PHYSICAL REVIEW A

STATISTICAL PHYSICS, PLASMAS, FLUIDS, AND RELATED INTERDISCIPLINARY TOPICS

THIRD SERIES, VOLUME 42, NUMBER 10 15 NOVEMBER 1990

Reconstructing equations of motion from experimental data with unobserved variables

Joseph L. Breeden and Alfred Hübler

*Center for Complex Systems Research–Beckman Institute and Department of Physics, University of Illinois, 405 North Mathews Avenue, Urbana, Illinois 61801**

(Received 7 June 1990; revised manuscript received 13 August 1990)

We have developed a method for reconstructing equations of motion for systems where all the necessary variables have not been observed. This technique can be applied to systems with one or several such hidden variables, and can be used to reconstruct maps or differential equations. The effects of experimental noise are discussed through specific examples. The control of nonlinear systems containing hidden variables is also discussed.

I. INTRODUCTION

Prior to the 1980s, researchers had always assumed that to study the dynamics of nonlinear systems with many degrees of freedom, time-series measurements of all the variables, or derivatives thereof, were necessary to generate state-space representations of the dynamics. For experimental systems, derivatives are particularly difficult to employ due to the noise problems. In 1980, Packard *et al.*[1] and Ruelle noted that a state-space representation of the dynamics could be reconstructed from a single time series through the use of delay coordinates. This delay-coordinate reconstruction would then be topologically equivalent to the dynamics of the true system. Whitney had shown much earlier[2] that any compact manifold with dimension m can be embedded in R^{2m+1}. Takens extended this[3] in proving that an embedding can be obtained for any system from only a single time series by using $2m+1$ delay coordinates.[4] While this combination of ideas thus far has been extremely useful in studying nonlinear systems, several difficulties arise in their application that we hope to address with an alternative method for reconstructing these *hidden* variables.

The most obvious difficulty in using delay coordinates is the issue of interpreting the results physically. If one is concerned only with forecasting, the method used for the modeling is irrelevant, as long as it works. Successful modeling techniques based upon delay coordinates and/or partitioning the state space to generate local fits have been developed.[5–7] However, relating these models back to physical principles or existing theories is often difficult if not impossible. Another important consideration is that Takens's theorem does not apply to systems with noise, i.e., experimental data. In the case of (noisy) experimental data, one must define what precisely is meant by an embedding. We will call an embedding any representation for which any two observations $\mathbf{X}_1(t)$ and $\mathbf{X}_2(t)$ within σ, the noise amplitude, of each other are followed by $\mathbf{X}_1(t+\delta t)$ and $\mathbf{X}_2(t+\delta t)$ within the propogated error of each other. Another way of stating this could be that for any region of the state space, the variance of the succeeding values is minimized. Casdagli has reported progress in extending Takens's theorem to find optimal embeddings in the low noise limit,[8] but at present there exists no general guarantee that a given reconstruction will be an embedding in the presence of noise.

In order to create a modeling technique in which existing information can be incorporated, the resulting model can be interpreted physically, and which is reasonably stable to noise; we base our technique upon the *flow method* developed by Cremers and Hübler[9] and Eisenhammer *et al.*[10] that is similar to that of Crutchfield and McNamara.[11] The flow method is a procedure for reconstructing a set of coupled maps (CM's) or ordinary differential equations (ODE's) from a trajectory of the system in state space. We will show that this may be easily adapted to the presence of hidden variables.

In Sec. II, we will review the flow method and demonstrate how hidden variables can be incorporated into this framework. Some limitations of these techniques are also addressed. Section III provides details on the implementation of our hidden variables technique with specific examples of its application to simulated data. The treatment of noisy experimental data is also discussed. The use of this modeling procedure in conjunction with nonlinear control theory to control systems where some variables are hidden from observation, control, or both is discussed in Sec. IV.

II. THEORY

A. Trajectory modeling without hidden variables

In the flow method as developed by Cremers and Hübler,[9] the dynamics throughout the state space is represented either with a single set of coupled maps

$$y_i(n+1)=f_i(\mathbf{y}(n)),\quad i=1,\cdots,N,\quad \mathbf{y}\in R^N$$

or a set of ordinary differential equations

$$\dot{y}_i(t)=f_i(\mathbf{y}(t)),\quad i=1,\ldots,N,\quad \mathbf{y}\in R^N .$$

The number of observed variables is assumed sufficient to embed the dynamics. The functions $\{f_i\}$ may be of any form, but are usually taken to be a series expansion in $\mathbf{y}$. This method has been successfully tested with Taylor- and Fourier-series expansions. In this manner, the modeling is done by finding the best expansion coefficients to reproduce the experimental data. Often the situation arises where the form of the functions $\{f_i\}$ is known, but the coefficients are unknown, e.g., this occurs frequently with rate equations for chemical processes. This added information greatly reduces the number of undetermined parameters, thus making the modeling computationally more efficient.

The modeling procedure begins by choosing some trial coefficients. The error in these parameters can be computed by taking each data point $\mathbf{x}(t_n)$ as an initial condition for the model equations. The predicted value $\mathbf{y}(t_{n+1})$ can then be calculated for CM's as

$$y_i(n+1)=f_i(\mathbf{x}(n)),\quad i=1,\ldots,N$$

or for ODE's as

$$y_i(t_{n+1})=x_i(t_n)+\int_{t_n}^{t_{n+1}} f_i(\mathbf{y}(t'))dt',\quad i=1,\ldots,N \qquad (1)$$

and compared to the experimentally determined value. Previous work[10] has shown that more stable models can often be obtained by comparing the prediction and the experimental data several time steps into the future. For the present analysis, we will predict the value only to the time of the first unused experimental data point. The error in the model is thus obtained by summing these differences

$$\chi_\nu^2=\frac{1}{N(M-1)-N_c}\sum_{i=1}^{M}\sum_{j=1}^{N}\frac{1}{\sigma_{ij}^2}[y_j(t_i)-x_j(t_i)]^2 , \qquad (2)$$

where N_c is the number of free coefficients, M is the number of data points, and σ_{ij} is the error in the j^{th} vector component of the i^{th} measurement. The task of finding the optimal model parameters has now been reduced to a χ^2 minimization problem. Thus, the best parameters are determined by

$$\frac{\partial\chi_\nu}{\partial p_i}=0,\quad \forall i .$$

Therefore the ability to determine these coefficients rests upon the strength of the algorithm employed to search through the space of parameters. Since this has been formulated as a standard χ_ν^2 problem, all of the normal statistical tests can be applied. Typically, $\chi_\nu\simeq 1$ implies that the modeling was successful; however, more sophisticated tests can be applied as well, e.g., F test, etc.[12] If the experimental uncertainties σ_{ij} are unavailable, this normalization factor can simply be removed from Eq. (2). This means that the χ_ν tests cannot be applied, but the best possible model can still be determined by locating the global minimum of χ_ν in the parameter space.

