Noise in the modeling and control of dynamical systems

Joseph L. Breeden, Friedrich Dinkelacker, 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 4 June 1990)

We demonstrate how noise can be an effective tool in modeling systems whose experimental data sets would normally be limited to a small region of the reconstructed state space. In fact, for systems with stable fixed points, using noise to extend the accessible state-space volume may be the only possibility for constructing a model. We find that noise can also be useful in modeling limit cycles when multiple systems generate the same closed trajectory and the model that represents the true dynamics is desired. We discuss the implications of our method for nonlinear control theory about which important questions on the effects of noise in real-time modeling have arisen.

I. INTRODUCTION

In order to describe nonlinear systems from their experimental time series, delay coordinates, derivatives, or some other state-space representation are frequently useful. This reconstruction of the experimental data is often topologically very simple compared to the possibly chaotic temporal representation.[1,2] The goal of this representation is to permit the dynamics to be modeled with few parameters.

Several techniques currently exist for constructing a model of the system based upon the state-space reconstruction. Implicit in all these methods is the assumption that to obtain a good model, the state space must be well sampled. In many scientific disciplines,[3] systems arise that exhibit limit-cycle or stable-fixed-point behavior where modeling would normally be considered ineffective or impossible. We will demonstrate how the presence of noise can actually be used advantageously in the construction of models for these systems. We must emphasize from the outset that we are concerned with noise within the dynamical system, and not just noise arising from imperfect experimental measurement techniques or related to simulated annealing procedures. Measurement noise does not have the modeling benefits which we will present in the following discussion.

This type of modeling can be extremely useful to the control of nonlinear dynamical systems.[4,5] Typically in nonlinear control theory, knowledge of the dynamics is required in a certain region of the state space. By using noise to increase the sampled state space, we should be able to generate improved models in these regions of interest.

In Sec. II we discuss modeling in the region about a stable fixed point. An analytical justification of our expectations on the effects of dynamical noise and specific examples thereof are presented. Modeling in the neighborhood of closed trajectories and chaotic attractors is discussed with examples in Sec. III. These ideas are then related to nonlinear control theory in Sec. IV, with concluding remarks in Sec. V.

II. MODELING FIXED-POINT DYNAMICS

Many experimental systems exist in a state of stable equilibrium. If the initial transient data are available, and the system is well isolated, the entire data set may consist of a single stable fixed point, or collection of fixed points. From this set of measurements, it is impossible to construct a model of the dynamics. The key to modeling this type of system is to increase the region of state space being observed. The ideal situation is to be able to supply a single sharp kick to the system while simultaneously eliminating all external noise, Fig. 1, thus providing a clean view of the state space for use in the modeling procedure. However, depending upon the specific system being studied, obtaining this type of transient data is often impossible. We intend to show that the continuous addition of random noise to the system can still be a powerful aid to

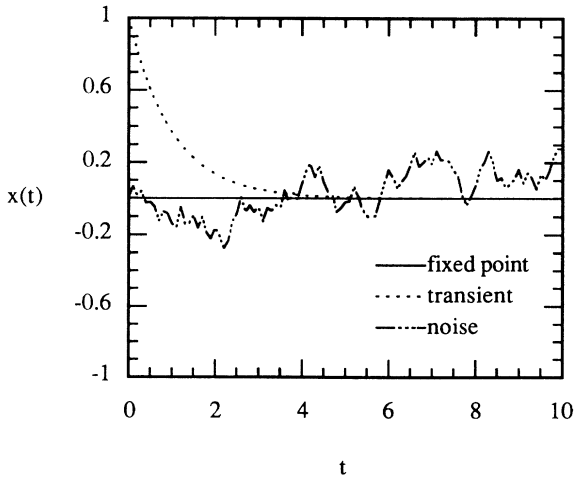

FIG. 1. This figure exhibits three typical situations: a stable fixed point, a single transient to that fixed point, and continuous added noise. These data were generated from Eq. (22) and are representative of the data sets used in the modeling.

modeling because of the small transients resulting from the noise.

Using this simple idea of increasing the observable state space through the addition of noise, we can relate the expected quality of the fit to the amplitude of the noise, ϵ, for some modeling procedures. We expect that, in general, as we gradually increase ϵ from $\epsilon=0$, the sampled state space should grow and the quality of the fit should improve. Alternatively, since we are forced to deal with finite length time series, we cannot expect to observe all the accessible state space with equal probability. Thus the quality of the fit should begin to degrade for large ϵ. Combining these two effects, we demonstrate that the model error is minimized for $\epsilon>0$. In some cases an optimum value of ϵ is shown to exist, while for others a noise threshold is obtained above which all models are reasonably accurate.

A. Noise in gradient dynamics

We can provide an analytic justification for our procedure if we restrict our investigations to the subset of dynamical systems exhibiting gradient dynamics. We begin by requiring that the system be written as a set of ordinary differential equations (ODE) such that

$$\frac{\partial \mathbf{q}}{\partial t}=\mathbf{K}(\mathbf{q}) \tag{1}$$

where $\mathbf{K}(\mathbf{q})$ represents the forces upon the system. When noise is added, we obtain the Langevin equation[6]

$$\frac{\partial \mathbf{q}}{\partial t}=\mathbf{K}(\mathbf{q})+\mathbf{F}(t) \tag{2}$$

where $\mathbf{F}(t)$ is a random force corresponding to noise, and therefore is assumed δ-correlated

$$\langle F_i(t)F_j(t')\rangle=Q_{ij}\delta(t-t'),\quad Q_{ij}\propto\epsilon_i\epsilon_j\ .$$

If we consider not just a single trajectory in state space, but rather a population of trajectories, the corresponding distribution function $f(t)$ is described by the Fokker-Planck equation

$$\frac{\partial f}{\partial t}=-\nabla\cdot(\mathbf{K}f)+\tfrac{1}{2}\sum_{i,j}Q_{ij}\frac{\partial^2 f}{\partial q_i\partial q_j}\ . \tag{3}$$

This is an approximation, keeping derivatives of the δ function only up to second order. In this way, Eq. (3) is a first-order approximation for small-amplitude noise of any type, so long as the noise is symmetric about the origin. In the case of Gaussian noise, Eq. (3) becomes exact.

