

## State and parameter estimation of spatiotemporally chaotic systems illustrated by an application to Rayleigh–Bénard convection

Matthew Cornick,<sup>1</sup> Brian Hunt,<sup>1</sup> Edward Ott,<sup>1</sup> Huseyin Kurtuldu,<sup>2</sup> and Michael F. Schatz<sup>2</sup>

<sup>1</sup>University of Maryland, College Park, Maryland 20742, USA

<sup>2</sup>Center for Nonlinear Science and School of Physics, Georgia Institute of Technology, Atlanta, Georgia 30332, USA

(Received 2 November 2008; accepted 29 December 2008; published online 5 February 2009)

Data assimilation refers to the process of estimating a system's state from a time series of measurements (which may be noisy or incomplete) in conjunction with a model for the system's time evolution. Here we demonstrate the applicability of a recently developed data assimilation method, the local ensemble transform Kalman filter, to nonlinear, high-dimensional, spatiotemporally chaotic flows in Rayleigh–Bénard convection experiments. Using this technique we are able to extract the full temperature and velocity fields from a time series of shadowgraph measurements. In addition, we describe extensions of the algorithm for estimating model parameters. Our results suggest the potential usefulness of our data assimilation technique to a broad class of experimental situations exhibiting spatiotemporal chaos. © 2009 American Institute of Physics.

[DOI: 10.1063/1.3072780]

**It is often desirable to predict the future state of a chaotic system, i.e., to forecast the system. Before one can hope to estimate the future state, one must know the current state. For many systems this is not always possible, either because measurements are too noisy, or because not all system variables can be measured. This is especially true in spatiotemporally chaotic systems (chaotic systems which are spatially extended) in which it may not be possible to achieve a measurement density high enough to reconstruct the state with sufficient accuracy. For example, in weather forecasting, measurements occur at weather stations which, in remote locations, are sparsely distributed. Traditional algorithms such as the extended Kalman filter “assimilate” current and previous measurements, using a model for the system dynamics, to estimate the current system state. However, these traditional algorithms do not scale well to high-dimensional systems with many degrees of freedom, a hallmark of spatiotemporal chaos. Recent developments in the field of numerical weather prediction have demonstrated algorithms capable of handling high-dimensional systems. Although originally developed for weather prediction, these algorithms can be applied to any spatiotemporally chaotic system. Here we present the first successful application of a recently developed data assimilation algorithm to a spatiotemporally chaotic laboratory experiment. We have chosen a commonly studied Rayleigh–Bénard convection laboratory experiment exhibiting a form of spatiotemporal chaos known as *spiral defect chaos*.**

### I. INTRODUCTION

Numerous systems exhibit spatiotemporal chaos. Examples of this complex behavior with many dynamical degrees of freedom occur in optics,<sup>1</sup> chemical and biological media,<sup>2</sup> and hydrodynamics,<sup>3,4</sup> including geophysical flows in the ocean and atmosphere. Estimation of the state of an

evolving dynamical system from measurements is often a prerequisite for prediction and control. However, obtaining the system state is a common experimental difficulty for many systems exhibiting spatiotemporal chaos, where available measurements may be incomplete and noisy. When an approximate model for the system is available, it can be used in conjunction with incoming measurements to estimate the evolving system state, a process referred to as “data assimilation.”

The Kalman filter<sup>5,6</sup> optimally solves the data assimilation problem for systems with linear dynamics (and Gaussian measurement noise). Several methods extending the Kalman filter methodology to nonlinear systems have been proposed, including the extended Kalman filter (EKF),<sup>7</sup> and the class of ensemble Kalman filters (EnKF).<sup>8</sup> Straightforward application of these methods to *large* spatiotemporally chaotic systems is often completely infeasible. In particular, the EKF requires inversion of  $N \times N$  matrices, where  $N$  is the number of model variables.<sup>9</sup> For spatiotemporally chaotic systems,  $N$  can be very large (e.g., in the millions) making such matrix inversions impossible in practice. Despite these difficulties, recent developments<sup>10–12</sup> from the field of numerical weather prediction<sup>13–18</sup> suggest the possibility of achieving good accuracy (as in a Kalman filter), but in a way that is computationally feasible for large spatiotemporally chaotic systems.

In this paper we test the efficacy of a new method, the local ensemble transform Kalman filter (LETKF).<sup>12</sup> Although originally motivated by application to weather prediction, the LETKF is potentially broadly applicable to any large spatiotemporally chaotic system. It is motivated by two observations: (i) When the state is examined in a local region that is small compared to the system size, it has been shown that it can be accurately described by a relatively few degrees of freedom;<sup>19</sup> and (ii) many spatiotemporally chaotic systems exhibit finite space/time correlation scales. Thus, by (ii) we

expect that the system state at the space/time coordinate  $(\mathbf{x}, t)$  is significantly correlated only in locations  $\mathbf{x}'$  at a previous time  $t - \Delta t$  that lie within some region, say  $|\mathbf{x} - \mathbf{x}'| \leq l_{\Delta t}$  (where  $l_{\Delta t}$  might be expected to increase with  $\Delta t$ ). With these points in mind, the LETKF uses a process that we refer to as *localization*. By this we mean that we employ many independent data assimilations in a set of overlapping local regions. We choose the size of these regions empirically, increasing them until our results are no longer further improved. (This procedure may be thought of as an operational means of estimating the effective average value of  $l_{\Delta t}$ .) Because these regions are relatively small, individual computations associated with them are not prohibitive. Furthermore, by use of a simple example<sup>10,11</sup> it was indicated that, by exploiting localization in this way, state estimates with accuracies virtually the same as those for a classical Kalman filter technique (thus presumably of near optimal accuracy) can be achieved (here we use the specific implementation described in Ref. 12).

The purpose of this paper is to test the localization methodology on a system which represents a realizable laboratory experiment. Assimilation schemes have been used on laboratory experiments before,<sup>20</sup> but never using a localization technique. We have chosen to investigate a common laboratory experiment that exhibits an especially high dimensionality, Rayleigh–Bénard convection. Flows such as spiral defect chaos<sup>3,21</sup> in the Rayleigh–Bénard problem are, perhaps, the best studied experimental examples of spatiotemporal chaos; nevertheless, many general aspects of spatiotemporal chaos remain poorly understood.

What follows is an introduction to Rayleigh–Bénard convection (Sec. II) and data assimilation (Sec. III), followed by tests of the accuracy of the LETKF (Secs. IV and V). We also investigate performance with extremely sparse/noisy measurements and test extensions of the LETKF for estimating model parameters. Details of the LETKF algorithm are described in the Appendix.

## II. RAYLEIGH–BÉNARD CONVECTION

In Rayleigh–Bénard convection, a horizontal fluid layer of thickness  $d$  is confined between a heated lower plate and a cooled upper plate. For a temperature difference  $\Delta T$  between the plates that is sufficiently small, the fluid is at rest and heat is transported by conduction (resulting in a temperature  $T$  which varies linearly with vertical distance). As  $\Delta T$  is raised above a critical value  $\Delta T_c$  there is an onset of fluid motion when buoyancy overcomes viscous dissipation and thermal diffusion.

Rayleigh–Bénard convection is typically modeled using the Boussinesq equations,<sup>22</sup> which are commonly nondimensionalized with temperature scaled by  $\Delta T$ , length scaled by  $d$ , and time scaled by the vertical diffusion time  $t_v = d^2/\kappa$ , where  $\kappa$  is the thermal diffusivity. This system of units is used throughout the paper. The temperature deviation from the conducting static solution is denoted as  $\theta$ . We solve the Boussinesq equations in the disk shaped region  $x^2 + y^2 \leq \Gamma^2$ ,  $|z| \leq \frac{1}{2}$ , with Dirichlet boundary conditions  $\mathbf{u} = 0$ ,  $\theta = 0$  on all walls.  $\Gamma$  is called the aspect ratio and denotes the radius of

the disk in units of  $d$ . In terms the fluid’s velocity  $\mathbf{u}$ , temperature deviation  $\theta$ , and pressure  $p$ , the Boussinesq equations take the form

$$\begin{aligned} \left(\frac{\partial}{\partial t} + \mathbf{u} \cdot \nabla\right) \mathbf{u} &= -\nabla p + \text{Pr} \nabla^2 \mathbf{u} + \text{Pr} R \theta \hat{\mathbf{z}}, \\ \left(\frac{\partial}{\partial t} + \mathbf{u} \cdot \nabla\right) \theta &= \nabla^2 \theta + \mathbf{u} \cdot \hat{\mathbf{z}}, \end{aligned} \tag{1}$$

$$\nabla \cdot \mathbf{u} = 0.$$

These equations have two dimensionless parameters, the Rayleigh number  $R$  and the Prandtl number  $\text{Pr}$ ,

$$R = \frac{g \alpha d^3 \Delta T}{\nu \kappa}, \quad \text{Pr} = \frac{\nu}{\kappa}. \tag{2}$$

