Observing chaos: Deducing and tracking the state of a chaotic system from limited observation

Paul So,[1] Edward Ott,[1,2] and W. P. Dayawansa[2,3]
[1]*Department of Physics and Laboratory for Plasma Research, University of Maryland, College Park, Maryland 20742*
[2]*Department of Electrical Engineering, University of Maryland, College Park, Maryland 20742*
[3]*Systems Research Center, University of Maryland, College Park, Maryland 20742*
(Received 7 July 1993)

A method is proposed whereby the full state vector of a chaotic system can be reconstructed and tracked using only the time series of a single observed scalar. It is assumed that an accurate mathematical description of the system is available. Noise effects on the procedure are investigated using as an example a kicked mechanical system which results in a four-dimensional dissipative map.

PACS number(s): 05.45.+b

I. INTRODUCTION

When one considers the situation where there is some experimental system behaving chaotically, one is able to accurately observe a single scalar measure of the system state. Formally, the system state is given by some vector $\mathbf{X}$ which is a function of time. The observed scalar can be expressed as some function of the system state $O=g(\mathbf{X})$. The question we ask is the following: Assuming that an accurate mathematical description of the system is available (e.g., for a discrete time system, the map is known), how can we deduce the system state $\mathbf{X}$ from measurements of O? As an example, say we have a mechanical system consisting of interconnected levers, gears, springs, etc., and this system is behaving chaotically on an attractor of a not too large dimension. Can we deduce the positions of all the parts of the system from observations of the time series of the position of just one of the levers?

One way of addressing this general problem is via the delay coordinate embedding technique. Takens [1] shows that, generically, a delay coordinate vector $(O_n, O_{n-1}, \ldots, O_{n-(N-1)})$ of sufficiently large N uniquely determines the system state $\mathbf{X}_n$. Thus, by using a computer to solve the known mathematical description of the system (assumed here to be a discrete time system), one can build up a mapping at each point on the attractor from an N-dimensional delay coordinate vector $(O_n, O_{n-1}, \ldots, O_{n-(N-1)})$ to the system state $\mathbf{X}_n$ at time n. This procedure could require the generation, storing, and searching of a large amount of data. Another way to address this problem is by utilizing the so called "extended Kalman filter," [2] which is a generalization to nonlinear systems of the usual linear Kalman filter for linear systems (see Sec. VI). By taking the statistics of noise into consideration, it can be shown that the Kalman Filter for a *linear* system is optimal in the sense that the error variance between the actual state and the estimated state is minimal. However, since the implementation of the extended Kalman filter requires, at each iterate, the manipulation of matrix equations which have the same dimension as the full dynamical system, the calculation can get quite cumbersome when the dimension of the system is large. In addition, when the Kalman filter is extended to a nonlinear system, the sense in which this method is optimal becomes unclear (c.f. Sec. V) [3]. Thus, while the embedding method and the extended Kalman filter may be useful for the purpose we address, they have drawbacks that motivate us to investigate other approaches [4]. Here, we propose a tracking technique for relatively high dimensional chaotic systems but with low dimensional attractors. The stability of our technique to the addition of small noise will also be investigated. Unlike the Kalman filter, which requires manipulation of matrix equations whose dimension is the full dimensionality of the system, the number of calculations in our method is of the order of the number of expanding directions, which may be much smaller than the system dimensionality.

The organization of this paper is as follows. In Sec. II, we will briefly review the construction of a linear observer in the standard feedback scheme. This provides the conceptual foundation for our nonlinear observer developed in Sec. III. In Sec. III, we derive the "full-order" observer for chaotic systems (the full-order observer estimates all components of the state vector from a given scalar time series). In Sec. IV, we introduce the kicked double rotor map example and use it to examine the characteristic convergence time and the basin of attraction for a single observer. Then, in Sec. V, we will introduce the "reduced-order" observer for chaotic systems. As its name suggests, the reduced-order observer is a more efficient special case of the full-order observer. We also apply this reduced-order observer to the kicked double rotor map example to demonstrate the effect of the addition of noise to the system and/or to the output function. Next, Sec. VI will compare the performance of the extended Kalman filter with our nonlinear observer technique. Section VII will provide a summary of our chaotic observer technique and a discussion of its advantages and drawbacks. Appendix A provides some generalizations of the discussion in the text. Finally, we note that, since we wish this paper to be understandable to researchers in chaotic dynamics, we have not assumed

prior knowledge of control theory or signal processing. The reader possessing such knowledge should skip the background material (e.g., Sec. II) provided in these fields.

II. LINEAR OBSERVERS

In *linear* control theory, it is possible to estimate unmeasured state variables using a "state observer." To be specific, consider a linear time independent d-dimensional system,

$$\mathbf{X}_{n+1}=\mathbf{A}\mathbf{X}_n, \quad O_n=\mathbf{G}\mathbf{X}_n \ , \tag{1}$$

where $\mathbf{X}$ is a d-dimensional column vector, $\mathbf{A}$ is a constant $d\times d$ matrix, and $\mathbf{G}$ is a constant d-dimensional row vector. The scalar function O_n is the observed physical output of the system. This system is observable at time n if it is possible to determine the system state $\mathbf{X}_n$ from the observation of outputs over a finite time interval. From a series of d measurements $(O_n,\ldots,O_{n+(d-1)})$, one can determine all the d components of the state vector $\mathbf{X}_n$ by solving the following matrix equation:

$$\begin{bmatrix} O_n \\ \vdots \\ O_{n+(d-1)} \end{bmatrix} = \mathbf{J} \begin{bmatrix} X_n^1 \\ \vdots \\ X_n^d \end{bmatrix}, \quad \text{where } \mathbf{J}= \begin{bmatrix} \mathbf{G} \\ \boldsymbol{GA} \\ \vdots \\ \boldsymbol{GA}^{n-1} \end{bmatrix} .$$

This equation has a unique solution if and only if the observability matrix $\mathbf{J}$ is of rank d. This is the observability condition introduced by Kalman for a linear time invariant system. Assuming the system to be observable, one can then reconstruct the actual state $\mathbf{X}_n$ of the system from a time series of the scalar output O_n using a state observer, defined by

$$\hat{\mathbf{X}}_{n+1}=\mathbf{A}\hat{\mathbf{X}}_n+\mathbf{C}[O_{n+1}-\hat{O}_{n+1}] \ , \quad \text{where } \hat{O}_{n+1}=\boldsymbol{GA}\hat{\mathbf{X}}_n \ . \tag{2}$$

The idea of this technique is to choose the control vector $\mathbf{C}$ such that the numerically generated state $\hat{\mathbf{X}}_n$ will converge to the actual state $\mathbf{X}_n$ with increasing n. To derive the necessary condition for this to happen, one can look at the dynamics of the error equation,

$$\mathbf{X}_{n+1}-\hat{\mathbf{X}}_{n+1}=[\,\mathbf{A}-\mathbf{C}\mathbf{G}\mathbf{A}\,](\mathbf{X}_n-\hat{\mathbf{X}}_n) \ , \tag{3}$$

obtained by subtracting (2) from (1). If the control vector $\mathbf{C}$ can be chosen so that the magnitudes of the eigenvalues of $[\,\mathbf{A}-\mathbf{C}\mathbf{G}\mathbf{A}\,]$ are all less than one, then the error will exponentially decrease to zero as n approaches infinity. A standard technique exists for choosing the control vector $\mathbf{C}$ to do this, and can be found in many control theory textbooks (e.g., see Ref. [5]).

III. CHAOTIC OBSERVERS—FULL ORDER

The general procedure for our chaotic observer technique is similar in spirit to the design of an observer in *linear* control theory. Conceptually, the observer is built upon a numerical copy of the actual system but with an additional time dependent correction term which compares the actual output of the chaotic system to the estimated output of the observer. Depending on the difference between the actual and the estimated output, the time dependent parameters in the correction term are adjusted so that the difference will exponentially decay to zero with time (Kalman filters also have this structure; see Sec. VI). In this section, we shall present a general procedure for doing this in the case of *nonlinear chaotic* systems.

We assume that the chaotic system that we want to observe is given by the following equations:

$$\mathbf{X}_{n+1}=\mathbf{M}(\mathbf{X}_n), \quad O_n=g(\mathbf{X}_n) \ . \tag{4}$$

Here $\mathbf{M}$ and g are *nonlinear* functions of the d-dimensional vector $\mathbf{X}_n$. The corresponding state observer is taken to be

$$\hat{\mathbf{X}}_{n+1}=\mathbf{M}(\hat{\mathbf{X}}_n)+\mathbf{C}_n[O_{n+1}-\hat{O}_{n+1}] \ , \tag{5}$$

where $\hat{O}_{n+1}=g(\mathbf{M}(\hat{\mathbf{X}}_n))$ and $\mathbf{C}_n$ is a *time-dependent* d-dimensional control column vector which we need to adjust at each iterate. Subtracting the equations for $\mathbf{X}_{n+1}$ and $\hat{\mathbf{X}}_{n+1}$ yields the error equation

$$\mathbf{X}_{n+1}-\hat{\mathbf{X}}_{n+1}=\mathbf{M}(\mathbf{X}_n)-\mathbf{M}(\hat{\mathbf{X}}_n) -\mathbf{C}_n[g(\mathbf{M}(\mathbf{X}_n))-g(\mathbf{M}(\hat{\mathbf{X}}_n))] \ . \tag{6}$$

