Observability of nonlinear dynamics: Normalized results and a time-series approach

Luis A. Aguirre, Saulo B. Bastos, Marcela A. Alves, and Christophe Letellier

Citation: Chaos **18**, 013123 (2008); doi: 10.1063/1.2885386
View online: http://dx.doi.org/10.1063/1.2885386
View Table of Contents: http://scitation.aip.org/content/aip/journal/chaos/18/1?ver=pdfcov
Published by the AIP Publishing

Observability of nonlinear dynamics: Normalized results and a time-series approach

Luis A. Aguirre,[1] Saulo B. Bastos,[1] Marcela A. Alves,[1] and Christophe Letellier[2]
[1]*Programa de Pós-graduação em Engenharia Elétrica, Universidade Federeal de Minas Gerais, Av. Antônio Carlos 6627, 31270-901 Belo Horizonte, MG, Brazil*
[2]*Groupe d'Analyse TOpologique et de MOdlisation de SYstmes Dynamiques, Université et INSA de Rouen — CORIA UMR 6614, Av. de l'Université, BP 12, F-76801 Saint-Etienne du Rouvray cedex, France*

(Received 19 November 2007; accepted 4 February 2008; published online 20 March 2008)

This paper investigates the observability of nonlinear dynamical systems. Two difficulties associated with previous studies are dealt with. First, a normalized degree observability is defined. This permits the comparison of different systems, which was not generally possible before. Second, a time-series approach is proposed based on omnidirectional nonlinear correlation functions to rank a set of time series of a system in terms of their potential use to reconstruct the original dynamics without requiring the knowledge of the system equations. The two approaches proposed in this paper and a former method were applied to five benchmark systems and an overall agreement of over 92% was found. © *2008 American Institute of Physics*. [DOI: 10.1063/1.2885386]

A key result in nonlinear time-series analysis is provided by embedding theory. The central idea is that an *m*-dimensional system can be studied from *any* generic single (scalar) time series, measured from it, as long as the dimension of the reconstructed space is sufficiently high. Such a result stands for infinitely long and infinitely clean data. However, it turns out that, due to practical limits, the analysis of nonlinear dynamical systems is affected by the variable chosen to observe the system. Observability theory provides some tools that enable us to define procedures in order to compare and quantify different variables as potential candidates to be measured for use by embedding and modeling techniques. This paper proposes two procedures that add to the current set of results available in this field.

I. INTRODUCTION

This paper addresses the problem of observability from the perspective of deciding if all the dynamical information of the full phase space $\mathbb{R}^m$ of an m-dimensional system can be retrieved by measuring one variable $s(t)=h(\boldsymbol{x})$, $h:\mathbb{R}^m \mapsto \mathbb{R}$, referred to as the observable, and h as the measuring function. Briefly, a system is observable if the full state can be found based on $s(t)$ only. Another way of seeing this problem is that for an observable system, it is possible to distinguish any two initial conditions in the original phase space from the recorded output only. Using poor observables can pose difficulties to global modeling,[1,2] Kalman filtering,[3] and synchronization schemes.[4]

In the case of linear systems, where $s(t)$ is a linear combination of the state variables, the determination of the system observability is a well-known problem and straightforward tests are available if the system equations are known.[5] Such tests provide "yes" or "no" answers; that is, the system is either observable or not. In the realm of linear systems, however, it was pointed out that although a similarity transformation will not change the rank of a particular matrix, it usually will change the respective determinant.[6] In other words, a linear system that is observable, when mapped to other space by a similarity transformation, will remain observable, but the practical investigation of the dynamics in the new space can turn out to be more (or less) difficult than before. That is, the degree of observability was changed by the referred transformation. The classical theory of observability for dynamical systems, either linear or nonlinear, do not provide ready-to-use ways of assessing the degree of observability, which is of concern only in practical problems.

Although the concepts of controllability and observability for linear systems were developed in the early 1960s, it was not until the 1970s that the nonlinear counterparts started to be developed for which the Lie algebra became an important tool.[7,8]

In the case of nonlinear dynamical systems there are at least two reasons that it is interesting to quantify the degree of observability. First, unlike in the linear case, the degree of observability of nonlinear systems varies throughout the phase space and it could turn out to be interesting to quantify the regions of greater or lesser observability (Ref. 9, Fig. 2). Second, the observable chosen to reconstruct the dynamics by means of embedding techniques plays an important role in various practical aspects of modeling and analysis.[10]

In previous works by the authors, two classes of observability matrices were used: (i) an observability matrix Q based on the Jacobian matrix of the flow evaluated along a given solution,[9] and (ii) an observability matrix $\mathcal{O}$ computed using Lie derivatives.[11] The advantage of the first approach is the relative ease with which Q can be computed. The advantages of the second approach is that the framework is better suited to analyze nonlinear system and $\mathcal{O}$ has an important and direct interpretation in terms of embedding theory.[11]

Both approaches, however, share two features. First, the measures obtained for the degree of observability are of relative value only. That is, results of different observables of the

same system can be compared but the values do not have an absolute meaning. Second, to build either Q or $\mathcal{O}$, the system equations must be known.

The aim of the present paper is to present two different approaches, which differ from the previous methods in terms of the features just mentioned. In particular, the first part of the paper will provide a way of measuring the degree of observability that yields normalized results either from Q or $\mathcal{O}$.

In the second part of the paper, a procedure is put forward by which it is possible to compare two observables of the same system from data, without the need of the system equations. This concern is motivated by a number of practical situations in which more than one observable is recorded from a system and a comparison is needed as a first step to further analysis.[12] The analytical approach to multivariable reconstruction has been dealt with elsewhere.[13] The approach put forward in this paper is based on omnidirectional nonlinear correlation functions[14,15] and although no mathematical proof is provided for the links between such functions and observability, it is argued, based on embedding theory and practice, that such a link nonetheless exists. It will be seen that the proposed time-series approach yields results that are similar to those provided using the system dynamical equations.

This paper is organized as follows. Section II provides some background material and notation. Section III describes a procedure that yields normalized measures of the degree of observability, and in Sec. IV the time-series approach is described. Numerical results are described in Sec. V. The main conclusions of the paper are provided in Sec. VI.