Here  $\alpha$  is the thermal expansion coefficient,  $\nu$  is the kinematic viscosity, and  $g$  is gravitational acceleration. The critical Rayleigh number for convective onset is  $R_c \approx 1707$ . The reduced Rayleigh number

$$\epsilon = \frac{R - R_c}{R_c} = \frac{\Delta T - \Delta T_c}{\Delta T_c} \tag{3}$$

measures the amount above onset. Fluid convection arises when  $\epsilon > 0$ . It is important to note that the Boussinesq equations are an approximation to the full Navier–Stokes equations. This approximation assumes small deviations of the density from its average value and neglects any temperature dependence of the transport coefficients. While the Boussinesq approximation is fairly good for the situation we will apply it to, it can be expected that it does lead to some model error.

We have investigated the parameter region near  $\epsilon = 1$ ,  $\text{Pr} = 1$ . At these values of  $\epsilon$  and  $\text{Pr}$ , the spatiotemporally chaotic state known as spiral defect chaos can arise;<sup>3,21</sup> however, in our studies using  $\Gamma \approx 20$ , the region is too small to support the large spirals typically seen in spiral defect chaos. Nevertheless, the convective flows in our studies exhibit complex behavior in both space and time (see Fig. 6 below, for an example of the spatial structure of the evolving state).

In experiments, Rayleigh–Bénard flows are visualized using the shadowgraph method,<sup>23</sup> an indirect measurement of the fluid’s spatial temperature variation. Time series of two-dimensional shadowgraph images are typically collected with sampling periods  $< 1t_v$  and with high spatial resolution ( $\sim 10^5 - 10^6$  pixels per image). Due to its difficulty, measurement of the fluid velocity is not performed in typical experiments. We note, however, that the so-called mean flow has been shown, through the use of simulations, to play a significant role in the dynamics.<sup>24</sup> Here we define the mean flow as  $\bar{\mathbf{u}}(x, y) \equiv \int \mathbf{u}_\perp(x, y, z) dz$ , where  $\mathbf{u} = \mathbf{u}_\perp + u_z \hat{\mathbf{z}}$ . Because of its physical importance, it would be desirable to be able to estimate the mean flow field  $\bar{\mathbf{u}}$ .

We connect the shadowgraph light intensity  $I(x, y)$  to the temperature field using the following relation:

$$I(x, y) = \frac{I_o(x, y)}{1 - a \nabla_{\perp}^2 \bar{\theta}(x, y)}. \tag{4}$$

Equation (4) is derived from geometric optics<sup>23,25</sup> under the approximation that  $|a \nabla_{\perp}^2 \bar{\theta}| \ll 1$ . In Eq. (4),  $\nabla_{\perp}^2 \equiv \partial^2 / \partial x^2 + \partial^2 / \partial y^2$  is the horizontal Laplacian, and the temperature field is vertically averaged:  $\bar{\theta}(x, y) \equiv \int \theta(x, y, z) dz$ .  $I_o(x, y)$  is the incident light intensity and  $a = 2z_1 |dn/dT|$ , where  $n$  is the index of refraction of the fluid,  $z_1$  is the optical path length from the midplane of the fluid layer to the image plane (in units of  $d$ ), and the temperature coefficient of the index of refraction  $|dn/dT|$  is evaluated at the average temperature of the fluid layer. We have checked the validity of the geometric optics approximation by computing  $I(x, y)$  for the simple state of straight convection rolls using both geometric optics and physical optics<sup>26</sup> for our setup under the conditions of our experiment. Under these conditions, we find the geometric optics approximation yields results that are in good agreement with the more exact results from physical optics.

### III. DATA ASSIMILATION

#### A. Outline of method

Our goal is to determine the full fluid state, given by the temperature and velocity fields  $[\theta(x, y, z)$  and  $\mathbf{u}(x, y, z)]$  from a time series of shadowgraph measurements, and we view this as a test case investigation of the general usefulness of the LETKF technique for laboratory experiments on spatiotemporal chaos. Moreover, we place particular emphasis on the ability of the state estimate to produce accurate forecasts.

We begin by considering a system state vector  $\xi$  with  $N$  components, for which we have a dynamical model,  $\xi_{j+1} = G(\xi_j)$ . Here,  $G$  is an integration of the Boussinesq equations (1) from a time  $t_j$  to  $t_{j+1} = t_j + \Delta t$ , where the  $t_j$  are the times at which we wish to construct an estimate of the system state (also the times at which measurements are assumed to be made). Our Boussinesq integration is performed using the pseudospectral method described in Ref. 27 and the state  $\xi$  consists of the variables  $\theta$  and  $\mathbf{u}$  defined on the grid points  $(r_m, \phi_n, z_l)$  of a cylindrical mesh;<sup>28</sup> symbolically,

$$\xi = \begin{bmatrix} \theta \\ \mathbf{u} \end{bmatrix}.$$

Most data assimilation algorithms are iterative, cycling between a predict and update step once every time interval  $\Delta t$ . In the update step, current measurements are used to update (or correct) the prediction. The prediction step then propagates the updated state, via the model, to the next measurement time (i.e., it is a short term forecast). The aim of this process is to synchronize the experiment and the model by coupling them via the measurements.

The LETKF assimilation method is based on the ensemble Kalman filter (EnKF), in which the update and predict steps take place for an ensemble of  $k$  system states.<sup>8</sup> This ensemble gives a finite sampling approximate representation of the probability distribution function (PDF) of the system

state. The updated ensemble  $\{\xi^{u,1} \dots \xi^{u,k}\}$  results from an update of the predicted ensemble  $\{\xi^{p,1} \dots \xi^{p,k}\}$ ,

update step:

$$\{\xi_j^{p,1} \dots \xi_j^{p,k}\} + \{\text{measurements}\} \rightarrow \{\xi_j^{u,1} \dots \xi_j^{u,k}\} \tag{5}$$

predict step:  $\xi_{j+1}^{p,i} = G(\xi_j^{u,i}) \quad i = 1 \dots k. \tag{6}$

The details of the update step are specific to the type of EnKF used, but in all cases it is based on the original Kalman filter equations. This iterative procedure begins with an initial predicted ensemble  $\{\xi_0^{p,1} \dots \xi_0^{p,k}\}$  consisting of states randomly sampled from the system attractor. The maximum likelihood estimate of the system’s state after an update step is the center of the updated ensemble,  $\bar{\xi}^u = 1/k \sum_i \xi_i^{u,i}$ .

When no localization is used, as the system size grows and the dynamical degrees of freedom increase, the necessary number of ensemble members  $k$  must increase so as to span the space of possible system states. This is a major drawback of ensemble methods, preventing their use for spatiotemporal chaos in large domains which would require an infeasibly large  $k$ . For example, in our numerical experiments we found for the Rayleigh–Bénard problem (with  $\Gamma = 20$ ) that, using the EnKF, it was not computationally feasible to use large enough ensembles to obtain results of any use.<sup>29</sup> The LETKF method, which localizes the update step, is advantageous since the number of ensemble members required is independent of the system size, making the method applicable to large domains.<sup>10,11</sup> An explanation of the LETKF’s update step (5), including the method of localization, is given in the Appendix.

#### B. Measurements

At times  $t_j (j=1, 2, \dots)$  we assume that several scalar measurements are taken, so that at each time we can represent the set of measurements by an  $s$ -component vector  $\mathbf{y}$ . In the context of Rayleigh–Bénard convection, the elements of the vector  $\mathbf{y}$  are the intensities of shadowgraph pixels,  $\mathbf{y} = [I(x_1, y_1) I(x_2, y_2) \dots I(x_s, y_s)]^T$ , where  $I(x_l, y_l)$  is the light intensity at the location  $(x_l, y_l)$  of pixel  $l$ . Note that the location of these intensity measurements need not occur on a uniform mesh; we assume that their location is fixed but arbitrary. Measurement noise is assumed to be normally distributed with mean zero and  $(s \times s)$  covariance matrix  $\mathbf{R}$ . We assume for simplicity that  $\mathbf{R} = \sigma^2 \mathbf{I}$ , i.e., a multiple of the identity matrix, so that measurement noise is homogeneous and uncorrelated with a standard deviation of  $\sigma$ .