Linearizing about $\hat{\mathbf{X}}_n$ gives

$$\delta\mathbf{X}_{n+1}=[\mathbf{D}\mathbf{M}(\hat{\mathbf{X}}_n)-\mathbf{C}_n\mathbf{D}_g(\mathbf{M}(\hat{\mathbf{X}}_n))\mathbf{D}\mathbf{M}(\hat{\mathbf{X}}_n)]\delta\mathbf{X}_n \ , \tag{7}$$

where $\mathbf{X}_n-\hat{\mathbf{X}}_n=\delta\mathbf{X}_n$ is a differential, and $\mathbf{D}\mathbf{M}(\hat{\mathbf{X}}_n)$ and $\mathbf{D}g(\mathbf{M}(\hat{\mathbf{X}}_n))$ are the derivatives of $\mathbf{M}(\hat{\mathbf{X}}_n)$ and $g(\mathbf{M}(\hat{\mathbf{X}}_n))$, respectively, with $\mathbf{D}\mathbf{M}$ a d by d matrix and $\mathbf{D}g$ a d-dimensional row vector. Recalling our discussion of observers from linear time independent systems, the matrix $[\,\mathbf{A}-\mathbf{C}\mathbf{G}\mathbf{A}\,]$ was a constant, and the long term evolution of the observer error is determined by $[\,\mathbf{A}-\mathbf{C}\mathbf{G}\mathbf{A}\,]^n$. This converges to zero with increasing n if the eigenvalues of $[\,\mathbf{A}-\mathbf{C}\mathbf{G}\mathbf{A}\,]$ have magnitudes less than 1. In the chaotic case, however, the long term behavior of the error is governed by the product of matrices of the form $[\mathbf{D}\mathbf{M}(\hat{\mathbf{X}}_n)-\mathbf{C}_n\mathbf{D}_g(\mathbf{M}(\hat{\mathbf{X}}_n))\mathbf{D}\mathbf{M}(\hat{\mathbf{X}}_n)]$ which change at each iterate,

$$\delta\mathbf{X}_{n+1}=\prod_{m=0}^{n}[\mathbf{D}\mathbf{M}(\hat{\mathbf{X}}_m) -\mathbf{C}_m\mathbf{D}_g(\mathbf{M}(\hat{\mathbf{X}}_m))\mathbf{D}\mathbf{M}(\hat{\mathbf{X}}_m)]\delta\mathbf{X}_0 \ . \tag{8}$$

While one can adjust each individual matrix at each iterate to have eigenvalues with magnitudes less than 1, this does not guarantee that the product goes to zero as n goes to infinity [6]. Below we give a procedure which yields convergence of our observer in the chaotic case.

For specificity of the discussion, we will assume that the chaotic attractor of our system is hyperbolic with two positive Lyapunov exponents and the rest negative. Thus, the tangent space at each point on the attractor can be decomposed into the sum of a two dimensional un-

stable subspace and a $(d-2)$-dimensional stable subspace. Noting that $\mathbf{DM}(\hat{\mathbf{X}}_n)$ maps the unstable subspace at $\hat{\mathbf{X}}_n$ into the unstable subspace at $\mathbf{M}(\hat{\mathbf{X}}_n)$ and similarly maps the stable subspace at $\hat{\mathbf{X}}_n$ into the stable subspace at $\mathbf{M}(\hat{\mathbf{X}}_n)$, we see that if $\mathbf{C}_n$ is chosen to lie in the unstable subspace at $\hat{\mathbf{X}}_n$, then the action of the $\mathbf{C}_n\mathbf{D}_g(\mathbf{M}(\hat{\mathbf{X}}_n))\mathbf{DM}(\hat{\mathbf{X}}_n)$ term in Eq. (8) will stay within the unstable subspace to $\mathbf{M}(\hat{\mathbf{X}}_n)$. Therefore, in a suitably chosen basis, the matrix representation of $[\mathbf{DM}(\hat{\mathbf{X}}_n)-\mathbf{C}_n\mathbf{D}_g(\mathbf{M}(\hat{\mathbf{X}}_n))\mathbf{DM}(\hat{\mathbf{X}}_n)]$ in Eq. (8) can be written in the following block form:

$$\begin{bmatrix} \mathbf{U}_n & \mathbf{W}_n \\ 0 & \mathbf{S}_n \end{bmatrix} .$$

In this representation, $\mathbf{U}_n$ is a 2×2 submatrix acting on the unstable subspace, $\mathbf{S}_n$ is a $(d-2)\times(d-2)$ submatrix acting on the stable subspace, and $\mathbf{W}_n$ is a $2\times(d-2)$ submatrix taking vectors from the stable subspace into the unstable subspace. One should note that $\mathbf{U}_n$ and $\mathbf{W}_n$ are functions of the yet to be determined control vector C_n, and that $\mathbf{S}_n$ is known [it is given by $\mathbf{DM}(\hat{\mathbf{X}}_n)$ to be restricted to the stable subspace].

We now can analyze the convergence of Eq. (8) by examining the product of n of these block matrices, which is given by the following formula:

$$\begin{bmatrix} \mathbf{U}_n\mathbf{U}_{n-1}\cdots\mathbf{U}_1 & \displaystyle\sum_{i=1}^{n}\left[\prod_{j=i+1}^{n}\mathbf{U}_j\mathbf{W}_i\prod_{k=1}^{i-1}\mathbf{S}_k\right] \\ 0 & \mathbf{S}_n\mathbf{S}_{n-1}\cdots\mathbf{S}_1 \end{bmatrix} . \tag{9}$$

Since the product $\mathbf{S}_n\mathbf{S}_{n-1}\cdots\mathbf{S}_1$ goes to zero as $n\to\infty$, the convergence of the product matrix, Eq. (9), depends on the product $\mathbf{U}_n\mathbf{U}_{n-1}\cdots\mathbf{U}_n$ and on the off diagonal term $\sum_{i=1}^{n}[\prod_{j=i+1}^{n}\mathbf{U}_j\mathbf{W}_i\prod_{k=1}^{i-1}\mathbf{S}_k]$. It is demonstrated in Appendix A that if one chooses the control vector $\mathbf{C}_n$ so that the product $\mathbf{U}_n\mathbf{U}_{n-1}\cdots\mathbf{U}_1$ goes to zero as $n\to\infty$, then the whole matrix (9) will go to zero as well.

To make the product $\mathbf{U}_n\mathbf{U}_{n-1}\cdots\mathbf{U}_1\to 0$ as $n\to\infty$, first consider each matrix $\mathbf{U}_n$ to be lower triangular (i.e., with $U_{ij}=0$ for $i<j$); then we have the following: (i) the product of two or more such matrices will still be lower triangular; (ii) the eigenvalues are just the diagonal elements; and, most importantly, (iii) the eigenvalues of the product of such matrices will be the product of their respective eigenvalues. Thus, if we choose $\mathbf{C}_n$ so that $\mathbf{U}_n$ is lower triangular with eigenvalues of magnitude less than 1, then the product of the $\mathbf{U}_n$'s will also be lower triangular with eigenvalues of magnitude less than 1. Most importantly, it can be shown that the product $\mathbf{U}_n\mathbf{U}_{n-1}\cdots\mathbf{U}_1\to 0$ as $n\to\infty$ (see Appendix A). Thus, we want to choose a basis for the unstable subspace, such that each of the $\mathbf{U}_n$ will be lower triangular, while simultaneously choosing $\mathbf{C}_n$, such that the eigenvalues of $\mathbf{U}_n$ will have magnitudes less than 1.

The first step in our procedure is to define two numbers, $\lambda_n^{(1)}$ and $\lambda_n^{(2)}$, and two basis unit column vectors, $\mathbf{e}_n^{(1)}$ and $\mathbf{e}_n^{(2)}$, for the unstable subspace at $\hat{\mathbf{X}}_n$ according to the following iterative procedure:

$$\lambda_n^{(1)}\mathbf{e}_{n+1}^{(1)}=[\mathbf{DM}(\hat{\mathbf{X}}_n)]\mathbf{e}_n^{(1)} , \tag{10a}$$

$$\lambda_n^{(2)}\mathbf{e}_{n+1}^{(2)}=[\mathbf{DM}(\hat{\mathbf{X}}_n) -\{C_n^{(1)}\mathbf{e}_{n+1}^{(1)}\}\mathbf{Dg}(\mathbf{M}(\hat{\mathbf{X}}_n))\mathbf{DM}(\hat{\mathbf{X}}_n)]\mathbf{e}_n^{(2)} , \tag{10b}$$

with $\mathbf{e}_0^{(1)}$ and $\mathbf{e}_0^{(2)}$ initialized from any two linearly independent column vectors in the unstable subspace of $\mathbf{DM}(\mathbf{X}_0)$. Here, $\lambda_n^{(1)}$ and $\lambda_n^{(2)}$ are the normalization factors associated with $\mathbf{e}_{n+1}^{(1)}$ and $\mathbf{e}_{n+1}^{(2)}$, respectively. We want to remind the reader that in this nonorthogonal basis, the components of $\mathbf{U}_n$ are given by the "inner product," $(\mathbf{U}_n)_{ij}=\mathbf{f}_{n+1}^{(i)}\mathbf{U}_n\mathbf{e}_n^{(j)}$, where the contravariant row vectors $\mathbf{f}_{n+1}^{(i)}$ are defined by $\mathbf{f}_{n+1}^{(i)}\mathbf{e}_{n+1}^{(j)}=\delta_{ij}$. With this basis representation, $\mathbf{U}_n$ will be in a lower triangular form with two free parameters $C_n^{(1)}$ and $C_n^{(2)}$:

$$\mathbf{U}_n=\begin{bmatrix} \lambda_n^{(1)}-C_n^{(1)}Dh_n^{(1)} & 0 \\ -C_n^{(2)}Dh_n^{(1)} & \lambda_n^{(2)}-C_n^{(2)}Dh_n^{(2)} \end{bmatrix} , \tag{11}$$

where

$$Dh_n^{(i)}=\mathbf{Dg}(\mathbf{M}(\hat{\mathbf{X}}_n))\mathbf{DM}(\hat{\mathbf{X}}_n)\mathbf{e}_n^{(i)}$$

and $C_n^{(i)}=\mathbf{f}_{n+1}^{(i)}\mathbf{C}_n$. If we adjust $C_n^{(1)}$ and $C_n^{(2)}$ so that the eigenvalues (i.e., the diagonal elements) of $\mathbf{U}_n$ are less than 1, then the product of the matrices in Eq. (8) will converge to zero as n increases.