B. Modeling with hidden variables

As stated, this method works when all the variables can be measured. Unfortunately, this is almost never the case in real experiments. Frequently, one or more of the variables is hidden, i.e., it cannot be directly measured. This requires that a new method be utilized to reconstruct the equations of motion for the dynamics.

To develop a technique for reconstructing the dynamics of systems with hidden variables, we assume that only one variable w is hidden, i.e., $N_h=1$; and our experimental data $\mathbf{x}$ contains N_o observables, $\mathbf{x}\in R^{N_o}$. The restriction on N_h is purely for illustration. The model equations are identical to the previous case, with

$$y_i(n+1)=f_i(\mathbf{y}(n)),\quad i=1,\ldots,N \qquad (3)$$

for a system of maps where $N=N_o+N_h$, $\mathbf{y}\in R^N$, and $\mathbf{y}(n)=\mathbf{x}(n),w(n)$. The predicted values $\mathbf{y}(n)$ are calculated from

$$y_i(n+1)=f_i(\mathbf{x}(n),w(n)),\quad i=1,\ldots,N \qquad (4)$$

thus requiring that we know $w(n)$. Since experimental data are available for the other N_o variables for all n, we can use those to solve for $w(n)$ as

$$f_i(\mathbf{x}(n),w(n))-x_i(n+1)=0,\quad i=1,\ldots,N . \qquad (5)$$

Having one hidden variable, only one of these equations is needed to solve for $w(n)$. If f_i were a Taylor-series expansion to l^{th} order, then solving for $w(n)$ produces l roots. In practice, since we do not expect our first guess for the model coefficients to be correct, each of the N_o equations is solved, thus generating lN_o possible solutions for $w(n)$. To be accepted, these roots are required to be real and to satisfy any known physical bounds upon the value of $w(n)$. Aside from these constraints, there is no *a priori* method to determine which root is correct. Therefore each is tried in turn with the best root chosen according to predictive accuracy.

Using one such $w(n)$, we determine $w(n+1)$ from Eq. (4). Now the predicted values $\mathbf{y}(n+2)$ can be calculated from $(\mathbf{x}(n+1),w(n+1))$ and compared to $\mathbf{x}(n+2)$. The error in the model is thus

$$\chi_\nu^2=\frac{1}{N_o(M-2)-N_c}\sum_{i=2}^{M}\sum_{j=1}^{N_o}\frac{1}{\sigma_{ij}^2}[y_j(i)-x_j(i)]^2 . \qquad (6)$$

The best value of $w(n+1)$ is stored for the next iteration. If at step $n+1$, no acceptable roots can be found, the previous best $w(n+1)$ is used to continue the calculation. Note that we do not use $w(n)$ to immediately calculate $\mathbf{y}(n+1)$. This is because the $\mathbf{x}(n+1)$ have already

been used in determining $w(n)$, thus causing Eq. (6) to improperly characterize the accuracy of the model.

This analysis can be extended simply to handle an arbitrary number of hidden variables. If we have at least as many observables as hidden variables $N_o \geq N_h$, then rather than solving one equation in one unknown, we must solve a system of equations in N_h unknowns.

$$\begin{aligned} &f_1(\mathbf{x}(n),\mathbf{w}(n))-x_1(n+1)=0\ , \\ &f_2(\mathbf{x}(n),\mathbf{w}(n))-x_2(n+1)=0\ , \\ &\vdots \\ &f_{N_h}(\mathbf{x}(n),\mathbf{w}(n))-x_{N_h}(n+1)=0\ , \end{aligned} \tag{7}$$

where $\mathbf{w}\in R^{N_h}$. Once the $\mathbf{w}(n)$ have been determined, the analysis proceeds exactly as before.

When $N_h > N_o$, we cannot generate enough equations from the first N_o model equations using only $\mathbf{x}(n)$ and $\mathbf{x}(n+1)$ to uniquely determine the N_h initial conditions of $\mathbf{w}(n)$. Therefore we create additional equations using more experimental data points:

$$\begin{aligned} &f_1(\mathbf{x}(n),\mathbf{w}(n))-x_1(n+1)=0\ , \\ &\vdots \\ &f_{N_o}(\mathbf{x}(n),\mathbf{w}(n))-x_{N_o}(n+1)=0\ , \\ &f_{N_o+1}(\mathbf{x}(n),\mathbf{w}(n))-w_1(n+1)=0\ , \\ &\vdots \\ &f_N(\mathbf{x}(n),\mathbf{w}(n))-w_{N_o}(n+1)=0\ , \\ &f_1(\mathbf{x}(n+1),\mathbf{w}(n+1))-x_1(n+2)=0\ , \\ &\vdots \\ &f_{N_o}(\mathbf{x}(n+1),\mathbf{w}(n+1))-x_{N_o}(n+2)=0\ . \end{aligned} \tag{8}$$

In Eqs. (8), the first N_o equations are the same ones used previously. However, the second N_h equations come from the model equations representing the dynamics of the hidden variables. These are needed to calculate $\mathbf{w}(n+1)$ as a function of $\mathbf{w}(n)$. The third set of equations determines N_o more of the $w_i(n)$ from $\mathbf{x}(n+2)$, etc. This continues until enough equations have been generated to determine $\mathbf{w}(n)$; $N_h([N_h/N_o]+1)$ equations are needed, where the square brackets indicate the greatest integer.

While the above method for handling $N_h > N_o$ is straightforward, an alternate approach exists which may be more efficient in some situations. In such cases, there will be $N_u = N_h - N_o$ unknown variables that cannot be reconstructed directly from the N_o observations at a single time. If the system is not chaotic, we can simply add one unknown parameter for each of the N_u unknown variables representing that variable's value at time t_0. This value can then be iterated at each time step and carried forward to the next step, exactly as is done when no acceptable roots can be found. This makes the optimization of the model coefficients more difficult due to an increase in the number of local minima in the χ_ν landscape. Fortunately, the correct minimum can still be determined by χ_ν satisfying the statistical tests. Foreknowledge of the proper form of the $\{f_i\}$ will be very useful for reducing the complexity of this search.

