J. Fluid Mech. (2011), *vol.* 682, *pp.* 289–303.
doi:10.1017/jfm.2011.222

State estimation in wall-bounded flow systems. Part 3. The ensemble Kalman filter

C. H. COLBURN[1]†, J. B. CESSNA[2] AND T. R. BEWLEY[1]

[1]Department of Mechanical Engineering, University of California, San Diego
La Jolla, CA 92093, USA

[2]Numerica Corporation, 4850 Hahns Peak Drive, Suite 200, Loveland, CO 80538, USA

(Received 16 September 2010; revised 1 March 2011; accepted 11 May 2011)

State estimation of turbulent near-wall flows based on wall measurements is one of the key pacing items in model-based flow control, with low-Re channel flow providing the canonical testbed. Model-based control formulations in such settings are often separated into two subproblems: estimation of the near-wall flow state via skin friction and pressure measurements at the wall, and (based on this estimate) control of the near-wall flow field fluctuations via actuation of the fluid velocity at the wall. In our experience, the turbulent state estimation sub-problem has consistently proven to be the more difficult of the two. Though many estimation strategies have been tested on this problem (by our group and others), none have accurately captured the turbulent flow state at the outer boundary of the buffer layer ($5 \leqslant y^+ \leqslant 30$), which is deemed to be an important milestone, as this is the approximate range of the characteristic near-wall turbulent structures, the accurate estimation of which is important for the control problem. Leveraging the ensemble Kalman filter (an effective variant of the Kalman filter which scales well to high-dimensional systems), the present paper achieves at least an order of magnitude improvement (in the near-wall region) over the best results available in the published literature on the estimation of low-Reynolds number turbulent channel flow based on wall information alone.

Key words: control theory, turbulence control

1. Introduction

In the past, estimation of chaotic fluid systems was motivated mostly by the need for accurate weather forecasts. Today, the prospects of potential implementation of real-time feedback control in manufacturing systems, or perhaps even aerodynamics systems, provide new motivation to study this fundamental problem. Bewley & Liu (1998), Bewley (2001) and Högberg, Bewley & Henningson (2003) developed optimal feedback kernels for the *control* of the linearized Navier–Stokes equation in a channel. The dependence of these feedback control kernels on the near-wall region, where the characteristic near-wall turbulent structures are located, emphasizes the importance of accurate state estimates in this region if effective feedback control is the ultimate aim.

State estimation is a problem that has been considered at length by researchers in many distinct communities. The 'controls' and 'dynamic systems' communities

† Email address for correspondence: ccolburn@ucsd.edu

have focused primarily, but not exclusively, on problems which (*a*) have numerically tractable solutions to the corresponding Riccati equations (Zhou, Doyle & Glover 1996), linear matrix inequalities (Scherer, Gahinet & Chilali 2002; Boyd *et al.* 1994) or dynamic programs (Bertsekas *et al.* 2001) (all three of these methods usually requiring a sufficiently low-state dimension or problems that can effectively be reduced to such via standard model reduction techniques) and (*b*) are characterized by uncertainties in the initial state, state disturbances and measurement noise that are well approximated as Gaussian. In most cases, these assumptions are not valid in estimation problems related to turbulent flows.

The weather forecasting community, on the other hand, has focused on estimation (a.k.a. 'data assimilation') strategies that are numerically tractable for high-dimensional discretizations of PDE systems. The two primary classes of data assimilation strategies which have been developed in this community and are available today for multi-scale uncertain systems are the ensemble Kalman filter (EnKF; see, Evensen 2003) and space/time variational (4DVar; see, Bouttier & Courtier 1999) methods.

The EnKF methods, which come in a few distinct variations, are particularly well suited for nonlinear multi-scale systems with substantial uncertainties. Even for some low-dimensional problems, the EnKF methods have been shown to provide significantly improved state estimates in certain nonlinear problems for which the more traditional extended Kalman filter (EKF) breaks down. The statistics of the estimation error in the EnKF are not propagated via a covariance matrix but are instead implicitly represented via the distribution of several perturbed trajectories ('ensemble members'), which themselves are propagated with the full nonlinear system model. On many problems, in practice, the collection of these ensemble members (itself called the 'ensemble') accurately captures the dominant directions of uncertainty of the estimation error (in phase space) even when a relatively small number of ensemble members are used. This is the key feature that lends the EnKF method its remarkable numerical tractability in high-dimensional problems.

The 4DVar methods propagate state and sensitivity ('adjoint') simulations back and forth across an optimization window of interest. An optimization is performed based on these iterative marches to minimize a cost function balancing: (*a*) a term accounting for the misfit of the estimate with the measurements over the optimization window, with (*b*) a 'background' term accounting for the 'old' estimate (that is, based on the measurements and statistics obtained prior to the present optimization window). Though such a retrospective analysis is certainly beneficial in certain ways in the estimation of nonlinear systems, 4DVar methods are not as natural as EnKF methods for representing the principal directions of uncertainty in the estimate, which is a critical ingredient of any effective state estimation strategy.