We are primarily interested in knowing, for a given noise amplitude ϵ, the width of the equilibrium distribution as a way of quantifying the average volume of the phase space observed. To do this, we simply let

$$\frac{\partial f}{\partial t}=0$$

in Eq. (2), yielding the steady-state equation

$$\nabla\cdot(\mathbf{K}f)=\tfrac{1}{2}\sum_{i,j}Q_{ij}\frac{\partial^2 f}{\partial q_i\partial q_j}\ . \tag{4}$$

If the forces $\mathbf{K}(\mathbf{q})$ are derived from a potential field, i.e.,

$$K_i=-\frac{\partial\phi(\mathbf{q})}{\partial q_i}\ , \tag{5}$$

then we have gradient dynamics.[7] If the diffusion coefficients Q_{ij} also satisfy

$$Q_{ij}=\delta_{ij}Q$$

then Eq. (4) has the general solution

$$f(\mathbf{q})=\mathcal{N}e^{-2\phi(\mathbf{q})/Q}\ ,\qquad \int_{-\infty}^{\infty}f(\mathbf{q})d\mathbf{q}=1 \tag{6}$$

where we assumed $\phi(\mathbf{q})$ ensures that $f(\mathbf{q})$ vanishes for $\mathbf{q}\to\infty$. For our subsequent investigations, we will take $Q=c\epsilon^2$, which indicates that the noise amplitude is equal along each axis in the state space and c is a parameter determined by the specific color of the noise. The following results will thus be generally applicable for any type of symmetric noise.

From Eq. (6), we may find the width of the equilibrium distribution along a specified axis as

$$\alpha_i^2=\langle q_i^2\rangle=\int_{-\infty}^{\infty}q_i^2 f(\mathbf{q})d\mathbf{q}\ . \tag{7}$$

The accessible volume of the state space is then

$$V=\prod_i\alpha_i\ .$$

The relationship between the accuracy of a given model and the volume of the state space observed is determined entirely by the specific modeling technique employed. For modeling techniques where this relationship is well defined, it becomes possible to predict the optimum noise amplitude using the analysis detailed above. In the current study, we are interested primarily in global models[8–10] as opposed to various local modeling techniques[11,12] since global models can be better utilized in the nonlinear control theory without feedback. In the following sections, we discuss two types of global modeling.

B. Noise in topological modeling

The first class of models we want to consider can be described as topological models. This technique is based upon obtaining a deterministic embedding of the dynamics in some representative state space. The topology of the dynamics is then modeled by making a transformation of the state space in a set of orthogonal functions. This method is effective even for cases of low dimensional chaos because we are modeling the state space, not the time series.

For the present analysis, we will consider one-dimensional systems and use Fourier transforms.[13] The state-space coordinates may be available observables, derivatives, or delay coordinates. For our one-dimensional system, we can often choose $\{x_n,x_{n+1}\}$ to reconstruct the dynamics. This is modeled as

$$x_{n+1}=h(x_n)=\sum_{k=1}^{N}b_k\sin(\pi k x_n)\ .$$

In order to use a fast Fourier transformation, the state space must be homogeneously distributed. This was achieved by employing a discrete equidistant grid x_i^g for the x_i. The values for $x_{n+1}^g = h(x_n^g)$ at the grid sites were calculated by binning the experimental data points $\{x_n^e\}$ and then averaging the points within each bin. Using the equidistant grid the experimental coefficients b_k^e can be calculated with

$$b_k^e = \frac{2}{N_g} \sum_{i=1}^{N_g} x_{i+1}^g \sin(\pi k x_i^g) .$$

One can show that this discrete Fourier transformation is similar to a maximum likelihood estimation.[14]

1. *Logistic map—period 1*

We begin by studying the simple case of the logistic map

$$x_{n+1} = \lambda x_n (1 - x_n) + F(\epsilon, n) , \tag{8}$$

where $\lambda \in [0,4]$ and $F(\epsilon,n)$ is an uncorrelated random force (bandlimited white noise between $-\epsilon$ and ϵ was applied at each time step). Since this is a mapping, it does not immediately satisfy the requirements of gradient dynamics necessary to utilize the results obtained in Sec. II A. However, we want to consider this map for $\lambda \in [1.0, 3.0]$ so that Eq. (8) has a single stable fixed point, x_0. In this case, a small amount of noise which takes x_n slightly away from x_0 will result in a transient decaying back to the fixed point. Therefore if we expand around this fixed point

$$x_n = x_0 + x_n'$$

we obtain

$$x_{n+1}' - x_n' = (\lambda - 1)x_0 - \lambda x_0^2 + (\lambda - 2\lambda x_0 - 1)x_n' - \lambda x_n'^2 .$$

The constant terms sum to zero so, if we are near the fixed point,

$$\frac{dx_n'}{dn} \simeq x_{n+1}' - x_n'$$
$$= (\lambda - 2\lambda x_0 - 1)x_n' - \lambda x_n'^2 = K(x_n'), \quad dn = 1$$

and

$$\phi(x_n') = -\int K(x_n') dx_n' . \tag{9}$$

In the following example, we take $\lambda = \frac{5}{2}$, so $x_0 = \frac{3}{5}$ and

$$\phi(x_n') = \tfrac{3}{4} x_n'^2 + \tfrac{5}{6} x_n'^3 .$$

Using Eqs. (6) and (7), we obtain

$$\alpha^2 = c\epsilon^2 \left[\frac{(2c\epsilon^2/3)^{1/2} + \frac{12}{5}}{(3c\epsilon^2/2)^{1/2} + \Gamma(\frac{1}{3})(12c\epsilon^2/5)^{1/3}} \right] . \tag{10}$$

This expression relates the accessible state-space volume $v(\epsilon)$ to ϵ. Since we are using band limited noise, the coefficient representing the color of the noise, c, is not easily determined analytically. However, if Eq. (10) is correct, we can obtain a value for c from the numerical experiment. The full width at half maximum measured from Fig. 2 as a function of ϵ is the experimental equivalent to α as defined in Eq. (10). We find that these data agree best with Eq. (10) for $c = 1.99$. The fit was made for $\epsilon \in (0.0, 0.2]$ with a confidence level of $R = 0.98$.

