Estimation of parameters in nonlinear systems using balanced synchronization

Henry D. I. Abarbanel*
Department of Physics and Marine Physical Laboratory (Scripps Institution of Oceanography), University of California, San Diego, La Jolla, California 92093-0402, USA

Daniel R. Creveling†
Department of Physics and Institute for Nonlinear Science, University of California, San Diego, La Jolla, California 92093-0402, USA

James M. Jeanne‡
Graduate Program in Computational Neurobiology and Institute for Nonlinear Science, University of California, San Diego, La Jolla, California 92093-0402, USA

(Received 30 July 2007; published 28 January 2008)

Using synchronization between observations and a model with undetermined parameters is a natural way to complete the specification of the model. The quality of the synchronization, a cost function to be minimized, typically is evaluated by a least squares difference between the data time series and the model time series. If the coupling between the data and the model is too strong, this cost function is small for any data and any model and the variation of the cost function with respect to the parameters of interest is too small to permit selection of an optimal value of the parameters. We introduce two methods for balancing the competing desires of a small cost function for the quality of the synchronization and the numerical ability to determine parameters accurately. One method of "balanced" synchronization adds to the synchronization cost function a requirement that the conditional Lyapunov exponent of the model system, conditioned on being driven by the data remain negative but small in magnitude. The other method allows the coupling between the data and the model to vary in time according to the error in synchronization. This method succeeds because the data and the model exhibit generalized synchronization in the region where the parameters of the model are well determined. Examples are explored which have deterministic chaos with and without noise in the data signal.

DOI: 10.1103/PhysRevE.77.016208 PACS number(s): 05.45.Xt

I. INTRODUCTION

Estimating parameters in a nonlinear dynamical model, given observed data, is an important aspect of developing predictive models of physical and biological systems [1–10]. Setting aside the issues of noise and experimental errors in acquiring the data and errors in the models themselves, one still has a significant challenge in estimating these parameters, especially when the dynamical behavior can be chaotic. As we recall below, the presence of positive Lyapunov exponents in this setting means that the numerical evaluation of a cost function representing parameter estimation quality, often a least squares metric, may suffer from sensitive dependence on initial conditions [11,12].

There have been a few investigations into the use of synchronization between the source of the experimental data and the model system, and these have shown substantial promise in successfully performing the desired model parameter estimation [6,7,9,10,13,14]. We will examine a class of such proposed synchronization methods for this problem and show that they suffer other deficiencies when the coupling, which we generically call K, between the observations and the model is too large. When K is large, synchronization is dictated by the manner of driving the model with the data, so the estimation metric is small. When K is too small, the data and the model do not synchronize, and information is not passed precisely between the data and the model whose parameters we wish to determine [6]. When K is too large, the desired sharp minimum in the estimation metric as a function of the parameters is removed.

In this paper we explore two methods for balancing these desired properties: small estimation metric for determining model parameter values and automatic synchronization of model and data signals without a loss of parameter estimation capability. We seek a balance between these behaviors, and we call the general method "balanced synchronization."

The problem we address can be cast into the following format. We imagine there is a physical or biological system whose state is determined by the N-dimensional vector $\mathbf{x}(t)=[x_1(t),x_2(t),\ldots,x_N(t)]$, where measurements are made on the system at time intervals τ. Starting with some initial time t_0, measurements are made at times $t_0+m\tau$; $m=0,1,2,\ldots,M$, resulting in observations $\mathbf{x}(m)=\mathbf{x}(t_0+m\tau)$. We assume this sampling is adequate to determine the frequencies of importance in the operation of the system of interest: τ is small enough and $M\tau$ is large enough.

One component of the state is now measured and stored for use in determining parameters in a model describing the physics of interest. This could be an arbitrary scalar function of the system's state $h[\mathbf{x}(t)]$ but we will take it as one of the state variables itself, namely, $x_1(t)$. The other $N-1$ state variables $\mathbf{u}(t)=[x_2(t),x_3(t),\ldots,x_N(t)]$ are unobserved. If they were, the analysis would proceed in a parallel fashion to what we present here.

*habarbanel@ucsd.edu
†dcreveli@physics.ucsd.edu
‡jjeanne@ucsd.edu

The observed (or "driver") system satisfies the differential equations

$$\frac{dx_1(t)}{dt} = F_1(x_1(t), \mathbf{u}(t), \mathbf{p}),$$

$$\frac{d\mathbf{u}(t)}{dt} = \mathbf{G}(x_1(t), \mathbf{u}(t), \mathbf{p}), \tag{1}$$

where the vector fields $F_1(\cdots)$ and $\mathbf{G}(\cdots)$ are presumed known, and the trajectories $\mathbf{x}(t)=[x_1(t), \mathbf{u}(t)]$ are determined by the initial conditions $\mathbf{x}(t_0)$ and the P parameters $\mathbf{p}=[p_1, p_2, \ldots, p_P]$.

The time series information $x_1(t)$ (the "data") is now passed to the receiver system which is our model for the process describing the observations. For this paper we take this model to be precisely that used in generating the data, so we know the dynamics, but assume we do not know the parameters $\mathbf{q}=[q_1, q_2, \ldots, q_P]$ of the receiver (the model). The state of the receiver is given by $\mathbf{y}(t)=[y_1(t), y_2(t), \ldots, y_N(t)]=[y_1(t), \mathbf{v}(t)]$. Without coupling to the data our model satisfies

$$\frac{dy_1(t)}{dt} = F_1(y_1(t), \mathbf{v}(t), \mathbf{q}),$$

$$\frac{d\mathbf{v}(t)}{dt} = \mathbf{G}(y_1(t), \mathbf{v}(t), \mathbf{q}), \tag{2}$$

and we seek to determine the $\mathbf{q}$ given a times series $x_1(t)$ [5].

For this purpose, we couple the receiver system to the input signal $x_1(t)$ using

$$\frac{dy_1(t)}{dt} = F_1(y_1(t), \mathbf{v}(t), \mathbf{q}) + K(x_1(t) - y_1(t)),$$

$$\frac{d\mathbf{v}(t)}{dt} = \mathbf{G}(y_1(t), \mathbf{v}(t), \mathbf{q}). \tag{3}$$

Had the signal $h[\mathbf{x}(t)]$ been measured, the coupling term would have been $K\{h[\mathbf{x}(t)]-h[\mathbf{y}(t)]\}$ [5]. The idea is that for some range of K the receiver (model) will synchronize to the data $y_1(t)\approx x_1(t)$, and this synchronization will be "most accurate" when we vary the parameters in the model and achieve $\mathbf{q}=\mathbf{p}$ [6].

The synchronization we consider here is "identity" synchronization which assumes that we know the model vector field $[F_1(\cdots), \mathbf{G}(\cdots)]$ exactly, that it matches the dynamics of the data source exactly, and that only the particular parameters $\mathbf{p}$ utilized in generating the data are not known. We will proceed with this scenario while acknowledging that this is unlikely to precisely be the case in practice. Instead we will have measured some quantity, call it $x_1(t)$, and have done our best to determine a model vector field—or differential equations—and wish to establish the parameters of that model. Further, we generally assume there is no uncertainty in either the data, expressed as noise in the data or inaccuracies in the measurements, or in the model. Our point of view is that the exploration we report here will be the starting point for more realistic situations where these and other complications, such as the model being wrong, are present. In two cases we explore the sensitivity of the methods to noise in the signal $x_1(t)$ transmitted from the data to the model.

Moving beyond these cautionary remarks, we need a principle to assist in estimating the $\mathbf{q}$. A natural choice similar to that used in Ref. [6] is a least squares criterion involving minimizing the cost function

$$C(\mathbf{q}) = \frac{1}{2}\int dt(x_1(t) - y_1(t;\mathbf{q}))^2 = \frac{\tau}{2}\sum_{m=1}^{M}(x_1(m) - y_1(m,\mathbf{q}))^2, \tag{4}$$

where we have made explicit the dependence of the receiver (model) solution on the parameters $\mathbf{q}$. $t=t_0+m\tau$ as above.

The minimization involves seeking a zero of

$$\frac{\partial C(\mathbf{q})}{\partial q_\alpha} = \int dt(y_1(t;\mathbf{q}) - x_1(t))\frac{\partial y_1(t;\mathbf{q})}{\partial q_\alpha}, \tag{5}$$

for $\alpha=1,2,\ldots,P$. The $\mathbf{q}$ giving these zeros will be our estimate of the model parameters $\mathbf{q}$ needed to match the settings $\mathbf{p}$ of the data source. Formally we also desire that the $P\times P$ matrix

$$\frac{\partial^2 C(\mathbf{q})}{\partial q_\alpha \partial q_\beta} \tag{6}$$

is positive definite.

The quantity

$$\frac{\partial y_1(t;\mathbf{q})}{\partial q_\alpha} \tag{7}$$

is determined, along with

$$\frac{\partial \mathbf{v}(t;\mathbf{q})}{\partial q_\alpha} \tag{8}$$

by

$$\frac{d}{dt}\begin{pmatrix} \dfrac{\partial y_1(t;\mathbf{q})}{\partial \mathbf{q}} \\ \dfrac{\partial \mathbf{v}(t;\mathbf{q})}{\partial \mathbf{q}} \end{pmatrix} = \left[\mathbf{D}\cdot\mathbf{F}(y_1(t;\mathbf{q}), \mathbf{v}(t;\mathbf{q})) - \begin{pmatrix} K & 0 \\ 0 & 0 \end{pmatrix}\right] \times \begin{pmatrix} \dfrac{\partial y_1(t;\mathbf{q})}{\partial \mathbf{q}} \\ \dfrac{\partial \mathbf{v}(t;\mathbf{q})}{\partial \mathbf{q}} \end{pmatrix} + \frac{\partial \mathbf{F}(\mathbf{y},\mathbf{q})}{\partial \mathbf{q}}, \tag{9}$$

where $\mathbf{D}\cdot\mathbf{F}$ is the $N\times N$ Jacobian matrix

$$\mathbf{D}\cdot\mathbf{F}_{ij}(\mathbf{y}) = \frac{\partial F_i(\mathbf{y})}{\partial y_j}; \quad i,j = 1,2,\ldots,N \tag{10}$$

of the model dynamics, and the matrix involving K is also $N\times N$ with only the upper left diagonal element nonzero. $\mathbf{F}(\mathbf{y})$ is the total N-dimensional vector field $\mathbf{F}(\mathbf{y})=(F_1(y_1,\mathbf{v}), \mathbf{G}(y_1,\mathbf{v}))$.