A possible concern with our method, as outlined above, is that, as time n increases, the vectors $\mathbf{e}_n^{(1)}$ and $\mathbf{e}_n^{(2)}$ might tend to become more and more nearly parallel. (This would invalidate our procedure since we assume that $\mathbf{e}_n^{(1)}$ and $\mathbf{e}_n^{(2)}$ span the two-dimensional unstable subspace.) We note, however, that we have the freedom of choosing the eigenvalues of $\mathbf{U}_n$, $\Lambda_n^{(i)}=\lambda_n^{(i)}-C_n^{(i)}Dh_n^{(i)}$. If we choose $\Lambda_n^{(1)}$ to be zero, then the collapse of $\mathbf{e}_n^{(1)}$ and $\mathbf{e}_n^{(2)}$ to a common direction can be prevented (see next paragraph), still leaving open one degree of freedom in choosing $\Lambda_n^{(2)}$. For definiteness and faster convergence, we set $\Lambda_n^{(2)}$ as well as $\Lambda_n^{(1)}$ to zero, in which case we have

$$\begin{aligned}\mathbf{C}_n&=C_n^{(1)}\mathbf{e}_{n+1}^{(1)}+C_n^{(2)}\mathbf{e}_{n+1}^{(2)}\\ &=(\lambda_n^{(1)}/Dh_n^{(1)})\mathbf{e}_{n+1}^{(1)}+(\lambda_n^{(2)}/Dh_n^{(2)})\mathbf{e}_{n+1}^{(2)} .\end{aligned} \tag{12}$$

The expression in Eq. (12) for the control vector $\mathbf{C}_n$ is valid as long as the denominators $Dh_n^{(i)}$ are not zero. In our numerical program, we set a minimum value such that whenever $Dh_n^{(i)}$ falls below that value, we set the control vector to zero. Thus, when $Dh_n^{(i)}$ is small, we do not attempt to bring $\mathbf{X}$ and $\hat{\mathbf{X}}$ together. But, if they were already close, they will still be close one iterate later. Hence, little is lost by turning the control off for one iterate, provided that this is done only infrequently.

To see that $\mathbf{e}_n^{(1)}$ and $\mathbf{e}_n^{(2)}$ do not typically approach a common direction as n increases, we assume that they are nearly parallel at time n, and then demonstrate that Eq. (10) and the expression for $C_n^{(1)}$ in Eq. (12) imply that they are not nearly parallel at time $n+1$. Setting $\mathbf{e}_n^{(2)}=\mathbf{e}_n^{(1)}+\delta\mathbf{e}_n$ with $|\delta\mathbf{e}_n|\ll 1$, Eqs. (10) and (12) yield

$$\lambda_n^{(2)}\mathbf{e}_{n+1}^{(2)}=\mathbf{DM}(\hat{\mathbf{X}}_n)\delta\mathbf{e}_n -\lambda_n^{(1)}\mathbf{e}_{n+1}^{(1)}\{[\mathbf{Dg}(\mathbf{M}(\hat{\mathbf{X}}_n))\mathbf{DM}(\hat{\mathbf{X}}_n)\delta\mathbf{e}_n]/Dh_n^{(1)}\} .$$

Thus $\mathbf{e}_{n+1}^{(2)}$ consists of two terms which are both typically of order $\delta\mathbf{e}_n$. The first term points in the direction $\mathbf{DM}(\hat{\mathbf{X}}_n)\delta\mathbf{e}_n$, which is different from the direction of $\mathbf{e}_{n+1}^{(1)}$ because $\delta\mathbf{e}_n$ is approximately perpendicular to $\mathbf{e}_n^{(1)}$ by assumption. The second term points in the direction $\mathbf{e}_{n+1}^{(1)}$ [which is the direction of $\mathbf{DM}(\hat{\mathbf{X}}_n)\mathbf{e}_n^{(1)}$]. Normally, since both terms are of order $\delta\mathbf{e}_n$, $\mathbf{e}_{n+1}^{(2)}$ will in general point in a different direction than $\mathbf{e}_{n+1}^{(1)}$. However, one can see a problem in the rare cases when $Dh_n^{(1)}$ is small. In these cases, the second term dominates the first term and we will have $\mathbf{e}_{n+1}^{(2)}$ nearly parallel to $\mathbf{e}_{n+1}^{(1)}$, but we avoid these "glitches" anyway by turning off the control vector, i.e., $\mathbf{C}_n \equiv 0$.

Although the above discussion is in the context of a two-dimensional unstable subspace, we note that no essential change is produced in the case where the unstable subspace has an arbitrary dimension d_u. In particular, Eq. (10) generalizes in a natural way to d_u equations for the d_u basis vectors $\mathbf{e}^{(i)}$. This generalization is given in Appendix B. Also treated is the generalization to the case where the observation is a vector $O=\mathbf{g}(\mathbf{X})$ of dimension $1<d_0<d$.

Since our method reconstructs the full state vector of the system, i.e., all d components of the d-dimensional vector $\mathbf{X}_n$, it is called the full-order observer. Computationally, since our method requires the manipulation of only d_u equations for the d_u basis vectors $\mathbf{e}^{(i)}$ at each iterate, the calculation for the control vector is relatively simple when $d \gg d_u$. In contrast, the extended Kalman filter requires the manipulation of $d \times d$ matrices at each iterate. See Sec. VI for a comparison of our technique with a extended Kalman filter.

IV. CONVERGENCE CHARACTERISTICS AND THE KICKED DOUBLE ROTOR EXAMPLE

To access the convergence characteristics of our chaotic observer technique, we will use a four dimensional map which describes the time evolution of a mechanical system, called the kicked double rotor [7]. As shown in the illustration, Fig. 1, the kicked double rotor consists of two massless rods of lengths L_1 and L_2 connected at the pivot P_2 and with the other end of rod 1 connected to a fixed pivot at P_1. Point masses m_1 and $m_2/2$ are attached at the end of rod 1 and the two ends of rod 2 as shown. At one of the ends of rod 2, an impulse force is applied at times $t=0, T, 2T, \ldots$, $\mathbf{f}(t) = f_0 \sum_{n=0}^{\infty} \delta(t-nT)\hat{\mathbf{y}}$. The kicked double rotor is governed by the following set of equations:

$$\mathbf{X}_{n+1} = \begin{bmatrix} \theta_{n+1} \\ \dot{\theta}_{n+1} \end{bmatrix} = \begin{bmatrix} \mathbf{K}\dot{\theta}_n + \theta_n \\ \mathbf{L}\dot{\theta}_n + \mathbf{G}(\theta_{n+1}) \end{bmatrix}, \tag{13}$$

where $\theta=(\theta_1,\theta_2)^{\dagger}$, $\dot{\theta}=(\dot{\theta}_1,\dot{\theta}_2)^{\dagger}$, and $\mathbf{G}(\theta)=(a_1\sin\theta_1, a_2\sin\theta_2)^{\dagger}$ (here $\dagger$ denotes transpose). θ_1 and θ_2 are angle variables giving the positions of the rotor arms, and $\dot{\theta}_1$ and $\dot{\theta}_2$ are the angular velocities of the rotor arms at the instant immediately after the nth kick. $a_{1,2}=(f_0/I)L_{1,2}$ are constants proportional to the strength of the periodic kick $\mathbf{f}(t)$. The moments of inertia about pivots 1 and 2 are chosen to be equal, $I=(m_1+m_2)L_1^2=m_2L_2^2$. $\mathbf{K}$ and $\mathbf{L}$ are constant matrices defined by

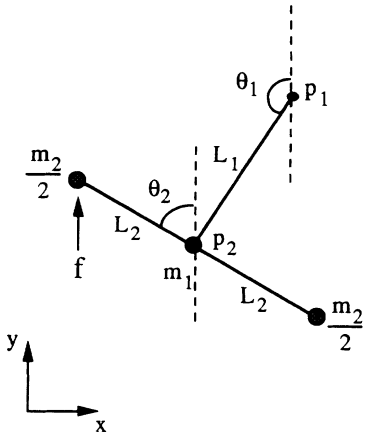

FIG. 1. The double rotor.

$$\mathbf{L}=\sum_{i=1}^{2}\mathbf{W}_i e^{\zeta_i T}, \quad \mathbf{K}=\sum_{i=1}^{2}\mathbf{W}_i\frac{e^{\zeta_i T}-1}{\zeta_i},$$

$$\mathbf{W}_1=\begin{bmatrix}\alpha & \beta\\ \beta & \sigma\end{bmatrix}, \quad \mathbf{W}_2=\begin{bmatrix}\sigma & -\beta\\ -\beta & \alpha\end{bmatrix},$$

$$\alpha=\frac{1}{2}\left[1+\frac{\nu_1}{\Delta}\right], \quad \sigma=\frac{1}{2}\left[1-\frac{\nu_1}{\Delta}\right], \quad \beta=-\frac{\nu_2}{\Delta},$$

$$\zeta_{1,2}=-\tfrac{1}{2}(\nu_1+2\nu_2\pm\Delta), \quad \Delta=(\nu_1^2+4\nu_2^2)^{1/2},$$

where ν_1 and ν_2 are the friction coefficients at the pivots (see Fig. 1 and Ref. [7]). In our numerical experiment, we used $g(\mathbf{X})=\theta_2$ and have chosen a particular set of values for the physical parameters [8]. The resulting chaotic attractor has one positive Lyapunov exponent ($\Lambda_L^{(1)}=0.670$, $\Lambda_L^{(2)}=-0.404$, $\Lambda_L^{(3)}=-1.192$, $\Lambda_L^{(4)}=-2.074$). Hence $\Lambda_L^{(1)}+\Lambda_L^{(2)}>0$ while $\Lambda_L^{(1)}+\Lambda_L^{(2)}+\Lambda_L^{(3)}<0$, and the Lyapunov dimension [9] of the attractor is thus given by

$$d_L=2+(\Lambda_L^{(1)}+\Lambda_L^{(2)})/\Lambda_L^{(3)},$$

which approximately yields 2.22. It is interesting to note that, although the attractor has only one positive Lyapunov exponent, there exist periodic orbits on the attractor with two dimensional unstable tangent spaces, and other periodic orbits on the attractor with one dimensional unstable tangent spaces [10]. Thus, the kicked double rotor map is not globally hyperbolic as assumed in the previous theoretical discussion. Nevertheless, we find that our method still works, and the performance is better if we use two basis vectors [as in Eqs. (10)–(12)].