We can now utilize the expression for $v(\epsilon)$ to predict the model coefficients arising from the experimental system. Fortunately, the logistic map is just a parabola in this state-space representation; so, in the ideal situation where we have noiseless data over the full range [0,1], the model coefficients are given by

$$b_k^t = 2\int_0^1 \lambda x(1-x)\sin(k\pi x)dx$$
$$= \frac{4\lambda}{(2\pi)^3}[1 + (-1)^{k+1}] .$$

If we now consider the case where the data are limited to a smaller region of the state space $[a,b] \subset [0,1]$, Fig. 3, we can predict what the coefficients for the experimental system should be. Given the following data:

$$f'(x) = \begin{cases} 0, & 0 \le x < a \\ \sum_{k=1}^{N} b_k^t \sin(k\pi x_i), & a \le x < b \\ 0, & b \le x < 1 \end{cases} \tag{11}$$

the corresponding Fourier transform is

$$f'(x) = \sum_{m=1}^{N} b_m^p \sin(m\pi x_i)$$

where the predicted coefficients are

$$b_m^p = 2\int_0^1 f'(x)\sin(k\pi x)dx ,$$

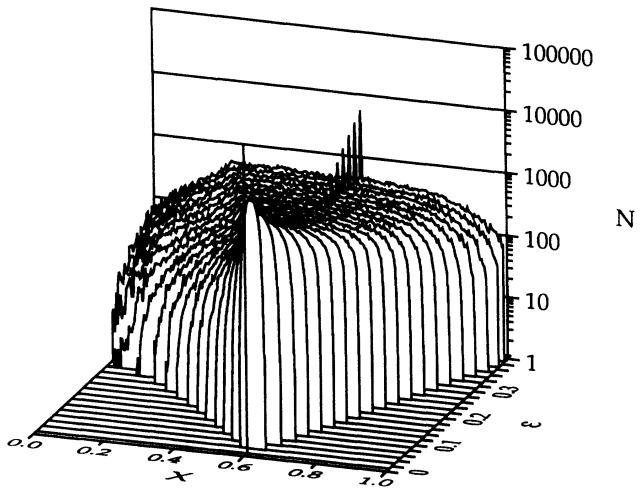

FIG. 2. The binning of 10 000 iterations of the logistic map, Eq. (8), is shown. We plot the number of iterations within each bin as a function of the noise amplitude ϵ. The ridge which appears at $x = 0.6$ for $\epsilon \gtrsim 0.2$ occurs because the large ϵ makes it possible for x to escape [0,1]. Thus the map must be restarted at x_0 causing the increased count at x_0.

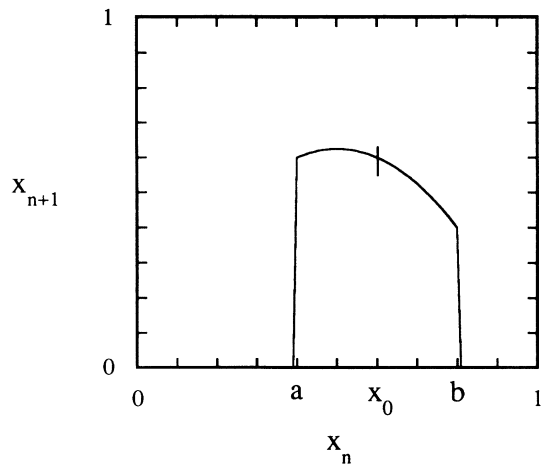

FIG. 3. To predict the coefficients of the experimental system, the dynamics are assumed to be perfectly known in the region $[a,b]$ where $0<a<x_0<b<1$.

$$b_m^p=\frac{1}{m}\sum_{\substack{n=1\\n\neq m}}^{\infty} b_n^t\left[\frac{\sin[(m-n)p]}{m-n}-\frac{\sin[(m+n)p]}{m+n}\right]_{p=a\pi}^{p=b\pi}$$
$$+\frac{1}{\pi}\left[\rho-\frac{\sin(4mp)}{2m}\right]_{p=a\pi}^{p=b\pi}. \tag{12}$$

When $a=0$ and $b=1$, Eq. (12) reduces to the proper orthonormality conditions. In all our Fourier transforms, we have made an odd periodic extension of the function so that only a sine series expansion is used.

If we let $a=x_0-v(\epsilon)$ and $b=x_0+v(\epsilon)$, we can use Eq. (12) to estimate the error in the experimental coefficients as a function of ϵ. The predicted average error is given by

$$\sigma_t^2=\sum_k (b_k^t-b_k^p)^2 . \tag{13}$$

We also define the average error for the numerical experiment as

$$\sigma_e^2=\sum_k (b_k^t-b_k^e)^2 . \tag{14}$$

Based upon the maximum likelihood estimation for this modeling of the logistic map, we can estimate the variance in b_k^e as

$$\eta_t^2=\frac{16\epsilon^2}{\pi^2 N} \tag{15}$$

where N is the total number of data points used in the model.[13] This assumes the data are homogeneously distributed over the full interval. Equation (15) agrees with our expectation that the models will degrade for large ϵ or small N. We will also determine the variance experimentally as

$$\eta_e^2=\frac{1}{M-1}\sum_{i=1}^{M}\sum_k (b_k^e-\overline{b_k^e})^2 \tag{16}$$

where the sum on i includes the number of independent attempts to model the system.

