Multi-parameter identification from scalar time series generated by a Malkus-Lorenz water wheel

Lucas Illing, Alison M. Saunders, and Daniel Hahs

Related Articles

Additional information on Chaos

Multi-parameter identification from scalar time series generated by a Malkus-Lorenz water wheel

Lucas Illing,[1,a)] Alison M. Saunders,[1] and Daniel Hahs[2]

[1]*Department of Physics, Reed College, Portland, Oregon 27708, USA*
[2]*Torch Technologies, Inc., Huntsville, Alabama 35802, USA*

We address the issue of multi-parameter estimation from scalar outputs of chaotic systems, using the dynamics of a Malkus water wheel and simulations of the corresponding Lorenz-equations model as an example. We discuss and compare two estimators: one is based on a globally convergent adaptive observer and the second is an extended Kalman filter (EKF). Both estimators can identify all three unknown parameters of the model. We find that the estimated parameter values are in agreement with those obtained from direct measurements on the experimental system. In addition, we explore the question of how to distinguish the impact of noise from those of model imperfections by investigating a model generalization and the use of uncertainty estimates provided by the extended Kalman filter. Although we are able to exclude asymmetric inflow as a possible unmodeled effect, our results indicate that the Lorenz-equations do not perfectly describe the water wheel dynamics. © *2012 American Institute of Physics*. [http://dx.doi.org/10.1063/1.3689441]

In this paper, we discuss the problem of estimating several unknown parameters on the example of angular velocity data from a Malkus water wheel. The Malkus water wheel is a mechanical experimental system whose dynamics is thought to be exactly described by the chaotic Lorenz equations. Although the measured time-series is proportional to one of the state variables of the model, the remaining state variables and the model parameters are unknown and need to be estimated. If one assumes that the Lorenz equations are the correct model, the main challenge is to guarantee that the estimator is stable and will converge to the correct parameter values. We utilize a novel globally convergent adaptive observer for the transformed Lorenz equations that can correctly identify all three unknown parameters. We also discuss how to judge whether the model does indeed perfectly describe the water wheel, as assumed.

I. INTRODUCTION

The dynamics of an ideal Malkus water wheel are described by the Lorenz equations.[1,2] Our wheel is close to an ideal one because all assumptions made in deriving the Lorenz equations as a proper model for the wheel's dynamics are met with good accuracy.[3] Nevertheless, the water wheel being a real system, small deviations from the ideal are present and noise affects both the measurements and system dynamics. Furthermore, the system parameters can only be determined approximately. In comparing the experimental output to the model, it would be desirable to know the parameters with higher precision. Even more ambitiously, one would like to obtain parameter values along with an uncertainty measure and an indication of the model fidelity,

whereby we mean an indication of how closely the Lorenz model matches the actual device dynamics.

The available data are the wheel's angle as a function of time. Taking a first derivative of the data, the wheel's angular velocity is determined, which is directly proportional to one of the variables of the Lorenz equations ($\tilde{x}_1$, see below). As is typical for many practical situations, no experimental output can be obtained that would directly correspond to the remaining state-variables of the model.[4,5] They have to be estimated along with the parameters, significantly increasing the difficulty of the estimation task.

Like the Lorenz equations, the water wheel dynamics are chaotic for large regions of parameter space. A challenging issue in the parameter estimation of chaotic systems is that, even in the absence of noise, the presence of highly unstable regions in state space may result in desynchronization, such that the dynamic variables of the actual system and the corresponding variables of the model diverge. Since in most methods the estimation of parameters is coupled to the estimation of states, such methods need to guarantee chaos synchronization.[6–8] Synchronization can be achieved by coupling the system output into the model. Thus, from the dynamics perspective, one deals with a situation where the system (the experimental device or plant) and the model (the observer or estimator) are a pair of unidirectionally coupled nonlinear systems. In such a setting, only the current (and potentially past) outputs are utilized, which has the advantage that these parameter estimators can be used on-line and as part of feedback-controllers.

In this paper, we discuss two strategies that fall within this approach to parameter estimation. One is based on an adaptive observer that utilizes the first and second derivative of the measured data. The purpose of adaptive observers is to guarantee global convergence of the state variables and parameters in the asymptotic limit. Their design is well

a)Electronic mail: illing@reed.edu.

established in the control theory literature for the case of noise-free linear time invariant systems[9] and has recently been extended to cover linear time varying (LTV) systems with single output as well as multiple outputs.[9–11] Adaptive observers for nonlinear systems are also known.[12] The guarantee of global convergence comes, in general, at the cost of assumptions that are difficult to check (e.g., the persistent excitation assumption) and the applicability of a particular adaptive observers design is limited to specific structures of the estimation problem. For example, although the adaptive observer proposed by Yu and Parlitz[12] deals with general nonlinear systems, when applied to the Lorenz system

$$
\begin{aligned}
\frac{d\tilde{x}_1}{ds} &= \sigma(\tilde{x}_2 - \tilde{x}_1), \\
\frac{d\tilde{x}_2}{ds} &= \rho\tilde{x}_1 - \tilde{x}_2 - \tilde{x}_1\tilde{x}_3, \\
\frac{d\tilde{x}_3}{ds} &= \tilde{x}_1\tilde{x}_2 - b\tilde{x}_3,
\end{aligned}
\tag{1}
$$

one finds that knowledge of the variable $\tilde{x}_1$ only allows one to estimate the σ parameter. To estimate ρ, one would have to know $\tilde{x}_2$. In order to determine ρ from knowledge of $\tilde{x}_1$, a very different observer design was proposed by Andrievskii *et al.*,[10] where the estimation problem for the Lorenz system with known $\tilde{x}_1$ is treated as an estimation problem for linear time-varying systems. In this paper, we show how this latter type of observer can be used to simultaneously determine ρ and σ (Sec. III).

As a second approach, we use an extended Kalman filter (EKF), which only requires the first derivative of the measured data.[13] One difference between Kalman filters and adaptive observers is that the fundamental consideration in designing Kalman filters is noise minimization not global convergence. In addition to providing an expectation value for the parameters, Kalman filters also estimate the corresponding statistical uncertainty. As discussed in Sec. IV, the uncertainty estimate provides an indicator of the model fidelity.

II. WATER WHEEL

A. Experiment

A well known mechanical implementation of the famous Lorenz equations is the Malkus water wheel.[1,2] We constructed a simple such system out of a bicycle wheel with attached syringes that is mounted to boards, such that it can be tilted at an angle α with respect to the horizontal, as shown in Fig. 1. At the wheel's top, the syringes are filled by a constant inflow of water described by the mass flux $Q(\theta)$ and water is leaking from all syringes at a rate proportional to the mass of water in each syringe. As a result of the water inflow and wheel tilt, gravity acts and drives rotational motion. Counteracting this is a magnetic brake that dampens rotations. Since the distribution of the water in the syringes determines the angular velocity and, conversely, the angular velocity determines how much water is added to the syringes, interesting motion of the wheel can arise. Our wheel is close to an ideal one because all assumptions made

in deriving the Lorenz equations as a proper model for the wheel's dynamics are met with good accuracy. For example, the syringes form a closed gap-less ring and no overflow takes place (for typical rotational motion), such that the total water mass reaches a constant and thus a constant moment of inertia is established after an initial transient. In addition, damping is almost entirely due to viscous friction by the magnetic brake and the leakage is well described by laminar pipe flow (Hagen-Poiseuille equation). As a result, we find steady state, periodic, and chaotic behavior of the wheel in agreement with predictions from the Lorenz model.[3]

B. The model

A model of the water wheel behavior can be obtained as is described by Strogatz[2] and Illing *et al.*[3] For an ideal water wheel, the description reduces to a simple three-dimensional system of ordinary-differential equations,

$$
\begin{aligned}
\frac{d\omega}{dt} &= -\frac{\kappa + Q_{\text{tot}}R^2}{I_{\text{tot}}}\omega(t) + \frac{\pi Rg\sin\alpha}{I_{\text{tot}}}a_1(t), \\
\frac{da_1}{dt} &= -ka_1(t) + \omega(t)b_1(t), \\
\frac{db_1}{dt} &= q_1 - kb_1(t) - \omega(t)a_1(t).
\end{aligned}
\tag{2}
$$

Here, $\omega(t)$ is the angular velocity and $a_1(t)$ and $b_1(t)$ are the amplitudes of the first Fourier-mode that is obtained through a series expansion of the function $m(\theta, t)$ describing the distribution of water around the wheel's rim. Alternatively, one may interpret $a_1(t)$ and $b_1(t)$ as the water wheel's center-of-mass coordinates in the plane of the water wheel.[14] The parameter I_{tot} is the total angular momentum of the wheel and water, κ characterizes the magnetic brake, Q_{tot} is the total mass flux of the water inflow, R the wheel's radius, g the gravitational constant, α the wheel's inclination angle, q_1 a coefficient related to the distribution $Q(\theta)$ of the water inflow, and k the leakage rate of the syringes. Although all parameters can be determined experimentally, some have considerable measurement uncertainties.