We need to face the issue of stability as the eigenvalues of the Jacobian $\mathbf{D}\cdot\mathbf{F}(\mathbf{y})$ iterated along the orbit $\mathbf{y}(t)$ may have

positive conditional Lyapunov exponents; conditioned on the driving signal $x_1(t)$. These are found in the usual way by concatenating products of the matrices $\mathbf{D}\cdot\mathbf{F}(\mathbf{y}(t))-\mathbf{K}$ and relying on the Oseledec theorem to assure us of the existence of the eigenvalues of the iterated product. If there are positive conditional Lyapunov exponents, then the synchronization manifold $y_1(t)=x_1(t)$ is not stable to small perturbations, and the evaluation of the derivatives of the cost function (5) is numerically quite uncertain [8,11,12].

The conditional Lyapunov exponents, however, can be made negative by increasing the magnitude of K. When the largest conditional Lyapunov exponent has become negative, then the synchronization manifold is stable to small perturbations, and evaluating the derivatives of the cost function is straightforward.

As K becomes large, the term $K(x_1(t)-y_1(t))$ dominates the right-hand side of the evolution equation for $y_1(t)$ unless $x_1(t)-y_1(t)\approx\frac{1}{K}$ or smaller. As this happens, all the derivatives in Eq. (5) approach zero, and we have little numerical variation of the derivatives as functions of the parameters $\mathbf{q}$. Simply stated the minimum we seek of the cost function becomes so flat in $\mathbf{q}$ space it is numerically extremely difficult to locate. So as we tame the numerical instability associated with chaos in the model dynamical system, we may send ourselves into a regime where the evaluation of the parameters $\mathbf{q}$ by minimizing the cost function becomes harder and harder.

In addition, when K is large, the dynamics of the model are entrained by the driving $K(x_1(t)-y_1(t))$ and any model is forced to follow $x_1(t)$. The ability to distinguish among models is thus lost.

We require a balance between these unacceptable limits, and to accomplish that, we need a way to choose a value of K that just leads to the largest conditional Lyapunov exponent being negative, yet is not such a large value of K that we lose the ability to see variations in the $\mathbf{q}$. We call this "balanced synchronization," and we now explore two ways to achieve this. Following an exploration of these ideas in the Lorenz system and in a model of the RF Colpitts oscillator, we report on laboratory experiments with analog circuits representing the Lorenz driver and "model" system.

II. BALANCED SYNCHRONIZATION

A. Controlling the largest conditional Lyapunov exponent

For synchronization of the experimental data $x_1(t)$ and the model system to be effective in estimating the $\mathbf{q}$, the largest conditional Lyapunov exponent must be negative [1]. Since we are depending on synchronization to drive the model dynamical variables $[y_1(t),\mathbf{v}(t)]$ to those taken on by the observed signal $[x_1(t),\mathbf{u}(t)]$, if synchronization fails, the foundation of the method would fail [5].

The conditional Lyapunov exponents are evaluated by perturbing the receiving or model system from the synchronization manifold $\mathbf{x}(t)=\mathbf{y}(t)$. Linearizing this perturbed dynamics we have for $\Delta(t)=\mathbf{y}(t)-\mathbf{x}(t)$

$$\frac{d\Delta(t)}{dt}=[\mathbf{D}\cdot\mathbf{F}(\mathbf{x}(t))-\mathbf{K}]\cdot\Delta(t), \tag{11}$$

where

$$\mathbf{K}=\begin{pmatrix}K & 0\\ 0 & 0\end{pmatrix}. \tag{12}$$

Recalling that we sample the data and the model at time intervals τ we may interpret this differential equation as a map between values of the perturbation at "time" n and time $n+1$:

$$\Delta(n+1)=\mathbf{D}\cdot\mathbf{H}(\mathbf{x}(n))\cdot\Delta(n), \tag{13}$$

where

$$\mathbf{D}\cdot\mathbf{H}(\mathbf{x}(n))=\mathbf{I}+\tau[\mathbf{D}\cdot\mathbf{F}(\mathbf{x}(n))-\mathbf{K}], \tag{14}$$

and $\mathbf{I}$ is the unit $N\times N$ matrix.

To find all of the conditional Lyapunov exponents, we need to iterate this map and evaluate the eigenvalues of the resulting product of matrices. We, however, do not require all of the conditional Lyapunov exponents, just the largest, and this entails a much easier calculation.

Take an arbitrary unit vector $\mathbf{w}$ in the N-dimensional space and multiply it by the iterated "effective" Jacobian

$$\mathbf{D}\cdot\mathbf{H}(\mathbf{x})^L=\mathbf{D}\cdot\mathbf{H}(\mathbf{x}(L))\cdot\mathbf{D}\cdot\mathbf{H}(\mathbf{x}(L-1))\cdots\mathbf{D}\cdot\mathbf{H}(\mathbf{x}(1)), \tag{15}$$

which carries the linearized perturbation at "time" 1 to its value at "time" $L+1$. Multiply the vector $\mathbf{D}\cdot\mathbf{H}(\mathbf{x})^L\mathbf{w}$ by its transpose

$$(\mathbf{D}\cdot\mathbf{H}(\mathbf{x})^L\mathbf{w})^T\cdot(\mathbf{D}\cdot\mathbf{H}(\mathbf{x})^L\mathbf{w}). \tag{16}$$

This grows as $e^{2L\lambda(\mathbf{q},K)}$ for large L. The quantity

$$\frac{1}{2L}\ln(\mathbf{D}\cdot\mathbf{H}(\mathbf{x})^L\mathbf{w})^T\cdot(\mathbf{D}\cdot\mathbf{H}(\mathbf{x})^L\mathbf{w}), \tag{17}$$

is just $\lambda(\mathbf{q},K)$, the largest conditional Lyapunov exponent. We wish to have $\lambda(\mathbf{q},K)$ slightly negative [11].

This suggests that we replace the least squares cost function with the balanced cost function

$$C(\mathbf{q},K)=\frac{1}{2}\int dtf[(x_1(t)-y_1(t;\mathbf{q}))^2]+\frac{1}{2}(\lambda(\mathbf{q},K)-\xi)^2, \tag{18}$$

where ξ is a small negative number and $f(x^2)$ a function which vanishes as x^2 near $x=0$. This enforces synchronization by asking that the first term be small, but does not allow K to be so large that $\lambda(\mathbf{q},K)$ is too negative. As an approximation to this cost function for K large, we may estimate the first term as being of order $\frac{1}{K^2}$, as $x_1(t)-y_1(t,\mathbf{q})\approx\frac{1}{K}$ while the second grows as K^2. If the total cost function is approximated as

$$\frac{A}{K^2}+BK^2, \tag{19}$$

with A and B constants, this has a minimum at approximately $K\approx\left(\frac{A}{B}\right)^{1/4}$. So K remains bounded, and there is a balance between the smallness of the synchronization error and the magnitude of the synchronization coupling strength K.

B. Balancing the magnitude of K: Dynamical coupling

Another idea is to allow the magnitude of the coupling K to vary in a manner which makes it grow when the data and the model are out of synchronization, while it decreases when they have synchronized. For this purpose we look at the following differential equations for $y_1(t)$, $\mathbf{v}(t)$, and $K(t)$

$$\frac{dy_1(t)}{dt} = F_1(y_1(t),\mathbf{v}(t),\mathbf{q}) + K(t)(x_1(t) - y_1(t)),$$

$$\frac{d\mathbf{v}(t)}{dt} = \mathbf{G}(y_1(t),\mathbf{v}(t),\mathbf{q}),$$

$$\frac{dK(t)}{dt} = -aK(t) + g[(\mathbf{x}(t) - \mathbf{y}(t))^2], \tag{20}$$

where $a>0$, $g(0)=0$, and for small argument $g(x^2)\approx x^2$. The solution to the $K(t)$ differential equation

$$K(t) = e^{-at}K(t=0) + \int_0^t dt' e^{-a(t-t')} g[(\mathbf{x}(t') - \mathbf{y}(t'))^2], \tag{21}$$

shows that the initial value of $K(t=0)$ is unimportant when $ta \gg 0$. If $g(x^2)$ is bounded by a constant C, $K(t) < \frac{C}{a}$. This means that the coupling is again balanced, but this time by the requirements of synchronization on the magnitude of $g(x^2)$ and the tendency for K to vanish. As time goes by there are intervals when the synchronization is lost and $(\mathbf{x}(t') - \mathbf{y}(t'))^2$ grows, this leads to a growth in the magnitude of $K(t)$ improving the synchronization. The balance between these two effects, decay of $K(t)$ to zero and growth of $K(t)$ to strengthen synchronization, is the embodiment of balanced synchronization in this method.

For our investigations we explored two different functions of the synchronization error $x_1(t) - y_1(t)$

$$\frac{dK(t)}{dt} = -aK(t) + \tanh[(\mathbf{x}(t) - \mathbf{y}(t))^2],$$

$$\frac{dK(t)}{dt} = -aK(t) + (\mathbf{x}(t) - \mathbf{y}(t))^2. \tag{22}$$

In the first $K(t) < \frac{1}{a}$, while for the second the perturbation to $K(t)$ can be quite large.