For the numerical experiment, the actual data used in the modeling process are shown versus ϵ in Fig. 4. Where no information was available, a value of zero was assumed to test the worst case scenario. This is the same assumption used in the analytic prediction, Eq. (11). Using Eqs. (14) and (16) we computed σ_e and η_e as functions of ϵ. These experimentally determined values are compared with the theoretical values in Fig. 5. In Fig. 5(a) we see that σ_t and σ_e have the same general behavior, but σ_e is actually more accurate for a given ϵ than predicted by Eq. (13). This is readily explained by the data in Figs. 2 and 4. The volume v obtained from the Fokker-Planck-based analysis was the full width at half the peak maximum. We assumed in Eq. (11) that no information was available outside the region $x\pm v$; but, since the system behaves as a random walk in a potential well, the experimental modeling will have more information available. This is clear from the figures. The full width at half maximum of Fig. 2, which we relate to v, is less than the sampled region in Fig. 4. Also note that near x_0, which according to Fig. 2 is where most of the x_n are binned, we obtain the best approximation to the true map. Far from x_0, the map is only a crude approximation since the histogram values are determined by only a few points, but this is still more information than was assumed in our prediction.

Thus we find that combining the Fokker-Planck analysis of the dynamics with a detailed knowledge of the modeling procedure does allow us to predict the accuracy of the experimental model for a given amount of noise, even though our predictions are somewhat impaired by an inaccurate assumption about the topological modeling. By comparing η_t and η_e in Fig. 5(b), we see again that some differences arise because our assumptions about the modeling process did not accurately represent

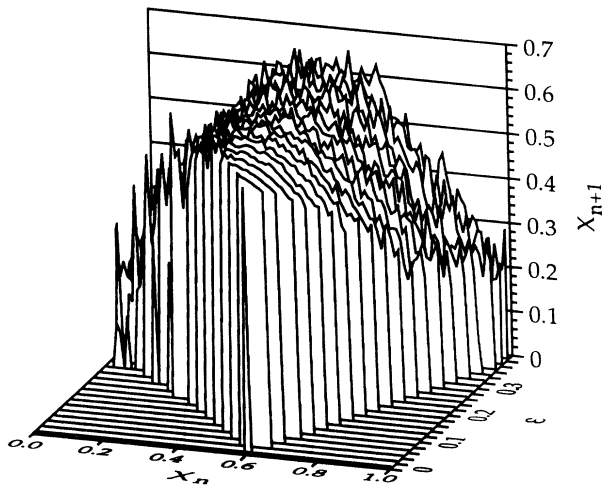

FIG. 4. In this figure we show the actual histograms used to compute some of the models as a function of ϵ. Note the increased variability occurring at the edges due to the poor statistics, cf. Fig. 2.

the averaging process involved in obtaining the grid values. However, as ϵ increases, these values do come into closer agreement.

To determine the optimum noise amplitude, we locate the minimum of the combined error, $(\sigma^2+\eta^2)^{1/2}$. This represents the point at which the noise-induced modeling error η becomes greater than the error arising from an insufficient sampling of the state space, σ. For the present example, we predict that the overall modeling error is minimized for $\epsilon \simeq 0.32$. From the numerical experiments, we see that as ϵ increases, η_e becomes significant to the total model error. Extrapolating σ_e and η_e, the optimal noise amplitude is estimated at $\epsilon \simeq 0.31$. Unfortunately, from Fig. 2 we determine that $\epsilon \gtrsim 0.2$ is unacceptable since this can take the dynamics outside the region on which the logistic map is defined. This is not a failure of the method; rather it is a constraint imposed by the system being studied. Thus, even though an extrapolation of the experimental results was required, we can get approximate theoretical and experimental confirmation of an optimum noise amplitude for topological modeling about a fixed point.

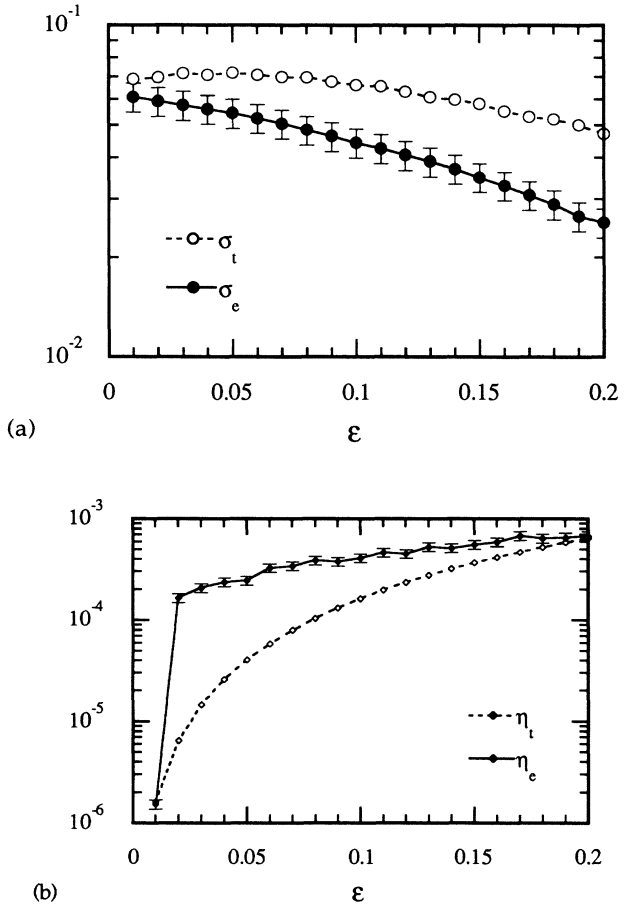

FIG. 5. (a) shows the predicted error σ_t and the experimentally determined error σ_e vs ϵ. In (b) the predicted variance η_t and experimental variance η_e are plotted vs ϵ. Experimenters commonly assume that reduced noise improves modeling, based upon the trend shown in (b). Since the error in (a) is initially much larger than the variance (b) but decreasing, we find instead that $\epsilon=0$ is not the optimum noise level when modeling dynamics about fixed points.