As is well known,[2] Eq. (2) can be mapped onto the Lorenz equations (1) with its third parameter equal one, $b = 1$. This mapping is achieved by introducing the dimensionless parameters

$$
\sigma = \frac{1}{k}\frac{\kappa + Q_{\text{tot}}R^2}{I_{\text{tot}}}, \quad \rho = \frac{q_1}{k^2}\frac{\pi Rg\sin\alpha}{\kappa + Q_{\text{tot}}R^2},
\tag{3}
$$

the dimensionless time $s = k\,t$, and the coordinates

$$
\tilde{x}_1 = \frac{\omega}{k}, \quad \tilde{x}_2 = \frac{\rho k}{q_1}a_1, \quad \tilde{x}_3 = \rho - \frac{\rho k}{q_1}b_1.
\tag{4}
$$

However, for our purposes, the above transformation is non-ideal. In the experiments, we measure the wheel's angle $\theta(t)$ as a function of time. Taking the first derivative of the data, one obtains the angular velocity $\omega(t)$. Taking a second derivative yields the angular acceleration, which, although already contaminated by noticeable amounts of noise is still a useful signal. In contrast, the third derivative is entirely
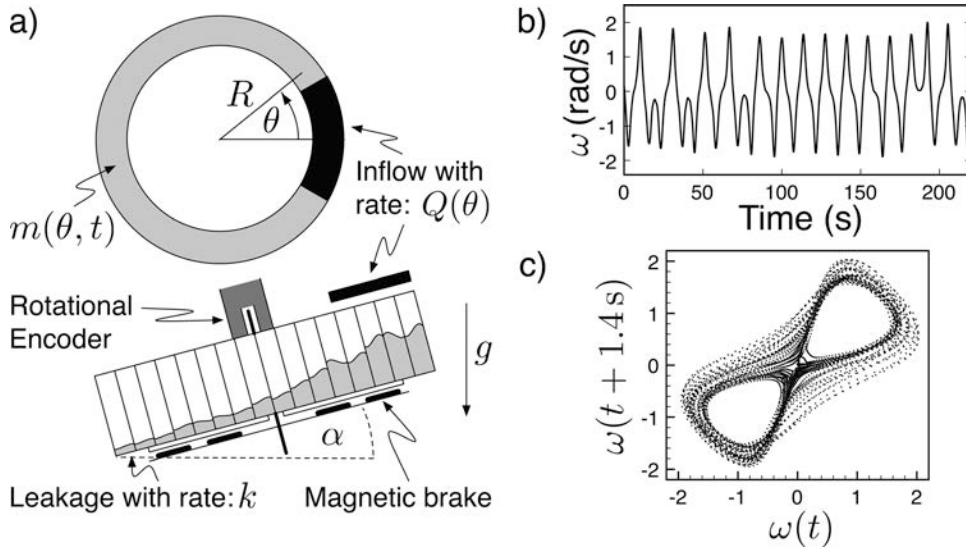
FIG. 1. (a) Implementation of the Malkus-Lorenz water wheel, constructed using a bicycle wheel, syringes, and rare earth magnets forming a magnetic brake. (b) Experimental chaotic time series and (c) delay embedding thereof.

dominated by noise and we do not use it for parameter estimation purposes. A more convenient transformation is, therefore, one that provides a model in terms of the measured angular velocity and acceleration and avoids the use of the uncertain parameter k in the variable transformation of observables. Thus, instead of Eq. (4), we use the coordinates

$$x_1 = \omega, \quad x_2 = \dot{\omega}, \quad x_3 = k^2 \rho \sigma \left(1 - \frac{k}{q_1} b_1 \right), \qquad (5)$$

to obtain an alternative form of the Lorenz equations

$$
\begin{aligned}
\dot{x}_1 &= x_2, \\
\dot{x}_2 &= k^2 \sigma(\rho - 1)x_1 - kx_2 - k\sigma x_2 - x_1 x_3, \\
\dot{x}_3 &= -kx_3 + x_1 x_2 + k\sigma x_1^2,
\end{aligned}
\qquad (6)
$$

where the overdot denotes the derivative with respect to the measured time t with units of seconds.

Ideally, one would like to be able to estimate all three uncertain parameters (k, σ, ρ) based on the measured time-series $x_1(t)$. This is a challenging problem and we will discuss for the rest of the paper how one can achieve such estimation and some trade-offs that are involved.

III. ADAPTIVE OBSERVER

A. General approach

The adaptive observer design approach that we use to estimate the parameters ρ and σ of the Lorenz system is applicable to nonlinear systems of the form

$$
\begin{aligned}
\dot{\mathbf{x}} &= \mathbf{A}(\mathbf{y})\mathbf{x} + \mathbf{F}(\mathbf{y})\mathbf{p}, \\
\mathbf{y} &= \mathbf{C}\mathbf{x},
\end{aligned}
\qquad (7)
$$

with variables $\mathbf{x} \in \mathbb{R}^n$, output $\mathbf{y} \in \mathbb{R}^\ell$, known output transfer matrix $\mathbf{C} \in \mathbb{R}^{\ell \times n}$, unknown constant parameters $\mathbf{p} \in \mathbb{R}^p$, matrix $\mathbf{A}(\mathbf{y}) : \mathbb{R}^\ell \to \mathbb{R}^{n \times n}$, and the nonlinear matrix function $\mathbf{F}(\mathbf{y})$ with components $\mathbf{F}(\mathbf{y}) = (\mathbf{f}_1, \mathbf{f}_2, ..., \mathbf{f}_p) \in \mathbb{R}^{n \times p}$, where $\mathbf{f}_j$ are nonlinear vector functions $\mathbf{f}_j(\mathbf{y}) : \mathbb{R}^\ell \to \mathbb{R}^n$ of the output $\mathbf{y}$. Although the dependence of $\mathbf{A}(\mathbf{y})$ and $\mathbf{F}(\mathbf{y})$ on $\mathbf{y}$ may

be nonlinear, $\mathbf{A}$ and $\mathbf{F}$ can be viewed as time dependent matrices because the signal $\mathbf{y}(t)$ is known. Therefore, Eq. (7) belongs to the class of LTV systems. It remains an LTV system even if the parameters $\mathbf{p}$ are appended as states by formally adding the equation $\dot{\mathbf{p}} = 0$. We note that Eq. (7) is not the most general form, as $\mathbf{C}$ does not have to be constant but may depend on time and one may add terms representing known functions of the output or functions of known external signals. Since these generalizations do not add any further difficulties to the observer design and because they are not relevant for the Lorenz system, we do not consider them.