In this approach to balanced synchronization, we can never have identity synchronization $\mathbf{x}(t)=\mathbf{y}(t)$, as the driving system (the data) is N dimensional $(\mathbf{x}(t))$ while the model system is $(N+1)$-dimensional $(\mathbf{y}(t),K(t))$. If the coupling parameter $K(t)$ were constant, identity synchronization would be possible. In our formulation of a criterion as given below for determining the parameters $\mathbf{q}$, we still use the notion of identity synchronization, and this may be approximately correct as the variation of $K(t)$ is bounded. Nonetheless, as we will show below, the determination of parameters in the model using this method works quite well. It is worth asking why the determination of parameters might be so accurate using this $K(t)$ method.

Now generalized synchronization between these two dynamical systems is possible [15,16]. To test for this we use the variant of the auxiliary system method as discussed by Tang *et al.* [19]. In this approach the signal from the driving system $x_1(t)$ is repeatedly presented to the model system $\mathbf{y}(t),K(t)$, Eq. (20). In each presentation the model system is taken to be in a different state by adjusting either the time at which $x_1(t)$ is presented or adjusting the initial conditions of the model system. In effect, the state of the model system is different for each presentation. If we call the model output from presentation $n=1,2,\ldots,;\mathbf{y}^{(n)}(t)$, then after a transient, these outputs from the various realizations of the model should agree: $\mathbf{y}^{(n)}(t)=\mathbf{y}^{(n')}(t)$; $n\neq n'$. We now look at these balanced synchronization methods in several examples.

III. EXAMPLES

A. The Lorenz model

Our first example is that of the Lorenz system. We have three dynamical equations for the driving oscillator

$$\frac{dx_1(t)}{dt} = \sigma(x_2(t) - x_1(t)),$$

$$\frac{dx_2(t)}{dt} = -x_2(t) + R_D x_1(t) - x_1(t)x_3(t),$$

$$\frac{dx_3(t)}{dt} = -bx_3(t) + x_1(t)x_2(t), \tag{23}$$

and three equations for the driven receiver

$$\frac{dy_1(t)}{dt} = \sigma(y_2(t) - y_1(t)) + K_{11}(x_1(t) - y_1(t)),$$

$$\frac{dy_2(t)}{dt} = -y_2(t) + Ry_1(t) - y_1(t)y_3(t),$$

$$\frac{dy_3(t)}{dt} = -by_3(t) + y_1(t)y_2(t). \tag{24}$$

1. Controlling the largest conditional Lyapunov exponent

We chose conventional values for the parameters σ and b [11]: $\sigma=16.0$ and $b=4.0$. In the driver oscillator we selected $R_D=45.92$. As described above, we chose a cost function which balances the least squares deviation of the driver input $x_1(t)$ and the output $y_1(t)$ against the deviation of the largest CLE $\lambda(K_{11},R)$ from a small negative number $\xi=-0.05$:

$$C(K_{11},R) = \frac{1}{2T}\int_0^T (x_1(t) - y_1(t))^2 dt + \frac{1}{2}(\lambda(K_{11},R) - \xi)^2, \tag{25}$$

and varied K_{11} and R. We plot the value of $C(K_{11},R)$ at its minimum for a given R as we vary K_{11}, $C_{\min}(R)$. Then we search over R for a minimum in the two variables R and K_{11}.

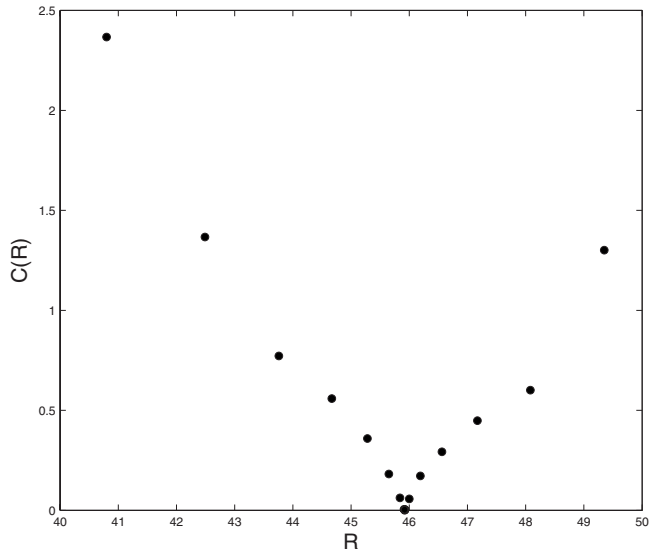

FIG. 1. The cost function $C(K_{11},R)$ Eq. (25) as a function of R at the minimum value of K_{11} for each R.

We selected $T=5000$ samples of the driving Lorenz model trajectory and the driven trajectory.

As we see in Fig. 1 there is a clear minimum in $C_{\min}(R)$ at $R=R_D$. In Fig. 2 we plot the cost function $C(K_{11},R)$. In the latter there is again a clear minimum for $K_{11}\approx 9$ and at $R=R_D$. This minimum results from the balance in the cost function between the two terms. Were the second term absent, we would see the cost function simply flatten out as K_{11} increased. Since $\lambda(K_{11},R)$ grows as K_{11} we find the desired balancing of the coupling as discussed above.

2. Balancing the magnitude of K

In this case we replace the driven system $(y_1(t),y_2(t),y_3(t))$ above with

$$\frac{dy_1(t)}{dt} = \sigma(y_2(t)-y_1(t)) + K(t)(x_1(t)-y_1(t)),$$

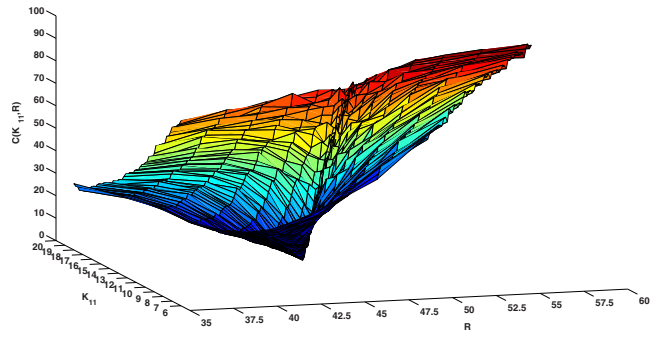

FIG. 2. (Color online) The cost function $C(K_{11},R)$ Eq. (25) as a function of R and K_{11}. The balancing effect of the "cost" for the requirement of the conditional Lyapunov exponent to be slightly negative leads to a clear minimum here.

$$\frac{dy_2(t)}{dt} = -y_2(t) + Ry_1(t) - y_1(t)y_3(t),$$

$$\frac{dy_3(t)}{dt} = -by_3(t) + y_1(t)y_2(t),$$

$$\frac{dK(t)}{dt} = -aK(t) + g(\delta^2(t)), \tag{26}$$

where $\delta^2(t)$ is the squared error $(x_1(t)-y_1(t))^2$ if we use just the scalar data. If we reconstruct state space using time delay variables, the scalar data is replaced by the D-dimensional vector $\mathbf{x}(t)=[x_1(t),x_1(t-T\tau),x_1(t-2T\tau),\ldots,x_1(t-(D-1)T\tau)]$ and the reconstructed state space vector of the model $\mathbf{y}(t)=[y_1(t),y_1(t-T\tau),y_1(t-2T\tau),\ldots,y_1(t-(D-1)T\tau)]$. Then $\delta^2(t)=|\mathbf{x}(t)-\mathbf{y}(t)|^2$. With our choice of time step in solving the Lorenz equations, we select $T\tau=0.1$, using the first minimum of average mutual information as a criterion. Using false nearest neighbors we find $D=3$.

3. Unbounded forcing function $g[\delta^2(t)]: g(z)=z^2$

We first choose as a forcing in the $K(t)$ equation, $g(z)=z^2$, and

$$\frac{dK(t)}{dt} = -aK(t) + \frac{(x_1(t)-y_1)^2}{A^2}, \tag{27}$$

with $A=a=0.05$.

We have evaluated the average squared error

$$\frac{1}{T}\int_{t_0}^{T+t_0}(x_1(t)-y_1(t))^2dt \tag{28}$$

and the "cost" function

$$\frac{1}{T}\int_{t_0}^{T+t_0}\tanh[(x_1(t)-y_1(t))^2]dt \tag{29}$$

for use in determining the parameter R in Eq. (24).

In Fig. 3 we show the integrated error Eq. (28) as a function of R when we use only the scalar values $x_1(t)$ and $y_1(t)$ in our error estimates. In Fig. 4 we show the same quantity when we use reconstructed time delay state space in the error estimate. In Fig. 5 we show the cost function Eq. (29) when only the scalars $x_1(t)$ and $y_1(t)$ are used in evaluating the error. In Fig. 6 we use time delay reconstruction of the phase space.

In each case there is a clear minimum at the value $R=R_D=45.92$.

4. Bounded forcing function $g[\delta(t)]: g(z)=4\tanh(z^2)$

We now choose as a forcing in the $K(t)$ equation, $g(z)=4\tanh(z^2)$, and

$$\frac{dK(t)}{dt} = -aK(t) + 4.0\tanh\left[(x_1(t)-y_1)\right)^2A^2, \tag{30}$$

with $A=a=0.05$.

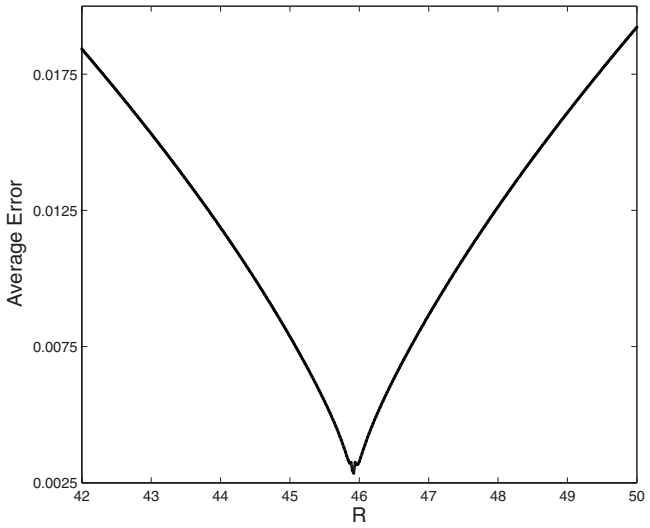

FIG. 3. Lorenz system $g(z)=z^2$ forcing in the $K(t)$ equation. The error is evaluated using only scalar signals.