However, if the system is chaotic, a single initial condition iterated through the entire data set cannot be expected to remain close to the experimental trajectory even for an accurate model. Thus several initial conditions will be needed for each unknown variable. The time between these initial conditions τ will depend in general upon the rate of information loss, i.e., the Lyapunov exponents.[13,14] An additional term can be added to χ_ν^2 representing the distance between the previous initial condition p_{ij} iterated up to the time of the next initial condition and that initial condition $p_{i,j+1}$ so that

$$\begin{aligned} \chi_n^2 = &\frac{1}{N_o(M-2)-N_c}\sum_{i=3}^{M}\sum_{j=1}^{N_o}\frac{1}{\sigma_{ij}^2}[y_j(i)-x_j(i)]^2 \\ &+\frac{1}{N_u(N_T-1)-N_T} \\ &\times\sum_{i=1}^{N_u}\sum_{j=1}^{N_T-1}\frac{1}{\sigma_{i,(j+1)\tau^2}}[p_{i,j+1}-f_i^T(p_{i,j})]^2\ , \end{aligned} \tag{9}$$

where N_T is the number of parameters needed for each undetermined variable. The added term in Eq. (9) is a simple end-point–matching condition that helps reduce the number of local minima in χ_ν^2. Because of the persistent complexity of the χ_ν^2 landscape, knowledge of the form of the model equations will be necessary, and they must be limited to only a few free parameters. This variation of our hidden variables reconstruction is generally the least robust of the options described, but it may be useful in some special situations.

While the discussion of hidden variable has thus far focused upon maps, it can easily be extended to reconstructing hidden variables in continuous systems (ODE's). For this, Eq. (5) becomes

$$f_i(\mathbf{x}(t_n),w(t_n))-\dot{x}_i(t_n)=0\ ; \tag{10}$$

so we must calculate the first derivatives of the observables $\dot{\mathbf{x}}$. The modeling process then proceeds exactly as in Eqs. (3)–(6) except that a modification of Eq. (1) replaces Eq. (4). The extension to multiple hidden variables is equally straightforward by simple modifications of Eqs. (5)–(9) and with the same restrictions as before. The only caveat to this process comes from the introduction of additional noise which occurs during the process of computing the derivatives of the experimental data sets $\dot{\mathbf{x}}(t_n)$ and the integration of the model equations necessary to calculate $\mathbf{w}(t_{n+1})$ and $\mathbf{y}(t_{n+2})$. To compensate for these errors, the σ_{ij} should be adjusted appropriately, thereby making it possible for the value of χ_ν to satisfy the statistical tests. This computational noise is discussed further in the examples of the Rössler and Lorenz systems, Sec. III.

C. Applicability

Unfortunately, the hidden variable reconstruction method detailed above is not a panacea. Further study has shown that it is not possible to construct models in some special cases. In this section, we first give an analytical example for how the model parameters are derived, and then provide tests to determine when a model will fail and possibly which terms cause the failure.

To illustrate how the parameters for a given model may be derived analytically, consider the example

$$\begin{aligned} x(n+1)&=2x(n)+\epsilon w(n)\ , \\ w(n+1)&=w(n)+\epsilon x(n)\ , \end{aligned} \tag{11}$$

where $x,w\in R^1$, all the $x(n)$ are known, and w is the hidden variable. To determine the free parameter ϵ we must also compute the initial condition $w(n)$, so we actually have two undetermined parameters. The solution may be obtained from the equations

$$\begin{aligned} x(n+1)&=2x(n)+\epsilon w(n)\ , \\ w(n+1)&=w(n)+\epsilon x(n)\ , \\ x(n+2)&=2x(n+1)+\epsilon w(n+1)\ , \end{aligned}$$

such that

$$w(n)=\frac{x(n+1)-2x(n)}{\epsilon}\ ,$$

$$\epsilon=\pm\left[\frac{x(n+2)-3x(n+1)+2x(n)}{x(n)}\right]^{1/2}.$$

Thus using $(x(n),x(n+1),x(n+2))$ we can determine $w(n)$ and ϵ uniquely. In general, $[(N_c+N_h)/N_o]+1$ data points will be required, where N_c is the number of free coefficients. Although this analytical method effectively demonstrates how the reconstruction of hidden variables works, it is impractical when the data contain any noise since we must then average the coefficients over the calculations for all the n experimental data points. Thus the statistical approach described in Sec. II is a more reliable technique.

Even though model (11) could be solved easily,

$$\begin{aligned} x(n+1)&=2x(n)+\epsilon w(n)\ , \\ w(n+1)&=w(n) \end{aligned} \tag{12}$$

is a simple example of a model which does not have unique solutions. Obviously, ϵ and $w(n)$ cannot be obtained independently. The χ^2_ν modeling procedure as currently described cannot distinguish these situations, and will arbitrarily select one solution from the possible solutions. Fortunately, prior to the modeling process we can determine if a given model has unique solutions, assuming that the model contains no extraneous terms. This is done by computing the Wronskian matrix for the equations required for an analytic solution, i.e., the Wronskian matrix for the system of equations

$$\{g_k\}=\begin{cases} x_i(n+j+1)-f_i(\mathbf{x}(n+j),\ \mathbf{w}(n+j);\mathbf{c})=0\ , & i=1,\ldots,N_o,\ \ j=0,\ldots,m-1 \\ w_i(n+j+1)-f_{i+N_o}(\mathbf{x}(n+j),\ \mathbf{w}(n+j);\mathbf{c})=0, & i=1,\ldots,N_h,\ \ j=0,\ldots,m-1 \\ x_i(n+m+1)-f_i(\mathbf{x}(n+m),\mathbf{w}(n+M);\mathbf{c})=0, & i=1,\ldots,N_c+N_h-N_o m \end{cases}$$

where $m=[(N_c+N_h)/N_o]$. We define the Wronskian matrix as the first derivative of each g_k with respect to the coefficients c_i and the values of the hidden variables $w_i(n+j)$. The determinant of this matrix is the Wronskian of the system W. If $W\neq 0$, the model will have unique solutions. This is illustrated by the Wronskian matrix for model (11)

$$\begin{bmatrix} \frac{\partial g_1}{\partial\epsilon} & \frac{\partial g_1}{\partial w(n)} & \frac{\partial g_1}{\partial w(n+1)} \\ \frac{\partial g_2}{\partial\epsilon} & \frac{\partial g_2}{\partial w(n)} & \frac{\partial g_2}{\partial w(n+1)} \\ \frac{\partial g_3}{\partial\epsilon} & \frac{\partial g_3}{\partial w(n)} & \frac{\partial g_3}{\partial w(n+1)} \end{bmatrix} = \begin{bmatrix} w(n) & \epsilon & 0 \\ x(n) & 1 & -1 \\ w(n+1) & 0 & \epsilon \end{bmatrix}$$

for which $W\neq 0$, thus confirming that unique solutions exist. However, $W=0$ for model (12), as expected.

