State and Parameter Estimation in Nonlinear Systems as an Optimal Tracking Problem

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

Philip E. Gill†
Department of Mathematics
University of California, San Diego,
La Jolla, CA 92093-0112, USA

Henry D. I. Abarbanel‡
Department of Physics and Marine Physical Laboratory (Scripps Institution of Oceanography)
University of California, San Diego, La Jolla, CA 92093-0402, USA
(Dated: December 2, 2007)

In verifying and validating models of nonlinear processes it is important to incorporate information from observations in an efficient manner. Using the idea of synchronization of nonlinear dynamical systems, we present a framework for connecting a data signal with a model in a way that minimizes the required coupling yet allows the estimation of unknown parameters in the model.

The need to evaluate unknown parameters in models of nonlinear physical, biophysical, and engineering systems occurs throughout the development of phenomenological or reduced models of dynamics. Our approach builds on existing work that uses synchronization as a tool for parameter estimation [1–9]. We address some of the critical issues in that work and provide a practical framework for finding an accurate solution. In particular, we show the equivalence of this problem to that of tracking within an optimal control framework. This equivalence allows the application of powerful numerical methods [13, 14] that provide robust practical tools for model development and validation.

PACS numbers: 05.45.Xt

The use of observed data to estimate parameters in a nonlinear dynamical model is an important component in the development of predictive models of physical and biological systems. This capability is required in many areas of research, but it is especially important in the development of reduced or phenomenological models of physical and biophysical phenomena, where parameters and even network connectivity may need to be specified as the model is being developed.

The synchronization of the experimental data with the model system has been suggested as a way to incorporate experimental information into the model, see, e.g., [1–9]. Synchronization shows promise in enabling the desired estimation, but there are important issues associated with the selection of the coupling strength of the data into the model. When this coupling is too small,

---
*Electronic address: dcreveli@physics.ucsd.edu
†Electronic address: pgill@ucsd.edu
‡Electronic address: hdia@jacobi.ucsd.edu

the data and the model do not synchronize, and information is not effectively exploited. When the coupling is too large, the desired minimum value of the synchronization cost function (defined as a function of the model parameters) is ill-defined. In addition, for large enough coupling, the data entrain any model, and the parameters in specific models cannot be determined at all.

The problem we address may be cast in the following form. We have a data-producing system described by the state vector $\mathbf{x}(t) = [x_1(t), x_2(t), \ldots, x_N(t)] = [x_1(t), \mathbf{x}_\perp(t)]$. Measurements are made at times $t_m = t_0 + m\tau$, resulting in $\mathbf{x}(m) = \mathbf{x}(t_0 + m\tau)$; $m = 1, \ldots, M$. One component of the state is now measured. This could be an arbitrary scalar function of the system state, but to simplify the discussion we assume that $x_1(t)$ is observed.

The observed system satisfies differential equations in $\mathbf{x}(t)$ that depend on a fixed set of parameters $\mathbf{p}$, i.e.,

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

$$\frac{d\mathbf{x}_\perp(t)}{dt} = \mathbf{F}_\perp(x_1(t), \mathbf{x}_\perp(t), \mathbf{p}).$$

The trajectories $\mathbf{x}(t)$ are determined by the initial conditions and the parameters $\mathbf{p}$.

The quantity $x_1(t)$ is now observed and passed to the model. For this discussion, we take the model to be the one used in generating the data, but we assume that the model parameters $\mathbf{q}$ are unknown. The state of the model is $\mathbf{y}(t) = [y_1(t), y_2(t), \ldots, y_N(t)] = [y_1(t), \mathbf{y}_\perp(t)]$.

In general the vector field $\mathbf{F}(\mathbf{y}, \mathbf{q})$ is *not* identical to that of the data source. One must explore *generalized synchronization* of the data and the model [7, 10, 11]. Strictly speaking, even when $\mathbf{q} \neq \mathbf{p}$ generalized synchronization is the best one can achieve. This important and somewhat complex issue will be explored in our larger paper [12] where space will allow a longer discussion.