In general  $\mathbf{y}$  is ideally (i.e., without noise) a function of the system state,  $\mathbf{y} = \mathbf{H}(\xi)$ ;  $\mathbf{H}$  is known as the observation operator.  $\mathbf{H}(\xi)$  outputs the vector of pixel intensities  $\mathbf{y}$  using a finite resolution approximation to Eq. (4), where for  $\nabla_{\perp}^2$  we use a finite difference on the cylindrical mesh. Note that, since we require  $|a \nabla_{\perp}^2 \bar{\theta}| \ll 1$  for Eq. (4) to be a good approximation,  $\mathbf{H}$  is only weakly nonlinear. Both  $\mathbf{R}$  and  $\mathbf{H}$  are critical components for the update step (5) of the ensemble-based methods.

### C. Parameter estimation

There is a straightforward extension of the ensemble methods for cases in which some model parameters are unknown. Consider the model

$$\xi_{j+1} = G(\xi_j, p), \tag{7}$$

where  $p$  is a vector of model parameters. We now formally extend the system state to include the model parameters,  $\gamma = \begin{bmatrix} \xi \\ p \end{bmatrix}$ , where the extended state evolves as

$$\gamma_{j+1} = \begin{bmatrix} \xi_{j+1} \\ p_{j+1} \end{bmatrix} = \begin{bmatrix} G(\xi_j, p_j) \\ p_j \end{bmatrix} = \hat{G}(\gamma_j). \tag{8}$$

Estimates of  $\gamma$  (and therefore of the parameters  $p$ ) result from an implementation in the same way as for  $\xi$ , but in the space of  $\gamma$  vectors. In general, observation operator parameters may also be estimated in exactly the same way as model parameters, by replacing  $H(\xi)$  by  $H(\xi, p) \equiv \hat{H}(\gamma)$ . Here  $p$  is a concatenation of model and observation operator parameters.

### D. Direct insertion

In order to assess how well the LETKF method is performing, we will compare it to a more naive approach that we call direct insertion (DI). With shadowgraph measurements, no state variables are measured directly; however, there is a one to one correspondence between a shadowgraph and the vertically averaged field  $\bar{\theta}(x, y)$ . With this in mind, the DI update step adjusts the ( $t=t_j$ ) predicted vertically averaged temperature field  $\bar{\theta}_j^u(x, y)$  to reflect the current measurement exactly.

At the time  $t_j$  of the shadowgraph measurement  $I_j(x, y)$ , the DI method updates the predicted temperature field  $\theta_j^p(x, y, z)$  by adding a correction  $\delta\theta_j(x, y, z)$  which is the unique field that is quadratic in  $z$ , matches the boundary conditions at  $|z|=\frac{1}{2}$ , and for which the updated field  $\theta_j^u(x, y, z) = \theta_j^p(x, y, z) + \delta\theta_j(x, y, z)$  satisfies

$$I_j(x, y) = \frac{I_o(x, y)}{1 - a\nabla_{\perp}^2 \bar{\theta}_j^u(x, y)}.$$

This gives the update

$$\delta\theta_j(x, y, z) = (\bar{\theta}_j^u(x, y) - \bar{\theta}_j^p(x, y))\left(\frac{3}{2} - 6z^2\right),$$

where  $\bar{\theta}_j^u(x, y)$  is found by solving a Poisson equation,

$$\nabla^2 \bar{\theta}_j^u(x, y) = \frac{1}{a} \left[ 1 - \frac{I_o(x_c, y_c)}{I_j(x_c, y_c)} \right], \tag{9}$$

and  $(x_c, y_c)$  is the location of the closest pixel to  $(x, y)$  that is observed. Note that with DI the velocity is not updated,  $\mathbf{u}_j^u(x, y, z) = \mathbf{u}_j^p(x, y, z)$ , rather it develops through coupling with the temperature during the simulation step,

$$\begin{bmatrix} \theta_{j+1}^p(x, y, z) \\ \mathbf{u}_{j+1}^p(x, y, z) \end{bmatrix} = G \left( \begin{bmatrix} \theta_j^u(x, y, z) \\ \mathbf{u}_j^u(x, y, z) \end{bmatrix} \right).$$

The  $z$ -dependence of the predicted temperature field is only slightly affected by the update since, if measurements are sufficiently frequent, the correction  $\delta\theta_j(x, y, z)$  is small. This method is the most successful data assimilation technique we have tested that does not use an update based on the Kalman filter. It is meant to represent what one might try when measurements are sufficiently dense and frequent, in which case DI is a reasonable alternative to more sophisticated data assimilation techniques.

## IV. RESULTS: PERFECT MODEL

### A. Setup of the numerical experiments

In this section we describe so-called *perfect model* tests in which a time series of states, generated from a Boussinesq simulation ( $\Gamma=20, \epsilon=1, Pr=1$ ) of one particular initial condition, serves as the “true” system. Simulated shadowgraph measurements of this time series are generated every  $\Delta t = 1/4$  by using Eq. (4) with the parameters  $a=0.08, I_o(x, y) = 0.5$ . By this technique we generate a situation in which the “true state” to be estimated and the model used to estimate it both evolve under exactly the same dynamical rules. In Sec. V we use real (not simulated) observations of a physical system for which the model dynamics is surely not an exact description.

To reproduce the effects of measurement noise we add to each pixel a small random error that is an uncorrelated normally distributed number with mean zero and standard deviation  $\sigma$ . Measurements are made sparse by removing shadowgraph pixels, leaving only those which lie on *observation locations*. We introduce the measurement density  $\rho \equiv s/(\pi\Gamma^2)$  as a measure of sparseness, where  $s$  is the number of observation locations. For  $\rho \geq 4$  we randomly and uniformly distribute observation locations over the disk, for  $\rho < 4$  the observation locations are placed on a Cartesian grid covering the disk (giving more repeatable results when using sparse measurements).

We apply the LETKF and DI methods to our simulated shadowgraphs to approximately reconstruct the original time series of true states. Here we document their performance as a function of measurement noise  $\sigma$  and measurement density  $\rho$ . Performance is quantified via the temperature and mean flow RMS relative error,

$$E_{\theta}(t) = \sqrt{\frac{\langle |\theta(x, y, z, t) - \theta^t(x, y, z, t)|^2 \rangle}{\langle |\theta^t(x, y, z, t)|^2 \rangle}},$$

$$E_{\bar{\mathbf{u}}}(t) = \sqrt{\frac{\langle |\bar{\mathbf{u}}(x, y, t) - \bar{\mathbf{u}}^t(x, y, t)|^2 \rangle}{\langle |\bar{\mathbf{u}}^t(x, y, t)|^2 \rangle}},$$

where  $\theta^t(x, y, z, t)$  and  $\bar{\mathbf{u}}^t(x, y, t)$  are the temperature and mean flow from the “true” time series of states, and  $\langle \cdot \rangle$  indicates a spatial average.

Simulated shadowgraphs are assimilated at times  $t_j, j = 1 \dots J$ . During this process we converge on an estimate of the system state ( $J$  chosen large enough to ensure convergence). At time  $t_j$  assimilation is turned off and the final updated state estimate is used as an initial condition for a long term forecast. Three measures of the quality of a state

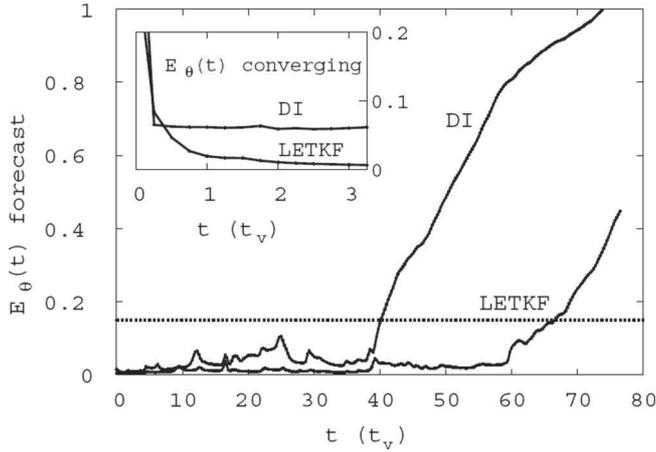


FIG. 1. Typical temperature error  $E_\theta(t)$  of forecasts with  $\sigma=0.01$  and  $\rho=127$ . The inset shows  $E_\theta(t)$  as each method converges on a state. Assimilation is turned off at time  $t_j=3.25$  in the small graph, corresponding to time  $t=0$  in the large graph. The dashed line is our chosen threshold,  $E_\theta(t) \leq 0.15$ , below which we consider the forecasts “good.” The fluid parameters are  $\Gamma=20$ ,  $\epsilon=1$ ,  $\text{Pr}=1$ .