First, we examine the behavior of the observer if the observer test orbit starts inside the linear region of the true orbit. The characteristic time for this case will, in principle, be dependent upon the chosen eigenvalues of $\mathbf{U}_n$. Ideally [i.e., if the dynamics were truely described by Eq. (7)], if we choose the eigenvalues of $\mathbf{U}_n$ to all be zero and neglect nonlinear effects, then the observer error along the unstable direction should vanish in two steps [11]. However, since the action of the U's is only a linear approximation of the true dynamics of the chaotic sys-

tem, the average number of iterates [12] needed to bring the separation between the observer orbit and the true orbit from 5.0×10^{-3} down to 1.0×10^{-7} in a normalized unit is about 10. Furthermore, when we increase the magnitude of the eigenvalue [13] $\Lambda_n^{(2)}$, the number of iterates needed to bring the separation down increases as expected. Figure 2 is a graph showing the average converging time of this chaotic observer as a function of $\Lambda_n^{(2)}$. For the fastest convergence time, it is in principal desirable to choose all eigenvalues of $\mathbf{U}_n$ to be zero as done in Eq. (12).

We now discuss the convergence characteristics of our observer technique for observer initial conditions outside the linear region. Typically, an observer orbit $\hat{\mathbf{X}}_n$ begins to track the true orbit $\mathbf{X}_n$ when $\hat{\mathbf{X}}_n$ is located within the linear region of the map $\mathbf{M}(\mathbf{X}_n)$. Furthermore, if we wait long enough, an initially nontracking observer orbit $\hat{\mathbf{X}}_n$ will typically and eventually fall within the linear region of $\mathbf{M}(\mathbf{X}_n)$ at some future time n. In our numerical experiments, we found that the average transient time before tracking sets in could be quite long. To remedy this problem, we use many observer test points with randomly chosen initial conditions on the attractor, and we continuously test each one to see if it has locked onto the true orbit $\mathbf{X}_n$. We do this by calculating $g(\mathbf{X}_n)-g(\hat{\mathbf{X}}_n)$ and declaring the orbits $\mathbf{X}_n$ and $\hat{\mathbf{X}}_n$ locked if this quantity is small for several successive iterates. We then take the observer state as $\hat{\mathbf{X}}_n$ for such a locked orbit. We can estimate the typical number of observer test orbits needed by determining the average convergence time $\langle\tau\rangle$ of a single observer. To determine this average time $\langle\tau\rangle$, we begin with a large number of randomly chosen observer test orbits. Then, a semilog plot of the number $\hat{N}$ of orbits which are still not tracking the true orbit after a time interval n is generated. Since the number of such orbits typically decays exponentially with n, i.e., $\hat{N}(n)=N_0e^{-n/\langle\tau\rangle}$, the inverse of the slope of this graph defines an average time $\langle\tau\rangle$ needed for an observer orbit to converge to the true orbit. In Fig. 3, we used 5000 randomly chosen observer test orbits and we estimated $\langle\tau\rangle\sim1500$. This value also gives a reasonable value for the number of observer test orbits needed so that at least one observer orbit will be tracking the true orbit after the first few time steps.

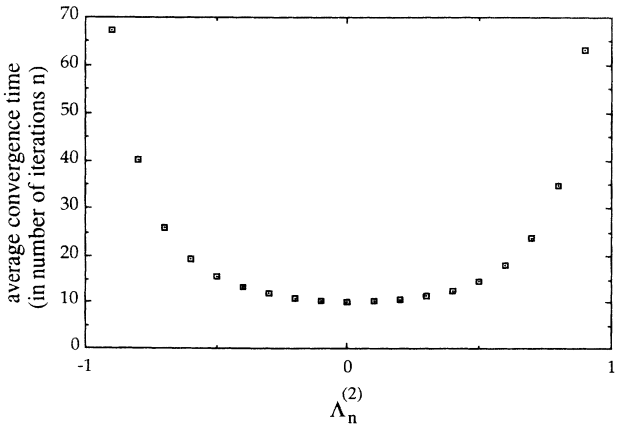

FIG. 2. Averaged convergence time (in numbers of iteration steps n) vs the magnitude of the second eigenvalue of $\mathbf{U}_n$, $\Lambda_n^{(2)}$. The convergence time is defined to be the number of iterates that the observer takes to reduce $|\mathbf{X}_n-\hat{\mathbf{X}}_n|$ 10 000 times (from 5.0×10^{-3} to 1.0×10^{-7}) and the average value is taken over 5000 randomly chosen observers.

V. CHAOTIC OBSERVER—REDUCED ORDER

A variant of our technique discussed in Sec. III can be formulated if there exists an invertible coordinate transformation

$$\mathbf{T}:\mathbf{X}\equiv(X^1,\ldots,X^{d-1},X^d)\rightarrow(\mathbf{Y}|Z)\equiv(Y^1,\ldots,Y^{d-1},Z),$$

where Z is the observed quantity $g(\mathbf{X}_n)$. In this case, the observer only needs to estimate a $(d-1)$-dimensional vector $\mathbf{Y}$ since Z is known from direct measurement. The state equation $\mathbf{X}_{n+1}=\mathbf{M}(\mathbf{X}_n)$ under this coordinate transformation $\mathbf{T}$ can be written as

$$\mathbf{Y}_{n+1}=\mathbf{M}_Y(\mathbf{Y}_n,Z_n) \text{ and } Z_{n+1}=M_Z(\mathbf{Y}_n,Z_n), \quad (14)$$

where $(\mathbf{M}_Y(\,),M_Z(\,))$ is the representation of $\mathbf{M}(\,)$ in the new coordinate system. The reduced-order observer for the unmeasured part of the state vector $\mathbf{Y}_n$ can then be defined as $\hat{\mathbf{Y}}_{n+1}=\mathbf{M}_Y(\hat{\mathbf{Y}}_n,Z_n)+\mathbf{C}_n[Z_{n+1}-M_Z(\hat{\mathbf{Y}}_n,Z_n)]$, where $\mathbf{C}_n$ is the $(d-1)$-dimensional control vector corresponding to the reduced-order observer. Forming the error equation, we have

$$\delta\mathbf{Y}_{n+1}=[\mathbf{D}\mathbf{M}_Y(\mathbf{Y}_n,Z_n)-\mathbf{C}_n\mathbf{D}M_Z(\mathbf{Y}_n,Z_n)]\delta\mathbf{Y}_n, \quad (15)$$

which can be treated using exactly the same techniques as already discussed. A similar expression for the control vector $\mathbf{C}_n$ can be derived as before [Eq. (12)], but with $\mathbf{D}_g(\mathbf{M}(\hat{\mathbf{X}}_n))\mathbf{D}\mathbf{M}(\hat{\mathbf{X}}_n)$ replaced by $\mathbf{D}M_Z(\mathbf{Y}_n,Z_n)$. The main obvious advantage achieved by doing this is that the dimensionality of the observer is reduced by 1.

From a different viewpoint, the reduced-order observer and the actual system can be considered as a pair of coupled systems with the actual system providing the driving signal and the observer as the response function. This pair of driven-response systems reduce to the one studied

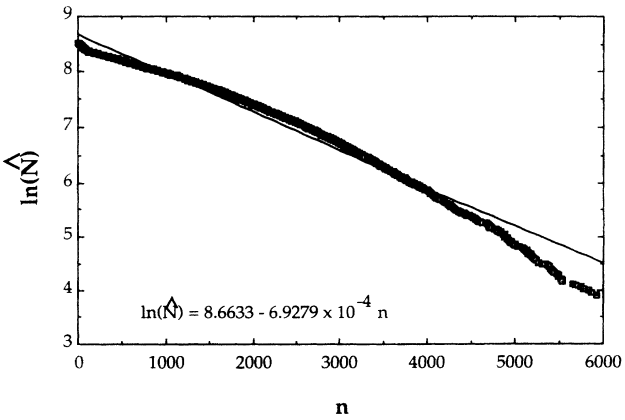

FIG. 3. $\ln(\hat{N})$ vs n with 5000 randomly chosen initial observer test points. An observer test orbit is said to be tracking the true orbit when $|\mathbf{X}_n-\hat{\mathbf{X}}_n|<1.0\times10^{-6}$.

by Pecora and Carroll [14] when the feedback control in our reduced-order observer is turn off ($\mathbf{C}_n \equiv \mathbf{0}$). The resultant observer in this case has also been called a "trivial reduced-order observer." The convergence of this trivial reduced order observer will obviously depend on the Lyapunov exponents of the system $\mathbf{Y}_{n+1}=\mathbf{M}_Y(\mathbf{Y}_n,\mathbf{Z}_n)$. Since $\mathbf{M}_Y(\ )$ is basically a subsystem of the full dynamical system $\mathbf{M}(\ )$, the number of positive Lyapunov exponents of $\mathbf{M}(\ )$ cannot be larger than the number of positive Lyapunov exponents of $\mathbf{M}_Y(\ )$. Pecora and Carroll [14] discuss the case where a physically constructed subsystem $\mathbf{M}_Y(\ )$ has no positive Lyapunov exponents, and it is then possible to synchronize $\mathbf{M}_Y(\ )$ with the chaotic signal generated by $\mathbf{M}(\ )$.

Returning now to our discussion of our reduced-order observer and using the double rotor map example (with the same set of parameters as in the full-order observer), we calculated the Lyapunov spectrums for the four different choices, $g(\mathbf{X}_n)=\theta_1$, θ_2, $\dot{\theta}_1$, and $\dot{\theta}_2$. In all these cases, $\mathbf{T}_Y(\mathbf{X})$ will simply be a projection onto the unmeasured components of $\mathbf{X}$. The results are given in Table I. For our numerical experiment, we have chosen $g(\mathbf{X}_n)=\theta_2$. In this case, the subsystem has one positive Lyapunov exponent so that synchronism as defined by Pecora and Carroll does not apply. However, as mentioned earlier, since the double rotor map is nonhyperbolic, we found that the performance of the reduced-order observer is better if we used two basis vectors (rather than one) in our calculation for the control vector $\mathbf{C}_n$.