II. OBSERVABILITY FOR NONLINEAR SYSTEMS

A. Concepts

Let $\dot{\boldsymbol{x}}=\boldsymbol{f}(\boldsymbol{x})$, with $\boldsymbol{f}:\mathbb{R}^m\mapsto\mathbb{R}^m$, and $s(t)=h(\boldsymbol{x})$, with $h:\mathbb{R}^m\mapsto\mathbb{R}$ be a nonlinear system and a scalar observable, respectively. From s, a reconstructed space can be formed, for instance by taking a sufficient number of successive derivatives $\boldsymbol{X}=(s(t),s^{(1)},\ldots,s^{(d-1)})$, where $s^{(j)}$ is the jth time derivative of s.[16] Therefore, there is Φ_s that maps the original phase space to the space spanned by the observable s and its successive derivatives; that is, $\Phi_s:\mathbb{R}^m(\boldsymbol{x})\mapsto\mathbb{R}^d(\boldsymbol{X})$. In this paper, the system in $\mathbb{R}^m(\boldsymbol{x})$ will be observable from $s(t)$ in dimension d, if it is possible to go back from $\mathbb{R}^d(\boldsymbol{X})$ to $\mathbb{R}^m(\boldsymbol{x})$ on a one-to-one basis; that is, if Φ_s defines a diffeomorphism. In other words, any two points in $\mathbb{R}^m(\boldsymbol{x})$ are also different in $\mathbb{R}^d(\boldsymbol{X})$. Observability of the system in $\mathbb{R}^m(\boldsymbol{x})$ from the space $\mathbb{R}^d(\boldsymbol{X})$ spanned by s and its successive derivatives is therefore directly related to the existence of singularities in Φ_s.

Observability theory for nonlinear systems provides an alternative way to check the invertibility of Φ_s. To see this, consider the jth-order Lie derivative of h in the direction of the vector field f,

$$\mathcal{L}_f^j h(\boldsymbol{x})=\frac{\partial\mathcal{L}_f^{j-1}h(\boldsymbol{x})}{\partial\boldsymbol{x}}\boldsymbol{f}(\boldsymbol{x}), \tag{1}$$

where $\mathcal{L}_f^0 h(\boldsymbol{x})=h(\boldsymbol{x})$. The time derivatives of s can be written in terms of Lie derivatives as $s^{(j)}=\mathcal{L}_f^j h(\boldsymbol{x})$. The observability matrix can be written as

$$\mathcal{O}_s(\boldsymbol{x})=\begin{bmatrix}\dfrac{\partial\mathcal{L}_f^0 h(\boldsymbol{x})}{\partial\boldsymbol{x}}\\ \vdots\\ \dfrac{\partial\mathcal{L}_f^{m-1}h(\boldsymbol{x})}{\partial\boldsymbol{x}}\end{bmatrix}. \tag{2}$$

Finally, if $\mathcal{O}_s^T\mathcal{O}_s$ (which has been called a *distortion matrix* in Ref. 16) is nonsingular, then the system is observable from the reconstructed space spanned by $s(t)$ and its successive derivatives up to order $m-1$. A key point is that $\mathcal{O}_s^T\mathcal{O}_s$ might be close to singularity but not mathematically singular. In that case, although the system is observable, in practice, the use of $s(t)$ for modeling and analysis is bound to be troublesome or, eventually, higher-order derivatives of $s(t)$ will be required. For that reason, the degree of observability can be measured by

$$\delta_s(\boldsymbol{x})=\frac{|\lambda_{\min}[\mathcal{O}_s^T\mathcal{O}_s,\boldsymbol{x}(t)]|}{|\lambda_{\max}[\mathcal{O}_s^T\mathcal{O}_s,\boldsymbol{x}(t)]|}, \tag{3}$$

where $\lambda_{\max}[\mathcal{O}_s^T\mathcal{O}_s,\boldsymbol{x}(t)]$ indicates the maximum eigenvalue of matrix $\mathcal{O}_s^T\mathcal{O}_s$ estimated at point $\boldsymbol{x}(t)$ (likewise for $\lambda_{\min}$). Then, $0\le\delta_s(\boldsymbol{x})\le 1$, and the lower bound is reached when the system is unobservable at point $\boldsymbol{x}$. Clearly, $\delta_s(\boldsymbol{x})$ is a type of condition number of the matrix $\mathcal{O}_s^T\mathcal{O}_s$.

In practice, it is usually convenient to average $\delta_s(\boldsymbol{x})$ along a solution $\boldsymbol{x}(t)$ during an interval of duration T as

$$\delta_s=\frac{1}{T}\sum_{t=0}^{T}\delta_s(\boldsymbol{x}(t)). \tag{4}$$

It is noted that the degree of observability is a local measure and that taking the average is useful inasmuch as it portrays an overall picture.

III. A NORMALIZED MEASURE OF DEGREE OF OBSERVABILITY

Although it is often impossible to check the global invertibility for general nonlinear maps, the Rank theorem provides a sufficient condition for local invertibility. The map Φ_s is locally invertible at $\boldsymbol{x}_0$ if the Jacobian matrix has full rank; that is, if

$$\operatorname{rank}\left(\left.\frac{\partial\Phi_s}{\partial\boldsymbol{x}}\right|_{\boldsymbol{x}=\boldsymbol{x}_0}\right)=m. \tag{5}$$

If condition (5) holds $\forall\,\boldsymbol{x}\in\mathbb{R}^m(\boldsymbol{x})$, the map Φ_s defines a global diffeomorphism from the original space $\mathbb{R}^m(\boldsymbol{x})$ to the reconstructed space $\mathbb{R}^d(\boldsymbol{X})$. This remark is used to define the following normalized measure of observability

$$\delta_s^*(\boldsymbol{x}) = \frac{\text{rank}[\mathcal{O}_s(\boldsymbol{x})]}{m}, \tag{6}$$

where the rank of a matrix is the dimension of its range or, in other words, the maximal number of independent rows or columns.[17]

The measure $\delta_s^*(\boldsymbol{x})$ is a sequence of numbers between 0 and 1. Notice that the numerator of (6) can only assume integer values from zero to m. For a global diffeomorphism, $\delta_s^*(\boldsymbol{x})=1$, $\forall\ \boldsymbol{x}\in\mathbb{R}^m(\boldsymbol{x})$. On the other hand, for maps that define diffeomorphisms only locally, there will be at least one region $\mathbb{R}_u(\boldsymbol{x})\subset\mathbb{R}^m(\boldsymbol{x})$ for which Φ_s will not be invertible. In such a region, $\mathcal{O}_s(\boldsymbol{x})$ will be rank deficient; therefore, $\delta_s^*(\boldsymbol{x})<1$ and $\boldsymbol{x}\in\mathbb{R}_u(\boldsymbol{x})$. Finally, a single representative value δ_s^* can be obtained by averaging along a solution of duration T, as in (4).

Definition (6) faces a challenge that is similar to that of the former definition; that is, to decide numerically when to declare a matrix rank deficient. In other words, a matrix could be practically rank deficient but not mathematically rank deficient. Therefore, most numerical algorithms require that a tolerance be given beyond which a matrix is considered rank deficient.