Again we evaluated the integrated error [Eq. (28)] and the cost function [Eq. (29)] as a function of R, and the results are essentially the same as for $g(z)=z^2$ above. We do not show them here.

To illustrate the generalized synchronization operating in the determination of the parameter R, we presented the same driving signal $x_1(t)$ to the model system in Eq. (26) with different initial conditions for $[y_1(t=0), y_2(t=0), y_3(t=0), K(t=0)]$. If the data system $\mathbf{x}(t)$ and the model with time-dependent coupling show generalized synchronization, then after a transient, the output $\mathbf{y}(t), K(t)$ should be the same for each presentation. In Fig. 7 we plot $y_1(t)$ for three presentations of the same driver data $x_1(t)$ when the receiver system $[y_1(t), y_2(t), y_3(t), K(t)]$ begins in three different initial conditions. In Fig. 8 we plot $K(t)$ for three presentations

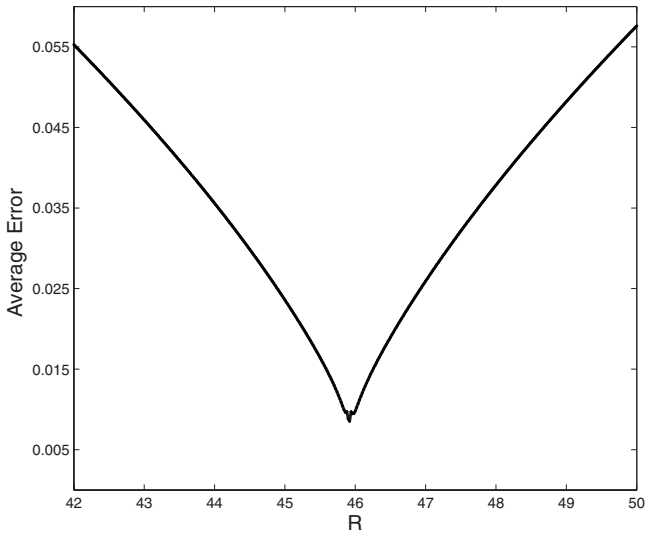

FIG. 4. Lorenz system $g(z)=z^2$ forcing in the $K(t)$ equation. The error is evaluated using time delay reconstructed phase space signals.

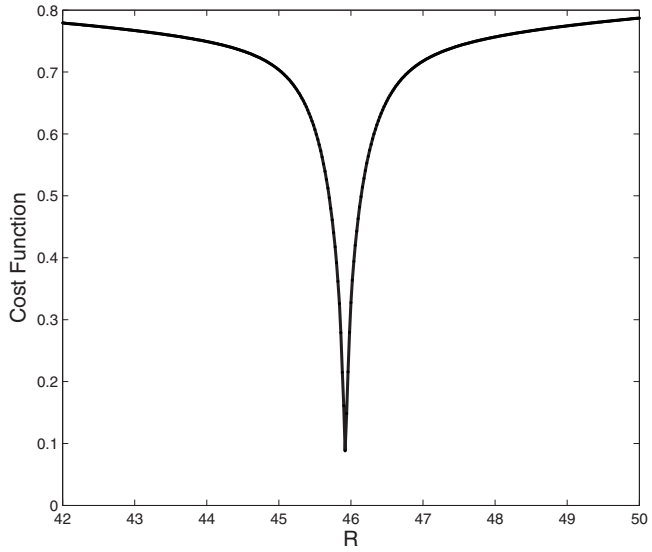

FIG. 5. Lorenz system $g(z)=z^2$ forcing in the $K(t)$ equation. The error is evaluated using only scalar signals.

of the same driver data $x_1(t)$ when the receiver system $(y_1(t), y_2(t), y_3(t), K(t))$ begins in three different initial conditions. This provides graphic, quantitative evidence for generalized synchronization between the three-dimensional driver system and the four-dimensional receiver system.

B. The Colpitts oscillator

The Colpitts oscillator is an electronic circuit using a bipolar junction transistor as the nonlinear gain element [17,18]. The three dynamical equations for this oscillator are

$$\frac{dx_1(t)}{dt} = \alpha_D x_2(t),$$

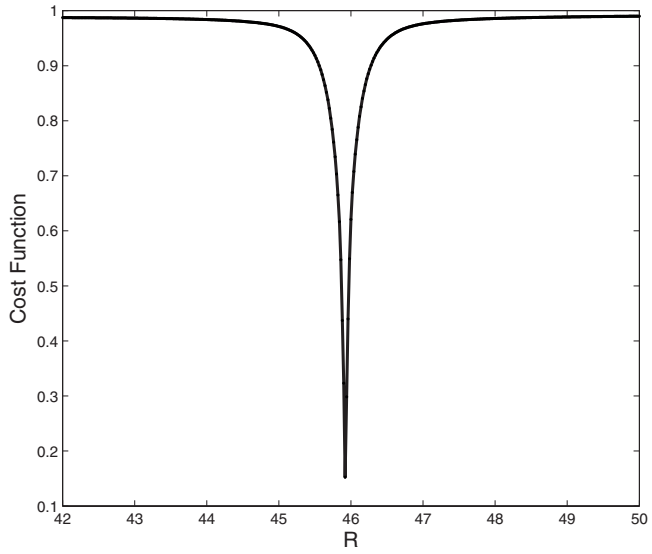

FIG. 6. Lorenz system $g(z)=z^2$ forcing in the $K(t)$ equation. The error is evaluated using time delay reconstructed phase space signals.

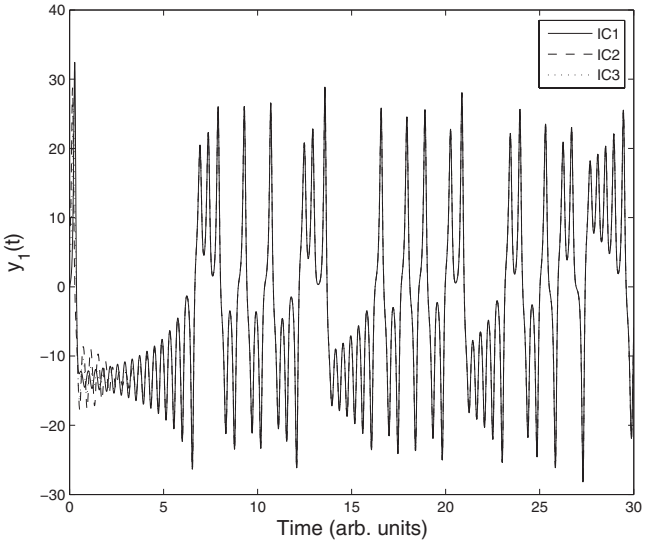

FIG. 7. $g(z)=\tanh(z^2)$ forcing in the $K(t)$ equation. Evidence for generalized synchronization of the three-dimensional Lorenz system $(x_1(t),x_2(t),x_3(t))$ with the four-dimensional system $(y_1(t),y_2(t),y_3(t),K(t))$ using the auxiliary system method. This shows the output $y_1(t)$ upon presentation of the same input signal $x_1(t)$ for three different initial conditions (IC) for the four-dimensional receiver system.

$$\frac{dx_2(t)}{dt} = -\gamma_D(x_1(t)+x_3(t)) - qx_2(t),$$

$$\frac{dx_3(t)}{dt} = \eta(x_2(t)+1-e^{-x_1(t)}). \tag{31}$$

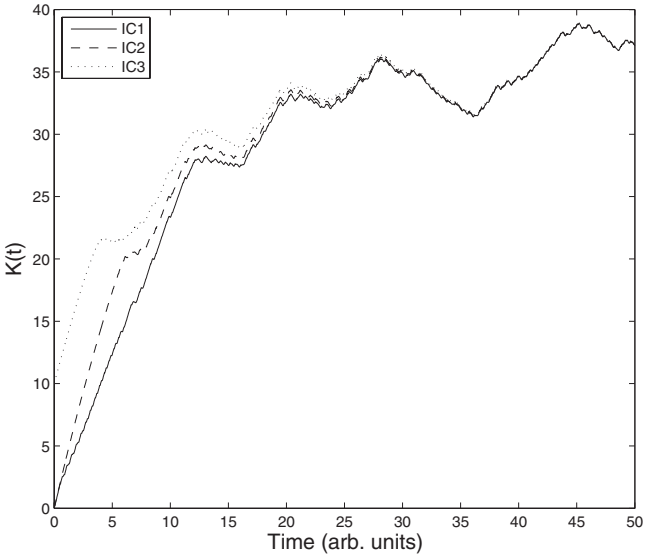

FIG. 8. $g(z)=\tanh(z^2)$ forcing in the $K(t)$ equation. Evidence for generalized synchronization of the three dimensional Lorenz system $(x_1(t),x_2(t),x_3(t))$ with the four-dimensional system $(y_1(t),y_2(t),y_3(t),K(t))$ using the auxiliary system method. This shows the output $K(t)$ upon presentation of the same input signal $x_1(t)$ for three different initial conditions for the four-dimensional receiver system.

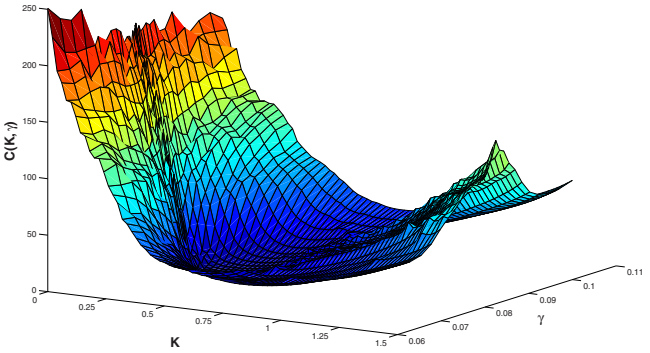

FIG. 9. (Color online) $C(K,\gamma)$ using the CLE method, Eq. (33), for Colpitts oscillator driven by $x_1(t)$ data to determine the coupling parameter K and the model parameter γ.