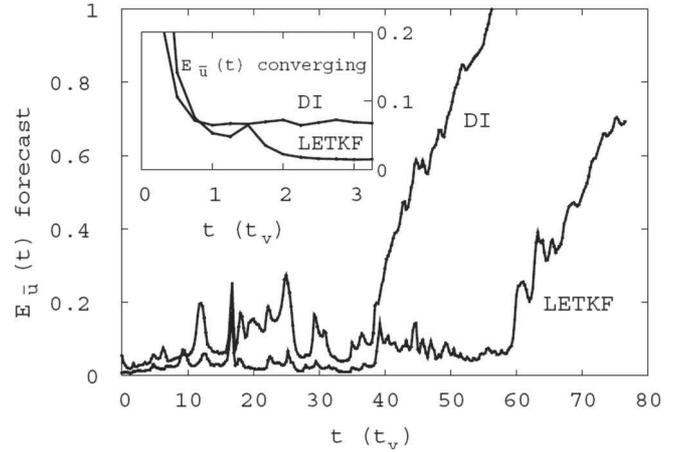


FIG. 2. Typical mean flow error  $E_{\bar{u}}(t)$  of forecasts with  $\sigma=0.01$  and  $\rho=127$ . The inset shows  $E_{\bar{u}}(t)$  as each method converges on a state. Assimilation is turned off at time  $t_j=3.25$  in the small graph, corresponding to time  $t=0$  in the large graph. The fluid parameters are  $\Gamma=20$ ,  $\epsilon=1$ ,  $\text{Pr}=1$ .

estimate are used: the *predictability time*  $\tau$ , defined as the time when  $E_\theta(t)$  first crosses the (somewhat arbitrary) value of 0.15, and the minimum values attained by  $E_\theta(t)$  and  $E_{\bar{u}}(t)$  during a forecast, denoted as  $E_\theta^{\min}$  and  $E_{\bar{u}}^{\min}$ . The latter two measures are used because the initial state estimate does not attain the minimum error, instead it occurs about  $1 t_v$  into the forecast. This is a result of the simulation rapidly balancing the fields by strongly suppressing field errors outside the Busse balloon. This effect is very slight in the LETKF forecasts, but can be quite strong in DI forecasts.

**B. Performance with noise/sparseness**

We define a “standard” *ideal* scenario as measuring a shadowgraph every  $t_v/4$  with  $\rho=127$  (corresponding to a  $451 \times 451$  shadowgraph image) and  $\sigma=0.01$  (this situation can be achieved in an experiment). Under these conditions the DI and LETKF (with  $k=18$  ensemble members) typically converge on a state estimate within  $\sim t_v$  and  $\sim 4t_v$ , respectively (observing  $\sim 4$  and  $\sim 16$  shadowgraphs, respectively). Under these conditions, measurements are sufficiently dense and frequent for DI to perform well; hence both DI and the LETKF are effective for estimation of the (unobserved) mean flow  $\bar{u}(x, y)$ . However, the LETKF typically achieves a minimum error  $E_{\bar{u}}^{\min}$  that is less than half that of DI. The forecast errors  $E_\theta(t)$  and  $E_{\bar{u}}(t)$  versus the forecast lead time  $t$  for a typical state estimate are shown in Figs. 1 and 2. The general character of the forecasts is an initial shadowing of the true state, followed by rapid divergence. When divergence begins, the spatial structure of the error is concentrated near defects. This behavior is expected, as the magnitude of the Lyapunov vector associated with the largest Lyapunov exponent has maximum magnitude at the location of defects.<sup>30</sup>

Under nonideal conditions the LETKF proves much more robust than DI. Results for sparse measurements, shown in Fig. 3, demonstrate the large range of  $\rho$  for which the LETKF converges. One can observe the existence of a

critical density of observations above which the LETKF does not substantially improve and below which it fails to converge. By adjusting the parameters of the LETKF’s update step (as described in the Appendix) we have been able to push the critical density as low as  $\rho=1.3$  without a significant loss of quality in the state estimate. DI on the other hand exhibits a steady increase in  $E_\theta^{\min}$  and  $E_{\bar{u}}^{\min}$  as  $\rho$  is decreased, as well as a rapidly deteriorating forecast when even a few observation locations are removed.

Just as there is a critical measurement density, we have also found evidence of a critical measurement frequency. This frequency lies somewhere between 1 and 2 shadowgraph images per vertical diffusion time for repeatable convergence of the LETKF under ideal conditions. This corresponds to about 1 Hz in a typical experiment.

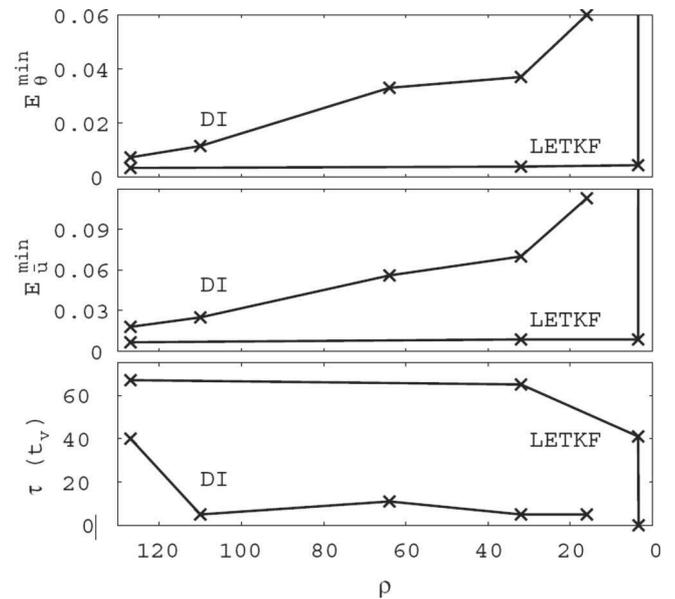


FIG. 3.  $E_\theta^{\min}$ ,  $E_{\bar{u}}^{\min}$ , and the predictability time  $\tau$  as the density of observations  $\rho$  is reduced with  $\sigma=0.01$ .

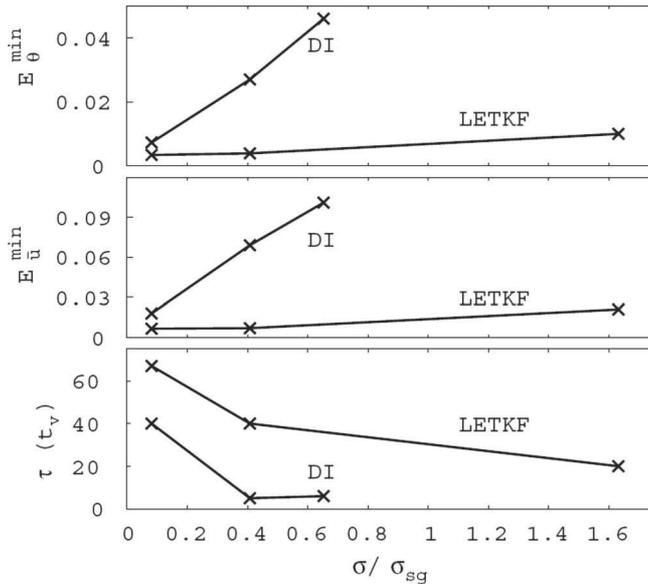


FIG. 4.  $E_{\theta}^{\min}$ ,  $E_u^{\min}$ , and the predictability time  $\tau$  as measurement noise is increased with  $\rho=127$ .