2. Logistic map—period 2

If we now take a $\in[3.0,3.449]$ in Eq. (8), the mapping becomes period 2 and thus we obtain two stable fixed points for $\epsilon=0$. For this example, we want simply to show that the general ideas developed for the period-1 case still hold. Thus we can proceed by investigating the map numerically in precisely the same manner as before.

For this example, a few important notes about the generation of the histogram are required. When accumulating statistics, gaps will naturally occur at the edges of the histogram where a cell with no counts will have two neighboring cells both containing information. Since we are using a Fourier series, these gaps can have a major effect upon the final coefficients. Therefore we smooth over these holes by interpolating between the neighbors. While this has a minor effect in the previous case, the same algorithm now connects the two fixed points with a straight line. This means that even with no noise, the model will be a rough approximation.

Using Eqs. (14) and (16) as before, we show the results of added noise for this situation, Fig. 6. We observe for this plot that even with the interpolation, a little added noise dramatically improves the accuracy of the model. For this example, the error falls precipitously until $\epsilon \simeq 0.14$. This is the point at which the noise is large enough to fill the region between the two fixed points, thus replacing the interpolation with dynamical information. Again, we cannot increase the noise past $\epsilon \simeq 0.2$, so we cannot determine if an optimal noise amplitude exists, but Fig. 6 gives a clear indication of the benefits of added noise for topological modeling.

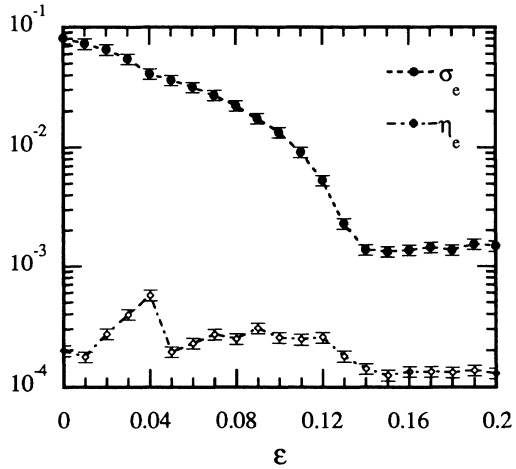

FIG. 6. This plot is identical to Fig. 5 except with $a=3.3$. For this case, the model improves dramatically up to $\epsilon \simeq 0.14$. Because of the presence of two fixed points, we do not have a theoretical prediction for the modeling error and any resulting optimal noise amplitude. However, such an optimal value may exist for $\epsilon > 0,2$ as in the period-1 case, but this cannot be confirmed from the experiment.

C. Noise in trajectory modeling

The other global modeling technique we wish to consider was developed by Cremers and Hübler[8] and is known as the trajectory method. In this case we try to reconstruct a coupled set of ordinary differential equations or maps, according to the nature of the dynamics. We proceed by making a series expansion of our model equations and then fitting the corresponding coefficients. The model can be based upon any type of series expansion since existence or lack of orthogonality never enters the modeling process. For the subsequent numerical experiments, we will use a Taylor-series expansion, since we can compare the resulting coefficients directly to the original equations. Fourier and Legendre series have also been used in other studies with an equal degree of success. This modeling technique can be applied for a state space with any number of dimensions, but two dimensions are used for illustration in the following discussion. In Eq. (17) is an example of an initial model for two dimensions based upon a Taylor-series expansion:

$$\begin{aligned}\dot{x}=f(x,y)&=a_{00}+a_{10}x+a_{01}y\\&\quad+a_{11}xy+\cdots+a_{ij}x^iy^j+\cdots,\\ \dot{y}=g(x,y)&=b_{00}+b_{10}y+b_{01}x\\&\quad+b_{11}xy+\cdots+b_{ij}x^jy^i+\cdots.\end{aligned}\tag{17}$$

We choose the coefficients a_{ij} and b_{kl} by taking each experimental data point $\mathbf{X}_e(t_i)$ as an initial condition to our model equations and calculating $\mathbf{X}_m(t_{i+1})$ from

$$\begin{aligned}x_m(t_{i+1})&=\int_{t_i}^{t_i+1}f(x_e(t),y_e(t))dt\ ,\\ y_m(t_{i+1})&=\int_{t_i}^{t_i+1}g(x_e(t),y_e(t))dt\ ,\end{aligned}$$

which we then compare to $\mathbf{X}_e(t_{i+1})$. In this manner, the quality of a given model is defined as

$$\mathcal{Q}=\sum_{i=1}^{N}\|\mathbf{X}_m(t_i)-\mathbf{X}_e(t_i)\|\ .\tag{18}$$

By minimizing $\mathcal{Q}$, we can determine the best coefficients in our expansion. This minimization process is essentially a search through the space of coefficients. This search and the numerical integrations necesssary for the associated ODE problem are performed using standard routines in the IMSL10 library.

Because of the search involved in the modeling process, we do not have a theory for how the model accuracy should scale in relation to the noise amplitude. However, we expect that the same counterbalancing effects of increasing the observable state space and increasing the variance in the data set will be important to selecting a noise amplitude which minimizes the error. In fact, most of the following examples will show that the trajectory method improves rapidly for small amounts of noise, but it then continues to generate high-quality models even for large amplitude noise. This is a reflection of the robustness of this modeling technique.

1. Coupled logistic maps

As a simple example of a two-dimensional system, we can take a pair of coupled logistic maps exhibiting a stable fixed point,

$$\begin{aligned}x_{n+1}&=ax_n(1-x_n)+\eta y_n+F(\epsilon,n)\ ,\\ y_{n+1}&=by_n(1-y_n)+\eta x_n+F(\epsilon,n)\ .\end{aligned}\tag{19}$$

Since the Fokker-Planck method developed in Sec. II A applies for any number of dimensions, we could also try it here using the gradient dynamic approximation developed in the one-dimensional case. However, since we do not know the exact correlation between the state-space volume observed and the quality of a model based upon a gradient search, this cannot be used effectively as a predictive tool.