Noise arises naturally in any real physical systems. For example, the strength of the kick in our double rotor might fluctuate because of nonuniformity of the motor output or there might be imprecision in the measuring device for the scalar time series. The simplest way to model noise into our system is to put an additive term $\varepsilon_a\delta_n$ to the right hand side of the double rotor map Eq. (4) to denote noise in the actual system and another additive term to the output function $O_n=g(X_n)+\varepsilon_b\nu_n$ to denote noise in the measurement of the observed quantity. Here, $\varepsilon_{a,b}$ are the maximum magnitudes of the noise. The components of δ_n and ν_n are uncorrelated random variables with a given noise statistics. In our numerical example, both of these random variables are chosen to have zero mean and uniform distribution in $|\nu|<1, |\delta|<1$. In addition, we assume that the observer system has no knowledge of the noise. We expect the method to work well when $|\varepsilon\delta_n|$ and $|\varepsilon\nu_n|$ are less than the typical radius of the linear region of the map. Since the chaotic system is nonlinear, the size of the linear region at each point $\mathbf{X}_n$ varies. Near those points where the linear region is smaller, the observer will be more sensitive to noise. In our numerical experiment, we found that for a sufficiently small value of $\varepsilon(<10^{-4})$, the observer was able to track the actual state continuously. However, as ε increases, the probability of the observer being kicked out of the linear region of the map $\mathbf{M}(\mathbf{X}_n)$ increases. When this happens, the observer orbit might temporary lose track of the true system state but we can quickly lock back onto the actual orbit again by first going back a few iterates to a point where the observer and the actual state are still close together. Then, we activate a set of N observer test points randomly chosen within a neighborhood centered on that past iterate of the observer. When one of these N observer test points begins to lock onto the actual state, we pick that particular observer test point as our new observed orbit and drop the rest of the N test points. On the other hand, if none of the N observer test points locks back onto the actual or-

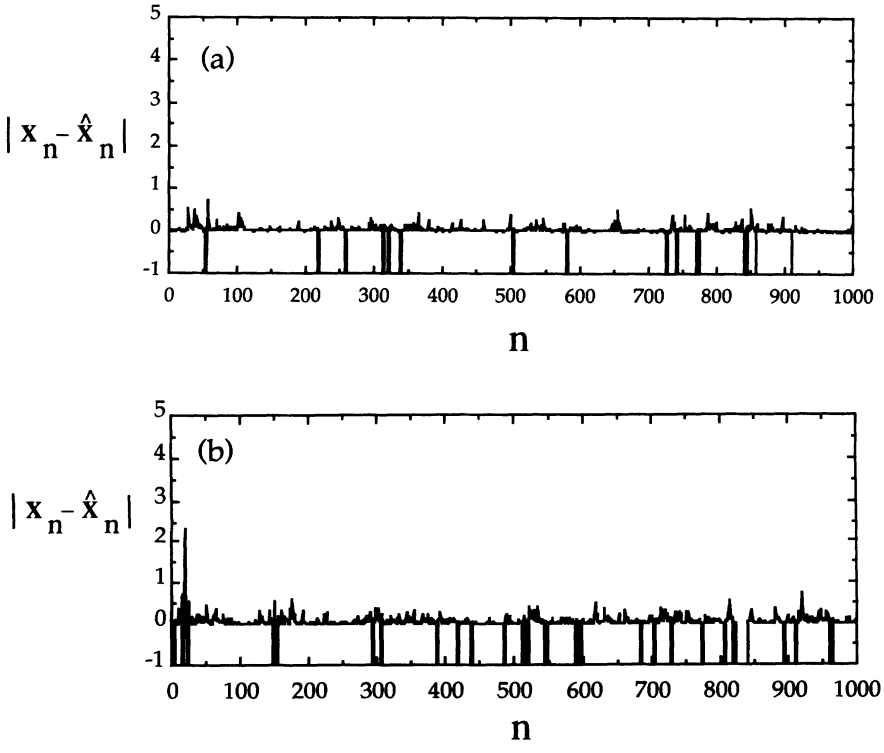

FIG. 4. $|\mathbf{X}_n-\hat{\mathbf{X}}_n|$ vs n with $N=1500$, $\varepsilon=2\times10^{-2}$. Spikes with negative magnitude indicate the moments when multiple observer test points were initiated. (a) Additive noise in actual system. (b) Additive noise in output measurements.

TABLE I. Table of sub-Lyapunov exponents for the reduced rotor map with the observed quantity $g(\mathbf{X}_n)=\theta_1$, θ_2, $\dot{\theta}_1$, or $\dot{\theta}_2$. The sub-Lyapunov exponents are calculated from a single orbit over 1 000 000 iterates.

Quantity	$\lambda_L^{(1)}$	$\lambda_L^{(2)}$	$\lambda_L^{(3)}$
$g(\mathbf{X})=\theta_1$	0.6716	−0.3414	−1.4209
$g(\mathbf{X})=\theta_2$	0.5916	−0.4437	−1.4473
$g(\mathbf{X})=\dot{\theta}_1$	0.6603	0.0000	−1.3258
$g(\mathbf{X})=\dot{\theta}_1$	0.0000	−0.3036	−1.1176

bit within a given short time limit, we reinitiate the procedure with another set of N randomly chosen observer test points. Figure 4 shows a plot of $|\mathbf{X}_n-\hat{\mathbf{X}}_n|$ versus n with $N=1500$ and $\varepsilon=2\times10^{-2}$. [Figure 4(a) is the result for additive noise to the actual system and Fig. 4(b) is the result for additive noise to the output function.] We see that our observer technique successfully tracks $\mathbf{X}_n$ even when ε is relatively large (signal-to-noise ratio ≈ 150) provided that N is sufficiently large.

One should note that our chaotic observer technique has an advantage over delay coordinate embedding in the situation where the system is driven by an observed time dependent variable input which may be temporally irregular. In the presence of a time dependent variable input, delay coordinate embedding will not work simply because the correlation between the delayed vector $(O_n, O_{n-1}, \ldots, O_{n-N})$ and the system state vector $\mathbf{X}_n$ will be lost. However, since the chaotic observer tracks the actual system in real time, the observer technique will still work. In order to examine the behavior of the chaotic observer for a system driven by a time dependent variable input, we replace the fixed strength of the kick to the double rotor at time $t=nT$ by $f_n=f_0+\beta_n$ where β_n defines the time dependent driving force to the rotor. Since we compute the control vector $\mathbf{C}_n$ in real time, the time variation in f_n will not affect our calculation provided that we have knowledge of the function β_n. In our numerical experiment, we choose $\beta_n=\varepsilon\sin[(2\pi/T)n]$, where ε is the amplitude of the perturbation and T is the period of the perturbation. When the observer orbit starts within the linear region of the map, it converges readily to the actual orbit as in the previous examples without time-dependent variable input [see Fig. 5(a)]. However, when the observer orbit begins outside the linear region, it typically takes a much longer time to converge to the actual orbit [see Fig. 6(a)]. This behavior is similar to the previous situation when the kick strength is a fixed value. In addition, we also varied the values of ε and T for a number of different trials but there were no significant differences in the convergence characteristic of the observer in all those cases [see Figs. 5(a) and 6(b)].

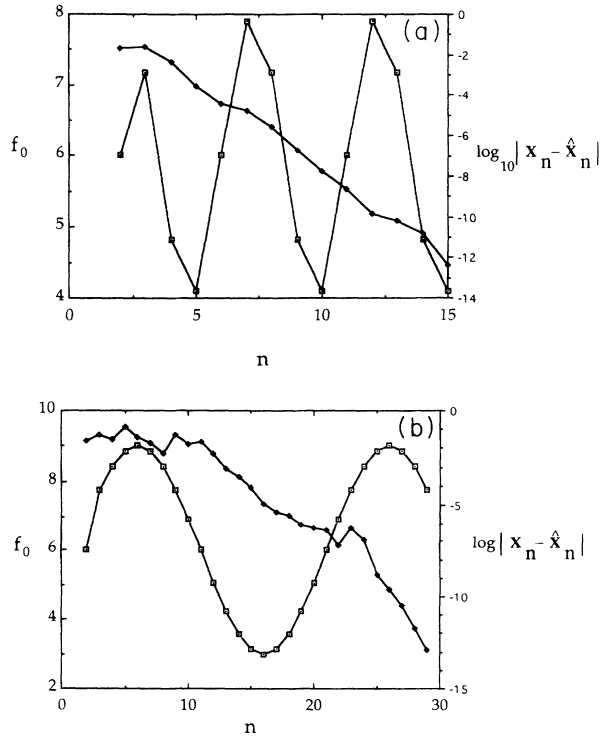

FIG. 5. $\log_{10}|\mathbf{X}_n-\hat{\mathbf{X}}_n|$ vs n with $\beta(n)=\varepsilon\sin[(2\pi/T)n]$ and $f_0=6.0$. The observer orbit begins within the linear region of the actual orbit. (a) $\varepsilon=3$; $T=20$. (b) $\varepsilon=2$; $T=5$. Solid squares denote $\log_{10}|\mathbf{X}_n-\hat{\mathbf{X}}_n|$ and empty squares denote f_0.

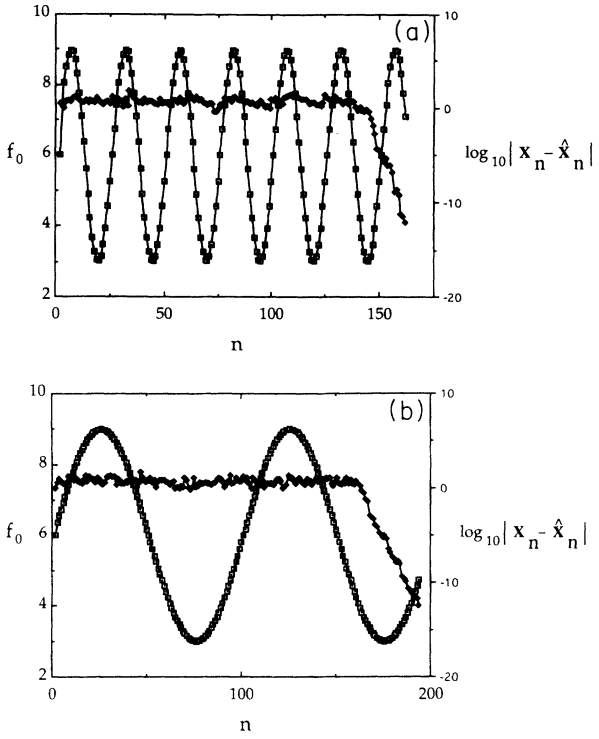

FIG. 6. $\log_{10}|\mathbf{X}_n-\hat{\mathbf{X}}_n|$ vs n with $\beta(n)=\varepsilon\sin[(2\pi/T)n]$ and $f_0=6.0$. The observer orbit begins outside the linear region of the actual orbit. (a) $\varepsilon=3$; $T=100$. (b) $\varepsilon=3$; $T=25$. Solid squares denote $\log_{10}|\mathbf{X}_n-\hat{\mathbf{X}}_n|$ and empty squares denote f_0.