A possible observer system is[9–11]

$$
\begin{aligned}
\dot{\mathbf{x}}_o &= \mathbf{A}(\mathbf{y})\mathbf{x}o + \mathbf{F}(\mathbf{y})\mathbf{q} - \mathbf{K}(\mathbf{y})(\mathbf{y}_o - \mathbf{y}) \\
&\quad - \mathbf{Z}\Gamma[\mathbf{C}\mathbf{Z}]^T(\mathbf{y}_o - \mathbf{y}), \\
\mathbf{y}_o &= \mathbf{C}\mathbf{x}_o, \\
\dot{\mathbf{Z}} &= [\mathbf{A}(\mathbf{y}) - \mathbf{K}(\mathbf{y})\mathbf{C}]\mathbf{Z} + \mathbf{F}(\mathbf{y}), \\
\dot{\mathbf{q}} &= -\Gamma[\mathbf{C}\mathbf{Z}]^T(\mathbf{y}_o - \mathbf{y}),
\end{aligned}
\qquad (8)
$$

with observer variables $\mathbf{x}_o \in \mathbb{R}^n$, estimated parameter vector $\mathbf{q} \in \mathbb{R}^p$, auxiliary variable vectors $\mathbf{z}_j \in \mathbb{R}^n$ $(j = 1, ..., p)$ forming the matrix $\mathbf{Z} = (\mathbf{z}_1, \mathbf{z}_2, ..., \mathbf{z}_p) \in \mathbb{R}^{n \times p}$, fixed gain matrix $\Gamma = \text{diag}(\gamma_1, \gamma_2, ..., \gamma_p) \in \mathbb{R}^{p \times p}$, and coupling matrix $\mathbf{K}(\mathbf{y}) \in \mathbb{R}^{n \times \ell}$.

To understand the structure of this observer, consider the equation describing the dynamics of the state-variable error, $\mathbf{e} = \mathbf{x}_o - \mathbf{x}$, and parameter error $\mathbf{r} = \mathbf{q} - \mathbf{p}$. Noting that $\dot{\mathbf{p}} = 0$ and introducing the effective matrix $\mathbf{G}(\mathbf{y}) \in \mathbb{R}^{n \times n}$ via $\mathbf{G}(\mathbf{y}) = \mathbf{A}(\mathbf{y}) - \mathbf{K}(\mathbf{y})\mathbf{C}$, we obtain

$$
\begin{aligned}
\dot{\mathbf{e}} &= \mathbf{G}(\mathbf{y})\mathbf{e} + \mathbf{F}(\mathbf{y})\mathbf{r} + \mathbf{Z}\dot{\mathbf{r}}, \\
\varepsilon &= \mathbf{C}\mathbf{e}, \\
\dot{\mathbf{Z}} &= \mathbf{G}(\mathbf{y})\mathbf{Z} + \mathbf{F}(\mathbf{y}), \\
\dot{\mathbf{r}} &= -\Gamma[\mathbf{C}\mathbf{Z}]^T\varepsilon.
\end{aligned}
\qquad (9)
$$

The crucial point of the observer structure in Eq. (8) is that the difference vector $\delta = \mathbf{e} - \mathbf{Z}\mathbf{r}$ evolves according to

$$\dot{\delta} = \mathbf{G}(\mathbf{y})\delta, \qquad (10)$$

as can be seen from Eq. (9). Equation (10) is entirely independent of the unknown parameter error $\mathbf{r}$. It results in an exponentially decaying solution $\delta(t)$, if an appropriate matrix $\mathbf{G}(\mathbf{y})$ can be constructed through the right choice of the coupling matrix $\mathbf{K}(\mathbf{y})$. Assuming exponential stability, one can then represent the error as $\mathbf{e} = \mathbf{Z}\mathbf{r}$, up to an exponentially decaying term, and therefore, $\varepsilon = \mathbf{C}\mathbf{Z}\mathbf{r}$. With this, the evolution of the parameter error is seen to be governed by the linear system

$$\dot{\mathbf{r}} = -\Gamma[\mathbf{C}\mathbf{Z}]^T \mathbf{C}\mathbf{Z}\mathbf{r} = -\mathbf{Q}(t)\mathbf{r}, \qquad (11)$$

where $\mathbf{Q}(t) = \Gamma[\mathbf{C}\mathbf{Z}]^T \mathbf{C}\mathbf{Z}$ is a time-varying positive-semidefinite matrix. It can be shown that the parameter error $\mathbf{r}(t)$ tends to zero exponentially fast if the signal $\mathbf{C}\mathbf{Z}(t)$ satisfies certain conditions (to be discussed below) in which case the signal is said to be persistently exciting. This is the essence of the observer and leads to the following statement about convergence:[11]

Theorem 1. *Assume that the state vector $\mathbf{x}(t)$ of the drive system is a bounded function of time for any bounded initial condition $\mathbf{x}(0)$ and any admissible constant parameter vector $\mathbf{p}$ and that the matrices $\mathbf{F}(\mathbf{y})$ and $\mathbf{A}(\mathbf{y})$ are piecewise continuous in time and bounded for any bounded $\mathbf{y}$. Let $\Gamma \in \mathbb{R}^{p \times p}$ in Eq. (8) be any symmetric positive-definite matrix.*

(1) If, for the matrix pair $(\mathbf{A}, \mathbf{C})$ in Eq. (7), there exists a bounded matrix $\mathbf{K}(\mathbf{y}) \in \mathbb{R}^{n \times n}$ such that

$$\dot{\delta} = [\mathbf{A} - \mathbf{K}\mathbf{C}]\delta \qquad (12)$$

is exponentially stable, then the solutions $\mathbf{y}_o = \mathbf{C}\mathbf{x}_o$ of Eq. (8) converge to the output of Eq. (7), $\mathbf{y}_o(t) \rightarrow \mathbf{y}(t)$, as $t \rightarrow \infty$.

(2) If, in addition, the $\mathbb{R}^{n \times n}$ matrix $\mathbf{Z}(t)$ that is generated by the ordinary differential equation (ODE)

$$\dot{\mathbf{Z}} = [\mathbf{A} - \mathbf{K}\mathbf{C}]\mathbf{Z} + \mathbf{F}(\mathbf{y}) \qquad (13)$$

is persistently exciting, which means that it has the property that there exist positive constants t_0, T_0, and ϵ such that

$$\lambda_i \left[\int_t^{t+T_0} \mathbf{Z}(\tau)^T \mathbf{C}^T \mathbf{C}\mathbf{Z}(\tau) d\tau \right] \geq \epsilon \quad \forall t > t_0, \qquad (14)$$

for $i = 1, \dots, n$, where $\lambda_i[\mathbf{M}]$ denotes the ith eigenvalue of $\mathbf{M}$, then Eq. (8) is a global exponential adaptive observer for system (7), i.e., $\mathbf{x}_o(t) \rightarrow \mathbf{x}(t)$ and $\mathbf{q}(t) \rightarrow \mathbf{p}$ exponentially fast as $t \rightarrow \infty$ for any initial conditions.

Above theorem highlights the importance of the persistent excitation condition, Eq. (14). Convergence of the measured output $\mathbf{y}$ and model output $\mathbf{y}_o$ does not guarantee that the parameters converge nor that the unobserved state variables converge. Unfortunately, the persistent excitation condition is often difficult to check a priori.

The meaning of the persistent excitation condition can be understood in several ways. Recognizing that the observable error $\varepsilon = \mathbf{y}_o - \mathbf{y}$ converges to zero if condition (1) in Theorem 1 is satisfied, one obtains the equation

$$0 = \mathbf{C}\mathbf{Z}\mathbf{r} = \sum_{i=1}^p r_i(t)\mathbf{C}\mathbf{z}_i(t). \qquad (15)$$

In this context, the persistent excitation condition is seen to be a statement of linear independence of the vector functions $\mathbf{C}\mathbf{z}_i(t)$.[12,15] However, since it is difficult to directly use the definition of linear independence to judge the linear correlation of a group of vector functions, this interpretation is of little help in proving persistent excitation. In the context of the LTV system (11), governing the evolution of the parameter error $\mathbf{r}(t)$, persistent excitation roughly says that the signal $\mathbf{C}\mathbf{Z}(t)$ should persist to guarantee that $\mathbf{r}(t)$ goes to zero asymptotically and that it should be complex enough, with the required minimum amount of complexity determined by the dimension of $\mathbf{r}$. In cases where $\mathbf{Q}$ in Eq. (11) is given in terms of a vector $\mathbf{u}(t)$ by $\mathbf{Q}(t) = \mathbf{u}(t)\mathbf{u}^T(t)$, it is sufficient that $\mathbf{u}(t)$ has $\lceil n/2 \rceil$ distinct frequencies in order to assure that $\mathbf{r} \in \mathbb{R}^n$ converges to zero.[9,16] Here, $\lceil x \rceil$ denotes the largest integer greater or equal to x. Thus, one expects a typical chaotic signal to be persistently exciting because it has a strictly positive power spectrum[17,18] and a simple sinusoidal signal to be sufficient to guarantee the convergence to zero of $\mathbf{r} \in \mathbb{R}^2$.[19]

Although we expect chaotic and periodic output of the water wheel to result in a persistently exciting signal $\mathbf{C}\mathbf{Z}(t)$, it is not trivial to connect the properties of the measured time-series $x_1(t)$ to those of the matrix function $\mathbf{C}\mathbf{Z}(t)$. Since $\mathbf{Z}(t)$ is explicitly computed as part of the observer system, we find it simpler to check numerically that the persistent excitation condition (14) is satisfied.