To investigate Eqs. (19) numerically, we employ the trajectory method to reconstruct maps rather than differential equations. This entails just the simple substitutions $x(t_i)\rightarrow x_n$ and $x(t_{i+1})\rightarrow x_{n+1}$. For the present two-dimensional case, we fit the coefficients a_{ij} and b_{ij} in the following series:

$$\begin{aligned}x_{n+1}&=a_{00}+a_{10}x_n+a_{01}y_n+a_{11}x_ny_n+\cdots\ ,\\ y_{n+1}&=b_{00}+b_{10}y_n+b_{01}x_n+b_{11}x_ny_n+\cdots\ .\end{aligned}\tag{20}$$

This expansion does not assume any foreknowledge of the system. All the terms are fitted, including those which will ultimately be found to be zero. These expansions are limited only through the imposition of a maximum order for the polynomial terms, as a means of improving computational efficiency. The best model is again constructed by minimizing $\mathcal{Q}$ defined for maps as

$$\mathcal{Q}=\sum_{i=1}^{N}\|\mathbf{X}_{i+1}^m-\mathbf{x}_{i+1}^e\|\ .\tag{21}$$

Once the best possible parameters have been obtained, the systematic error in the fit and the variance of the coefficients is computed as in Eqs. (14) and (16). For this example and the following ones, we have taken 2000 data points per model and 20 independent modeling attempts for each value of ϵ. These values are forced by computational limitations; and, as a result, η_e will be much larger than for the topological modeling.

We find experimentally that the modeling of Eq. (19) behaves very much like that of the simple logistic map example. As the noise is gradually increased, we see a drastic improvement in error of the model coefficients, Fig. 7. Also, note that the variance decreases simultaneously, indicating that the lack of unique solutions in the modeling process, a common shortcoming of the trajectory method, becomes less severe as the noise increases. This is simply because, as the state-space volume increases, the data better indicate to the search algorithm which coefficients represent the true dynamics. However, for large ϵ, the model is degraded just as rapidly due to the excessive noise in the experimental data. For this example, the numerical experiment indicates the presence of an optimal noise amplitude.

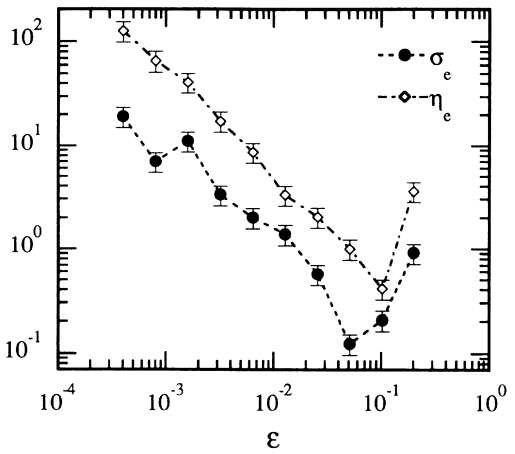

FIG. 7. These calculations are for the coupled logistic maps, Eqs. (19) with $a=2.0$ and $b=1.5$, and are based upon trajectory modeling. From the plot, we see that $\epsilon \simeq 0.05$ is the best choice for this example. Presumably, this minimum exists for the same general reasons outlined in Sec. II, although the existence of a single optimal noise amplitude need not be the general case for the trajectory modeling.

2. *Overdamped oscillator*

We can also apply the modeling with noise to the case of ordinary differential equations. As an example, we consider the overdamped oscillator equation

$$\dot{x} = -\gamma x + F(\epsilon, t) \tag{22}$$

where the inertial term $m\ddot{x}$ is negligible compared to $\dot{x}$. Numerically, we generated the data sets from Eq. (22) with a variable fifth-order Runge-Kutta integration scheme (from the IMSL10 libraries).

Upon examining the actual calculations, Fig. 8, we find that the variance is quite large, which is to be expected due to current computational limitations. For the present example, any additional information apparently makes a dramatic improvement in the quality of the model. This is probably a result of the strong damping, $\gamma=1$, used in this example. Also, we note that η_e again rises for increasing ϵ. By combining the systematic error and the variance, Fig. 8 suggests that there may still be an optimum noise level, although there is no way of obtaining analytic agreement. More importantly, we find here and in general that any noise makes a dramatic improvement in the models obtained via the trajectory method. Moreover, this method demonstrates remarkable robustness even under large amplitude noise.

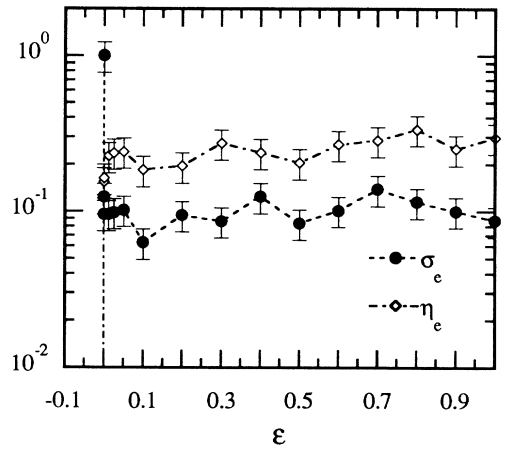

FIG. 8. The data used in these calculations were generated from the overdamped oscillator, Eq. (22), with $\gamma=1$. From this plot, we note that $\epsilon \simeq 0.1$ is an acceptable value to obtain a good model, but the modeling continues to be successful even for very large noise amplitudes.

3. Damped harmonic oscillator

As a final example for modeling dynamics about fixed points, we consider a damped harmonic oscillator

$$\ddot{x} + \mu\dot{x} + x = 0$$

where the second term is the damping term. We can break the equation into a pair of coupled ODE's as follows:

$$\begin{aligned} \dot{x} &= y \ , \\ \dot{y} &= -\mu y + x \ . \end{aligned} \tag{23}$$

Although these equations have the form of Eq. (1), a potential field cannot be found which will satisfy Eq. (5), so this does not represent gradient dynamics and the solution in Eq. (6) is not applicable. In cases such as this, it may be possible to solve the time-independent Fokker-Planck equation, Eq. (4), directly. We have chosen to investigate this system numerically since we are still using the trajectory model which impedes any further analytic considerations.