VI. EXTENDED KALMAN FILTER

The Kalman filter and the extended Kalman filter basically have the same mathematical structure as our nonlinear observer. This is to say that they all have the following form:

$$\hat{\mathbf{X}}_{n+1}=\mathbf{M}(\hat{\mathbf{X}}_n)+\mathbf{C}_n[O_{n+1}-g(\mathbf{M}(\hat{\mathbf{X}}_n))]\ , \qquad (16)$$

where $\hat{\mathbf{X}}_n$ is the estimated state of the system, $O_n=(g(\mathbf{X}_n)$ is the observed output of the system, and $\mathbf{C}_n$ is the control vector which will be adjusted according to the methods used. As we have stated earlier, the fundamental consideration in choosing $\mathbf{C}_n$ in our method is stability while the fundamental consideration in choosing $\mathbf{C}_n$ in the Kalman filter is *noise* minimization. To be concrete, let us say we have a noisy system,

$$\mathbf{X}_{n+1}=\mathbf{M}(\mathbf{X}_n)+\mathbf{H}_n\mathbf{w}_n,\quad O_n=g(\mathbf{X}_n)+\nu_n\ , \tag{17}$$

with $E[\mathbf{w}_n\mathbf{w}_m^\dagger]=R_n\delta_{nm}$ and $E[\nu_n\nu_m^\dagger]=Q_n\delta_{nm}$ ($E[\ ]$ denotes the expected value). We further assume that the two random processes $\{\mathbf{X}_n\}$ and $\{O_n\}$ are jointly distributed. Then, the extended Kalman filter is defined by the following set of equations:

$$\hat{\mathbf{X}}_{n+1}=\mathbf{M}(\hat{\mathbf{X}}_n)+\mathbf{C}[O_{n+1}-g(\mathbf{M}(\hat{\mathbf{X}}_n))]\ , \tag{18a}$$

$$\mathbf{C}_n=\Gamma_n\mathbf{D}g_n[\mathbf{D}g_n^\dagger\Gamma_n\mathbf{D}g_n+R_n]^{-1}\ , \tag{18b}$$

$$\Sigma_{n+1}=\Gamma_n[1-\mathbf{D}g_n(\mathbf{D}g_n^\dagger\Gamma_n\mathbf{D}g_n+R_n)^{-1}\mathbf{D}g_n^\dagger\Gamma_n]\ , \tag{18c}$$

$$\Gamma_n=\mathbf{D}\mathbf{M}_n\Sigma_n\mathbf{D}\mathbf{M}_n^\dagger+\mathbf{H}_nQ_n\mathbf{H}_n^\dagger\ . \tag{18d}$$

These equations in general are initialized by setting $\hat{\mathbf{X}}_0$ to the mean value of the initial orbit $\mathbf{X}_0$ and setting Σ_0 to the initial error covariance matrix $E[(\mathbf{X}_0-\hat{\mathbf{X}}_0)(\mathbf{X}_0-\hat{\mathbf{X}}_0)^\dagger]$. In the special case when the system is time invariant and linear, the corresponding linear Kalman filter can be directly derived by considering the time evolution of the means and covariance of the jointly distributed random variables $\{\mathbf{X}_n\}$ and $\{O_n\}$ [15]. Actually, $\hat{\mathbf{X}}_n=E[\mathbf{X}_n|O_n]$ is the conditional mean and $\Sigma_n=E[(\mathbf{X}_n-\hat{\mathbf{X}}_n)(\mathbf{X}_n-\hat{\mathbf{X}}_n)^\dagger|O_n]$ is the conditional error covariance of the state estimate. In the time invariant linear case, the control vector $\mathbf{C}_n$ and Σ_n are independent of the observed variable $\{O_n\}$ and there exist limiting values $\bar{\mathbf{C}}$ and $\bar{\Sigma}$ for Eqs. (18b) and (18c), respectively, if the system is completely observable (see Sec. II). Furthermore, it can be shown that the linear Kalman filter is optimal in the sense that the error covariance $\bar{\Sigma}$ for the linear Kalman filter is a minimum with respect to all other estimators in the form described by Eq. (16).

It is important to note that the extended Kalman filter Eqs. (18a) and (18b) is a nonlinear filter and its construction is based on the natural extension of the Kalman filter for the linearized system,

$$\mathbf{X}_{n+1}=\mathbf{D}\mathbf{M}(\hat{\mathbf{X}}_n)\mathbf{X}_n+\mathbf{H}_n\mathbf{w}_n,\quad O_n=\mathbf{D}g(\hat{\mathbf{X}}_n)\mathbf{X}_n+\nu_n\ , \tag{19}$$

at $\hat{\mathbf{X}}_n$. Recall that the question of optimality in the linear case is well defined because $\lim_{n\to\infty}\Sigma_n=\bar{\Sigma}$ exists and it is reasonable to say that $\bar{\Sigma}$ obtained from the Kalman filter is a minimum within a class of possible filters. However, when the extended Kalman filter is applied to the actual nonlinear system [Eq. (17)], $\lim_{n\to\infty}\Sigma_n$ in general does not exist and the question of optimality becomes less clear. Since the optimality of the extended Kalman filter with respect to noise is not a well-defined concept, we will try to compare the convergence characteristic of the extended Kalman filter and our nonlinear observer technique in the noiseless case. It should be noted that although noise is an essential part in the construction of the standard linear Kalman filter and the extended Kalman filter, they still function as estimators in the noiseless case. Setting $\mathbf{R}_n$ and Q_n to zero, Eq. (18) reduces to the following form:

$$\hat{\mathbf{X}}_{n+1}=\mathbf{M}(\hat{\mathbf{X}}_n)+\mathbf{C}_n[O_{n+1}-g(\mathbf{M}(\hat{\mathbf{X}}_n))]\ , \tag{20a}$$

$$\mathbf{C}_n=\Gamma_n\mathbf{D}g_n[\mathbf{D}g_n^\dagger\Gamma_n\mathbf{D}g_n]^{-1}\ , \tag{20b}$$

$$\Sigma_{n+1}=\Gamma_n[1-\mathbf{D}g_n(\mathbf{D}g_n^\dagger\Gamma_n\mathbf{D}g_n)^{-1}\mathbf{D}g_n^\dagger\Gamma_n]\ , \tag{20c}$$

$$\Gamma_n=\mathbf{D}\mathbf{M}_n\Sigma_n\mathbf{D}\mathbf{M}_n^\dagger\ . \tag{20d}$$

A discussion of the convergence characteristic for the nonlinear full-order observer can be found in Sec. III (see Figs. 2 and 3). We have determined the average convergence time $\langle\tau\rangle$ for the extended Kalman filter using the same method as in our nonlinear observer. We begin with a large number of randomly chosen initial points for the extended Kalman filter. Then, we determine the number $\hat{N}$ of filters which are still not following the true orbit after a time interval n. From the graph of $\ln(\hat{N})$ vs n, we estimate $\langle\tau\rangle$ to be approximately 500 (Fig. 7). Comparatively, for this parameter set of the double rotor map, the extended Kalman filter has a faster convergence rate than our nonlinear observer. However, the computation for the control vector $\mathbf{C}_n$ in the extended Kalman filter requires the manipulation of matrix equations with the same dimension as the full system while our nonlinear full-order observer technique requires only the manipulation of d_u basis vectors. In the case of the double rotor map ($d_u=2$), the number of computations required in our nonlinear observer technique is about half the number of computations needed in the extended Kalman filter. Another factor of 2 in the number of computations can be saved if we use only one basis vector in our reduced-order observer. While our example yields a factor of 3 in the average convergence rate between the two methods, the comparative saving in the number of computations using our method will in general improve linearly with the ratio d/d_u. We can see this by comparing Eqs. (B4) and (B5) for our nonlinear observer technique and Eq. (18) for the extended Kalman filter. As-

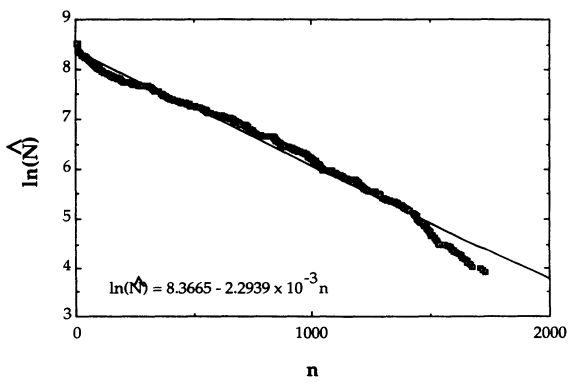

FIG. 7. $\ln(\hat{N})$ vs n with 5000 randomly chosen initial test points for the extended Kalman filter. A particular filter is said to be effectively working if $|\mathbf{X}_n-\hat{\mathbf{X}}_n|<1.0\times10^{-6}$.

suming that we are observing only a time series of scalar output (i.e., $d_0=1$), there will be d_u equations for the d_u basis vectors in Eq. (B4). Since each basis vector is d dimensional, the numbers of computations needed will roughly be proportional to $d\times d_u$. On the other hand, the iteration equation for the error covariance matrix Σ_n in Eqs. (18c) and (18d) involves manipulations of $d\times d$ matrices and the number of computation needs will roughly be proportional to $d\times d$. Lastly, Eqs. (B5) and (18b) for the control vector $\mathbf{C}_n$ in both methods are comparable in complexity. Therefore, the overall ratio of computation time between the two methods will roughly be d/d_u.

VII. CONCLUSION

In this paper, we have introduced a method for observing a chaotic system from a time series of a scalar function of the system state. Our observer technique will in general be more efficient than delay coordinate embedding in terms of computation time. This is also true as compared to an extended Kalman filter in cases where $d\gg d_u$. The reduction of computational steps further improves if a reduced-order observer can be constructed for the system. We have found in an example that our observer technique can be effective in a noisy environment (with signal to noise level ~ 150) provided we used the multiple observer technique discussed in Sec. V. In addition, our observer technique still applies when the input to the system is time dependent (this situation would in general prevent utilization of embedding). Although our discussion is limited to discrete time systems, we are currently working on extending this method to continuous time cases.