In this work, the following tolerance is proposed:

$$\epsilon = \frac{1}{m}\sum_{i=1}^{m}\epsilon_i[\boldsymbol{x}(t)], \tag{7}$$

where t spans the time interval over which the solution $\boldsymbol{x}(t)$ is considered and

$$\epsilon_i[\boldsymbol{x}(t)] = pm\frac{\overline{\sigma_{x_i(t)}} - \min[\sigma_{x_i(t)}]}{\max[\sigma_{x_i(t)}] - \min[\sigma_{x_i(t)}]} \tag{8}$$

if $\max[\sigma_{x_i(t)}]\neq\min[\sigma_{x_i(t)}]$ and $\epsilon_i[\boldsymbol{x}(t)]=pm$ otherwise, where $\sigma_{x_i(t)}$ are the set (one for each value of t) of maximum singular values of $\mathcal{O}_{x_i}(\boldsymbol{x})$, $\overline{\sigma_{x_i(t)}}$ is the average of such a set. Parameter p is used to define the tolerance, which should be small, of course. In this paper, values in the range $0\leq p\leq 0.05$ were used. For a square and real matrix as $\mathcal{O}_{x_i}(\boldsymbol{x})$, the singular values are defined as the square roots of the eigenvalues of $\mathcal{O}_{x_i}(\boldsymbol{x})^T\mathcal{O}_{x_i}(\boldsymbol{x})$.

The motivation behind (7) and (8) is as follows. In various numerical software (e.g., MATLAB), the tolerance used in the determination of the rank of a matrix A is basically composed as $\epsilon=p\times\max[\text{size}(A)]\times\text{norm}(A)$, where p is usually a very small constant, and the norm of A is usually taken as the $\|A\|_\infty$, which is defined as the largest singular value of A. From (8), it is readily seen that in the proposed method, $0\leq p\leq 0.05$, the size of $\mathcal{O}_{x_i}(\boldsymbol{x})$ is constant and equal to m. As for the norm, it must be noticed that in our case, at each time t within the window of data there will be m matrices $\mathcal{O}_{x_i}(\boldsymbol{x})$ (one for each observable). We desire *one* value of tolerance ϵ for all the matrices. One alternative would be simply to work with the average value of the set of largest singular values taken along the window of data: $\overline{\sigma_{x_i(t)}}$. This choice did work, but required very small values of p, which spanned a relatively wide window of possible values. In order to improve on this, the average was replaced by $\overline{\sigma_{x_i(t)}}-\min[\sigma_{x_i(t)}]$, which gives an idea of variability, divided by $\max[\sigma_{x_i(t)}]-\min[\sigma_{x_i(t)}]$, which normalizes the new "norm." This is done for each observable. Finally, (7) takes ϵ as the average over all the observables.

A. A simple example

To illustrate, consider the Rössler system[18]

$$\dot{x}=-y-z,\quad \dot{y}=x+ay,\quad \dot{z}=b+z(x-c), \tag{9}$$

with $(a, b, c)=(0.398, 2.0, 4.0)$. For the three dynamical variables of the Rössler system (9), the degrees of observability are $\delta_x=0.022$ (0.022), $\delta_y=0.133$ (0.133), and $\delta_z=0.0063$ (1.9×10^{-4}), where the values correspond to the definition used in Refs. 11 and 9, for the numbers in parentheses. Both definitions provide the observability order $y\triangleright x\triangleright z$, which can be considered the correct one for this simple system. However, although it is known that the phase space reconstructed with $(y,\dot{y},\ddot{y})$ is globally diffeomorphic to the original phase space, the degree of observability (3) is not maximal; that is, $\delta_y\neq 1$.

Using the definition (6), averaged as (4) with tolerance computed by (7) and (8) with $p=0.01$, the following degrees of observability were found: $\delta_x^*=0.8781$, $\delta_y^*=1.000$, and $\delta_z^*=0.3825$, which apart from yielding the same observability order as before, i.e., $y\triangleright x\triangleright z$, they indicate that the observable $s=y$ provides a global diffeomorphism in three dimensions (3D).

IV. TIME-SERIES APPROACH

Whether the degrees of observability are normalized or not, one thing remains as a drawback in situations where only data are available: the fact that the aforementioned results require the knowledge of the system equations.

The definition of some kind of observability measure from data is far from obvious. A naïve procedure would be to try to build from the data the set of governing equations. As often is the case, only one observable $s(t)$ is available and one would like to evaluate the observability of the system as seen through $s(t)$ at a certain dimension (model order). Instead of addressing this problem from an analytical point of view, as done in the previous section, here we will follow a more practical approach, for which we do not provide any mathematical proof; however, we do provide the motivation and numerical evidence.

It is claimed that if $s(t)$ is a good observable, the phase-space reconstructed using $s(t)$ should be robust to small fluctuations, such as those provoked by noise, small changes in embedding parameters, and so on. The rationale behind this is precisely the remark made in the previous section, where it was pointed out that lack of observability is associated with the impossibility to return from the reconstructed space to the original one. A bad observable is therefore characterized by regions in the reconstructed space where the trajectories squeeze together (at and close to singularities) in such a way that it is impossible to disentangle them and return to the original space on a one-to-one basis. This is actually also true for a good observable but a small embedding dimension. A good observable, on the other hand, is one for which a