We can also extend the use of the Wronskian matrix to help identify which parameters may be causing the difficulty. For a model with $W=0$, we can systematically fix the value of each of the c_i or $w_i(n+j)$. This is equivalent to removing the corresponding column and one row from the Wronskian matrix. If W is still 0, pairs of parameters may be fixed, etc., until $W\neq 0$, thus determining the rank of the matrix. More than this, if $W\neq 0$ for any choice of parameters to fix, then the system is underdetermined in the fashion common in linear systems. However, if only a few of the parameters can be fixed such that $W\neq 0$, then these are the parameters which cannot be determined independently. If a best guess can be made from physical considerations for all but one of these parameters, then the modeling can proceed normally.

Using the Wronskian matrix in this manner, we can construct the model equations such that unique solutions can be determined. If during modeling some of the terms are found to be unnecessary, $c_i\simeq 0$, the resulting model must be checked again to determine if the solutions are indeed unique. Thus we can have complete knowledge as to the success of the modeling procedure.

Up to now, all examples discussed have been simple polynomial expansions. The methods described are not dependent upon this and apply to any form of the model equations. There is a special class of models, though, which we should mention. It can occur that Eq. (5) or (10) can produce an infinite number of solutions, such as for

$$x(n+1)=\tan[\alpha w(n)] ,$$

$$w(n+1)=\beta w(n) ,$$

To treat this occurrence, we can begin by applying any known physical constraints on $w(n)$. However, if the quality of the model's predictions changes smoothly with $w(n)$, we can do the equivalent of a gradient search through the possible values of $w(n)$ to locate the correct one. This is not a perfect solution, but it is a reasonable approach.

III. EXAMPLES

Specific examples of the reconstruction of hidden variables will now be given for simulated data. This is an effective method of testing since the correct model parameters will already be known. The first example is for two coupled logistic maps,

$$x(n+1)=\lambda_1 x(n)[1-x(n)]+dy(n) ,$$
$$y(n+1)=\lambda_2 y(n)[1-y(n)]+dx(n) . \quad (13)$$

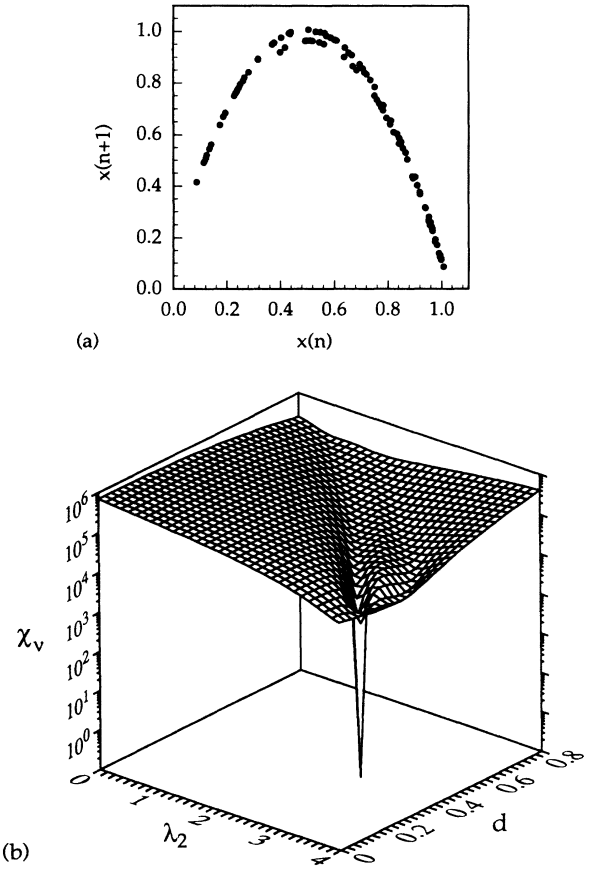

FIG. 1. A simple reconstruction for experimental data obtained from Eqs. (13) is shown in (a), and the resulting χ_ν landscape is in (b). From the representation in (a), we see that the data are clearly not one-dimensional, thereby necessitating a hidden variable reconstruction. The χ_ν surface in (b) has two local minima, but the correct solution, at $\lambda_2=3.5$ and $d=0.2$, is several orders of magnitude lower than the other. Even a simple gradient search is effective in cases such as this.

The experimental data were generated using $\lambda_1=2.0$, $\lambda_2=3.5$, and $d=0.2$. For the tests using Eqs. (13), we took y to be the hidden variable. A next return map for the observable data is shown in Fig. 1(a). We begin by assuming that the form of the equations of motion is known, and λ_2 and d are the only free parameters. In Fig. 1(b) we show the χ_ν landscape for these two parameters. This landscape has two minima: $\lambda_2=3.5$, $d=0.2$; and $\lambda_2=3.875$, $d=0.275$. According to the value of χ_ν, the first minimum qualifies as a solution, whereas the second minimum fails badly. In this case, the χ_ν landscape was sufficiently simple that even an unaided gradient search would have a high probability of locating the correct minimum.

For the previous test, the noise in the data was nothing more than roundoff error in the last significant digit. To better study the effects of noise, we have added band-limited noise in the range $-\epsilon$ to ϵ to the experimental data. We use the same model as above, so we still have λ_2 and d as our free parameters. In Fig. 2(a), we show the effects of dynamical noise. This means that the noise was added to the experimental system at each step in the mapping. This would be like making perfect measure-

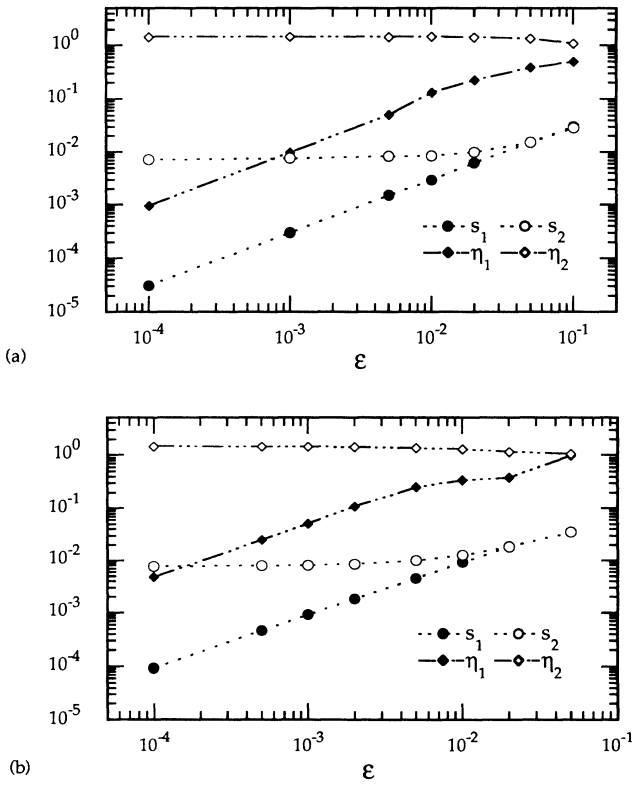

FIG. 2. These plots show s vs ϵ for the two minima in Fig. 1(b), where s is the error in the model without being normalized by σ_{ij}. In (a) the effects of noise added directly to the dynamical system are shown upon the depth s and accuracy η of the local minima. In (b) the measurements contained noise, but the dynamical system itself was noiseless.