B. Lorenz system

In order to write the transformed Lorenz system (6) as a linear time varying system of form (7) with unknown parameters σ and ρ, we need to assume that the variables $x_1 = d\theta/dt$ and $x_2 = d^2\theta/dt^2$ are available outputs of the experiments. As mentioned above, our data $\theta(t)$ are of sufficient quality that this is possible. In contrast, the determination of all three parameters σ, ρ, and k would require the use of $d^3\theta/dt^3$, which is not feasible. We, therefore, presume, for now, that k is known.

Introducing the parameter vector $\mathbf{p} = (p_1, p_2)^T$ with

$$p_1 = k\sigma \quad \text{and} \quad p_2 = k^2\sigma(\rho - 1), \qquad (16)$$

matrices

$$\mathbf{A}(\mathbf{y}) = \begin{pmatrix} 0 & 1 & 0 \\ 0 & -k & -y_1 \\ 0 & y_1 & -k \end{pmatrix}, \quad \mathbf{C} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

and

$$\mathbf{F}(\mathbf{y}) = \begin{pmatrix} 0 & 0 \\ -y_2 & y_1 \\ y_1^2 & 0 \end{pmatrix},$$

one can compactly write the transformed Lorenz system (6) as

$$\dot{\mathbf{x}} = \mathbf{A}(\mathbf{y})\mathbf{x} + \mathbf{F}(\mathbf{y})\mathbf{p},$$
$$\mathbf{y} = \mathbf{C}\mathbf{x}. \tag{17}$$

It is seen to be in the form required by Theorem 1. The first condition of the theorem can be satisfied by choosing the coupling matrix $\mathbf{K}$ and corresponding matrix $\mathbf{G} = \mathbf{A} - \mathbf{KC}$ as

$$\mathbf{K} = \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 0 \end{pmatrix} \quad \text{and} \quad \mathbf{G}(\mathbf{y}) = \begin{pmatrix} -1 & 1 & 0 \\ -1 & -k & -y_1 \\ 0 & y_1 & -k \end{pmatrix}, \tag{18}$$

respectively, and by making use of a the following theorem about necessary and sufficient conditions for exponential stability.

Theorem 2. *(Krasovskii[20,21]) The trivial solution* $\mathbf{x}(t) \equiv 0$ *of the system* $\dot{\mathbf{x}} = \mathbf{F}(\mathbf{x}, t)$ *with a continuously differentiable right-hand side is globally exponentially stable if and only if there exists a function* $V : \mathbb{R}^n \times [t_0, \infty) \to \mathbb{R}_+$ *and positive numbers* c_1, c_2, c_3, *and* c_4, *such that*

$$c_1 |\mathbf{x}|^2 \le V(\mathbf{x}, t) \le c_2 |\mathbf{x}|^2,$$
$$|\nabla_{\mathbf{x}} V(\mathbf{x}, t)| \le c_3 |\mathbf{x}|,$$
$$\frac{dV}{dt}(\mathbf{x}, t) \le -c_4 |\mathbf{x}|^2.$$

The desired result follows because the time derivative of the Lyapunov function $V(\delta) = \sum_{i=1}^{3} \delta_i^2 / 2$ evaluated along solutions of Eq. (10) with $\mathbf{G}$ as defined in Eq. (18) yields

$$\dot{V}(\delta) = -\delta_1^2 - k\delta_2^2 - k\delta_3^2 \le -\min(k, 1)|\delta|^2.$$

The second condition of Theorem 1 is checked numerically on-line while the observer is applied to the signal. To do so, we compute, along with observer (8) and for all times t larger than an initial transient time, the integral in Eq. (14) for some arbitrarily chosen integration interval T_0 ($T_0 = 10$ s) and check that the minimal eigenvalue of the resulting matrix is positive definite.

An example of the observer's convergence to the correct values is shown in Figure 2, where simulated chaotic data were generated through integration of the Lorenz system and then used as input to the observer. In this case, the input data were free of noise and the correct value of k was known. Both conditions are untrue for the actual data.

Due to the quality of our data, the first issue, noise, is not a fundamental problem. Although adaptive observers are not specifically optimized for noise, it is known that the observer (8) converges in the mean for input data corrupted by observational noise, if the noises are bounded and have zero mean.[11] In other words, despite the fact that neither the variable error $\mathbf{e}(t)$ nor the parameter error $\mathbf{r}(t)$ will converge to zero, their time average will. Therefore, averaging the parameter vector $\mathbf{q}(t)$ of the adaptive observer after a transient allows one to estimate the constant parameters in the presence of noise.
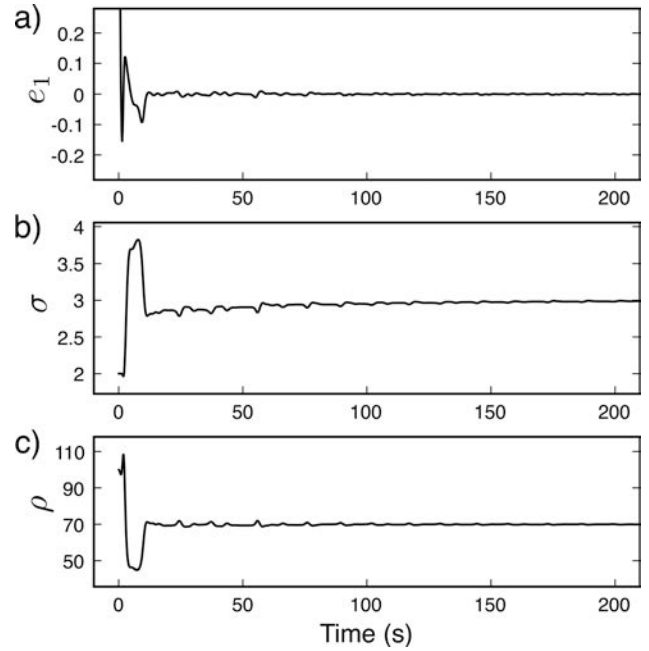

FIG. 2. Convergence of the observer for numerically generated noise-free data. (a) Difference of data and observer output. (b) Observer estimate of σ ($\sigma_{\text{true}} = 3$) and (c) ρ ($\rho_{\text{true}} = 70$).

The second issue is that the water wheel's leakage rate k is not known precisely. We do however know its order of magnitude from independent measurements that show k to be approximately 0.1 s^{-1}. The correct value of k can, therefore, be determined by running the observer for a range of k around this estimated value. Assuming that the water wheel dynamics is described by the Lorenz equations, the estimation error $e_1 = x_{o1} - x_1$, in the noise-free case, will converge to zero, if the value of k in the observer system is identical to the water wheel's true value k_{true}, but will not converge for incorrect k. This is demonstrated in Fig. 3(a) where noise-free simulated data were used as input to the observer. It is seen that the variance of the error e_1 appears to depend linearly on the difference $|k - k_{\text{true}}|$. In the presence of noise, both k-mismatch and noise contribute to the asymptotic oscillations of e_1. In this case, one expects the variance of the error e_1 to attain a nonzero minimum when the correct k is used.