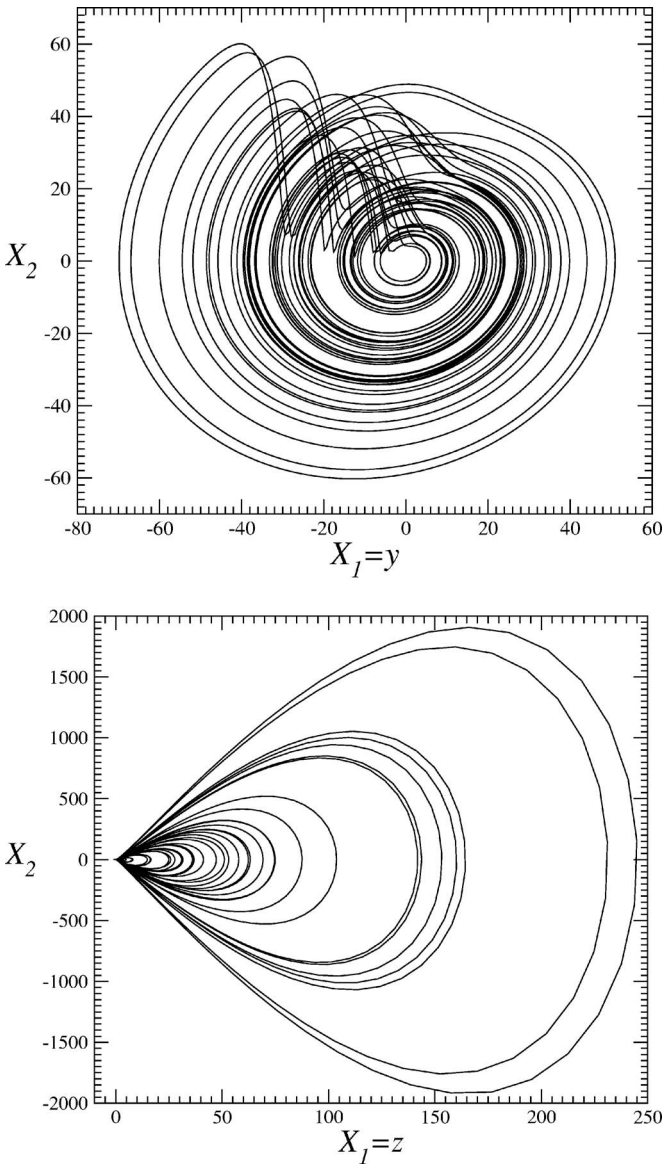

FIG. 1. Plane projections for the hyperchaotic Rössler system, reproduced from Ref. 11, Fig. 1. The top graph was produced using y as the observable variable and plotting $X_2 \times X_1$, with $X_1=y$ and $X_2=\dot{y}$. The bottom graph was produced using $X_1=z$ and $X_2=\dot{z}$, for which the origin is a singular point. In the neighborhood of the origin, the trajectories squeeze together, thus hindering the return to the original space. Therefore, z is a bad observable. In fact, as it will be seen in Table I, z is the worse observable of this system.

piece of the trajectory in the reconstructed space can be "perturbed" somewhat (by noise or small variations in some system or embedding parameter) and the return to the original space can still be performed. Examples of good and poor observables of the hyperchaotic Rössler system are shown in Fig. 1.

The question now is to decide which property of a reconstructed phase-space can be used to test for robustness and how to do it in a simple yet effective way. From what has been argued so far, it seems interesting to use some feature that is directly connected to the embedding process and that can be easily evaluated directly from a time series.

Two fundamental embedding parameters are the delay time (for time-delay coordinates) and the embedding dimension.[19] There are many possibilities for choosing these parameters independently. In Ref. 20, a method for choosing these parameters simultaneously has been proposed.

The method of false nearest neighbors[21,22] is used as a means to determine the embedding dimension. As for the delay time, the mutual information criterion is a well-known option.[23] In this respect, it has been argued that correlation functions (linear, nonlinear, and higher order) are useful tools and often yield excellent results at a lower computational cost than the mutual information.[24–26] Motivated by this, we investigate specific features of correlation functions computed using different observables for bench systems and derive observability information from them.

A set of correlation functions designed to "comprehensively detect nonlinear relationships"[14] will be used. The omnidirectional autocorrelation functions are defined for observable $s(k)$—the discrete version of $s(t)$—as[14]

$$
\begin{aligned}
r_{\alpha\alpha}(\tau) &= \frac{\Sigma_{k=\tau+1}^{N}(\alpha(k)-\bar{\alpha})(\alpha(k-\tau)-\bar{\alpha})}{\Sigma_{k=1}^{N}(\alpha(k)-\bar{\alpha})^2}, \\
r_{\alpha s'}(\tau) &= \frac{\Sigma_{k=\tau+1}^{N} s'(k)(\alpha(k-\tau)-\bar{\alpha})}{\left[\left(\Sigma_{k=1}^{N}(s'(k))^2\right)\left(\Sigma_{k=1}^{N}(\alpha(k)-\bar{\alpha})^2\right)\right]^{1/2}}, \\
r_{s's'}(\tau) &= \frac{\Sigma_{k=\tau+1}^{N} s'(k)s'(k-\tau)}{\Sigma_{k=1}^{N}(s'(k))^2}, \\
r_{s'\alpha}(\tau) &= \frac{\Sigma_{k=\tau+1}^{N}(\alpha(k)-\bar{\alpha})s'(k-\tau)}{\left[\left(\Sigma_{k=1}^{N}(\alpha(k)-\bar{\alpha})^2\right)\left(\Sigma_{k=1}^{N}(s'(k))^2\right)\right]^{1/2}},
\end{aligned}
\tag{10}
$$

where $\alpha(k)=|s'(k)|$ and the prime indicates that the average was removed; that is, $s'(k)'$ is a zero mean signal. Finally, these four functions (each of which detect a different type of nonlinearity) can be combined to form a single-valued function $\rho_{ss}(\tau)$ as follows:[14]

$$
\begin{aligned}
\rho_{ss}(\tau) &= \max(r_{\alpha\alpha}(\tau), r_{\alpha s'}(\tau), r_{s's'}(\tau), r_{s'\alpha}(\tau)) \\
&\quad \text{if } \frac{|\max(r_{\alpha\alpha}, r_{\alpha s'}, r_{s's'}, r_{s'\alpha})|}{|\min(r_{\alpha\alpha}, r_{\alpha s'} r_{s's'}, r_{s'\alpha})|} > 1, \\
\rho_{ss}(\tau) &= \min(r_{\alpha\alpha}(\tau), r_{\alpha s'}(\tau), r_{s's'}(\tau), r_{s'\alpha}(\tau)) \\
&\quad \text{if } \frac{|\max(r_{\alpha\alpha}, r_{\alpha s'}, r_{s's'}, r_{s'\alpha})|}{|\min(r_{\alpha\alpha}, r_{\alpha s'}, r_{s's'}, r_{s'\alpha})|} \leq 1,
\end{aligned}
\tag{11}
$$

where the argument τ was omitted in some functions for brevity.

The correlation function $\rho_{ss}(\tau)$ can be calculated for a given observable s for a range of values of τ. It has been noticed that for values of τ up to about 10, $\rho_{ss}(\tau)$ can be conveniently approximated by

$$
\rho_{ss}(\tau) \approx e^{-k\tau^2}, \tag{12}
$$

from which the parameter $k=-\ln(\rho_{ss}(\tau))/\tau^2$ can be obtained. Finally, the following average value is useful:

$$
k_m = \frac{1}{N_k}\sum_{\tau=1}^{N_k}\left(\frac{-\ln(\rho(\tau))}{\tau^2}\right). \tag{13}
$$

In this paper, $N_k=6$.