The magnitude of measurement noise is characterized by normalizing it to the RMS intensity variation of a typical shadowgraph, denoted  $\sigma_{sg}$ . In other words, the meaningful signal to noise ratio is  $\sigma_{sg}/\sigma$ . The variance  $\sigma_{sg}^2$  is obtained by averaging  $\langle |I_j(x,y) - \langle I_j(x,y) \rangle|^2 \rangle$  over many shadowgraph images [ $\sigma_{sg} \approx 0.12$  when  $a=0.08$  and  $I_o(x,y)=0.5$ ]. Results, shown in Fig. 4, indicate that DI forecasts are useful for low to moderate noise levels; whereas the LETKF operates up to and exceeding  $\sigma=\sigma_{sg}$ .

We note that all results are from one particular realization of the possible “true” time series, generated from one particular initial condition. These results are typical of what one can expect; however, variability can be expected (particularly in  $\tau$ ) for different data sets.

**C. Parameter estimation**

In our numerical parameter estimation experiments reported below we take  $\Gamma=20$ ,  $Pr=1$  and we estimated  $p = [R a]^T$ ,<sup>31</sup> where  $a$  is the observation operator parameter in Eq. (4). The initial ensemble  $\{\gamma_0^{p,1} \dots \gamma_0^{p,k}\}$  is constructed from states sampled from the attractor in the  $\xi$  component, while the  $p$  component is sampled from a normal distribution (with mean  $\bar{p} = [\bar{R} \bar{a}]^T$  and a diagonal covariance matrix with elements  $\sigma_R^2$  and  $\sigma_a^2$ ). A typical convergence process is demonstrated in Fig. 5. In this example the ensemble converges in  $8t_v$  on  $p = [R a]^T = [3414.26 \ 0.07979]^T \pm [1.61 \ 0.000072]^T$ , compared to the true value  $p = [3414.0 \ 0.08]^T$  (the error estimates for  $R$  and  $a$  are the standard deviations of the ensemble after the last update at  $t=t_f$ ). Remarkably, even when measurements are sparse ( $\rho=3.6$ , near the critical measurement density) the parameter estimates are very good,  $p = [3416.71 \ 0.07976]^T \pm [9.9 \ 0.00044]^T$ .

When estimating the state and parameters simultaneously, the eventual values of  $E_{\theta}^{\min}$  and  $E_u^{\min}$  are similar to those shown in Figs. 3 and 4. That is, the ability to estimate the system state is not adversely affected when parameters

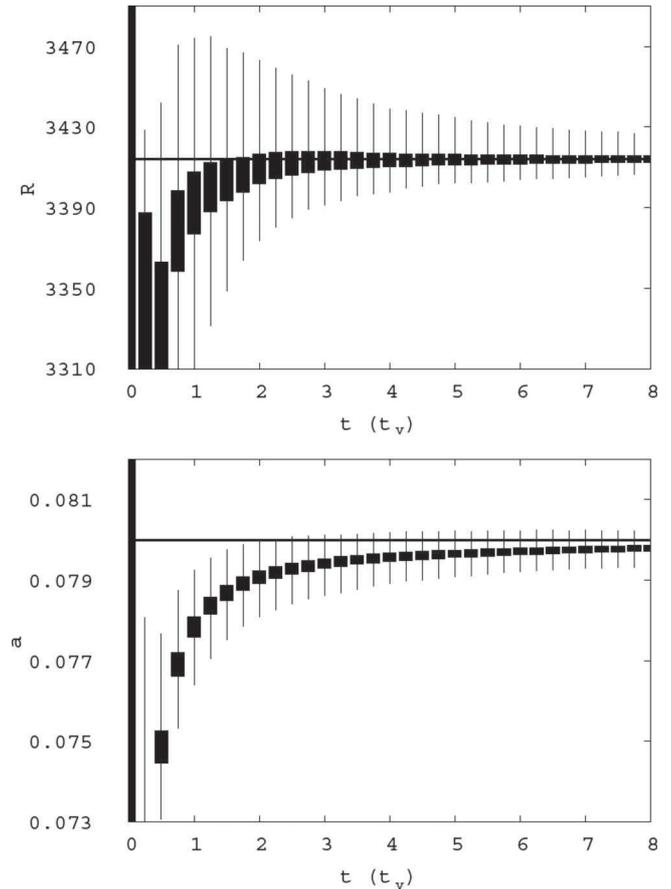


FIG. 5. Simultaneously estimating the parameters  $a$  (with true value 0.08) and  $R$  (with true value 3414). The error bars give a visual representation of the ensemble spread, extending one standard deviation up and down. The thick bars represent the case  $\rho=127$  and  $\sigma=0.01$ , while the thin lines represent the sparse measurement case,  $\rho=3.6$  and  $\sigma=0.01$ . The initial distribution was given mean ( $\bar{R}=3073$ ,  $\bar{a}=0.07$ ) and standard deviation ( $\sigma_R=683$ ,  $\sigma_a=0.02$ ).

are simultaneously estimated. It is important to note that estimating parameters (in  $\gamma$  space) requires more ensemble members than when parameters are known; thus parameter estimation tests were performed with  $k=20$ .

**V. RESULTS: EXPERIMENT**

The experiment differs from a perfect model scenario in that  $G$  and  $H$  are now approximations, requiring robustness to model error as well as observation operator error. In particular, the Boussinesq model is an approximation to the more exact Navier–Stokes equations and our geometric optics treatment is an approximation to a more involved physical optics treatment. For example, the Boussinesq equations do not treat the temperature dependence of the fluid viscosity, thermal expansion coefficient, or thermal conductivity; each of which varies by 5%–10% over the temperature range  $\Delta T$  of the experiment.

The geometry, parameter values, and boundary conditions are closely matched between experiments and simulations. For our experiments, the fluid is a thin ( $d = 0.602 \pm 0.002$  mm) layer of carbon dioxide gas compressed at a gauge pressure  $31.58 \pm 0.06$  bar. The layer is surrounded

by a circular boundary of radius  $12.50 \pm 0.02$  mm. In the experiment, the top, bottom, and lateral boundaries are composed of sapphire, aluminum, and polyethersulfone, respectively; the thermal conductivities of the boundaries exceed that of the gas by at least an order of magnitude. For this fluid, the critical temperature difference for convection onset is  $\Delta T_c = 6.02$  °C and the vertical diffusion time is  $t_v = 1.66 \pm 0.01$  s. A fixed temperature difference  $\Delta T = 10.23 \pm 0.09$  °C is imposed across the layer at a fixed mean temperature of  $22.6 \pm 0.1$  °C. These conditions correspond to  $R = 2902 \pm 26$  ( $\epsilon = 0.7$ ),  $Pr = 0.97$ , and  $\Gamma = 20.76 \pm 0.08$ . The temperature difference and pressure were stable to within the indicated uncertainties.

DI and the LETKF were used to assimilate shadowgraph images from the experiment. Images were taken every  $\Delta t = t_v/5$  (3.0 Hz) as  $395 \times 395$  bit maps ( $\rho = 90$ ) having  $\sigma_{sg} = 0.059$ . The measurement noise distribution was characterized by taking the difference between two images below onset. The distribution of pixel noise was normally distributed with a standard deviation of 0.0032, a signal to noise ratio of approximately 18.4 ( $\sigma/\sigma_{sg} \approx 0.054$ ).<sup>32</sup>

In experiments, the true fluid state is not available for directly ascertaining the accuracy of state estimates. Instead, we generate a forecast of the state estimate and compare the predicted shadowgraph sequence to subsequent measurements. We measure the forecast error by a technique which emphasizes the location of rolls and defects. Shadowgraphs are first filtered by removing high frequency components (wavelengths less than  $d/2$ ). We then threshold the image such that half the pixels are set to 1 (the remaining half are 0). This filtering/threshold procedure is applied to both the predicted and measured shadowgraph time series. The natural error measure is then the fraction of pixels incorrectly predicted, denoted  $E_I$ .

The LETKF was given as  $4t_v$  to converge on state and parameter estimates; this is sufficient for both ideal ( $\rho = 90$ ) and sparse observation ( $\rho = 4$ ) cases. Figure 6 shows a typical state estimate from the LETKF. In particular, Fig. 6(d) shows the vorticity potential from the extracted mean flow, a quantity not directly observed. Typical examples of the forecast error are shown in Fig. 7 for both methods. To the eye, DI state estimates look nearly identical to the LETKF estimates. However, DI forecasts are significantly worse than the LETKF forecasts, which use their respective  $R$  estimates. Forecasts demonstrate an approximately linear forecast error growth up to the saturation point near  $E_I = 0.5$ . To our knowledge, this is the first direct comparison of the Boussinesq equations (using accurate boundary conditions) with an experiment on a one-to-one forecast basis.