To test this strategy for determining all three unknown parameters, we used Eq. (6) to simulate data to which white Gaussian observation noise was added. The results of applying the adaptive observer to this data are shown in Figs. 3(b)–3d. It is seen that all three parameters are successfully determined with high accuracy.

The application of the observer and above strategy to the water wheel data results in the estimates shown in Fig. 4. We find that there is a clear minimum for the variance of e_1 and that the magnitude of the variance is comparable to those of the simulations with noise. In addition, the dependence of σ and ρ on k is similar to that in the simulations, as seen by comparing Figs. 4 and 3. The observer behaves as expected, indicating that the obtained parameter estimates ($k = 0.123$, $\sigma = 2.5$, and $\rho = 66$) are reasonable. Furthermore, as shown in Table I, the parameter values that we obtain by using the
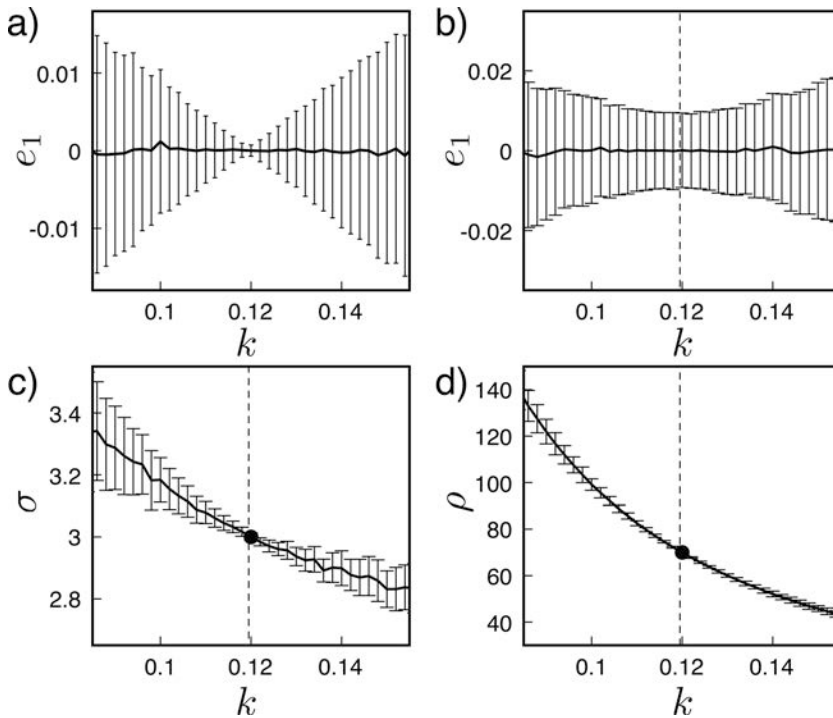
FIG. 3. Determination of all three parameters from numerically generated data ($k_{true} = 0.12$, $\sigma_{true} = 3$, and $\rho_{true} = 70$). Plotted are the mean values and standard deviation (errorbars). (a) No noise. The variance of the error e_1 converges to zero for the correct k (b) With noise. The minimum variance of e_1 is estimated to be at $k = 0.1195$ (dashed line). Parameters estimates: (c) $\sigma(k)$ with $\sigma(k = 0.1195) = 3.0$ and (d) $\rho(k)$ with $\rho(k = 0.1195) = 70.5$. The true values are shown by circles.

adaptive observer are consistent with the their estimated values obtained from direct measurements performed on the water-wheel device[3] and they are close to the estimates obtained using an extended Kalman filter (to be discussed below).

The relative smoothness of the variance values as k is scanned using the measured data [Figs. 4(c) and 4(d)] as compared to the numerically generated data [Figs. 3(c) and 3(d)] suggests that the estimates using the real data are not as sensitive to measurement noise. We believe that the variance of the latter estimates is dominated by the presence of unmodeled dynamics, as discussed below.

C. Generalized Lorenz model

The adaptive observer converges exponentially, if the system generating the data is exactly modeled by the Lorenz system and if there is no observational noise in the data nor dynamical noise in the system generating the output. Any actual system and measurement will contain noise. As a consequence, the observer error will never converge to zero. One is, therefore, confronted with the question of whether the remaining oscillations of the error and parameters [see Fig. 4(a)] are explained by observational noise, in which case taking the mean will provide a correct parameter esti-

TABLE I. Comparison of the parameter estimates obtained from measurements of the wheel parameters, the adaptive observer, and the extended Kalman filter.

Parameter	Experiment	Adaptive observer	EKF (averaged)
k	0.1	0.123	0.116
σ	2.7	2.5	2.65
ρ	69	66	78

mate, or whether these oscillations are caused by model insufficiencies.

One option to address this issue is to test more general models.[22] For the water wheel, many potential sources of non-ideal behavior are possible, including asymmetric inflow, kinetic friction, and slow parameter drifts such as those of the total mass flux Q_{tot}. As the model complexity increases considerably when one accounts for such effects, it is challenging to pin-point which of these effects, if any, is most relevant. Here, we consider just one possibility.

One of the simplest physically motivated extensions of the Lorenz model for the water wheel is one that takes into account effects due to an asymmetric inflow. In that case, one assumes that the mass flux of the added water is described by

$$Q(\theta) = \begin{cases} \dfrac{Q_{tot}}{2\delta\phi} & \theta \in [\phi - \delta\phi, \phi + \delta\phi], \\ 0 & \text{otherwise} \end{cases} \qquad (19)$$

where the parameter ϕ quantifies the asymmetry and $\delta\phi$ is a fixed constant ($\delta\phi = 26°$ in our experiments). The perfectly symmetric case, $\phi = 0$, results in the Lorenz model (6) that we have discussed so far. If one allows ϕ to be nonzero, the model changes to[3,23]

$$\begin{aligned} \dot{x}_1 &= x_2, \\ \dot{x}_2 &= p_2 x_1 - k x_2 - p_1 x_2 - x_1 x_3 + p_3 + C k p_3, \\ \dot{x}_3 &= -k x_3 + x_1 x_2 + p_1 x_1^2 - C p_3 x_1, \end{aligned} \qquad (20)$$

where the parameters p_1 and p_2 are given by Eq. (16) and the constant C,

$$C = \frac{Q_{tot}}{g\rho_w wR 2\delta\phi} \approx 0.002 \text{s}, \qquad (21)$$
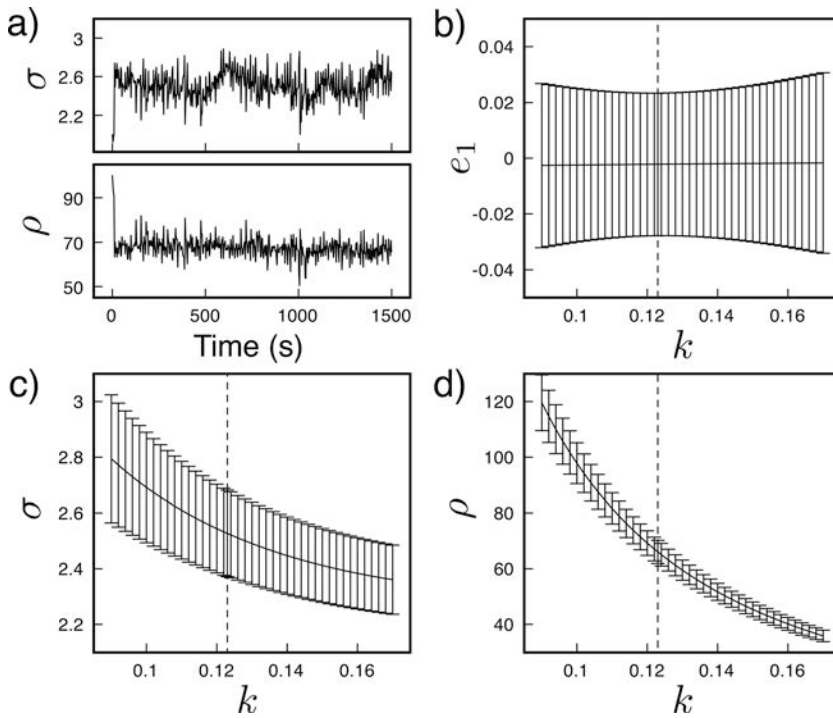
FIG. 4. Parameter estimation from water wheel data. (a) Evolution of observer's estimate for $k = 0.123$. (b) The standard deviation of e_1 is minimum at $k = 0.123$ (dashed line). Parameter estimates: (c) $\sigma(k = 0.123) = 2.5$ and (d) $\rho(k = 0.123) = 66$.