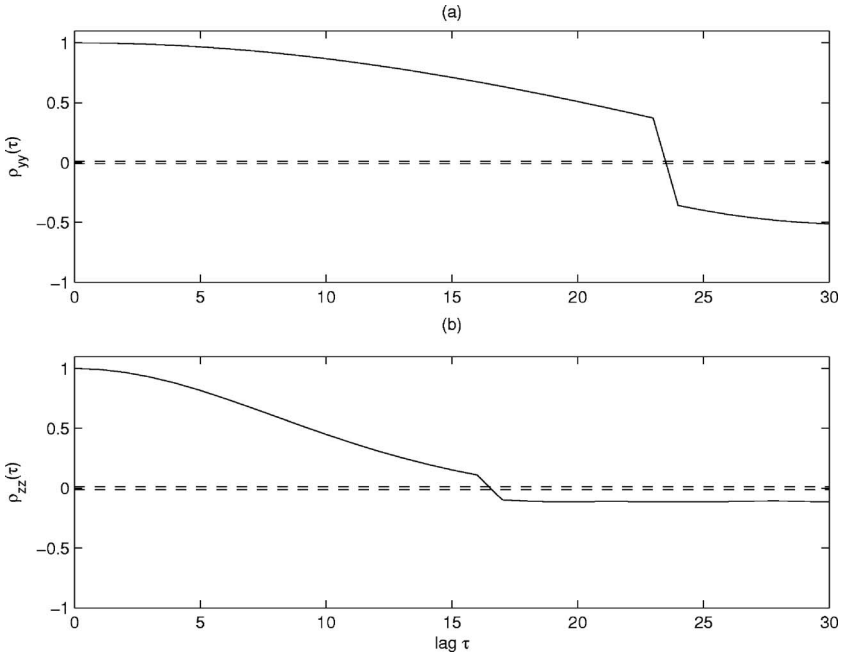

FIG. 2. Correlation function $\rho_{ss}(\tau)$ [see Eq. (11)] for the Rössler system. (a) $\rho_{yy}(\tau)$ and (b) $\rho_{zz}(\tau)$.

It is interesting to note that the correlation functions (10), and consequently (11) are normalized in such a way that $-1 \leq \rho_{ss}(\tau) \leq 1$, $\forall \tau$. Therefore, in principle, the k_m of observables of different systems can be compared. In practice, however, it should be noticed that the shape of $\rho_{ss}(\tau)$ [parametrized by Eq. (13) for small values of τ] depends on the sampling time used to collect the data. As each system has different time scales, it is often needed to sample the data at different rates, thus making comparison meaningless unless the sampling time is the same.

The rationale behind this calculation is as follows. The correlation function $\rho_{ss}(\tau)$ quantifies directly from the data a set of nonlinear autocorrelations present in the time series. If the data were random $\rho_{ss}(\tau)$ would be impulsive; that is, very (infinitely) fast varying. Time series that retain correlation for longer periods usually perform better as observables. Therefore, we need to quantify the "rate of nonlinear de-correlation" in the time series, and that is done by k_m. The greater $1/k_m$, the smaller the nonlinear de-correlation rate and the better the observable.

It should be noticed that a change in the sampling time by a fixed factor will affect $\rho_{ss}(\tau)$ by the same factor, and consequently, $1/k_m$ will also change. However, the relative values of $1/k_m$ for each time series of a given system will not change, therefore permitting comparison. Because $1/k_m$ is not a normalized value, comparison among different systems and for different sampling times should be avoided.

Unlike the definitions of the measures of observability, there is no direct mathematical proof that larger values of $1/k_m$ imply greater observability. However, this has been observed in many bench systems, as will be illustrated in Sec. V B. In what follows, a simple example is provided to illustrate the main points of the procedure.

A. A simple example

Here the Rössler system is considered anew. Equations (9) were simulated with $(a,b,c)=(0.398,2.0,4.0)$. The correlation functions for observables $s=y$ (the best case) and $s=z$ (the worse case) are shown in Fig. 2. The values obtained from the correlation function $\rho_{ss}(\tau)$ [see Eq. (11)] and a long $(N=3\times10^4)$ time series are $1/k_{mx}=625$, $1/k_{my}=714.3$, and $1/k_{mz}=122$, which indicates the expected observability order $y \triangleright x \triangleright z$.

As can be seen in Fig. 2, the correlation function $\rho_{ss}(\tau)$, defined in Eq. (11),[14] indicates that the time series z becomes uncorrelated faster than for the time series y. Moreover, after about 15 lags (the sampling time is $T_s=0.05$), there is much less nonlinear autocorrelation in z than in y. Other examples will be investigated in Sec. V B.

V. NUMERICAL RESULTS

This section will present some results obtained using the techniques proposed in this paper. First, the examples concern the use of normalized measures of observability. Second, the time-series approach is illustrated.

A. Normalized measure of the degree of observability

Table I shows the results obtained for five different systems. Whenever available, the observability measures obtained using other definitions are presented as well.

The same observability order was obtained, regardless of the definition used, for the systems: Rössler[18] and hyperchaotic Rössler.[27] For the Hénon–Heiles[28] system, the two definitions only disagree on variable u, but are mutually consistent in the other three variables. The results for the double-scroll attractor[29] are the first published by the authors.

For the Lorenz system[30]

$$\dot{x} = \sigma y - \sigma x, \quad \dot{y} = \rho x - y - xz, \quad \dot{z} = -\beta z + xy, \tag{14}$$

with $(\sigma, \rho, \beta)=(10.0, 28.0, 8/3)$, the following observability order was found in a previous work:[9] $z \triangleright y \triangleright x$ (see Table I). From a global point of view, variable z cannot possibly be the best observable because observing the system

TABLE I. Degree of observability: δ_s^* normalized ($p=0.01$).

System	s	δ_s^*	δ_s[a]	$1/k_m$	$1/k_m^{\text{noise}}$
Rössler	x	0.8781	0.022	625.0	249.9
	y	1.0000	0.133	714.3	258.1
	z	0.3825	0.0063	122.0	94.1
Hyperchaotic Rössler	x	0.2430	2.2×10^{-4}	742.3	260.8
	y	0.3138	9.0×10^{-4}	862.0	273.8
	z	0.0434	1.3×10^{-4}	15.7	15.2
	w	0.1394	2.1×10^{-4}	1581.9	320.3
Hénon-Heiles	x	0.1689	0.041	265.1	160.6
	y	0.1416	0.026	261.5	160.3
	u	0.3450	0.022	207.3	137.6
	v	0.0227	0.015	199.5	132.2
Lorenz	x	0.9234	6.5×10^{-6}[b]	666.7	256.3
	y	0.8162	8.2×10^{-6}[b]	370.4	195.8
	z	0.9536	2.2×10^{-5}[b]	232.6	148.3
Double-scroll[c]	x	1.0000	-	688.9	256.0
	y	1.0000	-	98.9	79.4
	z	1.0000	-	277.4	164.1

[a]Results published in Ref. 11.
[b]Results published in Ref. 9.
[c]See Table II.

through this variable eliminates the rotation symmetry present in Lorenz's attractor. However, as discussed in Ref. 9, the definition of the observability matrix $\mathcal{O}_s(\boldsymbol{x})$ is local and it fails to reveal that if the system is observed through $s=z$, all information about the symmetry will be lost. In practice, it means that two different points, one on each of the attractor wings, cannot be distinguished by looking at z. As for the other two variables, the respective degrees of observability are similar.

The results obtained by applying definition (6) to the Lorenz system are shown in Table I. A couple of remarks are worthwhile. First, the degree of observability of the system from the z variable has significantly decreased, compared to the other two variables. In fact, δ_z^* and δ_x^* are not too different. Second, the results suggest that x is a better observable than y, whereas the original definition would render y marginally superior to x.[9]

In order to investigate this further, a simple graphical procedure for assessing observability was used.[31] The results are shown in Fig. 3. First, the linear and nonlinear cause-effect couplings among the system variables are displayed [Fig. 3(a)]. The arrows leave the "cause" variable and point to the "affected" variable. Solid arrows indicate linear couplings and the dashed arrows indicate nonlinear couplings.