To determine $\mathbf{q}$ for a given $x_1(t)$, we couple the model to the input $x_1(t)$ via the system

$$
\begin{aligned}
\frac{dy_1(t)}{dt} &= F_1(y_1(t), \mathbf{y}_\perp(t), \mathbf{q}) + K(x_1(t) - y_1(t)) \\
\frac{d\mathbf{y}_\perp(t)}{dt} &= \mathbf{F}_\perp(y_1(t), \mathbf{y}_\perp(t), \mathbf{q}).
\end{aligned}
\tag{1}
$$

For some range of values of the scalar K, the model will synchronize to the data $\mathbf{y}(t) \approx \mathbf{x}(t)$, and the model will be "most accurate" when the model parameters realize $\mathbf{q} = \mathbf{p}$. The conditional Lyapunov exponent (CLE) [1] of the model system must be negative for the synchronization.

We need a principle to assist in estimating $\mathbf{q}$. A natural choice is to minimize the cost function $C(\mathbf{q}) = \frac{1}{2M} \sum_{m=1}^{M} g\big((x_1(m) - y_1(m))^2\big)$, where $g(u^2) \approx u^2$ for small u.

The determination of $\mathbf{q}$ involves seeking a zero of $\partial C(\mathbf{q})/\partial \mathbf{q}$. The equation for $\partial y_1(t)/\partial \mathbf{q}$ must confront the issue of stability, because the eigenvalues of the Jacobian, $\partial \mathbf{F}(\mathbf{y})/\partial \mathbf{y} - \mathbf{K}$, for $\mathbf{K} = \begin{pmatrix} K & 0 \\ 0 & 0 \end{pmatrix}$, iterated along the orbit $\mathbf{y}(t)$, may lead to positive CLEs for small K.

If there is a positive CLE, then the synchronization manifold $\mathbf{y}(t) = \mathbf{x}(t)$ is not stable with respect to small perturbations, and the evaluation of the derivatives of the cost function is not numerically stable. Typically, the landscape of $C(\mathbf{q})$ as a function of $\mathbf{q}$ is quite complex, with many local minima [15]. However, by increasing the magnitude of K, the CLEs may be forced to be negative and to behave as $-K$ for large K.

As K becomes large the term $K(x_1(t) - y_1(t))$ dominates the right-hand side of the equation for $y_1(t)$ unless $x_1(t) - y_1(t) \approx 1/K$ or smaller. As this happens, $C(\mathbf{q}) \approx 1/K^2$, and all the derivatives $\partial C(\mathbf{q})/\partial \mathbf{q}$ approach zero. The function $C(\mathbf{q})$ becomes so flat in $\mathbf{q}$ that the zero of $\partial C(\mathbf{q})/\partial \mathbf{q}$ is extremely difficult to locate numerically and the calculation of the parameters

$\mathbf{q}$ may not be possible.

We require a method that "balances" these extremes. We proceed by replacing the constant coupling K with a function of time (control) $u(t)$. In the cost function we add a penalty for large control values, i.e., we define $C(\mathbf{q}, \mathbf{u}) = \frac{1}{2M} \sum_{m=1}^{M} [g((x_1(m) - y_1(m))^2) + u(m)^2]$. For large $u(t)$ the first term in this function behaves as $1/u$, and this, combined with the growth of the second term, leads to a balanced magnitude for $u(t)$.

If the cost function is minimized subject to the constraint of satisfying the differential equations (1) with $K \rightarrow u(t)$, we obtain a classical tracking problem with optimal control [16]. The trajectory of the dynamical system for $\mathbf{y}(t)$ is controlled by $u(t)$ to track $x_1(t)$, or, in contemporary language, to *synchronize with the observed orbit* $x_1(t)$. Because of the properties of nonlinear systems, when the CLE of (1) with $K \rightarrow u(t)$ is negative, the unobserved components of the state, i.e., $\mathbf{y}_\perp(t)$, will also track the unobserved $\mathbf{x}_\perp(t)$.

To solve the optimization problem, we use a "direct transcription" method [13], which defines a finite-dimensional problem with variables given by the states $\mathbf{y}(m)$, the controls $u(m)$, and the fixed parameters $\mathbf{p}$. The cost function $\frac{1}{2M} \sum_{m=1}^{M} [g((x_1(m) - y_1(m))^2) + u(m)^2]$ is minimized subject to equality constraints that connect the $\mathbf{y}(m)$ and the $u(m)$ across each time interval $[t_m, t_{m+1}]$. These constraints are imposed in the finite-dimensional space of the variables $\{\mathbf{y}(m), u(m), \mathbf{q}\}$ and are characterized by an integration rule for the states $\mathbf{y}(m)$, and an interpolation rule for the control $u(t)$. Although the resulting finite-dimensional optimization problem has many variables, it is also smooth and has derivatives that may be calculated efficiently. Indeed, it is this transcription into a smooth large-scale constrained optimization problem that is the key to the robust and efficient estimation of the parameters.