For this study, we proceed with exactly the same method as for the overdamped oscillator with the simple generalization to two dimensions. The calculations are shown in Fig. 9. Due to the numerical integrations in-

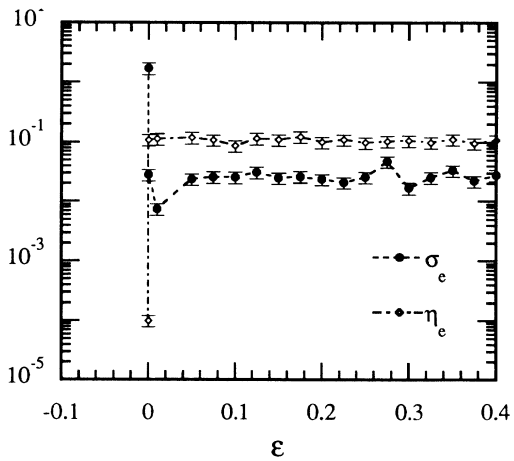

FIG. 9. These calculations use data from the damped harmonic oscillator, Eq. (23), with $\mu=1$. Any noise amplitude of $\epsilon \geq 0.01$ appears to be acceptable.

volved in the modeling procedure, this example is even more computationally restricted than the previous one-dimensional example. Correspondingly, in Fig. 9 we observe that the variance is very large and essentially constant versus ϵ. From the values of σ_e, there does appear to be an initial trend toward reduced error, which is quickly obscured by η_e. Thus these calculations show that once again the added noise is advantageous to the modeling process. It is remarkable to note how little noise is required to generate a model which approximates the true dynamics so accurately, and simultaneously how consistent these results are even for large noise levels.

III. LIMIT CYCLES AND CHAOTIC ATTRACTORS

Another class of systems which is often difficult to model are those exhibiting limit cycles. The problem is that when there are no transient data available, the rate of convergence to the limit cycle is unknown. In fact, a single closed trajectory may not even be a limit cycle. Unfortunately, the solution for modeling is not as clear here as for fixed-point dynamics. In general we expect the model accuracy to depend upon the noise amplitude in a manner similar to that discussed for the fixed-point modeling. However, some important exceptions do arise which we discuss later.

The trajectory method can be employed to obtain a good model for a closed orbit with no noise, but it is not necessarily accurate anywhere else in the state space, and thus cannot be expected to represent the true dynamics of the system. This ambiguity is caused by the possible existence of many different solutions which generate the same closed trajectory, but maintain different flow vector fields elsewhere in the state space. To choose the correct model, our technique of adding noise can become important.

As an example, consider a harmonic oscillator

$$\ddot{x}+x=0 \tag{24}$$

and a system similar to a van der Pol oscillator

$$\ddot{x}+(x^2+\dot{x}^2-A^2)\dot{x}+x=0\ . \tag{25}$$

Both of these systems can be reconstructed in an $(x,\dot{x})$ phase space. In this representation, the harmonic oscillator equation has a center point, so the phase space is a continuous set of closed circular orbits. However, Eq. (25) has an attractor, a stable limit cycle, which is also a circle in the state space, Fig. 10. If the given data are obtained just from this limit cycle, it becomes impossible to distinguish between the two systems. In this situation, it is clear that transients are necessary to select the proper model parameters, and adding noise to increase the accessible phase space is one way to accomplish this. Figure 10 also shows a trajectory with $\epsilon\neq 0$. Of the two data sets shown, the latter is far more effective for reconstructing the original dynamics.

When we perform the same modeling tests as was done for the previous examples, the results are similar to the earlier ones, except that now there is a clear trend toward more inaccurate models as ϵ increases. Figure 11 also displays a suggestion of the trend in σ_e noted previously,

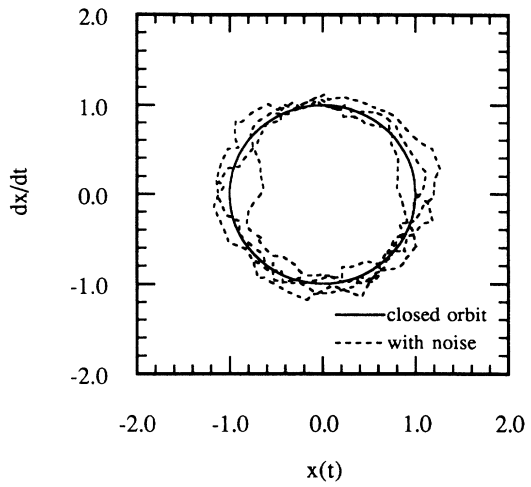

FIG. 10. This plot contains data generated from Eq. (25) with $A=1$ for two values of ϵ. The solid line is for $\epsilon=0$ and is indistinguishable from a trajectory of a harmonic oscillator, Eq. (24). The dashed line used $\epsilon=0.1$ and proved to be superior for modeling because of the increased stated-space sampling.

but further conclusions are difficult. The fact that only a slight amount of noise is necessary to obtain a reasonably good model further demonstrates the effectiveness of adding noise for the modeling dynamics around close orbits.