Second, each variable is considered as an observable. We start, say, from x, and travel contrary to the arrows (we want to know which variables affect x or, in other words, what kind of informations is encoded into x). So, starting from x in Fig. 3(a) we have two possibilities.

(i) Following the continuous arrow from x to itself, we return to x. Now we can only follow the continuous arrow from x to y (remember that the arrows are followed in the opposite directions). Because we are investigating observability in 3D, then only three stages are allowed. This first path is shown at the very top of Fig. 3(b) (a continuous arrow followed from x to x, and another continuous arrow followed from x to y).

(ii) Following the continuous arrow from x to y, we arrive at y. In turn, from y three possibilities exist: x via a dashed arrow, return to y via a continuous arrow, move to z through a dashed arrow. This is indicated in Fig. 3(b) together with the results for all the observables.

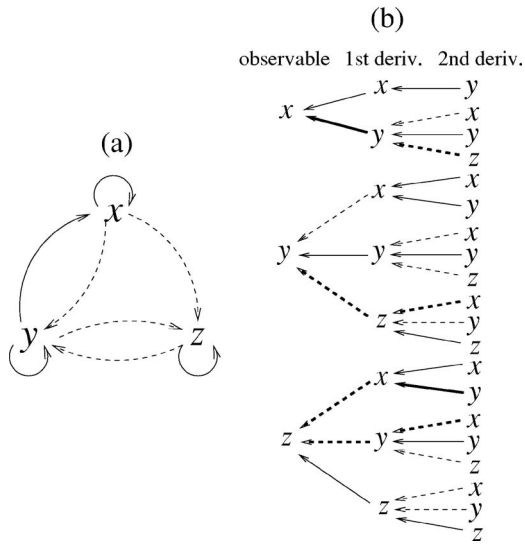

FIG. 3. (a) Couplings among the variables of the Lorenz system (14). Solid (dashed) arrows indicate linear (nonlinear) couplings. (b) Observability diagram. The paths indicated with boldface are those with information of the three variables x, y, and z.

Having built the observability diagram Fig. 3(b), how do we interpret it? First, we take notice of the information available in the observable and the successive derivatives, this corresponds to the three "columns" in Fig. 3(b). In particular, we search for a path that has information of the three variables x, y, and z. Such paths have been indicated in boldface in Fig. 3(b). Second, the type of couplings (linear or nonlinear) of the boldface paths are compared.

A global diffeomorphism is indicated by the three variables *only* connected with continuous arrows. It can be seen that there is no global diffeomorphism indicated in Fig. 3(b). For "mixed" paths, we then compare which have continuous arrows in the first stage, in general this represents a better observability scenario than to have a dashed arrow first and a continuous one after, or to have two dashed arrows, as in the case of the y variable. Thus, diagram Fig. 3(b) suggests that x is a better observable than y. In the case of the z variable, there are *two* boldface paths, confirming that, from a local point of view, z is a better observable than x. Therefore, the definition put forward in this paper is consistent with the graphical approach as applied to the Lorenz system.

Other two remarks are in order before we move on to the next section. Using the former definition, the analysis of the double-scroll attractor yielded measures of observability that

TABLE II. Normalized degrees of observability δ_s^* for the double-scroll attractor for increasing values of p and, therefore, of the tolerance ϵ. These results give the observability order $x \triangleright z \triangleright y$. The same order is given by the last column of Table I.

p	δ_x^*	δ_y^*	δ_z^*
0.01	1	1	1
0.02	1	0.68	1
0.10	1	0	1
0.25	0.68	0	0

were practically zero (smaller than 10^{-15}). This was difficult to interpret, especially because of the relative simplicity of this system[29]

$$\dot{x} = \alpha(y - x - f(x)), \quad \dot{y} = x - y + z, \quad \dot{z} = -\beta y - \gamma z, \quad (15)$$

where $f(x)=bx+0.5(a-b)(|x+1|-|x-1|)$ with $a=-8/7$, $b=-5/7$, and $(\alpha, \beta, \gamma)=(9, 100/7, 0)$. One of the difficulties with this system is the great difference in the velocities of convergence and divergence along the stable and unstable manifold. Another potential problem is the application of linearization techniques to discontinuous systems.

The set of equations (15) was integrated with step $\delta_t=0.05$ from the initial condition $(x_0, y_0, z_0)=(0.1, 0.1, 0.1)$. Applying definition (6) to this system, for small values of tolerance—obtained for $p=0.01$—the following results were obtained: $\delta_x^*=\delta_y^*=\delta_z^*=1$, which is consistent with the fact that *locally* the system is linear almost everywhere, and we would expect to have "global" diffeomorphisms almost everywhere as well. As the tolerance is increased δ_y^* quickly diminishes, whereas $\delta_x^*=\delta_z^*=1$ remain. Increasing the tolerance even further, δ_z^* drops, thus giving the observability order $x \triangleright z \triangleright y$ (see Table II). This result, and in particular the poorer observability properties of the observable y has been felt in the context of data sampling,[26] modeling,[32] and topological analysis.[33]

The results shown in Table II clearly point out that the case of the double-scroll attractor is numerically badly conditioned. Increasing p (and therefore increasing the tolerance ϵ) makes the rank quickly drop to zero. This scenario was observed for this system only. For the other systems investigated, the variation of the degree of observability with the tolerance is smooth (see Fig. 4 for an example).

Finally, the use of definition (6) permits a rough comparison of general observability properties among the systems. For instance, the modeling from data and observation of hyperchaotic systems is known to be more difficult than for 3D chaotic systems. This comparison was difficult to establish using the former definitions. For instance, the degrees of observability based on the condition number (3) for the Lorenz system are much smaller than those for the Hénon–Heiles system (see Table I). On the other hand, the normalized results can be compared. Granting that any comparison, though possible, is just an indication, still the values in Table I seem to suggest that overall, the Rössler hyperchaotic system and the Hénon–Heiles are indeed less observable than the 3D systems considered.