Here we report work using Simpson's rule for the integration and Hermite interpolation for the states and controls. Given the values of the states and controls at the mid-point $m2 = m + \frac{1}{2}$, Simpson's rule for integration defines a constraint in each interval of the form:

$$
\begin{aligned}
\mathbf{y}(m) + \frac{\tau}{6} [&\mathbf{G}(\mathbf{y}(m), u(m), \mathbf{q}) \\
&+ \mathbf{G}(\mathbf{y}(m+1), u(m+1), \mathbf{q}) \\
&+ 4\mathbf{G}(\mathbf{y}(m2), u(m2), \mathbf{q})] - \mathbf{y}(m+1) = 0,
\end{aligned}
$$

where $\mathbf{G}(\mathbf{y}, u, \mathbf{q}) = \mathbf{F}(\mathbf{y}, u, \mathbf{q}) + \mathbf{M}u(\mathbf{x} - \mathbf{y})$, $\mathbf{M}$ is an $N \times N$ time-independent matrix with only $\mathbf{M}_{11} = 1$ nonzero, and $\mathbf{u} = (u, 0, 0, \ldots, 0)$. In addition, $\mathbf{y}(m2)$, $\mathbf{y}(m)$, $u(m2)$ and $u(m)$ are required to satisfy the Her-

mite interpolation condition [17], giving

$$
\begin{aligned}
\mathbf{y}(m2) \;=\; & \frac{1}{2}(\mathbf{y}(m) + \mathbf{y}(m+1)) + \frac{\tau}{8}(\mathbf{G}(\mathbf{y}(m), u(m), \mathbf{q}) \\
& - \mathbf{G}(\mathbf{y}(m+1), u(m+1)), \mathbf{q})),
\end{aligned}
$$

and

$$
u(m2) = \frac{1}{2}(u(m) + u(m+1)) + \frac{1}{8}(du(m) - du(m+1)),
$$

where $du(m)$ is the slope parameter in the Hermite interpolation.

Next, we discuss the application of this transcription to an example of the Colpitts oscillator [18]. This is an electronic circuit that uses a bipolar junction transistor as the nonlinear gain element. The three dynamical equations for this oscillator are

$$
\begin{aligned}
\frac{dx_1(t)}{dt} &= \alpha_D x_2(t) \\
\frac{dx_2(t)}{dt} &= -\gamma_D(x_1(t) + x_3(t)) - q_D x_2(t) \\
\frac{dx_3(t)}{dt} &= \eta_D(x_2(t) + 1 - e^{-x_1(t)}).
\end{aligned}
$$

For fixed $\gamma_D = 0.08$, $q_D = 0.7$, and $\eta_D = 6.3$. For $\alpha_D \geq 3$, chaotic behavior was exhibited. The "data" $x_1(t)$ was collected for $\alpha_D = 5.0$.

Then we selected a model for the system that produced the data stream $x_1(t)$. The model system, including a control $u(t)$ to enforce synchronization, is

$$
\begin{aligned}
\frac{dy_1(t)}{dt} &= \alpha_M y_2(t) + u(t)(x_1(t) - y_1(t)) \\
\frac{dy_2(t)}{dt} &= -\gamma_M(y_1(t) + y_3(t)) - q_M y_2(t) \\
\frac{dy_3(t)}{dt} &= \eta_M(y_2(t) + 1 - e^{-y_1(t)}).
\end{aligned}
$$

The optimal tracking problem is to determine the state $\mathbf{y}(t)$, the control $u(t)$, and the parameters α_M, q_M, γ_M, and η_M.