Using the following parameter values $\gamma_D=0.08$, $q=0.6898$, and $\eta=6.2723$ [17] we explored a range of values of α. For $\alpha \geq 3$ or so, we found chaotic behavior of the oscillator, and we collected "data" $\mathbf{x}(t)$ from solving these equations for $\alpha_D=5.0$.

1. Controlling the largest conditional Lyapunov exponent

Using the output of this system we drove another Colpitts oscillator with $x_1(t)$

$$\frac{dy_1(t)}{dt} = \alpha_D y_2(t) + K(x_1(t)-y_1(t)),$$

$$\frac{dy_2(t)}{dt} = -\gamma(y_1(t)+y_3(t)) - qy_2(t),$$

$$\frac{dy_3(t)}{dt} = \eta(y_2(t)+1-e^{-y_1(t)}). \tag{32}$$

We again chose a cost function which balances the least squares deviation of the driver input $x_1(t)$ and the output $y_1(t)$ against the deviation of the largest CLE $\lambda(K,\gamma)$ from a small negative number $\xi=-0.05$:

$$C(K,\gamma) = \frac{1}{2T}\int_0^T (x_1(t)-y_1(t))^2 dt + \frac{1}{2}(\lambda(K,\gamma)-\xi)^2, \tag{33}$$

and varied K and γ.

In Fig. 9 we show $C(K,\gamma)$. As in the case for the Lorenz system, there is a clear minimum establishing a value for $K \approx 1$ and indicating that $\gamma=\gamma_D=0.08$.

2. Balancing the magnitude of K

Using the same time series $x_1(t)$, we then solved the following equations:

$$\frac{dy_1(t)}{dt} = \alpha y_2(t) + K(t)(x_1(t)-y_1(t)),$$

$$\frac{dy_2(t)}{dt} = -\gamma_D(y_1(t)+y_3(t)) - qy_2(t),$$

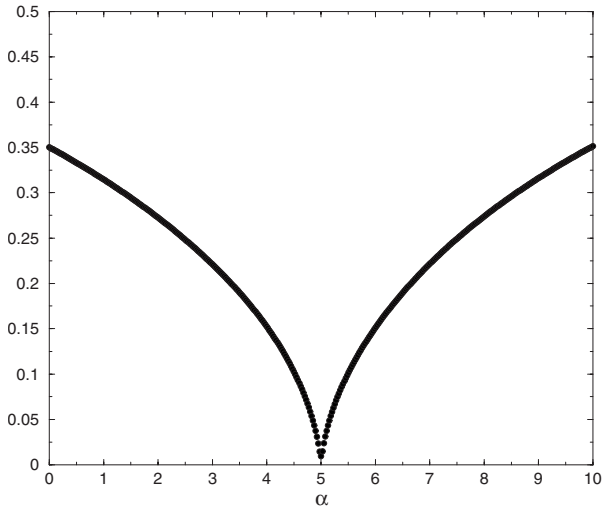

FIG. 10. Phase space estimation error with only knowledge of the data $x_1(t)$ and the equivalent component of the models state $y_1(t)$. Colpitts oscillator, $a=0.03$. No phase space reconstruction is used here.

$$\frac{dy_3(t)}{dt} = \eta(y_2(t) + 1 - e^{-y_1(t)}),$$

$$\frac{dK(t)}{dt} = -aK(t) + \tanh(\delta^2(t)/2), \tag{34}$$

with $\delta^2(t)=(x_1(t)-y_1(t))^2$ when we use scalar values of the data only and $\delta^2(t)=|\mathbf{x}(t)-\mathbf{y}(t)|^2$ when we use time delay reconstructed phase space values from $x_1(t)$ and $y_1(t)$ for the data and the model respectively. $a=0.03$ in our calculations.

Along with the time dependence of $K(t)$ we recorded the following phase space estimation error indicating the quality of the matching of the data and the model output

$$\frac{1}{T}\int_0^T dt \tanh(\delta^2(t)). \tag{35}$$

This metric is bounded above by unity, and when the parameters in the model are varied this should be a minimum when the model parameter α matches that chosen in producing the data $x_1(t)$. In Fig. 10 we plot the value of this phase space estimation error as a function of α using scalar data only for $\delta^2(t)$.

In Fig. 11 we show the time series of $K(t)$ when we use time delay reconstructed phase space vectors for error$^2(t)$. Note that $K(t)$ does not go to zero but remains bounded and remains much smaller than its upper bound $\frac{1}{a}\approx 33.3$. In this case we have a phase space error function shown in Fig. 12 which strongly indicates the correct value of the parameter α at $\alpha=5.0$.

We asked what constant value of K would produce identity synchronization between the two Colpitts oscillators when only the signal $x_1(t)$ was passed from the driver to the receiver. We found that $K\approx 0.75$ is where the largest conditional Lyapunov exponent becomes negative and identity synchronization is well established. Looking again at Fig. 11

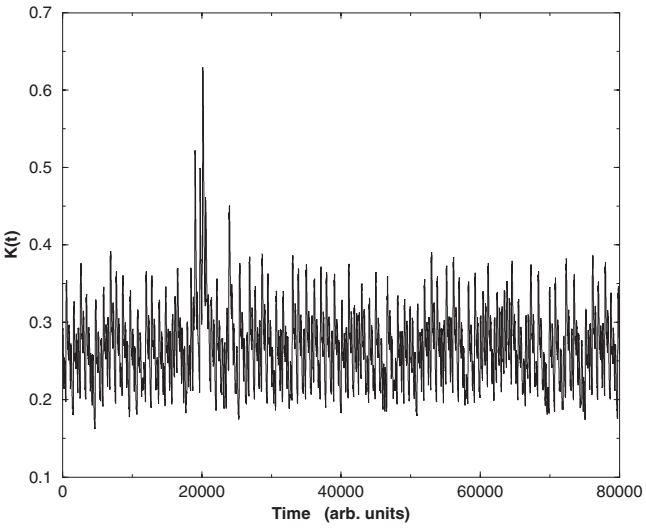

FIG. 11. $K(t)$ for reconstructed phase space for Colpitts oscillator $a=0.03$.

we see that the value of $K(t)$ varies around 0.25 not quite reaching unity at times indicating that there is, indeed, a balance between $K(t)$ decreasing towards zero, where synchronization is lost, and $K(t)$ rising due to the loss of synchronization.

We again examined the possibility of generalized synchronization in the case of the Colpitts oscillators coupled with varying $K(t)$. In Fig. 13 we show the output $K(t)$ and in Fig. 14, $y_1(t)$ in response to the same $x_1(t)$ presented three times to different initial states of Eq. (34). The transients are not shown.

Using the $K(t)$ method, we examined our ability to determine multiple parameters by evaluating the cost function over a set of values of η and γ in Fig. 15 we show Cost(η,γ) resulting from solving the four differential Eqs. (34) with $x_1(t)$. The minimum of this cost function determines both η and γ accurately.

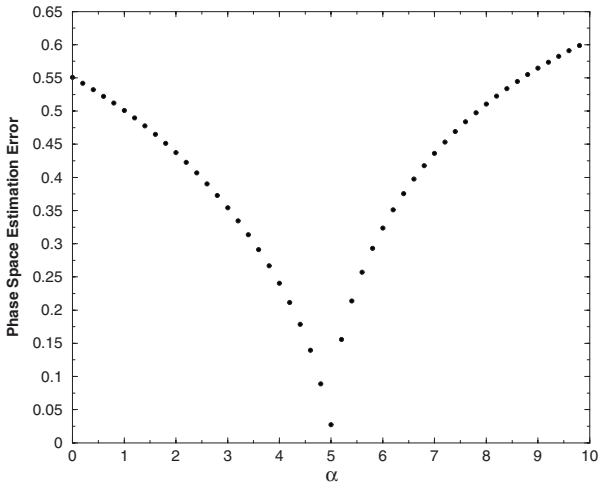

FIG. 12. Phase space estimation error for reconstructed phase space for Colpitts oscillator $a=0.03$.

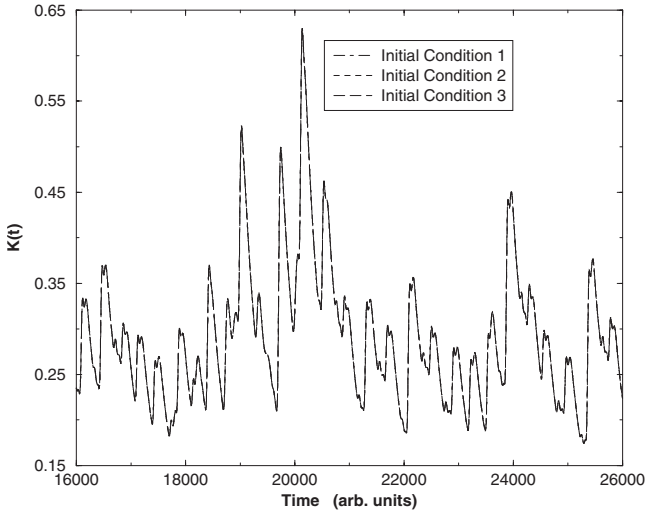

FIG. 13. $K(t)$ for the four-dimensional driven Colpitts oscillator (the "model") as a response to input $x_1(t)$ from a Colpitts "data" stream when the driven system is in three different states arising from three different initial conditions. This, via the auxiliary system method, indicates generalized synchronization between the data and the model.