ments of a dynamical system which was being constantly perturbed. The plot shows s,

$$s^2=\sum_{i=2}^{M}\sum_{j=1}^{N_o}[y_j(i)-x_j(i)]^2 ,$$

the error in the model as in Eq. (6) but not normalized with σ_{ij}, versus the maximum noise amplitude ϵ, for the two local minima shown in Fig. 1(b). We see that for very small noise levels, the error of the first minimum s_1 is $s_1\simeq\epsilon$, corresponding to $\chi_\nu\simeq 1$, and that the error of the second minimum s_2 is approximately constant and consistently fails any χ_ν tests. Since the experimental data no longer represent a trajectory of the true system with complete accuracy, the model coefficients will not be precise and will vary from one data set to another. The average difference between the model parameters and the correct values is also shown in Fig. 2(a), as η versus ϵ, where $\eta^2=(\lambda_1^m-\lambda_1)^2+(d^m-d)^2$ with the superscript m indicating the coefficients obtained from the model.

When the noise reaches approximately 2%, the trends just noted begin to break down and it becomes increasingly difficult to distinguish the two minima. For approximately 5% noise, both minima now satisfy our statistical tests, and it is no longer possible to determine which set of parameters is correct. The point at which this occurs is not generic, but rather depends specifically upon the system being studied. Clearly, the added noise has the effect of smoothing the s landscape, thus the first minima is effected most drastically. For the current case, to be able to determine which minimum is the correct solution for up to 5% noise is a significant achievement. Note that in all these simulations, Eq. (4) is used only to predict one time step forward. When the predictions are carried several time steps into the future, the false minima become more shallow, providing greater tolerance to noise.[10]

The reason this procedure eventually becomes noise limited is easily understood by noting that the local minima in the χ_ν landscape generally arise from the several-to-one nature of many maps. Each solution represents a different possible trajectory. Since we cannot say which will conform to the true trajectory, all solutions are treated equally, thereby allowing for a piecewise fit to the experimental trajectory. When the noise level is on the order of the errors in the piecewise fit, the correct minimum can no longer be distinguished based upon the χ_ν value. Fortunately, the models corresponding to false minima are typically unstable. Thus even in the presence of large amplitude noise, one should be able to distinguish the correct minimum based upon model stability.

For Fig. 2(b), we have added the noise after the data were generated. This is equivalent to making noisy measurements of a perfect system. The results here are essentially the same as in Fig. 2(a) except that the modeling is less tolerant to the noise, and the two minima become indistinguishable at a noise level of approximately 2%. This increased sensitivity to measurement noise was found to be due to the fact that dynamical noise actually increases the available information by portraying the dynamics in a larger region of the state space, whereas the measurement noise causes nothing but a loss of information.[15]

The case just shown had only two free parameters so that the χ_ν landscape could be represented graphically. This is, of course, not necessary and other tests were tried with more free parameters. When we assumed a model of the form

$$\begin{aligned} x(n+1)&=p_1x(n)=p_2x(n)^2+p_3y(n) ,\\ y(n+1)&=p_4y(n)+p_5y(n)^2+p_6x(n) , \end{aligned} \tag{14}$$

six free parameters were used, and we found many more local minima in the χ_ν landscape. However, 10% of the volume of this parameter space was within the basin of attraction of the correct minimum, and this minimum still satisfied χ_ν tests. Again, all of the other local minima had χ_ν values several orders of magnitude larger. Only roundoff noise is present here and for all of the following tests.

Next we consider data obtained from a set of ODE's. The first such example is the Rössler system,

$$\begin{aligned} \dot{x}&=-y-z ,\\ \dot{y}&=x+ay ,\\ \dot{z}&=b+(x-c)z , \end{aligned} \tag{15}$$

where we have used $a=0.343$, $b=1.83$, and $c=9.75$ to generate the data. For this modeling, we use the method for reconstructing ODE's and assume a model of the same form as Eqs. (15) with b and c unknown. The experimental data with $z(t)$ hidden is shown in Fig. 3(a). We have again calculated the χ_ν landscape for the free parameters, Fig. 3(b). In this case, there is only one minimum, and it occurs at the proper parameter values. In this case, however, the minimum is not as sharp as the previous example. This smoothing of the landscape comes from errors introduced through the calculation of the derivatives and integrals. The greatest error comes from the differentiation which was done simply as

$$\dot{x}(t_n)=\frac{x(t_{n+1})-x(t_{n-1})}{t_{n+1}-t_{n-1}} .$$

This error was not incorporated into χ_ν in Fig. 3 so as to illustrate the cumulative effect of these computational errors. The most remarkable aspect of this example is that the data in Fig. 3(a) only contain two excursions along the axis of the hidden variable, even though the full attractor has a fully developed funnel,[16] and yet the hidden variables were still reconstructed effectively. When a spline fit or other more accurate differentiation scheme is employed, the minimum in Fig. 3(b) becomes much more distinct.

We also consider the Lorentz system,

$$\begin{aligned} \dot{x}&=\sigma(y-x) ,\\ \dot{y}&=rx-y-xz ,\\ \dot{z}&=-bz+xy . \end{aligned}$$

The data were generated with $\sigma=10$, $r=50$, and $b=\frac{8}{3}$.

For this case, the time between data points is much larger than the previous example, Fig. 4(a). This tends to amplify differentiation errors, thus causing the χ_ν landscape to be even more strongly smoothed, Fig. 4(b). Again, we have not incorporated these errors into χ_ν so as to illustrate the magnitude of the effect. In fact, the minimum is still at the correct parameter values; but, if the differentiation errors are not symmetrically distributed, the minimum may wander some from the true values. The modeling in this case is remarkably successful, considering the coarseness of the experimental data.

In both of the examples where the models were ODE's only one minimum was found in the landscape. Because of the integration needed to predict the value of the observables at the next time, Eq. (1), we are no longer simply doing a one-step prediction as was done for the CM. Thus stable solutions are preferred and fewer local minima occur in the χ_ν landscape. This implies that it may be easier to generate models based on ODE's. We have, of course, chosen systems which could be modeled with a finite number of polynomial terms. If this fails, one may simply need to choose a different set of expansion functions. When appropriate expansion functions were selected, the only real limitations came from failings in the search algorithm employed to minimize χ_ν or from noise present in the data or introduced through the modeling technique.

