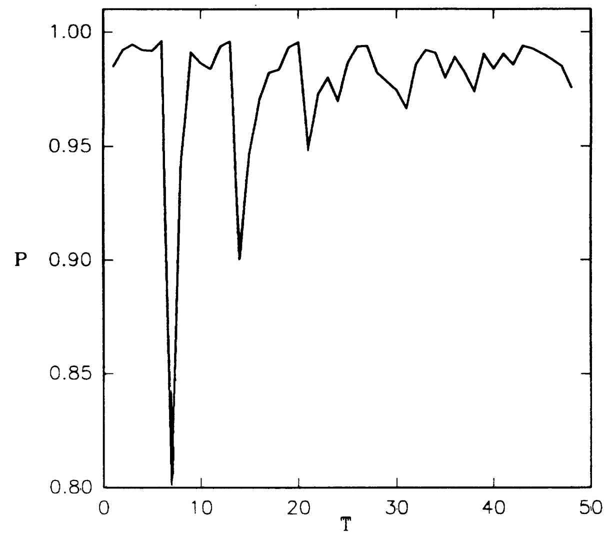

FIG. 7. P as function of T for $d = 6$ in the experimental example.

the plateau. We conclude in this example that $d = 5$ is a good estimation for the embedding dimension. The plot of P versus T presented in Fig. 4 has a maximum at $T \approx 30$, which coincides with the τ parameter used to integrate the equation.

C. Experimental time series

The noise-free studied time series, of which a short sample is plotted in Fig. 5, is the result of a measurement in a physics laboratory, and was proposed in the Santa Fe Institute time series competition [17]. We used the 1000 points available in the series, and $\epsilon = 0.2$ for the calculation of P.

In Fig. 6, P versus d for $T = 2$ is shown. The profile of this curve is not as simple as those obtained for model trajectories. However, an embedding dimension around 6 seems reasonable. The analysis on the optimal time delay is located between 1 and 6 steps as seen in Fig. 7.

IV. THE EFFECT OF NOISE

Often, experimental time series are contaminated with non-negligible noise. It is of practical relevance to characterize the robustness against noise of any procedure for state-space reconstruction. We have studied the effect of noise in the state-space reconstruction guided by the average scalar product. The Rössler attractor has been used for this purpose and white noise of 28 dB has been added to the $x(t)$ time series, used for the calculations in example A.

In Fig. 8 the plateau is reached for $d = 4$ when $T = 1$ and $\epsilon = 0.1$. The comparison of this result with the one

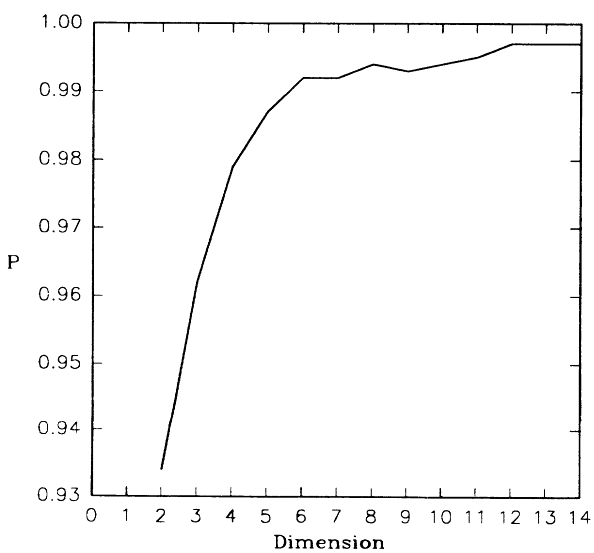

FIG. 6. Estimation of the dimension for an experimental time series, P as function of d for $T = 2$ and $\epsilon = 0.2$.

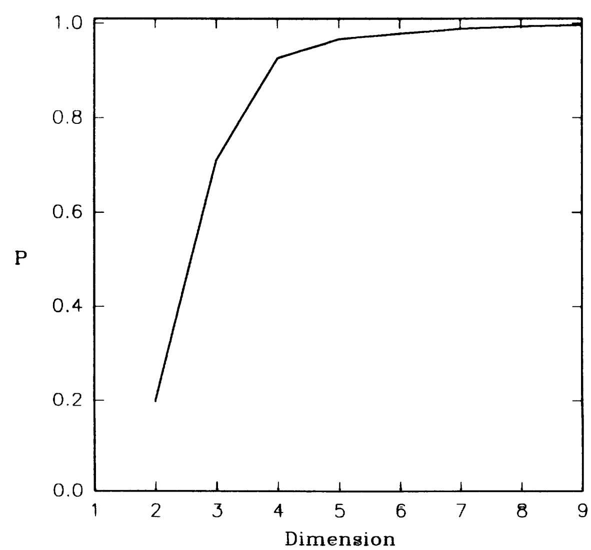

FIG. 8. P as function of d for Rössler attractor with 28 dB of white noise added with $T = 1$ and $\epsilon = 0.1$.