ACKNOWLEDGMENTS

The work of P. So and E. Ott was supported by the U.S. Department of Energy (Scientific Computing) and by the Office of Naval Research (Physics), and the work of W. P. Dayawansa was supported by the National Science Foundation (Engineering Research Center Program).

APPENDIX A: PROOF OF CONVERGENCE OF EQ. (9)

In this appendix, we will show that by choosing the magnitudes of all the eigenvalues of the matrix $\mathbf{U}_n$ to be less than 1 (recall that $\mathbf{U}_n$ is a lower triangular matrix), the product

$$\mathbf{J}_n=\prod_{i=1}^{n}\begin{bmatrix}\mathbf{U}_i & \mathbf{W}_i\\ 0 & \mathbf{S}_i\end{bmatrix}=\begin{bmatrix}\prod_{i=1}^{n}\mathbf{U}_i & \mathbf{K}_n\\ 0 & \prod_{i=1}^{n}\mathbf{S}_i\end{bmatrix}, \tag{A1}$$

where

$$\mathbf{K}_n=\sum_{i=1}^{n}\left[\prod_{j=i+1}^{n}\mathbf{U}_j\right]\mathbf{W}_i\left[\prod_{k=1}^{i-1}\mathbf{S}_k\right], \tag{A2}$$

will converge to zero with increasing n. To be specific, we will show that there exist positive constants $K<\infty$ and $0<\mu<1$ such that

$$\|\mathbf{J}_n\|\le K\mu^n . \tag{A3}$$

(Here, we have used the following matrix norm: $\|\mathbf{J}\|\equiv\sum_{i,j=1}^{d}|J_{ij}|$.)

First, we note that, since $\mathbf{S}_n$ is assumed to be stable, then by definition these exist constants $C_S<\infty$ and $0<\mu_S<1$, such that

$$\left\|\prod_{i=1}^{n}\mathbf{S}_i\right\|\le C_S\mu_S^n . \tag{A4}$$

In the following, we will show that there also exist positive constants $C_U<\infty$, $C_*<\infty$, $\mu_U<1$, and $\zeta<1$, such that

$$\left\|\prod_{i=1}^{n}\mathbf{U}_i\right\|\le C_u\mu_U^n \tag{A5}$$

and

$$\|\mathbf{K}_n\|\le C_*\zeta^n , \tag{A6}$$

provided that we choose the magnitudes of all the eigenvalues of $\mathbf{U}_i$ to be less than 1. From the definition of the matrix norm, we have from Eq. (A1),

$$\|\mathbf{J}_n\|=\left\|\prod_{i=1}^{n}\mathbf{U}_i\right\|+\|\mathbf{K}_n\|+\left\|\prod_{i=1}^{n}\mathbf{S}_i\right\| .$$

Now, utilizing Eqs. (A4)–(A6), we have

$$\|\mathbf{J}_n\|\le K\mu^n , \tag{A7}$$

where $K=3\max\{C_U,C_S,C_*\}$ and $\mu=\max\{\mu_U,\mu_S,\zeta\}$, which is the desired result. In what follows, we derive Eqs. (A5) and (A6).

We derive Eq. (A5) by using mathematical induction. First, we let Ω^{d_u} denote the set of all bounded d_u-dimensional lower triangular matrices whose eigenvalues all have magnitudes less than 1. For $d_u=1$, the matrices in Eq. (A5) are just scalars of magnitude less than 1. Thus, Eq. (A5) is trivially satisfied. Now assume that for a set of matrices $\{\mathbf{Z}_i:\mathbf{Z}_i\in\Omega^{d_u}\}$ there exist $K<\infty$ and $0<\beta<1$ such that Eq. (A5) is true,

$$\left\|\prod_{i=1}^{n}\mathbf{Z}_i\right\|\le K\beta^n . \tag{A8}$$

Then, we will use this assumption to show that there also exist $C<\infty$ and $0<\mu<1$ such that Eq. (A5) is true for $\{\mathbf{U}_i:\mathbf{U}_i\in\Omega^{d_u+1}\}$. Let us consider a particular matrix $\mathbf{U}_i\in\Omega^{d_u+1}$. We partition it into the following form:

$$\mathbf{U}_i=\begin{bmatrix}\mathbf{Z}_i & 0\\ \mathbf{b}_i & \Lambda_i\end{bmatrix},$$

where $\mathbf{Z}_i\in\Omega^{d_u}$ and Λ_i is the (d_u+1)th eigenvalue of $\mathbf{U}_i$. With this notation, the product of these $\mathbf{U}_i$ can be written as

$$\prod_{i=1}^{n}\mathbf{U}_i=\begin{bmatrix}\prod_{i=1}^{n}\mathbf{Z}_i & \mathbf{0}\\ \mathbf{L}_n & \prod_{i=1}^{n}\Lambda_i\end{bmatrix},$$

where

$$\mathbf{L}_n=\sum_{i=1}^{n}\left[\prod_{j=1}^{i-1}\Lambda_j\right]\mathbf{b}_i\left[\prod_{k=1}^{n-i}\mathbf{Z}_k\right].$$

Pick a value $\Lambda_{\max}$ satisfying $|\Lambda_i|\le\Lambda_{\max}<1$. Then,

$$\left\|\prod_{i=1}^{n}\Lambda_i\right\|\le\Lambda_{\max}^{n}. \tag{A9}$$

Now consider the norm of $\mathbf{L}_n$:

$$\begin{aligned}\|\mathbf{L}_n\|&=\left\|\sum_{i=1}^{n}\left[\prod_{j=1}^{i-1}\Lambda_j^{d_u+1}\right]\mathbf{b}_i\left[\prod_{k=1}^{n-i}\mathbf{Z}_k\right]\right\|\\&\le\sum_{i=1}^{n}\left\|\prod_{k=1}^{n-i}\mathbf{Z}_k\right\|\|\mathbf{b}_i\|(\Lambda_{\max})^{i-1}\\&\le\sum_{i=1}^{n}\left\|\prod_{k=1}^{n-i}\mathbf{Z}_k\right\|\|\mathbf{b}_{\max}\|(\Lambda_{\max})^{i-1}\\&\le\sum_{i=1}^{n}K\beta^{n-i}\|\mathbf{b}_{\max}\|(\Lambda_{\max})^{i-1}\\&\le\sum_{i=1}^{n}K'(\max\{\beta,\Lambda_{\max}\})^{n-1}\\&=nK'(\mu')^{n-1},\end{aligned}$$

where $\|\mathbf{b}_{\max}\|$ is the largest of the $\|\mathbf{b}_i\|$, $K'=K\|\mathbf{b}_{\max}\|$, and $\mu'=\max\{\beta,\Lambda_{\max}\}$. In order to put the above inequality in the right form, we pick a γ such that $\mu'<\gamma<1$. We then choose a constant K^* such that

$$nK'(\mu')^{n-1}\le K^*\gamma^n,$$

for $n\ge1$ where the constant K^* is any number bigger than the maximum over n of $nK'(\mu')^{n-1}\gamma^{-n}$, (since $\gamma>\mu'$ this maximum is finite). Thus

$$\|\mathbf{L}_n\|\le K^*\gamma^n. \tag{A10}$$

With the bounds given by Eqs. (A8)–(A10), one can choose $C_U=3\max\{K,K^*,1\}$ and $\mu_U=\max\{\beta,\Lambda_{\max},\gamma\}$ so that

$$\left\|\prod_{i=1}^{n}\mathbf{U}_i\right\|\le C_U\mu_U^n.$$

To complete our demonstration of Eq. (A7), we now derive Eq. (A6).

$$\begin{aligned}\|\mathbf{K}_n\|&=\left\|\sum_{i=1}^{n}\left[\prod_{j=i+1}^{n}\mathbf{U}_j\right]\mathbf{W}_i\left[\prod_{k=1}^{i-1}\mathbf{S}_k\right]\right\|\\&\le\sum_{i=1}^{n}\left\|\prod_{k=i+1}^{n}\mathbf{U}_k\right\|\|\mathbf{W}_i\|\left\|\prod_{k=1}^{i-1}\mathbf{S}_k\right\|\\&\le\sum_{i=1}^{n}C_U\mu_U^{n-i}\|\mathbf{W}_{\max}\|C_S\mu_S^{i-1}\\&\le\sum_{i=1}^{n}C'\eta^{n-1}\\&=nC'\eta^{n-1},\end{aligned}$$

where $C'=\|\mathbf{W}_{\max}\|\max\{C_U,C_S\}$ and $\eta=\max\{\mu_U,\mu_S\}<1$. Choosing $\eta<\zeta<1$ and C^* such that

$$nC'\eta^{n-1}\le C^*\zeta^n,$$

for $n\ge1$, we then yield Eq. (A6).