IV. NOISE ANALYSIS

We next asked how robust our method was to noise. Here we consider only the effects of transmission noise, that is, noise in the data signal $x_1(t)$. Noise $n(t)$ was represented as a normally distributed, zero mean signal added to the data signal. We quantified the amount of noise using the traditional signal-to-noise ratio

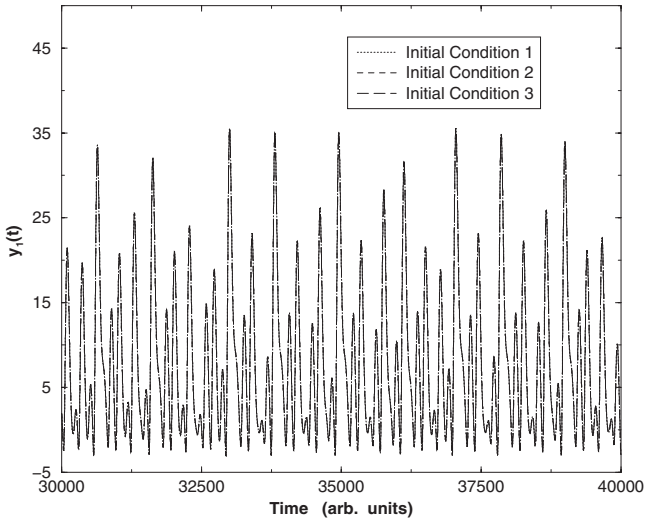

FIG. 14. $y_1(t)$ for the four-dimensional driven Colpitts oscillator (the "model") as a response to input from a Colpitts "data" stream when the driven system is in three different initial states. This, via the auxiliary system method, indicates generalized synchronization between the data and the model.

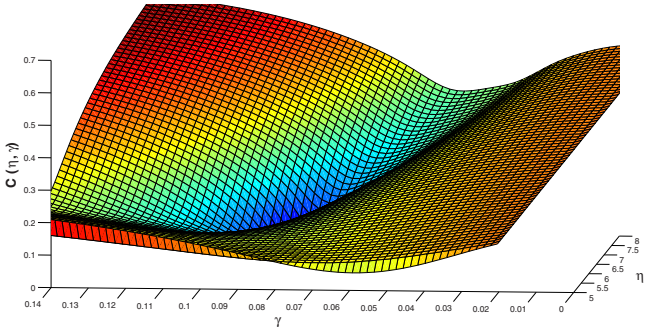

FIG. 15. (Color online) $C(\eta, \gamma)$ using the $K(t)$ method, Eq. (34), for the Colpitts oscillator driven by $x_1(t)$ data to determine the parameters η and γ. Here, η=6.723 and γ=0.0797.

$$(\mathrm{SNR}) = 10 \log_{10} \frac{\sum x_1^2(t)/T}{\sum n^2(t)/T}, \tag{36}$$

and computed the cost function using the original non-noisy signal. We tested a range of signal to noise ratios and then estimated the SNR threshold at which the cost function no longer had a global minimum at the correct value of the data parameter. This analysis was performed on the Colpitts oscillator using both synchronization methods described above. For the method of controlling the largest CLE, we found the minimum SNR to be roughly 10 dB (see Fig. 16). For the method of balancing the magnitude of K, we found the minimum SNR to be roughly 20 dB (see Fig. 17). It should be noted, however, that as the noise increases in a time series, the CLEs become increasingly difficult to compute, since the noise effectively increases the dimensionality of the system. Thus, the method of controlling the largest CLE is likely to be less efficient in noise than in the absence of noise.

V. DISCUSSION

Determining unknown model parameters from observed data is one of the critical and traditional steps in developing predictive models [7]. The idea of using synchronization of the data source and the model to establish unknown parameters is not at all a new idea, and in one manner or another has always been used in this context [6]. The "synchronization" in an automated fashion as explored in this paper has not always been the established procedure where "fits good to the eye" or other qualitative methods have been used.

In trying to implement formal synchronization of the experimental system producing the data and a proposed model for use in predicting the future behavior of the system, we encounter a problem that if the coupling of the data into the dynamical equations of the model is too large, the variation of a traditional, least squares cost function suffers from very weak variations in the desired parameters, thus reducing the value of the method for determining those parameters. Also if the coupling is too weak and the system is chaotic, the very instabilities that lead to the chaos interfere with numerically stable parameter determination [6].

This dilemma has led us to suggest two forms of "balanced synchronization" which we have discussed and ex-

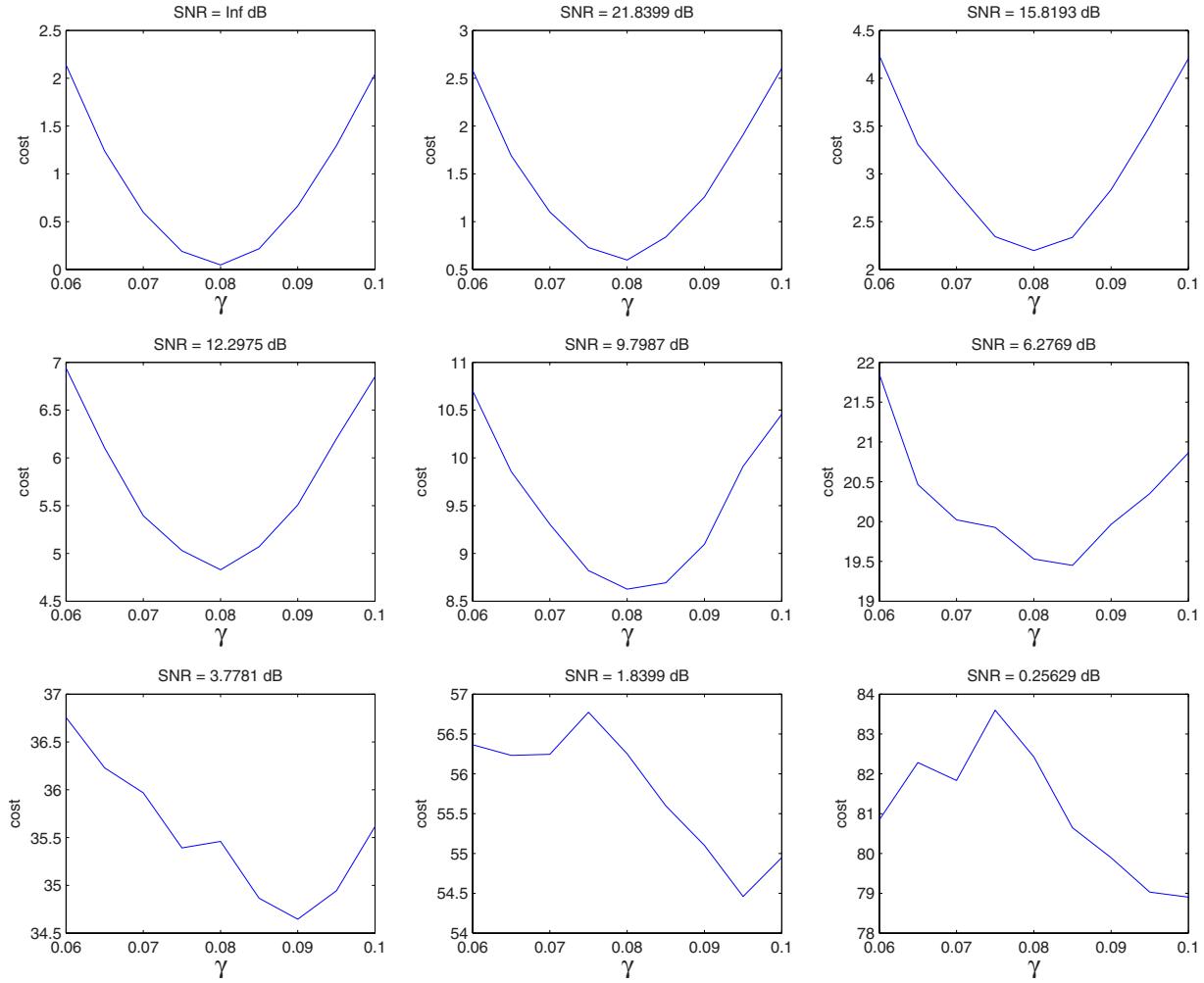

FIG. 16. (Color online) Parameter scan over γ in the Colpitts oscillator using the CLE method with $\gamma_{\text{data}}=0.08$ and $K=1$. This method correctly identifies a minimum in the cost function as long as the signal-to-noise ratio is greater than about 10 dB.

plored here in three examples. The balance between too large and too small a coupling of the data with the model has been explored by adding to the least square or other comparison of the data input and equivalent model output an additional cost associated with the conditional Lyapunov exponent (CLE) of the model, conditioned on being driven by the data. If one requires this CLE to be slightly negative, this bounds the value of the coupling while assuring the data and the model remain synchronized. When they are synchronized, the information about the interesting parameters contained in the data is efficiently passed to the model. If they are not synchronized, that does not occur.

The data has also been explored by altering the model system by adding a temporal variation of the coupling between the data and the model. The proposed equation for the coupling then balances the driving of the coupling to zero against the mismatch of the data and model signals. This method works in our examples because the enlarged model system and the data source exhibit generalized synchronization, and information transmission in that setting is also efficient and complete.

We have shown in two examples that the methods work well and exhibit the needed generalized synchronization. We have primarily assumed in this paper that the systems are deterministic, namely, no noise interferes with the information in the data about the parameters we wish to establish. Further we have assumed that the model is correct since our "data" was generated from that model and only numerical parameters need be determined.

We have also explored two cases where the data stream was contaminated by noise of various magnitudes. In these examples, the use of the CLE was shown to present some difficulty. When noise is present, the dimension of the active variables of a dynamical system can become quite large in practice. In such a case determining the largest CLE presents significant challenges. We suspect that while the cost function with a largest CLE addition has theoretical appeal, it may be difficult to implement in practice.