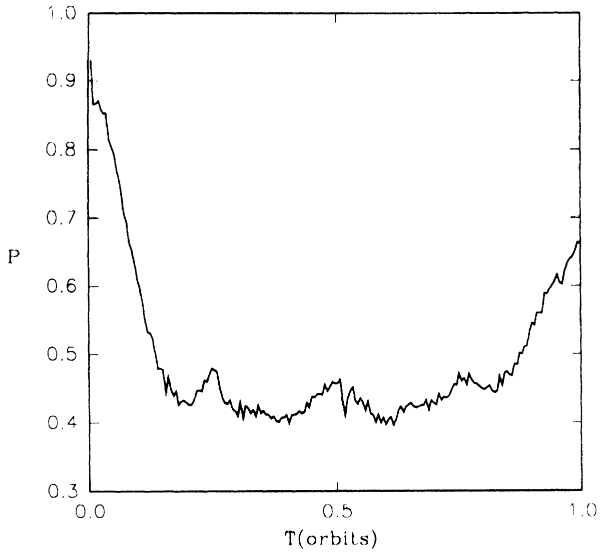

FIG. 9. P as function of T for $d = 4$ in Rössler attractor with 28 dB of white noise added with $\epsilon = 0.1$.

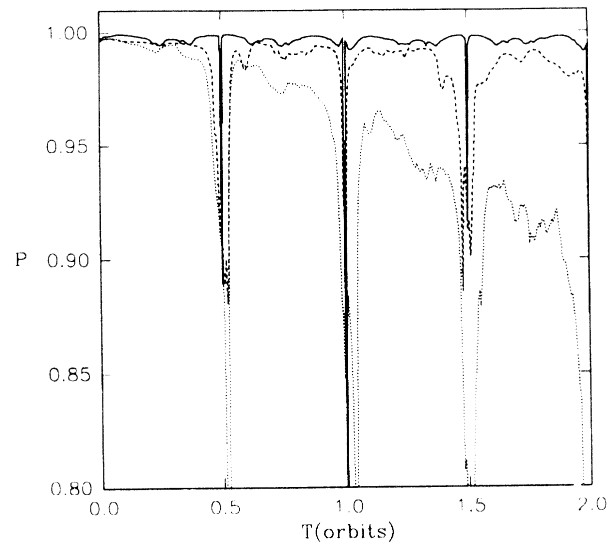

FIG. 10. Variation of P as function of T for different Lyapunov exponents of the Rössler attractor.

obtained in example A shows that the presence of noise leads to a greater dimension than before. To understand this it is important to remember the work of Broomhead and King [14]. In this work when a rotation in the state space is performed by means of PCA, the deterministic subspace is separated from the stochastic subspace. Therefore in Fig. 8, since a rotation by PCA has been performed on the original state space we find that for $d < 4$ the stochastic subspace is not separated from the desterministic subspace and the state-space reconstruction is not good. The dimension $d = 4$ is the minimum for a good state-space reconstruction in the Rössler attractor. As shown in Fig. 8, our method gives precisely 4 as an adequate dimension for a good state-space reconstruction. We can also see in Fig. 9, as in the case of the noise-free time series, that the smallest T values are the best for reconstructing the state space.

V. RELATIONSHIP BETWEEN THE DEGENERACY OF P AND THE LYAPUNOV EXPONENT

It is worth mentioning the observed relation between the slope of the average scalar product when the time delay is varied and the Lyapunov exponent in Fig. 10. In the Rössler attractor, the control parameter [c in Eq. (3.1)] allows a change in the Lyapunov exponent of the attractor trajectories. We calculate the average scalar product P for different values of the Lyapunov exponent and dimension $d = 3$. We have seen that for nonchaotic dynamical systems there is no degeneracy in the average scalar product for increasing values of T. When we are dealing with chaotic dynamical systems, we experimentally observe how P degenerates when increasing the maximun Lyapunov of the system (see Fig. 10). The mathematical equation which explains this behavior has not been figured out, but this phenomena is closely related to the concept of irrelevance, that is, the degeneracy of the state-space reconstruction with increasing values of time delays for chaotic dynamical systems, coined by Casdagli *et al.* [18].