As a final example, we study the standard map,

$$p_{n+1}=p_n-\frac{\alpha}{2\pi}\sin(2\pi q_n)\mathrm{mod}1\ ,$$

$$q_{n+1}=q_n+p_{n+1}\mathrm{mod}1\ ,$$

in the case where $\alpha=0.97$, and $p_0=0.5$, and $q_0=0.97$, Fig. 5(a). The numerical experiment was performed with $\{q_n\}$ assumed hidden and the trial model

$$p_{n+1}=p_n-\frac{\alpha}{2\pi}\sin(2\pi q_n)\mathrm{mod}1\ ,$$

$$q_{n+1}=q_n+\beta q_n^2+p_{n+1}\mathrm{mod}1\ ,$$

where βq_n^2 is obviously an extraneous term. Although the model is not polynomial, the analysis proceeds exactly as

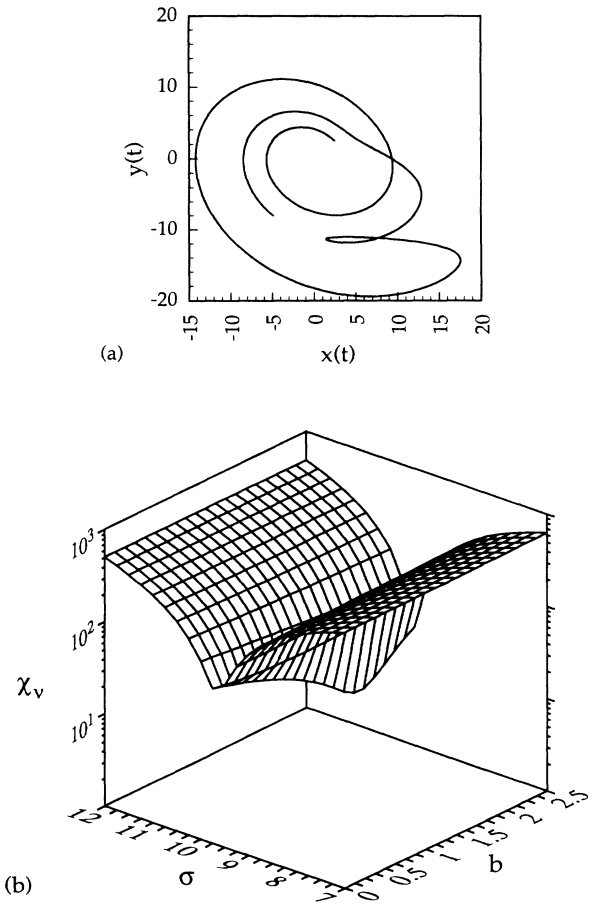

FIG. 3. Observable data are shown in (a) with the corresponding χ_ν landscape in (b). The data set was very planar and has only two loops which require a third dimension for explanation. Nonetheless, it is sufficient to produce a clearly defined minimum in χ_ν.

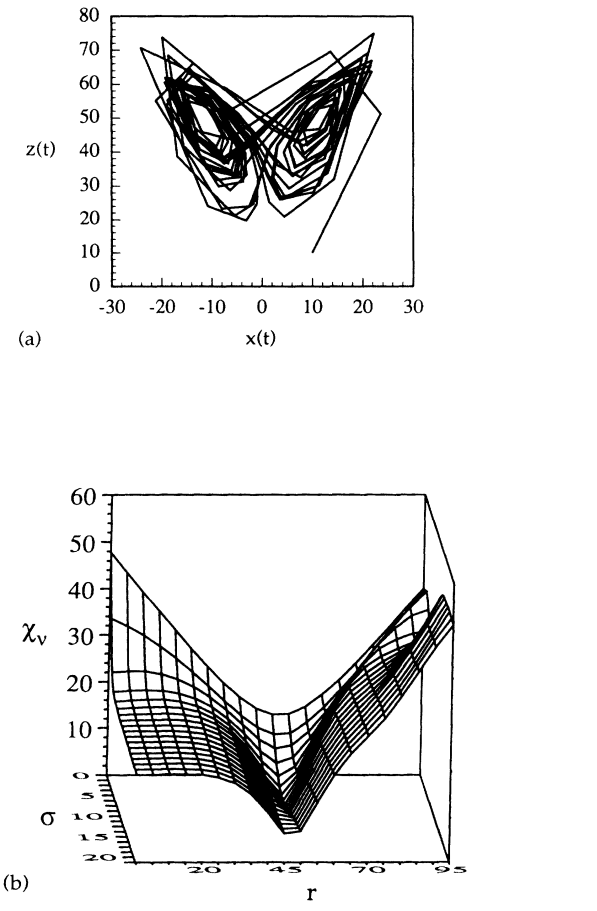

FIG. 4. In (a) the observable data are shown. Note that these data are very coarse in time, thus making the calculation of $\dot{\mathbf{x}}$ very noisy. Still, (b) shows that a minimum is produced in χ_ν corresponding to the correct parameter values. $\chi_\nu\neq 1$ at the minimum, because we have not incorporated the computational noise into σ_{ij} so as to illustrate the effect.

for the other map examples. Because of the arcsine function used to solve for q_n, two possible solutions are obtained at each iteration. The resulting χ_ν landscape is shown as a function of α and β, Fig. 5(b). The minimum does occur at the correct value of α, but more importantly β is recognized as a spurious parameter with a preferred value of zero. We have found in further examples that as more extraneous terms are added, the χ_ν landscape has an increasing number of local minima, but the correct minimum is still apparent.