The Hermite-Simpson transcription was applied with 150 time steps. The cost function

$$
\begin{aligned}
C \;=\; & \frac{1}{2} \sum_{m=1}^{M} [(x_1(m) - y_1(m))^2 + u(m)^2 + du(m)^2 \\
& + (x_1(m2) - y_1(m2))^2 + u(m2)^2],
\end{aligned}
$$

includes the synchronization term, a term $u(m)^2 + u(m2)^2$ that limits the magnitude of the control, and a term $du(m)^2$ that serves to provide a smooth control.
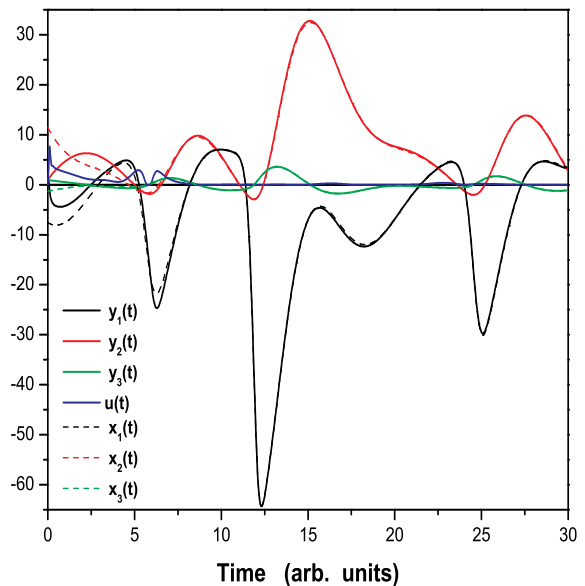

FIG. 1: The SNOPT solution of the optimal tracking problem for the Colpitts oscillator. The model system is given the "data" $x_1(m)$, and the state variables $(y_1(m), y_2(m), y_3(m))$, the control $u(m)$, and the parameters $\alpha_M, \gamma_M, q_M,$ and η_M are determined. After an initial transient, all state variables accurately track the data. The model parameters were determined to be $\alpha_M = 4.99$, $\gamma_M = 0.082$, $q_M = 0.704$, and $\eta_M = 6.26$. The parameters used to generate the data $x_1(t)$ were $\alpha_D = 5.0$, $\gamma_D = 0.08$, $q_D = 0.7$, and $\eta_D = 6.3$.

The midpoint quantities $\mathbf{y}(m2)$ and $u(m2)$ are found using cubic Hermite interpolation; the value of $x_1(m2)$ was presumed measured and provided in the data file. If $x_1(m2)$ is not known then it may be defined using linear interpolation. These relations were imposed as additional equality constraints in the optimization. The finite-dimensional problem was solved using the SNOPT constrained optimization package [13].

Figure 1 gives the state variables $\mathbf{y}(m)$, the transmitted data $x_1(m)$, and the unknown state variable data $x_2(m)$ and $x_3(m)$. Also shown is the control $u(m)$. We purposely integrated the model equations with fixed initial conditions in order the demonstrate the ability of SNOPT to track data with a transient coming from an unknown initial condition.

We also explored the ability of the method to track time-varying parameters. The time dependence is determined externally, and we represent the parameter variation by a cubic polynomial over a time interval $N\tau$,
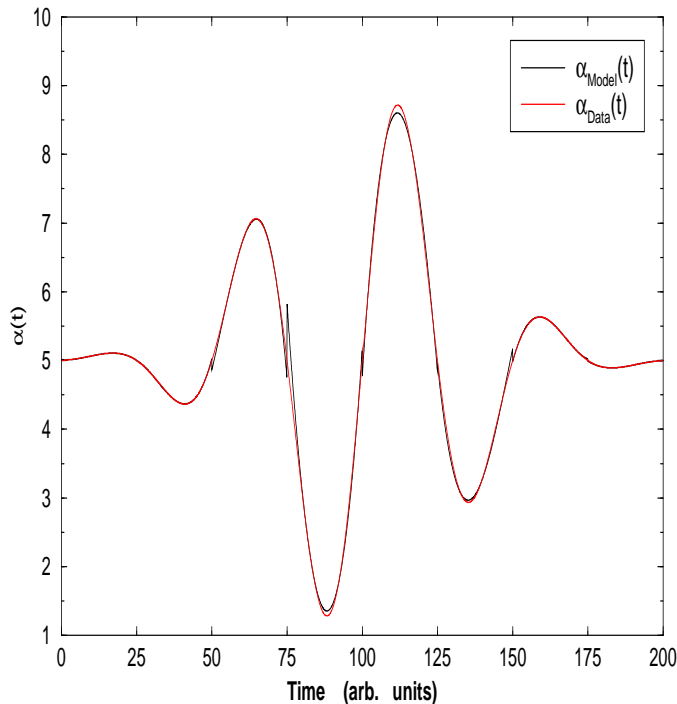
FIG. 2: The SNOPT solution for the optimal tracking problem of a Colpitts oscillator with a time-dependent parameter. All the state variables track accurately over the time interval used. The fixed parameters γ, q, and η were as before.

where N was between 25 and 200, to provide an appropriate smoothness to the time dependence.