This last observation is underlined by another example we worked out but do not present here. We studied the three dimensional dynamical system associated with a reduced Hodgkin-Huxley model as developed by Fan and Chay [21]. This represents a spiking neuron with relatively long quiet periods between the spikes. The $K(t)$ method was very successful in determining both the maximal conductances and the kinetic parameters in this model given a chaotic data

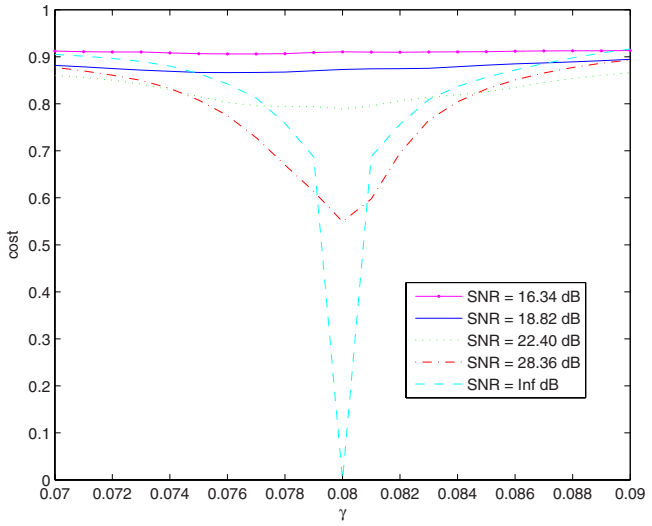

FIG. 17. (Color online) Parameter scan over γ in the Colpitts oscillator using the $K(t)$ method with $\gamma_{\text{data}}=0.08$. This method correctly identifies a minimum in the cost function as long as the SNR is greater than about 20 dB.

stream from the model itself. However, the use of the largest CLE approach did not succeed because the determination of the CLE when two very distinct time scales were present proved numerically infeasible. There are likely methods to deal with this difficulty, but we have not pursued them further.

In the framework of balanced synchronization there is much further work to be done supported by the success in this first investigation of the idea. At least the following issues need to be explored.

How well does the method work when the data is taken from an experimental system and the model is our best idea of the underlying dynamics? This is a case when the model itself may be wrong and the data may be noisy either from noise impinging on the dynamical evolution of the experiment or entering the use of the instruments in measuring the data values.

What is the general role of noise as it affects both the dynamics of the data source and the instruments measuring the data values?

If the model is "wrong" how do we account for the use of synchronization as a tool? Since all models are wrong in some aspect, this requires exploration both in a general sense and within explicit examples.

When one uses this method to determine many, many parameters can we develop an efficient algorithm for the required search in multidimensional parameter space? If one can do that, then extension of the use of balanced synchronization to networks, biological and physical, is likely possible.

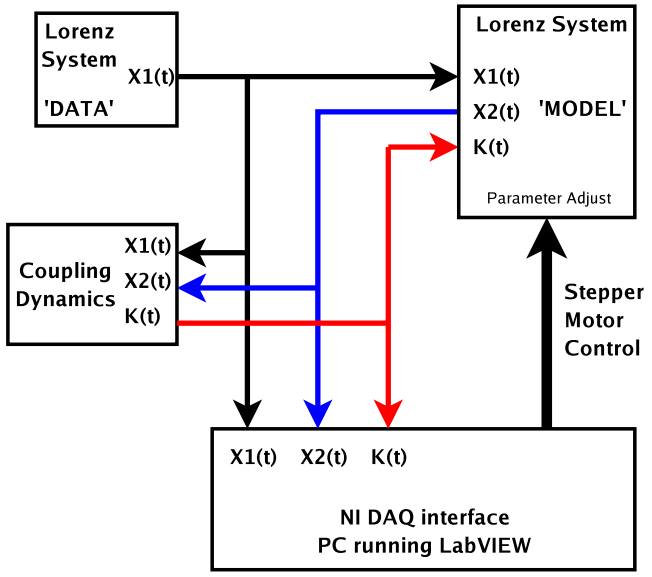

FIG. 18. (Color online) Diagram of experimental setup.

ACKNOWLEDGMENTS

This work was partially funded by a grant from the National Science Foundation, Grant No. NSF PHY0097134, and by a grant from the National Institutes of Health, NIH R01 NS40110-01A2. HA is partially supported by the NSF sponsored Center for Theoretical Biological Physics at UCSD.

APPENDIX: ELECTRONIC CIRCUIT REALIZATION OF PARAMETER ESTIMATION IN THE LORENZ SYSTEM

An electrical circuit implementation of the dynamical coupling $[K(t)]$ method of parameter estimation was built to explore this method in a more realistic setting. The experimental setup is as shown in Fig. 18. The data signal $x_1(t)$ is generated from a Lorenz system with fixed parameters (σ_1, b_1, R_1). The model is a Lorenz system with two parameters (σ_2 and b_2) identical to the data system and an adjustable third parameter R_2. A data acquisition (NI-DAQ) card connected to a PC running LABVIEW software acquires $x_1(t)$, $x_2(t)$, and $K(t)$ at a rate of 50 kHz. The PC is able to control the value of R_2 by turning a stepper motor that is connected to a 10-turn potentiometer. A typical parameter scan consists setting R_2, taking 2 s of data, computing the value of the cost function and repeating over the available range of R_2 (in particular, $23.5 \leq R_2 \leq 50.6$).

The basic design of the Lorenz portion of the electrical circuit is adapted from Cuomo [20]. In order to keep voltages in the range of the available power supplies, all state variables are scaled by a factor of 20. Time is scaled by a factor of 1000. With these scalings, the Lorenz system that needs to be realized in the electrical circuit is the following:

$$\dot{x} = 1000[\sigma(y - x)],$$

$$\dot{y} = 1000[Rx - y - 20xz],$$

$$\dot{z} = 1000[20xy - bz].$$

Figure 19 shows the schematic of the Lorenz systems and

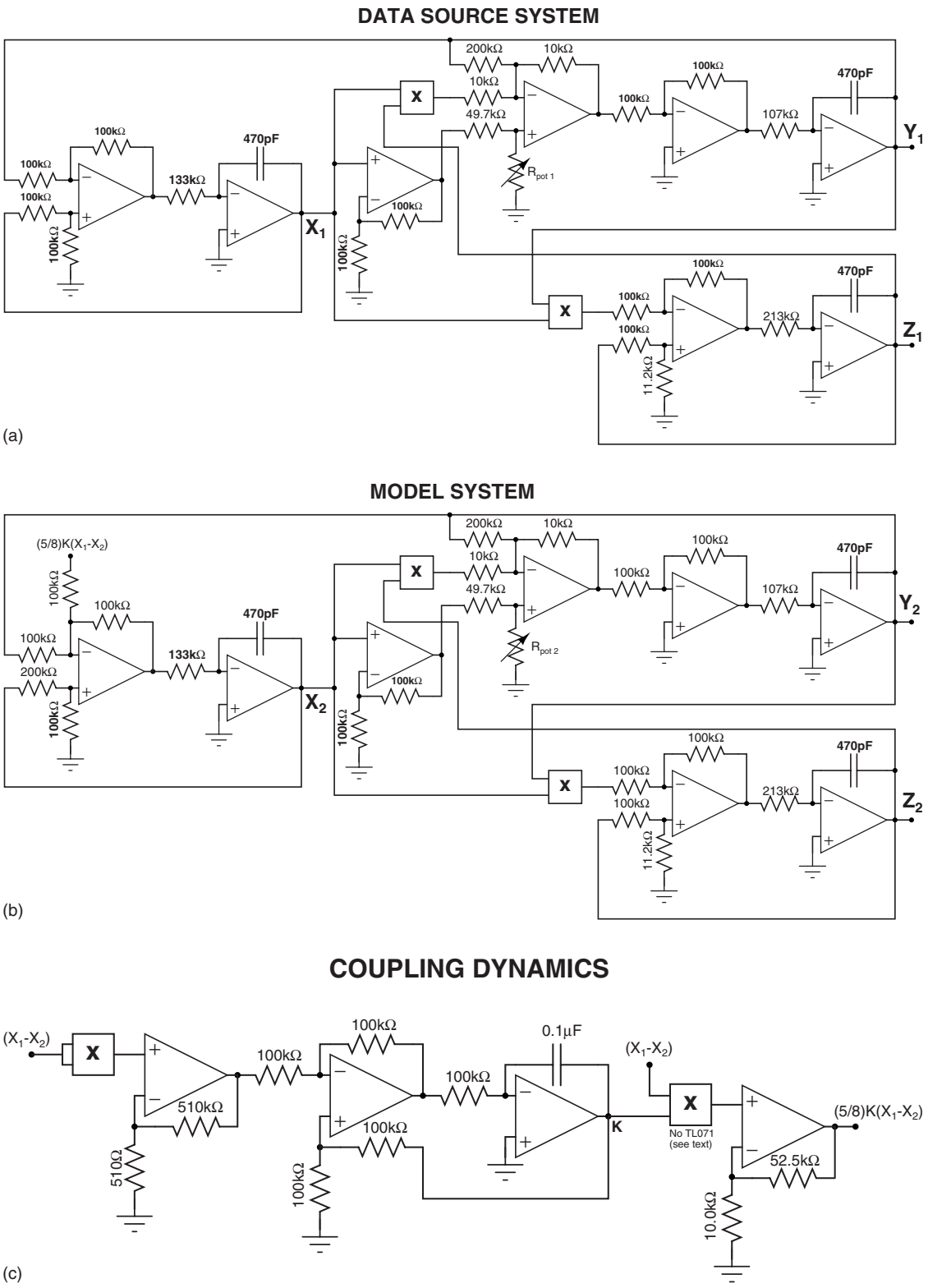

FIG. 19. Schematic of Lorenz circuits and coupling.

the coupling dynamics. The components are standard 5% tolerance resistors and 20% tolerance tantalum capacitors. No care was taken to match each component between the two systems, so there inevitably are differences between the two Lorenz systems at the component level. The Op Amps are general purpose TL084's and the multipliers are Analog Devices AD633's. Since the output of the AD633 is internally divided by 10, the multiplication symbol shown in the schematic is actually an AD633 with a TL071 Op-Amp used for a gain of 10.