is defined in terms of the previously introduced constants Q_{tot}, g, R, and $\delta\phi$, together with the density of water ρ_w and width w of the syringes holding the water. The degree of asymmetry of the inflow is captured by a new fourth parameter

$$p_3 = k^3 \rho\sigma \tan\phi. \tag{22}$$

This model is still in the form of Eq. (7) and, therefore, is amenable to parameter estimation by the adaptive observer. Performing such a three parameter estimate for model Eq. (20) for various values of k, we find that the optimal k is reduced from $k = 0.123$ to $k = 0.11$ (see Table II) and that the estimates of the parameters ρ and σ for this k also change when compared to the estimate for the symmetric case. The estimated asymmetry angle ϕ is small with a mean value of $1.3°$ and a standard deviation of $10°$, i.e., its value is consistent with zero. Although the average amplitude of fluctuation of the error e_1 around its mean is reduced for the three parameter estimate of the asymmetric model (20) as compared to the two parameter estimate of the symmetric Lorenz model (6), this reduction from 0.025 to 0.020 is due to the fact that the observer has additional degrees of freedom and

not because the angle is truly $1.3°$. This is seen when fixing p_3 in model (20) by setting the asymmetry angle ϕ to its estimated value of $1.3°$. In this case, the observer can vary two parameters only and we find that the optimal k, the standard deviation of the error e_1, and parameters ρ and σ are identical to those of the two parameter estimate for the symmetric case (see Table II). We conclude that asymmetry of the inflow is not a relevant issue in our experimental system.

D. Summarizing the adaptive observer estimation

We have discussed an adaptive-observer based strategy that allows the determination of all three unknown parameters in the Lorenz model (6). The estimated parameter values are consistent with the values obtained from direct measurements. A drawback of the adaptive observer estimation procedure is the necessity to utilize a one-dimensional search over the k-parameter as well as higher-order derivatives of the data, because one needs to know x_2 in Eq. (6).

An *ad-hoc* solution to this problem is to construct an adaptive observer for all three parameters (σ, ρ, and k), which would require input-data x_1, x_2, and x_3, and to replace, subsequently, in the observer all x_2 and x_3 by the observer variables x_{o2} and x_{o3}. The thereby obtained set of differential equations is found to converge to the correct values of the parameters when driven by numerically generated noise-free chaotic data for x_1. Yet, although we do observe correct parameter estimates for the randomly chosen initial conditions of our numerical test, the main advantage of the adaptive observer approach is lost, we can no longer prove global convergences. In this case, other approaches might offer greater advantages. One such approach is Kalman filtering.

IV. EXTENDED KALMAN FILTER

For linear systems and Gaussian noise, Kalman filters are optimal in the sense that they minimize the mean square

TABLE II. Comparison of the parameter estimates obtained by considering the Lorenz model, Eq. (6), and a generalized model that allows asymmetric inflow, Eq. (20). For the parameters ϕ, σ, and ρ, the quoted mean and ± 1 standard deviation values are obtained for data cycles after filter convergence.

Parameter	Eq. (6)	Eq. (20)	Eq. (20)
k	0.123	0.11	0.123
ϕ	—	$1.3° + 10°$	$1.3°$ (fixed)
$\text{std}(e_1)$	0.025	0.020	0.025
σ	2.5 ± 0.2	2.7 ± 0.1	2.5 ± 0.2
ρ	66 ± 4	77 ± 4	66 ± 4

error. It is not clear that they are noise-optimal for nonlinear systems, nevertheless, these filters have advantages. For example, Kalman filters inherently take noise into account and consequently provide estimates with confidence intervals. In addition, there are fewer restrictions on the model structure than is the case for adaptive observers, e.g., a filter can be constructed for any ODE-based model. Such advantages have led to the widespread use of Kalman filters, such as in atmospheric climate models where Ensemble Kalman Filters[24] are being applied.[25,26]

A. Implementation

For the waterwheel, which is a simple analogue of an atmospheric convection cell, an EKF was implemented to estimate the three kinematic states and three parameters. The three parameters are obtained via estimates of states in an augmented state space, where the three additional degrees of freedom are

$$x_4 = k \quad x_5 = k\sigma \quad x_6 = k^2\sigma(\rho - 1) \qquad (23)$$

yielding the following equations of motion corresponding to Eq. (6),

$$
\begin{aligned}
\dot{x}_1 &= x_2, \\
\dot{x}_2 &= x_6 x_1 - x_4 x_2 - x_5 x_2 - x_1 x_3, \\
\dot{x}_3 &= -x_4 x_3 + x_1 x_2 + x_5 x_1^2, \\
\dot{x}_4 &= \dot{x}_5 = \dot{x}_6 = 0.
\end{aligned}
\qquad (24)
$$

The filter processes discrete-time measurements $y_k = y(t_k) = x_1(t_k) + \nu(t_k)$, where $\nu_k = \nu(t_k)$ is assumed to be a zero-mean white Gaussian noise process modeling the measurement noise with variance R. The implemented EKF was a continuous-discrete filter as, for example, summarized in Gelb.[27] To compactly express the EKF algorithm, let the system and measurement model be given by

$$
\begin{aligned}
\dot{\mathbf{x}}(t) &= \mathbf{f}[\mathbf{x}(t)] + \mathbf{w}(t), & \mathbf{w}(t) &\sim N[0, \mathbf{Q}(t)], \\
y_k &= \mathbf{H}\mathbf{x}_k + \nu_k, & \nu_k &\sim N[0, R_k],
\end{aligned}
\qquad (25)
$$

where $\mathbf{f}$ is the RHS of Eq. (24), $\mathbf{H} = [100000]$ (since only state vector component x_1 is measured), and $\mathbf{Q}$ models zero-mean white Gaussian process-noise. It is further assumed that the initial condition is distributed normally and that the process and measurement noise terms are uncorrelated:

$$
\begin{aligned}
\mathbf{x}(0) &\sim N[\widehat{\mathbf{x}}_0, \mathbf{P}_0], \\
E[\mathbf{w}(t)\nu_k^T] &= 0 \qquad \forall k, t.
\end{aligned}
\qquad (26)
$$

The EKF utilizes a Bayesian recursive predictor/corrector structure to compute an approximation of the conditional mean $\widehat{\mathbf{x}}(t_k) \simeq E[\mathbf{x}(t_k)|y_1, y_2, \ldots, y_k]$.

At the beginning of each data cycle, the previous $(t = t_{k-1})$ state vector and error covariance matrix estimates are projected ahead to yield estimates valid at the current data cycle $(t = t_k)$. These estimates of the prediction step are denoted as $\widehat{\mathbf{x}}_k^{(-)}$ and $\mathbf{P}_k^{(-)}$, respectively, where the $(-)$ indicates that the current measurement, y_k, has not yet been

used. Subsequently, a correction step is performed that results in the estimates $\widehat{\mathbf{x}}_k^{(+)}$ and $\mathbf{P}_k^{(+)}$, where the $(+)$ indicates that the measurement from the kth data cycle, y_k, has been utilized in the computation. The resulting estimates are then used to initialize the next data cycle.