VI. DISCUSSION

Flow analysis on state spaces reconstructed from a time series is a useful perspective for representing time series. In particular, we may conclude that the average scalar product along trajectories can be used to select near-optimal time delays and dimension for state-space reconstruction.

For time series originating from a system of ODE's, the proposed method provides empirical evidence for the fact that small time delays are the best selection for state-space reconstruction. Also, when delay differential equations are used to generate the time series, the time delay used in the integration is singled out by the analysis.

We also find it quite useful that flow analysis gives a measure of time series chaoticity. In particular, P decreases with the time delay with a slope related to the Lyapunov exponent. An interesting problem for further study is the quantification of the Lyapunov exponent by means of the degeneracy of P.

ACKNOWLEDGMENTS

We are grateful for the support of CICYT (PB91-0045 and TAP92-117). One of us (R.H.) also acknowledges M.E.C for financial support. We are also very grateful to Professor Edward Ott for reviewing the manuscript.

APPENDIX

We estimate here the ϵ value for a uniform distribution of points in the state-space reconstruction. We imagine that all points, N, are within a sphere of radius R whose volume is given by

$$V(R) = C(d)R^d,$$

where $C(d)$ is a constant. Therefore, the ratio of the number of points, n_0, within a sphere of radius ϵ to the total number of points can be obtained by using

$$\frac{n_0}{N} = \left(\frac{\epsilon}{R}\right)^d . \qquad \text{(A1)}$$

The value of R can be determined from the variance along any axis and, as we are considering an isotropic distribution, then the probability distribution depends on the $(d-1)$ surface of the d sphere

$$\lambda^2 = \frac{\int_0^R r^2 r^{d-1} dr}{\int_0^R r^{d-1} dr} = R^2 \left(\frac{d}{d+2}\right) . \qquad \text{(A2)}$$

Since the state-space reconstruction was scaled to $\lambda = 1$, the R value can be obtained from Eq. (A2) and substituted into Eq. (A1) from which ϵ can be determined,

$$\epsilon = \sqrt{\frac{2+d}{d}} \left(\frac{n_0}{N}\right)^{\frac{1}{d}} .$$

[1] N. H. Packard, J. P. Crutchfield, J. D. Farmer, and R .S. Shaw, Phys. Rev. Lett. **45**, 712 (1980).
[2] F. Takens, in *Dynamical Systems and Turbulence,* edited by D. Rand and L. S. Young (Springer, Berlin, 1981).
[3] J. P. Eckman and D. Ruelle, Rev. Mod. Phys. **57**, 617 (1985).
[4] T. Sauer, J. A. Yorke, and M. Casdagli, J. Stat. Phys. **65**, 579 (1991).
[5] P. Grassberger and I. Procaccia, Phys. Rev. Lett. **50**, 346 (1983).
[6] J. Theiler, Phys. Rev. A **34**, 2427 (1986).
[7] Mingzhou Ding, Celso Grebori, Edward Ott, Tim Sauer, and James A. Yorke, Phys. Rev. Lett. **70**, 25, 3872 (1993).
[8] A. I. Mees, P. E. Rapp, and L. S. Jennings, Phys. Rev. A **36**, 340 (1987).
[9] A. M. Albano, J. Muench, and C. Schwartz, Phys. Rev. A **38**, 3017 (1988).
[10] A. M. Fraser and H. L. Swinney, Phys. Rev. A **33**, 1134 (1986).
[11] Mathew B. Kennel, Reggie Brown, and Henry D. I. Abarbanel, Phys. Rev. A **45**, 3403 (1992).
[12] M. Casdagli, Physica D **35**, 335 (1989).
[13] D. T. Kaplan and Leon Glass, Phys. Rev. Lett. **68**, 427 (1992).
[14] D. S. Broomhead and G. P. King, Physica D **20**, 217 (1986).
[15] O. E. Rössler, Phys. Lett. A **57**, 397 (1976).
[16] M. C. Mackey and L. Glass, Science **197**, 287 (1977).
[17] N. A. Gershenfeld and Andreas S. Weigend, in *Time Series Prediction*, edited by Andreas S. Weigend and Neil A. Gershenfeld (Addision-Wesley, Reading, MA, 1993).
[18] M. Casdagli, S. Eubank, J. D. Farmer, and J. Gibson, Physica D **51**, 52 (1991).