Exceptions to the results in Fig. 11 occur in systems where one or more of the model coefficients is determined explicitly by the precise shape of the limit cycle. One well-known example of this is the standard form of the van der Pol oscillator:[15]

$$\ddot{x}-\mu(1-x^2)\dot{x}+x=0\ . \tag{26}$$

The constant μ determines the shape of the limit cycle for the van der Pol equation. If one adds a small amount of noise, some additional transient data are then available,

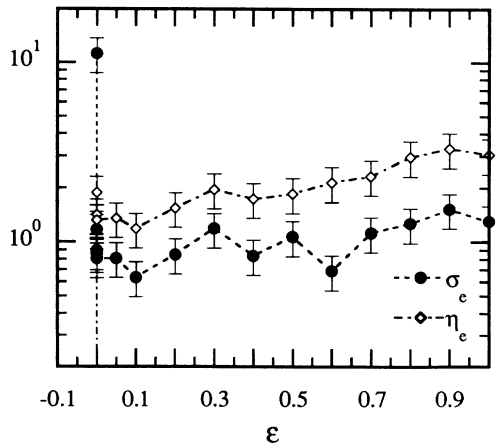

FIG. 11. The error and variance are shown for the models obtained from Eq. (25). This example shows a stronger sensitivity to large-amplitude noise so an optimal noise amplitude may exist for $\epsilon\simeq 0.1$.

but the shape of the limit cycle may become slightly distorted. In practice, this small distortion causes the addition of any noise to degrade the system. Eisenhammer *et al.*[9] have shown that for systems where the noise cannot be excluded, effective models are still possible. They overcome this problem by including both long-term and short-term predictions into the quality function, Eq. (18), to encourage the construction of stable models. In effect, this averages over the noise in an attempt to regain the shape of the original noiseless trajectory. Obviously, the ideal data set for modeling the van der Pol equation would be one which includes a single, long, noiseless transient settling onto the limit cycle. From our analysis and previous studies, we expect that for modeling systems with closed orbits, no single method will give the best results in all situations, but the method of adding noise we have described should be applicable in a majority of them.

When attempting to model a chaotic attractor, many of the same advantages and disadvantages revealed in the case of limit cycles will again apply. For many systems, some added noise should again increase the state space and allow for better models off the attractor. However, some attractors will have coefficients which depend sensitively upon the shape of the attractor, and thus are best modeled without noise, just as was discovered for the van der Pol oscillator, Eq. (26). Therefore our modeling technique should, as a general rule, be helpful when dealing with a chaotic attractor; but each case must be considered individually, as with the closed trajectories.

IV. NOISE AND NONLINEAR CONTROL WITHOUT FEEDBACK

The methods discussed here have important implications for the subject of nonlinear control. Nonlinear control, as described by Hübler and Lüscher[4] and Jackson and Hübler,[16] is a technique for controlling the dynamics of a nonlinear system without feedback from the experimental system. This type of control is entirely dependent upon the ability to obtain a global model of the dynamics. Once such a model exists, the system can be controlled by using this model instead of the feedback loops frequently utilized in control theory. If the system being controlled is at a fixed point or closed orbit, it is imperative that the model employed represent the true behavior of the system not just at these few points, but throughout the state space. If the model excluded such information, any small perturbation which occurs during the control process could take the dynamics to a region where the model was inaccurate and the control would fail. These problems can be avoided by the process we have demonstrated for modeling with noise. Small noise-induced inaccuracies in the model should not disrupt the entrainment of the system.[17]

This nonlinear control without feedback can also be used in a modified version where the system being controlled is continuously monitored to determine if the original model is failing and to regenerate a new model in real time.[13] The presence of noise also has the same important advantages for this type of control, both in aiding the continuous monitoring of the model accuracy and the generation of new models. In effect, this implies that for a large class of systems, control and adaptive control become more robust if the system is slightly noisy, because of the corresponding improvements in the model-building process.

V. CONCLUSIONS

Scientists, almost instinctively, strive to reduce or eliminate noise. In general, this is a worthwhile and often critical objective, but we have shown that for dynamical systems in stable equilibria, noise can be an essential tool in modeling and control. Ironically, the perfectly noiseless system is the worst possible case for these systems. Since the experimenter may not be able to introduce a single clean transient into the system for observation, we have demonstrated that even continuous random noise is an asset in modeling the dynamics. When noise cannot be manually injected into the system, such as astrophysical systems, the best choice for modeling may be that system which contains noise in the dynamics naturally.

In our studies, noise is always beneficial for modeling systems at stable fixed points, but is not always so helpful for other situations and must, in fact, be considered on a case-by-case basis. If the noise is found not to be useful, the detrimental effects from naturally occurring noise can be minimized by employing the smoothing technique developed previously.[8] Still, the method of using noise to increase the state-space sampling opens, for modeling and control, dynamical systems in many scientific fields which would otherwise have been unapproachable.

ACKNOWLEDGMENTS

This work was supported in part by Office of Naval Research, Grant No. N00014-88-K-0293 and National Science Foundation, Grant No. PHY86-58062.

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

[2]*Dimensions and Entropies in Chaotic Systems,* Vol. 32 of *Springer Series in Synergetics,* edited by G. Meyer-Kress (Springer-Verlag, New York , 1986).

[3]J. A. Lichtenberg, and M. A. Lieberman, *Regular and Stochastic Motion* (Springer, New York, 1982).

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

[5]B. A. Huberman and E. Lumer (unpublished).

[6]H. Haken, *Synergetics, an Introduction* (Springer-Verlag, New York, 1983).

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

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

[9]T. Eisenhammer, A. Hübler, N. Packard, and J. A. S. Kelso, Center for Complex Systems Research Technical Report No.

CCSR-89-7, 1989 (unpublished).
[10]J. P. Crutchfield and B. S. McNamara, Complex Syst. **1**, 417 (1987).
[11]N. Packard, Center for Complex Systems Research Technical Report No. CCSR-89-10, 1989 (unpublished).
[12]J. D. Farmer and J. J Sidorowich, Los Alamos Technical Report No. LA-UR-88-901, 1988 (unpublished).
[13]A. Hübler, Helv. Phys. Acta **62**, 286 (1989).
[14]R. Georgii, W. Eberl, E. Lüscher, and A. Hübler, Helv. Phys. Acta **62**, 290 (1989).
[15]B. van der Pol and J. van der Mark, Nature (London), **120**, 363 (1927).
[16]E. A. Jackson and A. Hübler, Center for Complex Systems Research Technical Report No. CCSR-90-3, 1989 (unpublished).
[17]K. Chang, A. Hübler, and N. Packard, Center for Complex Systems Research Technical Report No. CCSR-89-5, 1989 (unpublished).