To perform the prediction step, the following differential equations are integrated from $t = t_{k-1}$ to $t = t_k$ using initial conditions $\widehat{\mathbf{x}}(t_{k-1}) = \widehat{\mathbf{x}}_{k-1}^{(+)}$ and $\mathbf{P}(t_{k-1}) = \mathbf{P}_{k-1}^{(+)}$, respectively,

$$
\begin{aligned}
\dot{\widehat{\mathbf{x}}}(t) &= \mathbf{f}[\widehat{\mathbf{x}}(t)], \\
\dot{\mathbf{P}}(t) &= \mathbf{Df}[\widehat{\mathbf{x}}(t)]\mathbf{P}(t) + \mathbf{P}(t)\mathbf{Df}^T[\widehat{\mathbf{x}}(t)] + \mathbf{Q}(t).
\end{aligned}
\qquad (27)
$$

The outputs of the integration are then identified as $\widehat{\mathbf{x}}_k^{(-)} = \widehat{\mathbf{x}}(t_k)$ and $\mathbf{P}_k^{(-)} = \mathbf{P}(t_k)$. In Eq. (27), the differential equation for the covariance matrix is based on a linearization of the covariance rate equation around the predicted state, that is

$$\mathbf{Df}_{ij}[\widehat{\mathbf{x}}(t)] = \left.\frac{\partial f_i[\mathbf{x}]}{\partial x_j}\right|_{\mathbf{x}=\widehat{\mathbf{x}}(t)}. \qquad (28)$$

This linearization is the main approximation of the EKF method. It implies that the noise distribution remains Gaussian; a basic assumption of the Kalman filter.

In the correction step, the Kalman gain $\mathbf{K}_k$ is used for updating state and covariance estimates.

The Kalman gain is a vector that determines how much weight is given to the model prediction versus the new data. Its value at each iteration is set according to the predicted estimate of the covariance matrix.

$$
\begin{aligned}
\mathbf{K}_k &= \mathbf{P}_k^{(-)}\mathbf{H}^T\left[\mathbf{H}\mathbf{P}_k^{(-)}\mathbf{H}^T + R_k\right]^{-1}, \\
\widehat{\mathbf{x}}_k^{(+)} &= \widehat{\mathbf{x}}_k^{(-)} + \mathbf{K}_k\left[y_k - \mathbf{H}\widehat{\mathbf{x}}_k^{(-)}\right], \\
\mathbf{P}_k^{(+)} &= [\mathbf{1} - \mathbf{K}_k\mathbf{H}]\mathbf{P}_{k-1}^{(+)}.
\end{aligned}
\qquad (29)
$$

In order to obtain correct estimates from the above EKF, the initial states and initial covariance matrix have to be given values that are reasonably close to the true ones. In addition, their process and measurement noise variances have to be specified. Their choice provides tuning parameters for the filter.

For the EKF of the waterwheel, the first three filter states were initialized using estimates for x_1, x_2 and x_3 computed from the first three measurements y_1, y_2, and y_3 by simple differencing and substitution into the dynamical equations. The initial parameter estimates (filter states four, five, and six) were deduced from physical measurements of the waterwheel components. The initial covariance matrix was computed by performing a Monte Carlo experiment in which the state initialization process was repeated for 10 000 trials using the same, fixed, initialization, and simulated measurements. At each trial the fixed, "true" initialization was set equal to the initial first three states computed from the real measurement data. After completion of the last trial, the error covariance of the set of 10 000 initial estimates was computed and subsequently employed in the filter as the initial estimate $\mathbf{P}_0$.

The measurement noise was chosen as $R = (0.0016)^2$, reflecting the digitization noise of the rotational encoder used to measure the angle. In terms of the process-noise, two filter tunings were evaluated to gain insight into the estimation problem:

$$\mathbf{Q}_1 = \mathrm{diag}(10^{-8}\,10^{-8}\,10^{-8}\,10^{-8}\,10^{-8}\,10^{-8}) \qquad (30)$$

and

$$\mathbf{Q}_2 = \mathrm{diag}(10^{-8}\,10^{-8}\,10^{-8}\,10^{-3}\,10^{-3}\,10^{-3}). \qquad (31)$$

B. Results

Approximatelym 1500 s of data were processed at a sampling interval of 0.101 s. Using $\mathbf{Q} = \mathbf{Q}_1$, the time history of the estimates for the parameter $x_4 = k$ are shown in Fig. 5(a). The final estimates for the waterwheel parameters computed from the final estimates of the x_i values ($i = 4, 5, 6$) were $k = 0.100$, $\sigma = 2.61$, and $\rho = 105$. The mean estimates of the parameters using the x_i estimates obtained from 500 s to the end of the epoch yielded $k = 0.116$, $\sigma = 2.65$, and $\rho = 77.9$. From this and from Fig. 5(a), it is seen that the filter does not converge to steady constant values. When the x_4 estimates are plotted along with the standard deviation error bars obtained from the covariance matrix at each data cycle [see Fig. 5(b)], it is found that the error bars do not extend across the variation of the parameter estimates observed after filter convergence.

As a potential remedy, it was thought to increase the value of the $\mathbf{Q}$ components corresponding to the parameter states (elements 4, 5, and 6 on the main diagonal) to indicate greater uncertainty in the stationarity of the parameters ($\mathbf{Q} = \mathbf{Q}_2$). When this tuning was employed, the effect of the increase in filter bandwidth was to allow greater variation in the parameter estimates during the estimation epoch [see Fig. 5(c)]. It was also seen that the error bars still did not encompass the variability of the estimates [see Fig. 5(d)]. These results suggest that the parameters being estimated are, in fact, not constants, or equivalently, that the model fidelity is not sufficient to obtain statistically significant estimates of its parameters.

To test the idea that the filter's performance is due to model insufficiencies, a simulated measurement record was generated using the identical model employed in the filter and three assumed parameter values. When this record was processed, the filter converged promptly and obtained nearly constant estimates with consistent error bars [see Figs. 5(e) and 5(f)].

V. DISCUSSION

The Malkus waterwheel is a physical system that exhibits chaotic behavior qualitatively modeled by the Lorenz equations. In this work, we have demonstrated how two disparate estimation techniques—an adaptive observer and an EKF—can yield waterwheel state and parameter estimates that are in good agreement. What also emerges from our results is some insight into how model sufficiency may be assessed using these two techniques. To this end, the adaptive observer and the EKF have the following relative merits/demerits: (1) The adaptive observer provides for provable global convergence based on the assumed model equations being correct. Hence, a lack of convergence in practice indicates the model is (strictly) incorrect. However, the binary quality of the information thus obtained limits its utility for quantifying the degree of insufficiency. (2) The EKF provides state estimate error variance estimates that can indicate model insufficiency in a more quantitative way. However, it lacks a guaranteed region of convergence.

The EKF's finite region of convergence is not usually problematic when the experimental device (plant) is well characterized and the measurement noise is "small." The filter relies on linearization and thus for highly oversampled data, as in the case at hand, one expects the linearization to be a reasonable approximation and the computed variance estimates to be trustworthy. This expectation was verified by simulating waterwheel measurements using numerical data generated by integration of Eq. (25) and processing the measurements in the EKF [results shown in Figs. 5(e) and 5(f)]. To illustrate the EKFs region of convergence, we display, in Fig. 6, the approximate region of convergence for the waterwheel parameters σ and ρ (corresponding to x_5 and x_6). To generate the figure, measurement data were simulated using Eq. (24) assuming true values $(k, \sigma, \rho) = (0.13, 1.92, 62.5)$, respectively. A fixed error covariance matrix was used to initialize the filter. All the
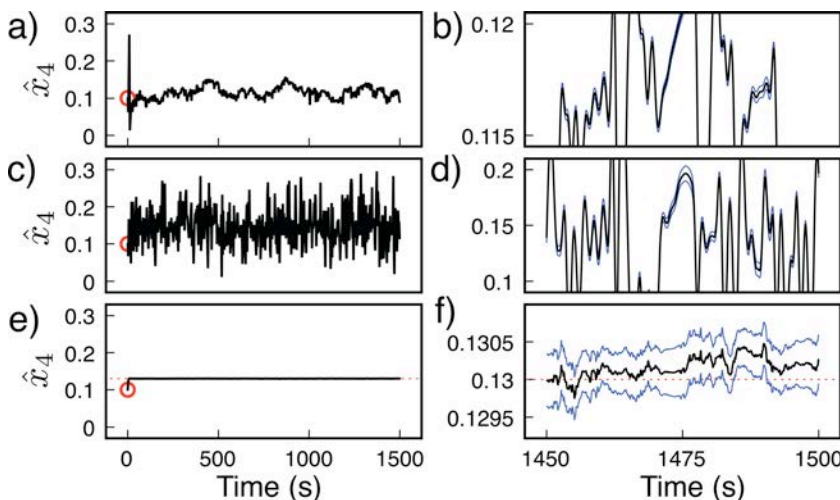