The Rayleigh number can be accurately measured directly in experiments; thus parameter estimation is unnecessary for the purpose of determining  $R$ . However, we place an emphasis on the ability of state and parameter estimates to generate good forecasts. Thus we allowed the LETKF to estimate  $R$ , as the model error can typically be compensated for, to some extent, by adjustment of model parameters off their measured values. In the dense measurement case ( $\rho = 90$ ), the LETKF converges on the parameter estimate  $R = 2625$  (the experimentally measured value is  $R = 2902 \pm 26$ ).

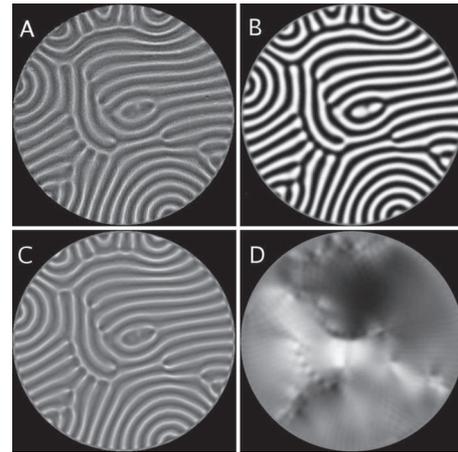


FIG. 6. An estimate of the fluid state after assimilating for  $4t_v$  ( $J = 20$  frames). (a) The  $t = t_f$  shadowgraph measurement indicating columns of warm rising fluid (dark) and cold descending fluid (light). (b) Temperature profile  $\bar{\theta}(x, y)$  of the state estimate. (c) The modeled shadowgraph  $H(\bar{\theta}(x, y))$  of the state estimate for comparison to (a). (d) The inferred vorticity potential  $\phi(x, y)$  which solves  $\nabla^2 \phi(x, y) = -\hat{z} \cdot (\nabla \times \bar{u})$  and indicates regions of clockwise rotating (dark) and counterclockwise rotating (light) mean flow.

When  $\rho = 4$  the LETKF converges on the estimate  $R = 2491$ . These estimates are obtained consistently (with slight variation) throughout the experimental data set. In fact, forcing the LETKF to use the measured  $R$  value harms the forecast, bringing it up to the level of the DI forecast (which uses the true value  $R = 2902$ ). This indicates that the advantage of the LETKF in this case lies in its ability to estimate parameters which are optimal (in the sense of producing the best forecasts when the forecast model is not exact).

Figure 8 shows how the forecasts of Fig. 7 compare with typical perfect model forecasts using the same parameters as the experiment ( $R = 2902$ ,  $\Gamma = 20.8$ ,  $Pr = 0.97$ ) as well as the same measurement frequency ( $\Delta t = t_v/5$ ), density ( $\rho = 90$ ), and approximately the same noise level ( $\sigma/\sigma_{sg} = 0.083$ ).

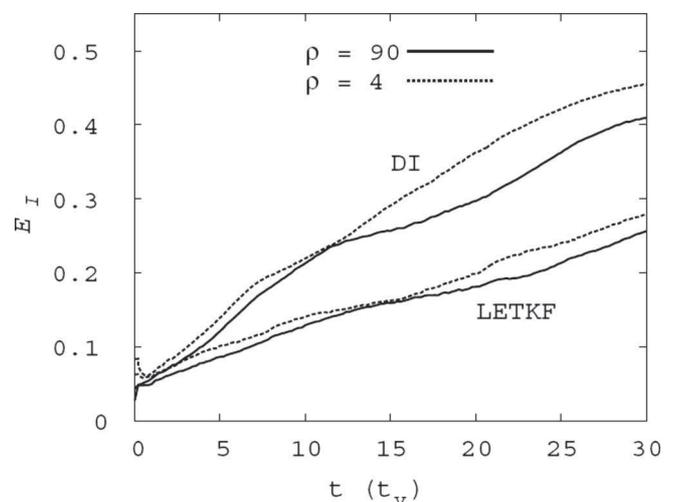


FIG. 7. Forecast error  $E_I$  for DI and LETKF methods are shown for high and low measurement densities. The LETKF forecast uses its parameter estimate ( $R = 2625$  for  $\rho = 90$ ,  $R = 2491$  for  $\rho = 4$ ) while the DI forecast uses the measured value  $R = 2902$ .

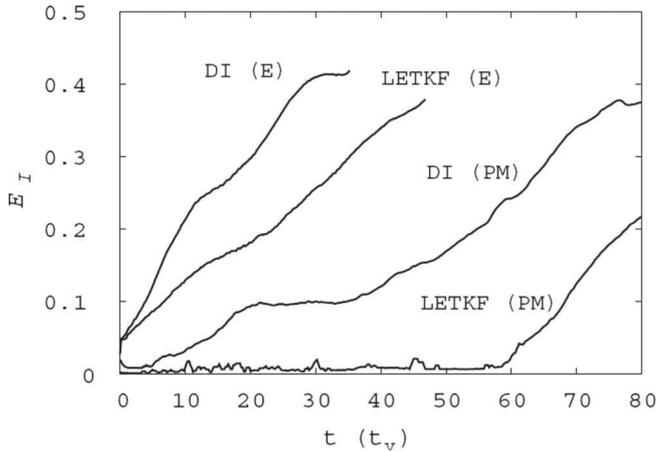


FIG. 8. Forecast error  $E_I$  for DI and LETKF methods in perfect model (PM) tests and when using experimental data (E). The LETKF forecasts use the estimated value of  $R$ , while DI forecasts use the true value. The parameters in all cases are  $R=2902$ ,  $\Gamma=20.8$ , and  $Pr=0.97$ . Noise levels are low ( $\sigma/\sigma_{sg}=0.083$  [PM] and  $\sigma/\sigma_{sg}=0.054$  [E]), and the density of measurements is high ( $\rho=90$ ).

Forecasts in the experimental situation are seen to be less accurate than those in our perfect model tests. We note that, among the uncertainties in the experiment, it is the uncertainty in the aspect ratio which has the largest potential to produce forecast error. Sensitivity tests with the model showed that the measured uncertainties in all quantities, including the aspect ratio, were too small to account for a significant portion of the forecast error. Hence, we attribute the discrepancy between the perfect model and experimental results to non-Boussinesq effects.<sup>33</sup>

**VI. CONCLUSIONS**

We have investigated two methods for estimating the fluid state in Rayleigh–Bénard convection experiments, DI and the LETKF. Both methods are effective for this purpose, with the LETKF outperforming DI both when using experimental data and in perfect model tests, especially when data are sparse/noisy. One purpose of this paper is the introduction of data assimilation methodology to a community for which DI-type techniques are the only techniques known, and to demonstrate that more involved techniques can be worth the effort. We have demonstrated that techniques developed for weather forecasting (such as the LETKF) can be successfully applied to a real laboratory system. We believe this is important as a means of introducing consideration of data assimilation techniques to the large community of researchers investigating spatiotemporal chaos in laboratory experiments.

**ACKNOWLEDGMENTS**

We are grateful to Laurette Tuckerman for the use of her FORTRAN code for simulating the Boussinesq equations in a cylindrical geometry. Support for this work by National Science Foundation Grant Nos. 04-34225 and 04-34193 and the Office of Naval Research (Physics) is gratefully acknowledged.

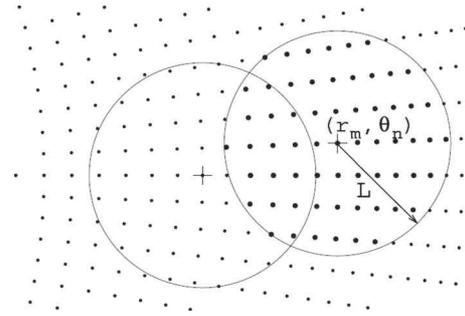


FIG. 9. Two local regions are shown on a reduced resolution mesh. Every grid point  $(m,n)$  is the center of a local region. Associated with each local region  $(m,n)$  is the local state vector  $\xi_{mn}$  consisting of state variables on the indicated horizontal grid points and all vertical grid points associated with them.