APPENDIX B: GENERALIZATION OF THE FULL-ORDER OBSERVER TO SYSTEMS WITH ARBITRARY NUMBER OF UNSTABLE DIRECTIONS AND WITH VECTORAL OUTPUT FUNCTION $O=g(X)$

Consider a d-dimensional system,

$$\mathbf{X}_{n+1}=\mathbf{M}(\mathbf{X}_n),\quad O_n=\mathbf{g}(\mathbf{X}_n), \tag{B1}$$

where the observed output O_n is a d_0-dimensional vector. Let us say that the unstable subspace of this dynamical system is d_u dimensional. Then, the full-order nonlinear observer can be defined as

$$\hat{\mathbf{X}}_{n+1}=\mathbf{M}(\hat{\mathbf{X}}_n)+\mathbf{C}_n[\mathbf{O}_{n+1}-\hat{\mathbf{O}}_{n+1}]. \tag{B2}$$

Here, $\mathbf{C}_n$ is a $d\times d_0$ time dependent control matrix. The error equation corresponding to this observer and the true system is given by

$$\delta\mathbf{X}_{n+1}=[\mathbf{DM}(\hat{\mathbf{X}}_n)-\mathbf{C}_n\mathbf{Dg}(\mathbf{M}(\hat{\mathbf{X}}_n))\mathbf{DM}(\hat{\mathbf{X}}_n)]\delta\mathbf{X}, \tag{B3}$$

where $\mathbf{X}_n-\hat{\mathbf{X}}_n=\delta\mathbf{X}_n$ is a differential, and $\mathbf{DM}(\hat{\mathbf{X}}_n)$ and $\mathbf{Dg}(\mathbf{M}(\hat{\mathbf{X}}_n))$ are the derivatives of $\mathbf{M}(\hat{\mathbf{X}}_n)$ and $\mathbf{g}(\mathbf{M}(\hat{\mathbf{X}}_n))$, respectively, with $\mathbf{DM}$ a $d\times d$ matrix and $\mathbf{Dg}$ a $d_0\times d$ matrix. Now, we will restrict the action of $\mathbf{C}_n$ to the unstable subspace by the following construction:

$$\mathbf{C}_n=\sum_{i=1}^{d_u}\boldsymbol{\beta}_n^{(i)}\text{ with }\boldsymbol{\beta}_n^{(i)}=C_1^{(i)}\mathbf{e}_{n+1}^{(i)}\mathbf{V}_1+\cdots+C_{d_0}^{(i)}\mathbf{e}_{n+1}^{(i)}\mathbf{V}_{d_0},$$

where $\{\mathbf{e}_n^{(i)}\}$ [defined in Eq. (A4)] is a set of column vectors which spans the unstable subspace at $\hat{\mathbf{X}}_n$ and $\{\mathbf{V}_j\}$ can be any complete set of row vectors which spans $\mathbb{R}^{d_0}$. In this case, $[\mathbf{DM}(\hat{\mathbf{X}}_n)-\mathbf{C}_n\mathbf{Dg}(\mathbf{M}(\hat{\mathbf{X}}_n))\mathbf{DM}(\hat{\mathbf{X}}_n)]$ can again be put in a block form:

$$\begin{bmatrix}\mathbf{U}_n & \mathbf{W}_n\\ \mathbf{0} & \mathbf{S}_n\end{bmatrix},$$

with $\mathbf{U}_n$ as a $d_u\times d_u$ submatrix acting on the unstable subspace, $\mathbf{S}_n$ as a $(d-d_u)\times(d-d_u)$ submatrix acting on the stable subspace, and $\mathbf{W}_n$ as a $d_u\times(d-d_u)$ submatrix taking vectors from the stable subspace into the unstable subspace. Then, we define d_u numbers $\lambda_n^{(1)},\ldots,\lambda_n^{(d_u)}$, and d_u basis column vectors $\mathbf{e}_n^{(1)},\ldots,\mathbf{e}_n^{(d_u)}$, for the unstable subspace at $\hat{\mathbf{X}}_n$ according to the following procedure:

$$\lambda_n^{(1)}\mathbf{e}_{n+1}^{(1)}=[\mathbf{DM}(\hat{\mathbf{X}}_n)]\mathbf{e}_i^{(1)}$$
$$\lambda_n^{(2)}\mathbf{e}_{n+1}^{(2)}=[\mathbf{DM}(\hat{\mathbf{X}}_n)-\beta_n^{(1)}\mathbf{Dg}(\mathbf{M}(\hat{\mathbf{X}}_n))\mathbf{DM}(\hat{\mathbf{X}}_n)]\mathbf{e}_n^{(2)}\ ,$$
$$\vdots \tag{B4}$$
$$\lambda_n^{(d_u)}\mathbf{e}_{n+1}^{(d_u)}=\left[\mathbf{DM}(\hat{\mathbf{X}}_n)-\sum_{i=1}^{d_u-1}\beta_n^{(i)}\mathbf{Dg}(\mathbf{M}(\hat{\mathbf{X}}_n))\mathbf{DM}(\hat{\mathbf{X}}_n)\right]\mathbf{e}_n^{(d_u)}\ .$$

In this set of basis, $\mathbf{U}_n$ will be in lower triangular form with $d_u\times d_0$ free parameters, $C_j^{(i)}, 1\le i\le d_u, 1\le j\le d_0$ [see Eq. (A3)]:

$$\mathbf{U}_n=\begin{bmatrix}\lambda_n^{(1)}-\mathbf{f}_{n+1}^{(1)}\beta_n^{(1)}\mathbf{D}h_n^{(1)} & 0 & \cdots & 0\\ -\mathbf{f}_{n+1}^{(2)}\beta_n^{(2)}\mathbf{D}h_n^{(1)} & \lambda_n^{(2)}-\mathbf{f}_{n+1}^{(2)}\beta_n^{(2)}\mathbf{D}h_n^{(2)} & \cdots & 0\\ \vdots & \vdots & \vdots & \vdots\\ -\mathbf{f}_{n+1}^{(d_u)}\beta_n^{(d_u)}\mathbf{D}h_n^{(1)} & -\mathbf{f}_{n+1}^{(d_u)}\beta_n^{(d_u)}\mathbf{D}h_n^{(2)} & \cdots & \lambda_n^{(d_u)}-\mathbf{f}_{n+1}^{(d_u)}\beta_n^{(d_u)}\mathbf{D}h_n^{(d_u)}\end{bmatrix},$$

where $\mathbf{D}h_n^{(d_u)}=\mathbf{Dg}(\mathbf{M}(\hat{\mathbf{X}}_n))\mathbf{DM}(\hat{\mathbf{X}}_n)\mathbf{e}_n^{(d_u)}$ is a d_0-dimensional column vector and $\{\mathbf{f}_{n+1}^{(i)}\}$ is the corresponding set of dual vectors for $\{\mathbf{e}_{n+1}^{(1)}\}$. We can adjust these $d_u\times d_0$ free parameters $C_j^{(i)}$ such that the eigenvalues of $\mathbf{U}_n$, $\Lambda_n^{(i)}=\lambda_n^{(i)}-\mathbf{f}_{n+1}^{(i)}\beta_n^{(i)}\mathbf{D}h_{(j)}$ will all be less than 1. Since $\mathbf{U}_n$ is d_u dimensional, by setting all its eigenvalues to be less than 1 will only provide d_u conditions. In other words, we will have $d_u\times(d_0-1)$ parameters left for our disposal. This is an advantage that is not possible in the case when the observed output is a scalar (i.e., $d_0=1$). In the simplest case, when we set all eigenvalues of $\mathbf{U}_n$ and the rest of the $d_u\times(d_0-1)$ parameters (i.e., $C_j^{(i)}=0\forall j>1$) to be zero, the control matrix $\mathbf{C}_n$ is given by the following equation:

$$\mathbf{C}_n=\sum_{i=1}^{d_u}\frac{\lambda_n^{(i)}}{\mathbf{V}_1\mathbf{D}h_n^{(i)}}\mathbf{e}_{n+1}^{(i)}\mathbf{V}_1\ . \tag{B5}$$

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

[2] D. D. Anderson and J. B. Moore, *Optimal Filtering* (Prentice-Hall, Englewood Cliffs, NJ, 1990), pp. 193–211.

[3] We will initially be interested in noiseless situations, and it should be noted that the Kalman filter procedure also applies in these situations.

[4] Another common situation arises when the system is acted upon by external time-dependent inputs β_n. In that case one seeks $\mathbf{X}_n$ given knowledge of β_n and O_n. Now, the embedding technique becomes inapplicable, but the alternative which we shall discuss still applies.

[5] K. Ogata, *Modern Control Engineering* (Prentice-Hall, Englewood Cliffs, NJ, 1990), pp. 677–886.

[6] As an example, the product of the following sequence of matrices:

$$\begin{bmatrix}\frac{1}{2} & 2\\ 0 & \frac{1}{2}\end{bmatrix}\begin{bmatrix}\frac{1}{2} & 0\\ 2 & \frac{1}{2}\end{bmatrix}\begin{bmatrix}\frac{1}{2} & 2\\ 0 & \frac{1}{2}\end{bmatrix}\cdots$$

will be infinite while the eigenvalues of each individual matrix are less than 1, i.e., $\frac{1}{2}$.

[7] F. J. Romeiras, C. Grebogi, E. Ott, and W. P. Dayawansa, Physica D (to be published); C. Grebogi, E. Kostelich, E. Ott, and J. A. Yorke, Phys. Lett. A **118**, 448 (1986); **120**, 497(E) (1987).

[8] In terms of the notation used in Fig. 1 (see also Romeiras *et al.*), the parameters used in our numerical example are: $L_1=1, L_2=1.1, m_2=(1.1)^{-2}, m_1=1-m_2, \nu_1=\nu_2=1$; the time interval between successive kicks is $T=1$; the impulse strength of a kick is $f_0=6$ (which gives $a_1=6$ and $a_2=6.6$). For this case, synchronism as defined by Pecora and Carroll [14] does not occur.

[9] J. D. Farmer, E. Ott, and J. A. Yorke, Physica D **7**, pp. 153 (1983).

[10] F. J. Romeiras, E. Ott, C. Grebogi, and W. P. Dayawansa, in *Proceedings of the American Control Conference* (IEEE, Piscataway, NJ, 1991), pp. 1113–1119; and Romeiras *et al.* [7].

[11] In our numerical experiment, $\mathbf{U}_n$ is in the following form, $\begin{bmatrix}0 & 0\\ \mathbf{x} & 0\end{bmatrix}$. The product of two of these matrices is identically zero. In general, when the dimension of $\mathbf{U}_n$ is d_u (Appendix B), if $\mathbf{U}_n$ is lower triangular with zeros on the diagonal, then the product of these $\mathbf{U}$'s will go to zero in d_u iterates.

[12] The average is determined from 5000 randomly chosen initial observer test orbits within the linear region of the true orbit.

[13] As mentioned earlier, we are only free to choose $\Lambda_n^{(2)}$ because we need to keep the basis vectors in Eq. (10) from collapsing by setting $\Lambda_n^{(1)}=0$. Recall that in Eq. (12), we set $\Lambda_n^{(2)}=0$ also for faster convergence rate. (Also, we use $\Lambda_n^{(2)}=0$ in our numerical experiments, Figs. 3–7.)

[14] L. Pecora and T. Carroll, Phys. Rev. Lett. **64**, 821 (1990).

[15] B. D. Anderson and J. B. Moore, *Optimal Filtering* (Ref. [2]), pp. 39–45.