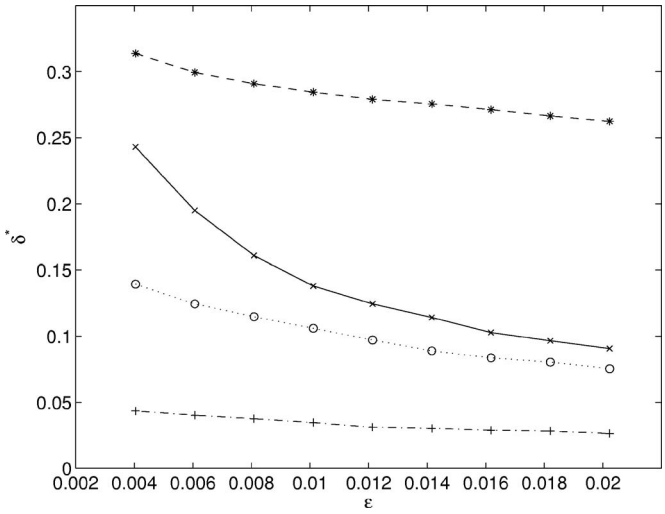

FIG. 4. Normalized degrees of observability for the hyperchaotic Rössler system. The smallest tolerance ϵ corresponds to $p=0.01$, and the largest to $p=0.05$. $(-\times-)$: δ_x^*, $(-*-)$: δ_y^*, $(\cdots\circ\cdots)$: δ_w^*, and $(\cdot-+-\cdot)$: δ_z^*.

B. Time series approach

The results presented in this section are reported in the two last column of Table I, and were obtained from long and stationary time series ($N=3\times10^4$) to avoid numerical artifacts. As before, the sampling time is $T_s=0.05$. The last column of Table I reports results obtained from data with 10% of white Gaussian noise added to the noise-free time series. As can be seen, the only effect of the noise is to de-correlate the data, but the observability order remained unchanged in every case. All the other columns in the table refer to noise-free scenarios. It should be noticed, however, that much shorter time series can be confidently used.

The Lorenz system (14) with $(\sigma, \rho, \beta)=(10.0, 28.0, 8/3)$ was considered anew. For this system, the following was obtained $1/k_{mx}=666.7$, $1/k_{my}=370.4$, and $1/k_{mz}=232.6$. Two remarks are in order for this system, as for the normalized results, the time-series approach also indicated that x is a better observable than y (see Sec. V A for a detailed discussion). However, unlike all previous definitions of degree of observability investigated by the authors, the time-series approach does indicate the z as the worse observable. This is known to be true based on symmetry considerations.[9] However, the previous definitions are all local, in character, and cannot take the symmetry into account. Because the correlation function is computed from data only, and from the z variable no symmetry can be seen, the low value of $1/k_{mz}=232.6$ is not due specifically to the symmetry, but rather to the morphology of the z time series.

To test this last hypothesis the proto-Lorenz system[34] was investigated (see Ref. 9 for results on this system using previous definitions). The main idea of the proto-Lorenz is that the dynamics are topologically the same but with no symmetry. The proto-Lorenz system is related to the Lorenz system by the following coordinate transformation

$$u = x^2 - y^2, \quad v = 2xy, \quad w = z. \quad (16)$$

This system was then simulated using the same parameters as in Ref. 9, which in turn are equivalent to those of the

Lorenz system simulated in this paper. Moreover, the integration step and the sampling time were the same as for the Lorenz system, thus making the parameter k_m comparable. For the proto-Lorenz system the following values were obtained: $1/k_{mu}=63.1$, $1/k_{mv}=116.1$, and $1/k_{mw}=235.4$, where u, v, and w correspond to x, y, and z, respectively. It should be noticed that for both the Lorenz and proto-Lorenz systems, the results for the z (w, respectively) variables are very similar $1/k_{mz}=232.6\approx 235.4=1/k_{mw}=235.4$, as expected since $z=w$. The great difference appears in the other variables, namely, in the case of the Lorenz system (that has an inversion symmetry), the variables that "see" the symmetry (x and y) have their observability increased *relative to* z. This does not happen with the proto-Lorenz, for which the w variable remains the most observable. This is an indication that the new time-series approach seems to be much more sensitive to the complexity of the couplings between the dynamical variables than to symmetries, as it operates on a window of data, which includes information of the two wings of the Lorenz system. Indeed, the couplings between dynamical variables of the proto-Lorenz system are far more complicated from the u and/or v point of view than from the w variable. It is also worth pointing out that the time-series approach yielded the observability order $w \triangleright v \triangleright u$ for the proto-Lorenz system, which is exactly the same as the order found in Ref. 9 for this system.

For the double-scroll attractor, the time-series approach proposed in this paper, based on omnidirectional nonlinear correlation functions, also suggested the same observability order as did the normalized measures of observability. Therefore, for the three 3D systems investigated (Rössler, Lorenz, and Chua's double-scroll), the results obtained with the three methods were in total agreement with the exception that only the time-series approach indicated that the z variable of the Lorenz system is the worse observable, which, in fact, is known to be true due to the symmetry.

As for the four-dimensional hyperchaotic systems investigated, some disagreement was observed. For the Hénon–Heiles system, the former (non-normalized) results and the time-series approach yielded exactly the same observability order; namely, $x \triangleright y \triangleright u \triangleright v$. On the other hand, the normalized results indicate u as the best observable, but otherwise they agree with the other methods as regards the order of the observables x, y, and v.

For the hyperchaotic Rössler system, the normalized and non-normalized results provide the same observability order $y \triangleright x \triangleright w \triangleright z$. On the other hand, the time-series approach indicates w as the best observable, maintaining the order of the other variables unchanged. In the case of this system, a glance at the time series (Fig. 5) helps understand the problem.

Variable w behaves in a way that resembles a sawtooth motion. This remark explains both why variable w is not a good observable and why the correlation-based index $1/k_m$ is incorrectly large. The long ramplike windows of w clearly induce autocorrelation, hence unduly increasing $1/k_m$. On the other hand, the w time series displays very short windows where the signal drops very quickly. This behavior accounts for the rather low observability. As the mathematical measures of the degree of observability cannot see symmetries, which typically would significantly reduce the quality of an observable, the used correlation function [see Eq. (11)] cannot distinguish a poor observable whenever it may have long ramplike windows, as for this case.

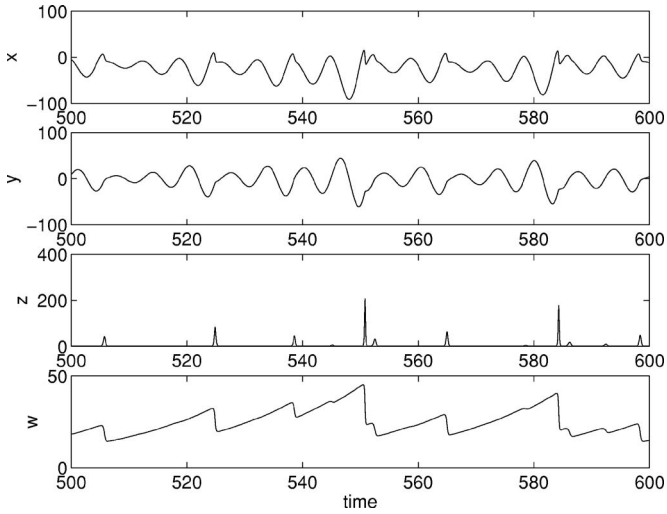

FIG. 5. Windows of the time series of the hyperchaotic Rössler system.