**APPENDIX: THE LETKF ALGORITHM**

We now describe the LETKF’s update step (5). This appendix is an adaptation to our Rayleigh–Bénard problem of the technique developed in Ref. 12. Because of the measurement noise we cannot know the system state exactly. Thus, we seek the PDF for  $\xi$ . In order to apply the Kalman filter methodology, we assume that this PDF is Gaussian, i.e., it is proportional to  $\exp\{-(\xi-\bar{\xi})^T P^{-1}(\xi-\bar{\xi})/2\}$ . The center of this distribution  $\bar{\xi}$  is the most likely state, while the error covariance matrix  $P$  characterizes the uncertainty of that estimate. For a given ensemble  $\xi^i$  the mean and covariance are, respectively, estimated by

$$\bar{\xi} \equiv \frac{1}{k} \sum_i \xi^i, \tag{A1}$$

$$P \equiv \frac{1}{k-1} Z(Z)^T, \tag{A2}$$

where the columns of  $Z$  are the ensemble perturbations,  $Z \equiv [\delta\xi^1 | \delta\xi^2 | \dots | \delta\xi^k]$ , with  $\delta\xi^i = \xi^i - \bar{\xi}$ . The ensemble is to be constructed to represent the Gaussian PDF with mean (A1) and covariance (A2).

Let  $\xi_{mn}$  be a vector whose components consist of the collection of all elements of  $\xi$  that lie on grid points within a horizontal distance  $L$  of the point  $(r_m, \phi_n)$  of the mesh used by the model. We call  $\xi_{mn}$  a local state and  $L$  the local region radius. There are as many local regions as horizontal grid points  $(r_m, \phi_n)$ , hence these regions are heavily overlapping (see Fig. 9). When the center of the local region  $(r_m, \phi_n)$  is near the radial boundary, the local region is the intersection of a disk having radius  $L$  centered at  $(r_m, \phi_n)$  and the disk-shaped domain with radius  $\Gamma$  centered on the origin. Note that, since the problem of interest is essentially two dimensional, local regions are indexed by two indices  $(m,n)$ . The three-dimensional nature of the system is reflected in the fact that, for each horizontal grid point, the vector  $\xi_{mn}$  contains the state at all  $z$  levels. Associated with the updated and predicted global ensemble members  $\xi^{u,i}$  and  $\xi^{p,i}$  are the local ensemble members  $\xi_{mn}^{u,i}$  and  $\xi_{mn}^{p,i}$  (all local states, global states, and ensemble states have an implied time index  $j$ ).

The predicted observation ensemble of shadowgraphs

$\{y^{p,1} \dots y^{p,k}\}$  is defined as  $y^{p,i} = \mathbf{H}(\xi^{p,i})$  (the projection of the predicted ensemble into the observation space). Let  $y_{mn}^{p,i}$  be all elements of  $y^{p,i}$  within the local region  $(m, n)$ . If there are  $s_{mn}$  measurements made within the local region  $(m, n)$ , then the vector  $y_{mn}^{p,i}$  has dimension  $s_{mn}$ . We form the matrix  $\mathbf{Y}_{mn}^p \equiv [\delta y_{mn}^{p,1} | \delta y_{mn}^{p,2} | \dots | \delta y_{mn}^{p,k}]$  where  $\delta y_{mn}^{p,i} = y_{mn}^{p,i} - \bar{y}_{mn}^p$  and  $\bar{y}_{mn}^p$  is defined as in Eq. (A1). The local measurements  $y_{mn}$  have an associated local  $s_{mn} \times s_{mn}$  covariance matrix  $\mathbf{R}_{mn}$ , which is equal to  $\sigma^2$  multiplied by the  $s_{mn} \times s_{mn}$  identity matrix. We modify this matrix by forming the tapered diagonal covariance matrix  $\mathbf{Q}_{mn}$  having  $i$ ,  $i$ th element  $Q_{mn}^{ii} \equiv [\sigma e^{-(r/r_f)^2}]^2$ , where  $r$  is the (horizontal) distance from the grid point  $(m, n)$  to the measurement location associated with the  $i$ th element of  $y_{mn}$ , and  $r_f$  is some falloff distance. This modification effectively weighs measurements further from the grid point  $(m, n)$  less heavily when estimating the state at the point  $(m, n)$ . This type of distance-dependant modification to covariance matrices has been investigated previously.<sup>17</sup> We also weigh current measurements more heavily than prior ones by the method of multiplicative variance inflation in which the predicted covariance matrix is inflated by a factor  $\Omega^2 > 1$ , to lessen the influence of prior measurements on the current state, and to compensate in some rough way for model error and nonlinearities.<sup>12,16</sup> Ordinarily  $\Omega$  is chosen empirically. The perfect model tests reported used variance inflation  $\Omega = 1.0-1.1$ , whereas results from experimental data in Sec. V used an inflation factor of  $\Omega \approx 1.4$ .

We proceed to compute the updated ensemble. As derived and discussed in Refs. 11 and 12, the procedure listed below is followed. The inputs are the global predicted ensemble  $\xi^{p,i}$  and the measurement  $\mathbf{y}$ . The output is the global updated ensemble  $\xi^{u,i}$ .

Compute each  $y^{p,i} = \mathbf{H}(\xi^{p,i})$  and  $\bar{y}^p$ .

Form the matrix  $\mathbf{Y}^p$  with columns  $\delta y^{p,i}$ .

Compute  $\bar{\xi}^p$  and form the matrix  $\mathbf{Z}^p$  with columns  $\delta \xi^{p,i}$ .

For each grid point  $(m, n)$  perform steps (1)–(7):

- (1) Form  $y_{mn}$  from the elements of the current measurement  $\mathbf{y}$ , along with the tapered covariance matrix  $\mathbf{Q}_{mn}$ .
- (2) Form  $\bar{y}_{mn}^p$  and  $\mathbf{Y}_{mn}^p$  from the relevant elements of  $\bar{y}^p$  and  $\mathbf{Y}^p$ .
- (3) Compute the updated  $k \times k$  covariance matrix,

$$\tilde{\mathbf{P}}_{mn}^u = [(k-1)\Omega^{-2}\mathbf{I} + (\mathbf{Y}_{mn}^p)^T \mathbf{Q}_{mn}^{-1} \mathbf{Y}_{mn}^p]^{-1}. \quad (\text{A3})$$

- (4) Next compute

$$\mathbf{w}_{mn} = \tilde{\mathbf{P}}_{mn}^u (\mathbf{Y}_{mn}^p)^T \mathbf{Q}_{mn}^{-1} (\mathbf{y}_{mn} - \bar{y}_{mn}^p). \quad (\text{A4})$$

- (5) Calculate the matrix

$$\mathbf{W}_{mn} = [(k-1)\tilde{\mathbf{P}}_{mn}^u]^{1/2} + \mathbf{w}_{mn}, \quad (\text{A5})$$

where, by adding a vector to a matrix we mean adding it to each column of the matrix. The  $1/2$  power here indicates taking the positive symmetric matrix square root.

- (6) Form the matrix  $\mathbf{Z}_{mn}^p \equiv [\delta \xi_{mn}^{p,1} | \delta \xi_{mn}^{p,2} | \dots | \delta \xi_{mn}^{p,k}]$  from the relevant elements of  $\mathbf{Z}^p$ . Also form  $\bar{\xi}_{mn}^p$  from  $\bar{\xi}^p$ .
- (7) Finally, compute the local updated ensemble perturbations,

$$\mathbf{Z}_{mn}^u = \mathbf{Z}_{mn}^p \mathbf{W}_{mn}. \quad (\text{A6})$$

As before,  $\mathbf{Z}_{mn}^u \equiv [\delta \xi_{mn}^{u,1} | \delta \xi_{mn}^{u,2} | \dots | \delta \xi_{mn}^{u,k}]$ , and the local updated ensemble is given by  $\xi_{mn}^{u,i} = \bar{\xi}_{mn}^p + \delta \xi_{mn}^{u,i}$ .

To complete the update step, components of the global updated ensemble member  $\xi^{u,i}$  at each horizontal grid point  $(m, n)$  are taken to be equal to the elements of  $\xi_{mn}^{u,i}$  at the center of local region  $(m, n)$ . Note that each local region is assimilated independently, allowing for massive parallelization.