We considered data $\alpha_D(t) = 5 + 4e^{(t-100)^2/2000} \sin(2\pi t/50)$. This variation crosses the bifurcation boundaries between fixed point, limit cycle and chaotic behavior of the Colpitts solution. Figure 2 gives the results for this case. The initial conditions $\mathbf{y}(0)$, $u(0)$, $\alpha_M(0)$ were determined by the optimization, and all state variables $\mathbf{y}(m)$ tracked the data $\mathbf{x}(m)$ very accurately. $N = 25$.

Many methods have been explored for parameter estimation in nonlinear systems; two quite interesting ones are discussed in detail in [15]. The multiple shooting and extended Kalman filter approaches considered there show good results when applied to simple systems, as here. In [12] we will explore models where parameters enter in a nonlinear fashion. We will also consider the role of both observational and dynamical noise.

We call the proposed method a "dynamical microscope". It combines a balanced dynamical synchronization of the observed data with a model of the system that produced the data. The method should allow the use of one or more dynamical variables from an observed network of nonlinear oscillators to determine other parameters associated with nodes or links of the network. When they refer to network links, they determine the connectivity of the network. We believe that the method will provide a useful tool for study and validation of complex models. Of course, the method does not replace the need to develop models based on insights into the physics or biophysics of the processes involved.

[1] Pecora, L. M. and T. L. Carroll, *Physical Review Letters* **64**, 821 (1990).
[2] Parlitz, U., L. Junge, and L. Kocarev, *Physical Review E* **54**, 6253 (1996).
[3] Dedieu, H. and M. J. Ogorzalek, *IEEE Transactions, Circuits and Systems I* **44**, 948 (1997).
[4] Maybhate, A. and R. E. Amritkar, *Physical Review E*, **59**, 284 (1999).
[5] Pikovsky A., M. Rosenblum, and J. Kurths, *Synchronization. A universal concept in nonlinear sciences*, Cambridge University Press (2001).
[6] Sakaguchi, H., *Physical Review E*, **65**, 027201 (2002).
[7] Tokuda, I, et al, Proceedings of the 7th Experimental Chaos Conference, San Diego, CA, USA (2002)
[8] Konnur, R., *Physical Review E*, **67**, 027204 (2003).
[9] Huang, D., *Physical Review E*, **69**, 067201 (2004).
[10] Rulkov, N. F., M. M. Sushchik, L. S Tsimring, and H. D. I. Abarbanel, *Physical Review E* **51**, 980-994 (1995).
[11] Smith, L. and P. McShary, *Phys. Rev. Lett.* **83**, 4285-4288 (1999).
[12] Abarbanel, H. D. I., D. R. Creveling, P. E. Gill, J. Rofeh, B. Tyler, and Z. Singer, *Physical Review E*, to be submitted, Fall, 2007.
[13] Gill, P. E., W. Murray, and M. A. Saunders, *SIAM Review*, **47**, 99–131 (2005).
[14] Barclay, A., P. E. Gill, and J. B. Rosen, in R. Bulirsch, et al, eds, *Variational Calculus, Optimal Control, and Applications*," vol. 124 of *International Series on Mathematics*, 207–222, Birkhäuser; Basel, Boston, and Berlin (1998).
[15] Voss, H. U., J. Timmer, and J. Kurths, *Int. J. Bif. Chaos*, **14**, 1905–1933 (2004).
[16] Kirk, D. E., "Optimal Control Theory: An Introduction," Prentice-Hall, Englewood Cliffs, New Jersey (1970); see section 5.2.
[17] Strang, G., *Introduction to Applied Mathematics*, p. 175, Wellesley-Cambridge Press, (1986).
[18] Kennedy, M. P. *IEEE Trans. Cir. Syst. I*, **41**, 771 (1994).
[19] Hargraves, C. R. and S. W. Paris, *J. of Guidance, Control, and Dynamics* **10**, 338–348, (1987).