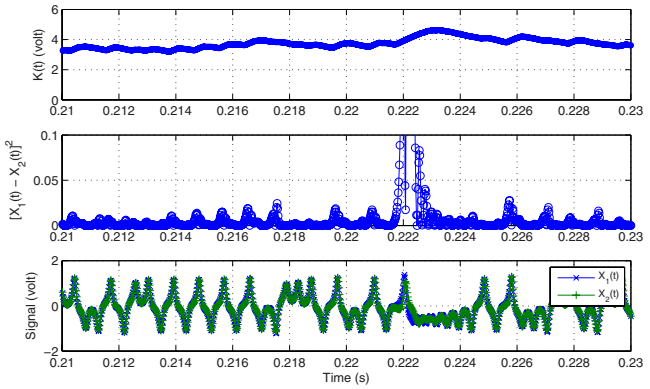

FIG. 20. (Color online) Sample time series. $R_{pot1}=R_{pot2}=50\ \mathrm{k}\Omega$.

When taking data, the potentiometer in the data system is set manually to a fixed desired value, the potentiometer in the receiver system is set by the PC with the use of a stepper motor. The values of the parameter R may be calculated from the potentiometer values using the following equation:

$$R = \frac{82R_{\mathrm{pot}}}{R_{\mathrm{pot}} + 49.7\ \mathrm{k}\Omega}.$$

The potentiometers are both 10-turn 100 kΩ. Parameter scans use the range $20\ \mathrm{k}\Omega \le R_{\mathrm{pot}} \le 80\ \mathrm{k}\Omega$ which translates to the parameter range $23.5 \le R_2 \le 50.6$ mentioned above. The equations describing the dynamics of the above electrical circuit are the following:

$$\dot{x_1} = 1000[16(y_1 - x_1)],$$

$$\dot{y_1} = 1000[R_1 x_1 - y_1 - 20x_1 z_1],$$

$$\dot{z_1} = 1000[20x_1 y_1 - 4z_1],$$

$$\dot{x_2} = 1000[16(y_2 - x_2) + 10K(x_1 - x_2)],$$

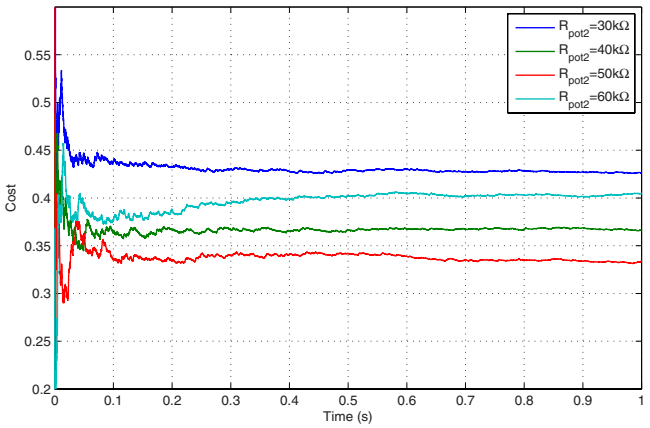

FIG. 21. (Color online) Cost function convergence. $R_{pot1}=50\ \mathrm{k}\Omega$.

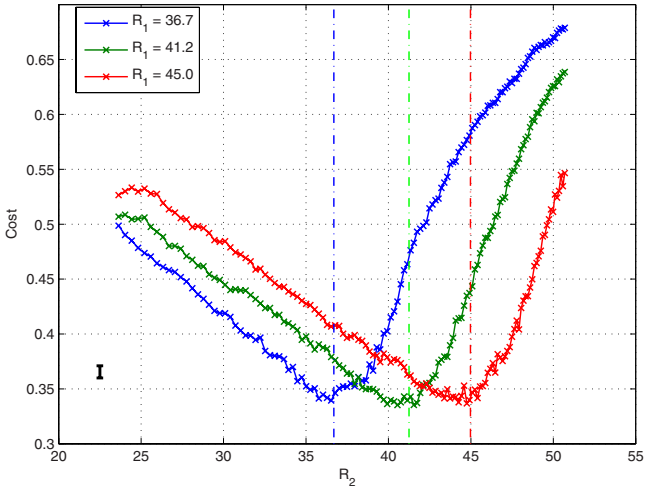

FIG. 22. (Color online) Parameter scans for several R_1.

$$\dot{y_2} = 1000[R_2 x_2 - y_2 - 20x_2 z_2],$$

$$\dot{z_2} = 1000[20x_2 y_2 - 4z_2],$$

$$\dot{K} = 1000[-0.10K + 110(x_1 - x_2)^2].$$

Note that in this appendix the receiver system has dynamical variables with subscript 2, so the coupling in the "model" is $10K(x_1 - x_2)$.

Figure 20 shows a section of the time series of $x_1(t)$, $x_2(t)$, and $K(t)$ as measured by the PC. The Lorenz systems are creating the expected time series wave forms for $x_1(t)$ and $x_2(t)$ and these wave forms are for the most part synchronized. The $K(t)$ wave form is tending to decay toward zero with growth dependent on the spiking in the error signal $(x_1(t) - x_2(t))^2$. After acquiring 2 s of data sampled at 50 kHz, $N = 100\,000$ samples, the computer calculates the associated cost function based on the following formula:

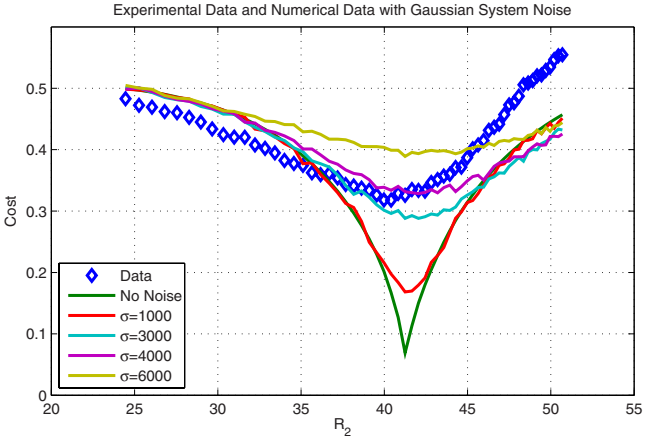

FIG. 23. (Color online) Experimental data and simulation with Gaussian noise added to the state equations.

$$C = \frac{1}{N}\sum_{j=1}^{N} \tanh[200 \cdot (x_1(j) - x_2(j))^2]$$

The convergence of this cost function is verified in Fig. 21 at four positions in the parameter scan ($R_{\mathrm{pot},2} = 30, 40, 50, 60\ \mathrm{k}\Omega$) with $R_{\mathrm{pot},1} = 50\ \mathrm{k}\Omega$. In each case the cost function converges to a steady value within the fist second of the 2 s sample window. The steady value of this cost function is automatically recorded over the range of R_2 producing the parameter scan plots shown in Fig. 22.

In the case shown in Fig. 22, three different values of the data system parameter R_1 were used. In each case there is clearly a minimum in the cost function in the neighborhood of $R_2 = R_1$. There are many sources of noise in this lab setup and it is not surprising that the minimum in the cost function is broader and shallower than the theoretical (noiseless) result—see Fig. 23. Also, since the components used are of low tolerance, it is very likely that the system parameters that are assumed to be identical between the data system and the model (σ, b) are in fact not the same. Also, the value of R depends on several other components in addition to the potentiometer, so that equal settings on the potentiometers does not necessarily imply $R_2 = R_1$. Still, even with these imperfections, this experiment demonstrates that the dynamical coupling method $[K(t)]$ does a good job of identifying a small neighborhood in which the presumably unknown parameter lies.

[1] L. M. Pecora and T. L. Carroll, Phys. Rev. Lett. **64**, 821 (1990).
[2] U. Parlitz, L. Junge, and L. Kocarev, Phys. Rev. E **54**, 6253 (1996).
[3] H. Dedieu and M. J. Ogorzalek, IEEE Trans. Circuits Syst., I: Fundam. Theory Appl. **44**, 948 (1997).
[4] A. Maybhate and R. E. Amritkar, Phys. Rev. E **59**, 284 (1999).
[5] A. Pikovsky, M. Rosenblum, and J. Kurths, *Synchronization. A Universal Concept in Nonlinear Sciences* (Cambridge University Press, Cambridge, 2001).
[6] H. Sakaguchi, Phys. Rev. E **65**, 027201 (2002).
[7] I. Tokuda, U. Parlitz, L. Illing, M. B. Kennel, and H. D. I. Abarbanel, (unpublished).
[8] H. U. Voss, J. Timmer, and J. Kurths, Int. J. Bifurcation Chaos Appl. Sci. Eng. **14**, 1905 (2004).
[9] R. Konnur, Phys. Rev. E **67**, 027204 (2003).
[10] D. Huang, Phys. Rev. E **69**, 067201 (2004).
[11] H. D. I. Abarbanel, *Analysis of Observed Chaotic Data* (Springer-Verlag, New York, 1996).
[12] H. Kantz and T. Schreiber, *Nonlinear Time Series Analysis*, 2nd ed. (Cambridge University Press, Cambridge, 2003).
[13] D. Yu, M. Righero, and L. Kocarev, Phys. Rev. Lett. **97**, 188701 (2006).
[14] G. S. Duane, D. Yu, and L. Kocarev (unpublished).
[15] N. F. Rulkov, M. M. Sushchik, L. S. Tsimring, and H. D. I. Abarbanel, Phys. Rev. E **51**, 980 (1995).
[16] H. Abarbanel, N. F. Rulkov, and M. M. Sushchik, Phys. Rev. E **53**, 4528 (1996).
[17] H. B. Fotsin and J. Daafouz, Phys. Lett. A **339**, 304 (2005).
[18] M. P. Kennedy, IEEE Trans. Circuits Syst., I: Fundam. Theory Appl. **41**, 771 (1994).
[19] D. Y. Tang and N. R. Heckenberg, Phys. Rev. E **55**, 6618 (1997).
[20] K. M. Cuomo, A. V. Oppenheim, and S. H. Strogatz, IEEE Trans. Circuits Syst., II: Analog Digital Signal Process. **40**, 626 (1993).
[21] Y.-S. Fan and T. R. Chay, Biol. Cybern. **71**, 417 (1994).