To estimate parameters, simply replace  $\xi$  with  $\gamma$  everywhere in the above steps. This formulation assumes state variables are spatially extended. Thus, when adding global parameters to the state space we must assume that they are spatially dependant. That is, when estimating both the Rayleigh number and  $a$  of Eq. (4) ( $\mathbf{p} = [R \ a]^T$ ), the state  $\gamma$  is the concatenation of  $\xi$  and  $\hat{\mathbf{p}}$ , where  $\hat{\mathbf{p}} = [R_{11} \dots R_{mn} \dots a_{11} \dots a_{mn} \dots]^T$ . The LETKF is then augmented by averaging these parameter values over the grid, after the update step, to form global parameters. This average is performed for each global ensemble member  $\gamma^{u,i}$  by setting its  $\hat{\mathbf{p}}$  component to  $\hat{\mathbf{p}}^{u,i} = [\bar{R}^i \ \bar{R}^i \ \dots \ \bar{a}^i \ \bar{a}^i \ \dots]^T$ , where  $\bar{R}^i$  and  $\bar{a}^i$  are the spatial averages of  $R$  and  $a$  for the  $i$ th ensemble member. The model  $\mathbf{G}$  and observation operator  $\mathbf{H}$  then use the parameters  $\bar{R}^i$  and  $\bar{a}^i$  when applied to ensemble member  $i$ . If the model allows for spatially dependent parameters, then this last averaging step is not necessary.

The LETKF formulation is advantageous since the number of ensemble members required for convergence is independent of the system size,<sup>10</sup> making the method applicable to large domains. The number of ensemble members will presumably scale with the number of dynamical degrees of freedom in a local region. We used a local region radius of  $L = 2.6d$  and a falloff distance of  $r_f = 1.4d$  in the perfect model section. For the experimental data, we found that a local region radius of  $L = 2.6d$  and a falloff distance of  $r_f = 1.0d$  worked well. In both cases,  $L$  is comparable to the correlation length of spiral defect chaos of  $2.7d$  when  $\epsilon = 0.7$  and  $2.3d$  when  $\epsilon = 1.0$ .<sup>34</sup> As the ensemble converges, it tends to confine itself to a space of dimension lower than  $k$ , indicating that one could optimize by “pruning” the ensemble size as it converges. All the results in this paper are for a constant  $k = 18$  (or  $k = 20$  when estimating parameters), but we have found that starting with  $k = 18$  and reducing to  $k = 8$  linearly within 10 measurement times gives similar results with a significant reduction in computation time. In addition, the strength of the model nonlinearities is largest when the ensemble spread is large (during the first few assimilation steps), thus one can begin assimilation with a large  $\Omega$  and reduce it linearly to speed convergence. This procedure was found to be successful, but was not performed in the results reported here.

<sup>1</sup>F. T. Arecchi, S. Boccaletti, and P. Ramazza, *Phys. Rep.* **318**, 1 (1999).

<sup>2</sup>H. L. Swinney and V. I. Krinsky, *Waves and Patterns in Chemical and Biological Media* (MIT Press, Cambridge, 1991).

<sup>3</sup>S. W. Morris, E. Bodenschatz, D. S. Cannell, and G. Ahlers, *Phys. Rev. Lett.* **71**, 2026 (1993).

- <sup>4</sup>M. C. Cross and P. C. Hohenberg, *Rev. Mod. Phys.* **65**, 851 (1993).
- <sup>5</sup>The Kalman filter is a form of recursive Bayesian estimation. Essentially, it uses the model to find an optimal fitting of the current system state to current and past measurements.
- <sup>6</sup>R. Kalman, *J. Basic Eng.* **82**, 35 (1960).
- <sup>7</sup>D. Simon, *Optimal State Estimation: Kalman, H Infinity, and Nonlinear Approaches* (Wiley-Interscience, New York, 2006).
- <sup>8</sup>G. Evensen, *Data Assimilation: The Ensemble Kalman Filter* (Springer, Berlin, 2006).
- <sup>9</sup>For example, for a discretized partial differential model evolving  $M$  scalar spatial fields in time, the number of model variables is  $M$  multiplied by the number of grid points.
- <sup>10</sup>E. Ott, B. R. Hunt, I. Szunyogh, A. V. Zimin, E. J. Kostelich, M. Corazza, E. Kalnay, D. J. Patil, and J. A. Yorke, *Phys. Lett. A* **330**, 365 (2004).
- <sup>11</sup>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, Ser. A* **56**, 415 (2004).
- <sup>12</sup>B. Hunt, E. Kostelich, and I. Szunyogh, *Physica D* **230**, 112 (2007).
- <sup>13</sup>J. S. Whitaker and T. M. Hamill, *Mon. Weather Rev.* **130**, 1913 (2002).
- <sup>14</sup>M. K. Tippett, J. L. Anderson, C. H. Bishop, T. M. Hamill, and J. S. Whitaker, *Mon. Weather Rev.* **131**, 1485 (2003).
- <sup>15</sup>C. H. Bishop, B. J. Etherton, and S. J. Majumdar, *Mon. Weather Rev.* **129**, 420 (2001).
- <sup>16</sup>J. Anderson, *Mon. Weather Rev.* **129**, 2884 (2001).
- <sup>17</sup>J. W. T. M. Hamill and C. Snyder, *Mon. Weather Rev.* **129**, 2776 (2001).
- <sup>18</sup>P. L. Houtekamer and H. L. Mitchell, *Mon. Weather Rev.* **129**, 123 (2001).
- <sup>19</sup>D. J. Patil, B. R. Hunt, E. Kalnay, J. A. Yorke, and E. Ott, *Phys. Rev. Lett.* **86**, 5878 (2001).
- <sup>20</sup>M. Galimiche, J. Sommeria, E. Thivolle-Cazat, and J. Verron, *Meccanica* **12**, 331 (2003).
- <sup>21</sup>E. Bodenschatz, W. Pesch, and G. Ahlers, *Annu. Rev. Fluid Mech.* **32**, 709 (2000).
- <sup>22</sup>F. H. Busse, *Rep. Prog. Phys.* **41**, 1929 (1978).
- <sup>23</sup>W. Merzkirch, *Flow Visualization* (Academic, New York, 1974), p. 258.
- <sup>24</sup>K.-H. Chiam, M. R. Paul, M. C. Cross, and H. S. Greenside, *Phys. Rev. E* **67**, 056206 (2003).
- <sup>25</sup>S. Rasenat, G. Hartung, B. L. Winkler, and I. Rehberg, *Exp. Fluids* **7**, 412 (1989).
- <sup>26</sup>S. Trainoff and D. S. Cannell, *Phys. Fluids* **14**, 1340 (2002).
- <sup>27</sup>L. S. Tuckerman, *J. Comput. Phys.* **80**, 403 (1989).
- <sup>28</sup>The grid points are located at the Chebyshev points in  $r$  and  $z$ ,  $r_m = \Gamma \cos(\pi(m-1)/2(N_r-1))$ ,  $\phi_n = 2\pi n/N_\phi$ ,  $z_l = 1/2 \cos(\pi(l-1)/(N_z-1))$ . In the investigated parameter region we find  $N_r=151$ ,  $N_\phi=451$ ,  $N_z=9$  (for a total of roughly  $6 \times 10^5$  grid points), and a simulation time step of  $t_v/100$  sufficient.
- <sup>29</sup>We also investigated the smaller domain  $\Gamma=6$  with  $\epsilon=2.0$  and found that the use of the EnKF required  $k \approx 100$  ensemble members.
- <sup>30</sup>D. A. Egolf, I. V. Melnikov, W. Pesch, and R. E. Ecke, *Nature (London)* **404**, 733 (2000).
- <sup>31</sup>Due to limitations in our simulation, we were unable to estimate  $Pr$  in the full  $\Gamma=20$  system. Instead we estimated  $Pr$ ,  $R$ , and  $a$  together in small aspect ratio tests ( $\Gamma=4$ ,  $R=8540$ ,  $Pr=1$ ). The results of these tests indicated that all three parameters can be simultaneously estimated with high accuracy.
- <sup>32</sup>We found empirically that using larger values for  $R$  ( $\sigma \approx 0.01-0.02$ ) in the LETKF's update step (rather than the measured noise level of  $\sigma=0.0032$ ) resulted in better forecasts.
- <sup>33</sup>We attempted to model some non-Boussinesq effects. However, there are several ways the fluid can be non-Boussinesq, many of which are not easy to model. For example, our numerical technique relies heavily on the assumption of incompressibility, and deviations from this behavior are beyond our modeling capabilities. Some non-Boussinesq effects were included (temperature dependence of conductivity and specific heat), but did not significantly affect forecast error when compared with experimental data.
- <sup>34</sup>S. Morris, E. Bodenschatz, D. Cannell, and G. Ahlers, *Physica D* **97**, 164 (1996).