FIG. 5. (Color online) EKF estimates of $\hat{x}_4$ ($= k$) versus time: (a), (c), and (e) full time interval and (b), (d), and (f) a zoom (black center line) with ± 1 standard deviation errors (blue upper and lower line). Estimates obtained using (a) and (b) data and filter tuning matrix $\mathbf{Q}_1$, (c) and (d) data and filter tuning matrix $\mathbf{Q}_2$, (e) and (f) simulated data and filter tuning matrix $\mathbf{Q}_1$. Initial estimates: red circles and true value for simulation: red dotted line.
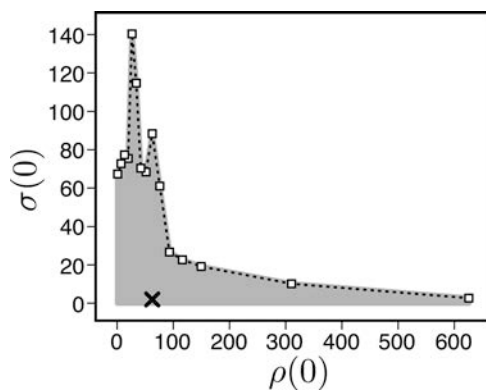
FIG. 6. Approximate region of EKF convergence (gray) in first quadrant with respect to variations in initial estimates of parameters σ and ρ from the true parameter value (cross).

filter states were initialized to their true values except for the states x_5 and x_6. By varying these initial estimates within the first quadrant, re-running the filter and noting the filter behavior at each run, the approximate region of convergence with respect to the two parameters was mapped out. For the water-wheel, even with the model mismatch, we see that the initial parameter uncertainties are not problematic for convergence.

For conditions in which parametric uncertainty is problematic for convergence, there is the option to run the EKF using a variety of initial parameter value guesses. Also, in this regard, it is worth noting that the two estimator characteristics may be beneficially combined to yield a practical hybrid approach as follows. First, the adaptive observer, with its global convergence feature, is used to generate initial estimates within the EKF region of convergence. Next, the measurement data are re-processed using the EKF to obtain state and parameter estimates with corresponding error covariance statistics.

Theoretically, the waterwheel's motion is governed by the Lorenz equations, and our experimental system supports this statement qualitatively: The phase portrait obtained using the state estimates has the typical "butterfly" morphology. However, despite this similarity, shortcomings in the model are evident when real-world data are processed by the EKF to estimate model parameters. The key observation is that when the model is insufficiently matched to the plant, model parameter estimates may not converge to constant values. In such a case, the true error variances will not match variance estimates computed by the EKF. Comparing Fig. 5(f) (in which there is no model mismatch) with Fig. 5(d) (real data), it is clear that the parameter estimates and associated error bars in the real data do not comport with the assumption that the parameter is constant; the estimates at some time points are too many standard deviations away from the estimates at other time points. The reason that the parameter estimates vary in such a statistically unlikely way is that the EKF has to account for the unmodeled dynamics in some way, if it does not diverge. Thus, the EKF steers the parameter estimates to accommodate these unmodeled dynamics, and it does this to an extent that is inconsistent with the null hypothesis that the parameters are constants. This conclusion is further supported by the observation that when the bandwidth of the filter is increased [compare Figs. 5(a) and 5(b)], the parameter estimate variations become dramatically more pronounced. That is, when the

filter is allowed to adjust the "constant" to a greater degree, it does so considerably. We believe that the most likely cause of these effects is a non-stationary mass flux of the water inflow.

Of course, no model of a real system can be perfect and estimates of model parameters can be expected to contain biases and other systematic errors. The EKF, while lacking the global convergence property of the adaptive observer, is a useful tool since it provides error statistics which allow the sufficiency of the model to be assessed. An ongoing direction of research is to quantify the degree of model sufficiency and develop a sound statistical procedure that can be used to test for unmodeled dynamics.

ACKNOWLEDGMENTS

[1]W. V. R. Malkus, Mem. Soc. R. Sci. Liege, Collection 6, **IV**, 125 (1972).
[2]S. H. Strogatz, *Nonlinear Dynamics and Chaos* (Addison-Wesley, Reading, MA, 1994).
[3]L. Illing, R. Fordyce, A. Saunders, and B. Ormond, "Experiments with a malkus-lorenz water wheel: Chaos and synchronization," Am. J. Phys. (to be published).
[4]I. Tokuda, U. Parlitz, L. Illing, M. Kennel, and H. Abarbanel, AIP Conf. Proc. **676**, 251 (2003).
[5]A. Raue, V. Becker, U. Klingmuller, and J. Timmer, Chaos **20**, 045105 (2010).
[6]P. So, E. Ott, and W. P. Dayawansa, Phys. Rev. E **49**, 2650 (1994).
[7]H. D. I. Abarbanel, D. R. Creveling, and J. M. Jeanne, Phys. Rev. E **77**, 016208 (2008).
[8]H. D. I. Abarbanel, M. Kostuk, and W. Whartenby, Q. J. R. Meteorol. Soc. **136**, 769 (2010).
[9]K. S. Narendra and A. M. Annaswamy, *Stable Adaptive Systems* (Prentice Hall, Englewood Cliffs, New Jersey, 1989).
[10]B. R. Andrievskii, V. O. Nikiforov, and A. L. Fradkov, Autom. Remote Control **68**, 1186 (2007).
[11]Q. H. Zhang, IEEE Trans. Autom. Control **47**, 525 (2002).
[12]D. C. Yu and U. Parlitz, Phys. Rev. E **77**, 066221 (2008).
[13]In point of fact, the extended Kalman filter will perform the estimation using the raw measured data (does not need the first derivative), but here the first derivative is employed in keeping with the input to the adaptive filter.
[14]L. E. Matson, Am. J. Phys. **75**, 1114 (2007).
[15]F. Sun, H. Peng, Q. Luo, L. Li, and Y. Yang, Chaos **19**, 023109 (2009).
[16]S. Boyd and S. S. Sastry, Automatica **22**, 629 (1986).
[17]D. Ruelle, Phys. Rev. Lett. **56**, 405 (1986).
[18]I. Melbourne and G. A. Gottwald, Nonlinearity **21**, 179 (2008).
[19]For systems converging to a fixed point, transients can be exploited, as discussed, for example, in H. Peng, L. Li, Y. Yang, and F. Sun, Phys. Rev. E **83**, 036202 (2011).
[20]N. N. Krasovskii, *Stability of Motion* (Stanford University Press, Stanford, CA, 1963), pp. 58–62 (translated by J. L. Brenner).
[21]A. L. Fradkov and A. Y. Progromsky, *Introduction to Control of Oscillations and Chaos* (World Scientific, Singapore, 1998), p. 51.
[22]F. Sorrentino and E. Ott, Chaos **19**, 033108 (2009).
[23]A. A. Mishra and S. Sanghi, Chaos **16**, 013114 (2006).
[24]C. H. Bishop, B. J. Etherton, and S. J. Majumdar, Mon. Weather Rev. **129**, 420 (2001).
[25]E. Ott, B. R. Hunt, I. Szunyogh, A. V. Zimin, E. J. Kostelich, M. Corazza, E. Kalnay, D. J. Patil, and J. A. Yorke, Tellus **56A**, 415 (2004).
[26]J. D. Annan, D. J. Lunt, J. C. Hargreaves, and P. J. Valdes, Nonlinear Processes Geophys. **12**, 363 (2005).
[27]*Applied Optimal Estimation*, edited by A. Gelb (MIT, Cambridge, MA, 1974), pp. 188.