IV. CONTROL

Much progress has been made in recent years in the area of controlling nonlinear systems without the continuous feedback required by traditional methods.[17,18] The success of this nonlinear control theory hinges upon the creation of a good model for the system. To control a system, we consider the following equations:

$$\dot{\mathbf{x}}=\mathbf{f}(\mathbf{x})+\mathbf{F}(t) \quad \text{(experimental dynamics)} ,$$

$$\dot{\mathbf{q}}=\mathbf{h}(\mathbf{q}) \quad \text{(model dynamics)} ,$$

$$\dot{\mathbf{u}}=\mathbf{g}(\mathbf{u}) \quad \text{(goal dynamics)} .$$

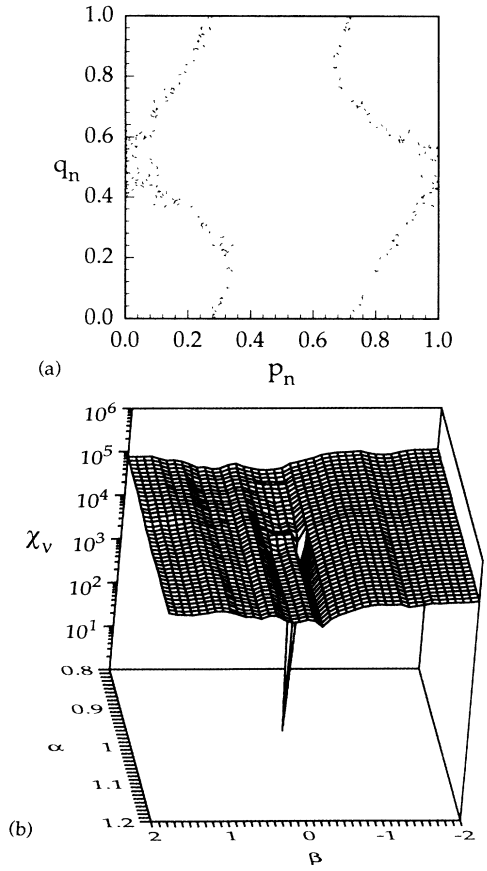

FIG. 5. Both variables of the system are shown in (a), but the $\{q_n\}$ were assumed to be hidden during modeling. (b) shows the χ_ν landscape in α and β with one deep minimum at $\alpha=0.97$, $\beta=0.0$, satisfying the statistical tests and one shallow minimum near by.

The equation for the experimental dynamics represents the actual system to be controlled. The goal dynamics represents the system to which we wish to drive the original system. The driving force needed to control the system is

$$\mathbf{F}(t)=\mathbf{g}(\mathbf{u})-\mathbf{h}(\mathbf{u}) .$$

Previous work has shown that $\mathbf{g}(\mathbf{u})$ cannot be chosen arbitrarily, but is subject to certain stability constraints and restrictions on the initial conditions when the driving force is applied. These issues have been detailed elsewhere,[18] and we will assume that all these conditions are met.

For effective entrainment of the experimental system to the goal dynamics, a model must be constructed for which $\mathbf{h}\simeq\mathbf{f}$. Chang, Hübler, and Packard[19] have shown that when the model is not perfect the distance between the experimental system and the goal dynamics scales linearly with the error in the model coefficients. This indicates that the control is stable despite possible small modeling errors produced during the hidden variable reconstruction.

One case that can arise is a system in which some of the variables are hidden from observation, thus hidden during modeling, but all the variables may be controlled. For a numerical experiment, we consider

$$\dot{\mathbf{x}}=\mathbf{f}(\mathbf{x})+\mathbf{F}(t), \quad \mathbf{f}(x)=\begin{cases}\sigma(y-x)\\ rx-y-xz\\ -bz+xy\ ;\end{cases}$$

$$\dot{\mathbf{u}}=\mathbf{g}(\mathbf{u}), \quad \mathbf{g}(\mathbf{u})=\begin{cases}\sigma'(v-u)\\ r'u-v-uw\\ -b'w+uv\ .\end{cases}$$

Here $\mathbf{x}(t)$ is the experimental system and $\mathbf{u}(t)$ is the goal system. We have considered the specific case where

$$\sigma=10.0, \quad \sigma'=\sigma ;$$

$$r=50.0, \quad r'=125.0 ;$$

$$b=\tfrac{8}{3}, \quad b'=b .$$

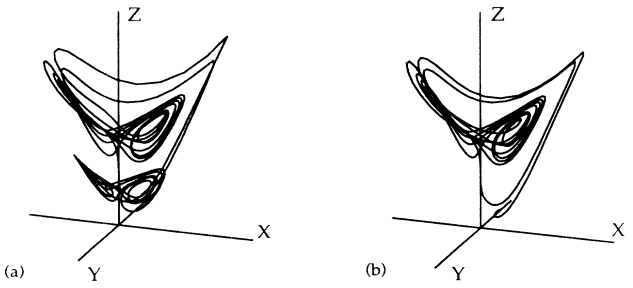

FIG. 6. The trajectory of the experimental system is shown in (a) and the goal system in (b). The driving force (and thus the goal dynamics) does not begin at the same time as (a). Thus the experimental system begins on the attractor in the lower half of (a) before driving, but rapidly entrains to the goal dynamics that correspond to the upper attractor.

During the modeling process, we assumed that $y(t)$ was hidden, but the original parameters were reconstructed without difficulty. Thus this case is not significantly different from previous studies.[17] For the goal system, we have chosen simply to drive the system to a higher Reynold's number. In Fig. 6 we show the results of this control. Figure 6(a) shows the trajectory of the experimental system, and Fig. 6(b) shows the trajectory of the goal system (turned on at a later time) which was used to drive the experimental system. In the experimental system, the dynamics begin in the lower half of Fig. 6(a) where the trajectory is on the attractor for the undriven system. When the driving is turned on, the experimental system rapidly entrains to the goal system, as expected.

This example was shown because there may be important situations where some variables are hidden during the modeling process, but can still be controlled. One such case might be in chemical systems where the reactants are known, but only a subset of these can be continuously monitored. By using the techniques of Sec. II, we can reconstruct the rate equations for the chemical reactions. To control the system may then be simple since one might easily be able to add reactants to the system even though the concentrations of those reactants are hidden from observation. Thus the combination of hidden variable reconstruction and nonlinear control may have important applications in chemical systems.

The second example we wish to consider differs from the previous case in that we wish to control the system even when some of the variables are not controllable, i.e., a driving force cannot be applied. Consider the following example:

$$\dot{\mathbf{x}}=\mathbf{f}(\mathbf{x})+\mathbf{F}(t), \quad \mathbf{f}(\mathbf{x})=\begin{cases}\sigma(y-x)\\ rx-y-xz\\ -bz+xy\ ;\end{cases}$$

$$\dot{\mathbf{u}}=\mathbf{g}(\mathbf{u}), \quad \mathbf{g}(\mathbf{u})=\begin{cases}\sigma'(v-u)\\ r'u-v-uw\\ f_z(\mathbf{u})\ ;\end{cases}$$

$$\mathbf{F}(t)=\mathbf{g}(\mathbf{u})-\mathbf{f}(\mathbf{u})\ .$$

In this case, we again took $y(t)$ to be the hidden variable, but we also chose to make $z(t)$ an uncontrollable variable, so $\mathbf{F}$ applied only to the x and y components. (There is no reason to assume that the hidden variables are also uncontrollable variables, or vice versa.) Since we have an uncontrollable variable, we can no longer choose to drive the system to an arbitrary goal. Instead, we must restrict our goals just to altering the controllable variables. This is similar to previous work on controlling nonlinear systems from Poincaré maps.[20] Since the driving there could be applied only in the plane of the Poincaré map, the goal was also required to be a Poincaré map. For our test, we again chose to drive the system to a different r value:

$$\sigma=10.0, \quad \sigma'=\sigma\ ;$$

$$r=50.0, \quad r'=125.0\ ;$$

$$b=\tfrac{8}{3}\ .$$

It is important to understand the differences between the driving force here and the fully controlled case. When all the variables were controllable, even though our goal did not attempt to alter the dynamics of either $x(t)$ or $z(t)$, we could have done so had we chosen to. In the present case, we cannot drive $z(t)$ directly, so we must restrict our goals to those systems which leave the z dynamics unchanged. (See *Note added in proof.*) The distance between the experimental and goal systems after the driving was turned on is shown in Fig. 7. We see that the experimental system still entrains rapidly to the goal system. Obviously then, not all systems can be entrained by this method, but it will be possible in many cases. Well-chosen goal dynamics is important for the entrainment.

The problem of having uncontrollable variables is an interesting one which is still under investigation. Current results indicate that it may be possible to pick specific goals such that the uncontrollable variable can be entrained indirectly to a new dynamics. The types of entrainment that are possible in systems with one or many uncontrolled variables is an area of continuing research.

V. DISCUSSION

In the preceding analysis, we have demonstrated that it is possible to reconstruct hidden variables in a dynamical system directly from the observables by simultaneously reconstructing the equations of motion for the system. The greatest advantage of this technique is that the re-

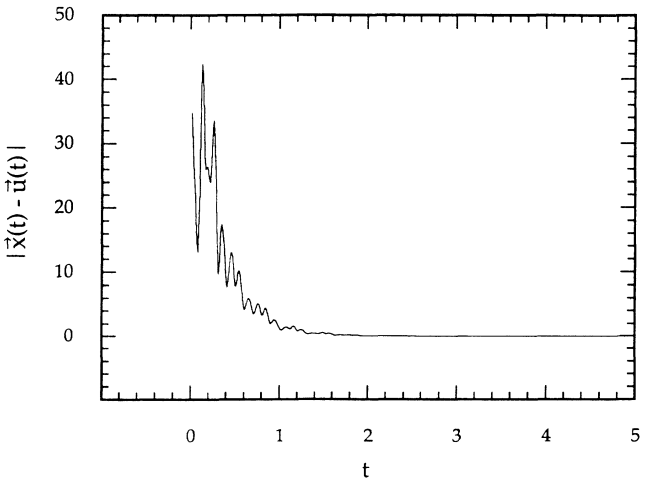

FIG. 7. The distance between the trajectory of the experimental system $\mathbf{x}$ and the goal dynamics $\mathbf{u}$ is shown as a function of time. The driving force was turned on at $t=0$.

sulting models can be interpreted physically so as to understand, as well as forecast, the dynamics. This also has interesting implications in relation to Takens's theorem. If the modeling fails, no statement can be made about the proper embedding of the system. However, since the models being generated are inherently deterministic, when the model is a true representation of the dynamics according to the statistical tests we must have found an embedding for the system. Moreover, unlike Takens's theorem, this can be applied even when noise is present in the experimental data.

Note added in proof. Due to an unfortunate choice of examples, the fully controlled system with the z dynamics left unchanged is identical to the example where $z(t)$ is considered uncontrollable.

ACKNOWLEDGMENTS

This research was funded in part by U. S. Office of Naval Research Grant No. N00014-88-K-0293 and National Science Foundation Grant No. PHY86-58062. Computing support was supplied by the National Center for Supercomputing Applications and the Beckman Institute, both at the University of Illinois, Urbana-Champaign. One of us (J.L.B.) wishes to thank E. A. Jackson and P. Newton for many informative discussions.

*Electronic address: [breeden,alfred]@complex.ccsr.uiuc.edu.

[1]N. Packard, J. Crutchfield, J. Farmer, and R. Shaw, Phys. Rev. Lett. **45**, 712 (1980).

[2]H. Whitney, Ann. Math. **37**, 645 (1936); **45**, 220 (1936) **45**, 247 (1936).

[3]F. Takens, in *Dynamical Systems and Turbulence,* Vol. 898 of *Lecture Notes in Mathematics,* edited by D. A. Rand and L. S. Young (Springer-Verlag, Berlin, 1981), pp. 366–381.

[4]G. Meyer-Kress, *Dimensions and Entropies in Chaotic Systems,* Vol. 32 of *Springer Series in Synergetics* (Springer-Verlag, Berlin, 1986).

[5]J. Farmer and J. Sidorowich, Los Alamos National Laboratory Technical Report No. LA-UR-88-901, 1988.

[6]T. P. Meyer, F. C. Richards, and N. H. Packard, Phys. Rev. Lett. **63**, 1735 (1989).

[7]N. Packard, Center for Complex Systems Research Technical Report No. CCSR-89-10, 1989.

[8]M. Casdagli, (unpublished).

[9]J. Cremers and A. Hübler, Z. Naturforsch, **42a**, 797 (1986).

[10]T. Eisenhammer, A. Hübler, N. Packard, and J. A. S. Kelso, Center for Complex Systems Research Technical Report No. CCSR-89-7, 1989.

[11]J. P. Crutchfield and B. S. McNamara, J. Complex Sys. **1**, 417 (1987).

[12]P. R. Bevington, *Data Reduction and Error Analysis for the Physical Sciences* (McGraw-Hill, St. Louis, 1969).

[13]E. A. Jackson, *Perspectives of Nonlinear Dynamics* (Cambridge University Press, New York, 1989), Vol. 1.

[14]A. Wolf, J. B. Swift, H. L. Swinney, and J. A. Vastano, Physica D **16**, 285 (1985).

[15]J. L. Breeden, F. Dinkelacker, and A. Hübler, Center for Complex Systems Research Technical Report No. CCSR-90-1, 1990.

[16]J. L. Breeden, N. Packard, and H. N. Cohn, Center for Complex Systems Research Technical Report No. CCSR-90-2, 1990.

[17]A Hübler and E. Lüscher, Naturwissenschaften, **76**, 67 (1989).

[18]E. Jackson and A. Hübler, Center for Complex Systems Research Technical Report No. CCSR-90-3, 1989.

[19]K. Chang, A. Hübler, and N. Packard, Center for Complex Systems Research Technical Report No. CCSR-89-5, 1989.

[20]R. Georgii, W. Eberl, E. Lüscher, and A. Hübler, Helv. Phys. Acta **62**, 290 (1989).