VI. CONCLUSION

This paper has discussed issues relating to the observability of nonlinear dynamical systems. Two approaches were presented, illustrated, and discussed. In the first part of the paper, a normalized measure of the degree of observability was defined based on the rank of the observability matrix. That matrix is the same as defined in other works. In the second part of the paper, a time-series approach to observability was proposed, based on a recently defined set of nonlinear correlation functions.

Concerning the normalized results, that were computed for five systems, the main conclusions are as follows. (i) Four out of those five systems were previously analyzed by the authors.[9,11] In two cases (Rössler and hyperchaotic Rössler) the normalized results indicate the same observability order, thus indicating consistency. (ii) The only discrepancy was for the Lorenz system for which the normalized results indicate that x is a better observable than y. As argued in the paper, this result is confirmed by the system observability diagram. (iii) The normalized results produced consistent results for the double-scroll attractor, which could not be obtained with the previous definition. (iv) The normalized results, unlike the former definition, permit a rough comparison among different systems. In this way, it is now evident that hyperchaotic systems are overall less observable than three-dimensional chaotic systems.

In the second part of the paper, a procedure based on omnidirectional nonlinear correlation functions was proposed as a way to assess the nonlinear autocorrelation of the time series. It has been argued from the literature that there is a link between correlation functions in general and the quality of an embedding. The procedure is simple and only requires a time series.

The proposed time-series approach was applied to the same five systems as before with the following main conclusions.

(i) For the 3D systems, the results were in full agreement with the results obtained for the measures of the degree of observability, including the double-scroll attractor, which is known to be particularly difficult.
(ii) For the Hénon–Heiles system, the time-series approach and the former measure of the degree of observability indicated exactly the same order. There was disagreement in the order of only one variable (u) when compared to the normalized result. The graphical approach indicates that u should not be ranked first in terms of observability order.
(iii) For the hyperchaotic Rössler system, the two definitions of observability degree gave the same results. The time approach, however, suggested the w variable as the best observable, which is hard to accept based on the graphical approach. As pointed out in the paper, this failure is a consequence of the sawtooth waveform of the time series. It is conjectured that the observability properties of time series with such a waveform will be difficult to evaluate using the time-series approach of this paper.
(iv) Because of the nature of the correlation functions used, the time-series approach is very robust to noise.
(v) In the two examples that investigated systems with symmetry (Lorenz and double-scroll), the time-series approach did rank the observables correctly. To what extent this is, in fact, due to the symmetry is a matter for further research.

Overall, the three methods agreed very well (over 92% of the calculated values are in agreement with respect to observability order). The normalized measures of the degree of observability require the knowledge of the system equations, but enables comparisons among different systems. In fact, this enabled us to verify that, as suspected before, hyperchaos is less observable in general than chaos. The time-series approach does not require any equations and does take into account symmetries in the data.

ACKNOWLEDGMENTS

This work has been partially supported by CNPq and CAPES (Brazil) and CNRS (France).

[1] J. Maquet, C. Letellier, and L. A. Aguirre, J. Theor. Biol. **228**, 421 (2004).
[2] J. Madár, J. Abonyi, and F. Szeifert, Ind. Eng. Chem. Res. **44**, 3178 (2005).
[3] A. Sitz, U. Schwarz, and J. Kurths, Int. J. Bifurcation Chaos Appl. Sci. Eng. **14**, 2093 (2004).
[4] L. A. B. Tôrres, Physica D **228**, 31 (2007).
[5] T. Kailath, *Linear Systems* (Prentice Hall, Englewood Cliffs, NJ, 1980).
[6] L. A. Aguirre, IEEE Trans. Educ. **38**, 33 (1995).
[7] H. J. Sussman and V. J. Jundjeric, J. Differ. Equations **12**, 95 (1972).
[8] A. J. Krener, SIAM J. Control **12**, 43 (1974).
[9] C. Letellier and L. A. Aguirre, Chaos **12**, 549 (2002).
[10] C. Letellier, J. Maquet, L. Le Sceller, G. Gouesbet, and L. A. Aguirre, J. Phys. A **31**, 7913 (1998).
[11] C. Letellier, L. A. Aguirre, and J. Maquet, Phys. Rev. E **71**, 056202 (2005).
[12] L. M. Hively, V. A. Protopopescu, and P. C. Gailey, Chaos **10**, 864 (2000).
[13] L. A. Aguirre and C. Letellier, J. Phys. A **38**, 6311 (2005).
[14] L. F. Zhang, Q. M. Zhu, and A. Longden, Int. J. Syst. Sci. **38**, 47 (2007).
[15] Q. M. Zhu, L. F. Zhang, and A. Longden, Automatica **43**, 1519 (2007).
[16] M. Casdagli, S. Eubank, J. D. Farmer, and J. Gibson, Physica D **51**, 52 (1991).
[17] G. H. Golub and C. F. Van Loan, *Matrix Computations*, 2nd ed. (Johns Hopkins, London, 1989).
[18] O. E. Rössler, Phys. Lett. **57A**, 397 (1976).
[19] T. Buzug and G. Pfister, Physica D **58**, 127 (1992).
[20] L. M. Pecora, L. Moniz, J. Nichols, and T. L. Carroll, Chaos **17**, 013110 (2007).
[21] M. B. Kennel, R. Brown, and H. D. I. Abarbanel, Phys. Rev. A **45**, 3403 (1992).
[22] L. Cao, Physica D **110**, 43 (1997).
[23] A. M. Fraser and H. L. Swinney, Phys. Rev. A **33**, 1134 (1986).
[24] A. M. Albano, A. Passamante, and M. E. Farrell, Physica D **54**, 85 (1991).
[25] G. Kember and A. C. Fowler, Phys. Lett. A **179**, 72 (1993).
[26] L. A. Aguirre, Phys. Lett. A **203**, 88 (1995).
[27] O. E. Rössler, Phys. Lett. **71A**, 155 (1979).
[28] M. Hénon and C. Heiles, Astron. J. **69**, 73 (1964).
[29] T. Matsumoto, L. O. Chua, and K. Tokumasu, IEEE Trans. Circuits Syst. **33**, 828 (1986).
[30] E. Lorenz, J. Atmos. Sci. **20**, 130 (1963).
[31] C. Letellier and L. A. Aguirre, Phys. Rev. E **72**, 056202 (2005).
[32] L. A. Aguirre, G. G. Rodrigues, and E. M. Mendes, Int. J. Bifurcation Chaos Appl. Sci. Eng. **7**, 1411 (1997).
[33] C. Letellier, G. Gouesbet, and N. Rulkov, Int. J. Bifurcation Chaos Appl. Sci. Eng. **6**, 2531 (1996).
[34] R. Miranda and E. Stone, Phys. Lett. A **178**, 105 (1993).