1.1. *Related background on state estimation*

Estimation, in general, involves the determination of a probability distribution. This probability distribution describes the likelihood that any particular point in phase space accurately represents the truth state. That is, without knowing the actual state of a system, estimation strategies attempt to represent the probability of any given state using only a time history of noisy observations and an approximate dynamic model of the system of interest. Given this statistical distribution, estimates can be inferred about the 'most likely' state of the system, and how much confidence should be placed in that estimate. Unfortunately, in this most general form, the estimation problem is intractable for most systems. However, given certain justifiable

assumptions about the nature of the model and its associated disturbances and uncertainties, simplifications can be applied with regards to how the probability distributions are modelled. In linear systems with Gaussian uncertainty of the initial state, Gaussian state disturbances and Gaussian measurement noise, it can be shown that the probability distribution of the optimal estimate is itself Gaussian (see, e.g. Anderson & Moore 1979). Mathematically, for any linear system

$$\boldsymbol{x}_k = \boldsymbol{F}_{k-1}\boldsymbol{x}_{k-1} + \boldsymbol{w}_{k-1} \qquad \boldsymbol{w} \sim N(\boldsymbol{0}, \boldsymbol{Q}) \tag{1.1}$$

$$\boldsymbol{y}_k = \boldsymbol{H}\boldsymbol{x}_k + \boldsymbol{v}_k \qquad \boldsymbol{v} \sim N(\boldsymbol{0}, \boldsymbol{R}), \tag{1.2}$$

the entire distribution of the estimate in phase space can be represented exactly by its mean $\bar{\boldsymbol{x}} = E[\boldsymbol{x}]$ and its second moment about the mean (that is, its covariance) $\boldsymbol{\Sigma}$ where

$$\boldsymbol{\Sigma} = E[(\boldsymbol{x} - \bar{\boldsymbol{x}})(\boldsymbol{x} - \bar{\boldsymbol{x}})^H]. \tag{1.3}$$

In this paradigm, the notation $\boldsymbol{x} \sim N(\bar{\boldsymbol{x}}, \boldsymbol{\Sigma})$ denotes explicitly that the random variable $\boldsymbol{x}$ has a normal (Gaussian) distribution about its mean and covariance. This is the essential piece of theory that leads to the traditional Kalman filter (KF), first introduced by Kalman (1960) and Kalman & Bucy (1961).

Sequential data assimilation methods (e.g. the KF) provide a method to propagate $\hat{\boldsymbol{x}}$ and $\boldsymbol{P}$ (that is, estimates of $\bar{\boldsymbol{x}}$ and $\boldsymbol{\Sigma}$, respectively) forward in time, making the appropriate updates to both upon the receipt of each new measurement. It is useful to think of these quantities, at any given time t_k, as being conditioned on a subset of the available measurements. The notation $\hat{\boldsymbol{x}}_{k|j}$ represents the highest likelihood estimate at time t_k given measurements up to and including time t_j. Similarly, $\boldsymbol{P}_{k|j}$ represents the covariance corresponding to this estimate. In particular, $\hat{\boldsymbol{x}}_{k|k-1}$ and $\boldsymbol{P}_{k|k-1}$ are often called the prediction estimate and the prediction covariance, whereas $\hat{\boldsymbol{x}}_{k|k}$ and $\boldsymbol{P}_{k|k}$ are often called the current estimate and the current covariance. The linear, discrete-time evolution equations for the KF are

$$\hat{\boldsymbol{x}}_{k|k-1} = \boldsymbol{F}_{k-1}\hat{\boldsymbol{x}}_{k-1|k-1}, \tag{1.4}$$

$$\boldsymbol{P}_{k|k-1} = \boldsymbol{F}_{k-1}\boldsymbol{P}_{k-1|k-1}\boldsymbol{F}_{k-1}^H + \boldsymbol{Q}_{k-1}, \tag{1.5}$$

$$\hat{\boldsymbol{x}}_{k|k} = \hat{\boldsymbol{x}}_{k|k-1} + \boldsymbol{P}_{k|k-1}\boldsymbol{H}^H(\boldsymbol{H}\boldsymbol{P}_{k|k-1}\boldsymbol{H}^H + \boldsymbol{R})^{-1}(\boldsymbol{y}_k - \boldsymbol{H}\hat{\boldsymbol{x}}_{k|k-1}), \tag{1.6}$$

$$\boldsymbol{P}_{k|k} = \boldsymbol{P}_{k|k-1} - \boldsymbol{P}_{k|k-1}\boldsymbol{H}^H(\boldsymbol{H}\boldsymbol{P}_{k|k-1}\boldsymbol{H}^H + \boldsymbol{R})^{-1}\boldsymbol{H}\boldsymbol{P}_{k|k-1}, \tag{1.7}$$

where (1.4) and (1.5) propagate the state and covariance, respectively, between measurement updates, and (1.6) and (1.7) update the state and covariance, respectively, at measurement times. Note that $\hat{\boldsymbol{x}}_{k|k+K}$, for some $K > 0$, is often called a smoothed estimate and may be obtained in the sequential setting by a Kalman smoother (see, Rauch, Tung & Striebel 1965; Anderson & Moore 1979).

For nonlinear systems with relatively small uncertainties, a common variation on the KF known as the EKF has been developed in which the mean and covariance are propagated about a linearized trajectory of the full system. Essentially, if a Taylor-series expansion for the nonlinear evolution of the covariance is considered and all terms higher than quadratic are dropped, what is left is the differential Riccati equation associated with the EKF covariance propagation. Though this approach gives acceptable estimation performance for nonlinear systems when uncertainties are small as compared with the fluctuations of the state itself, EKF estimators often diverge when uncertainties are more substantial and other techniques are needed.

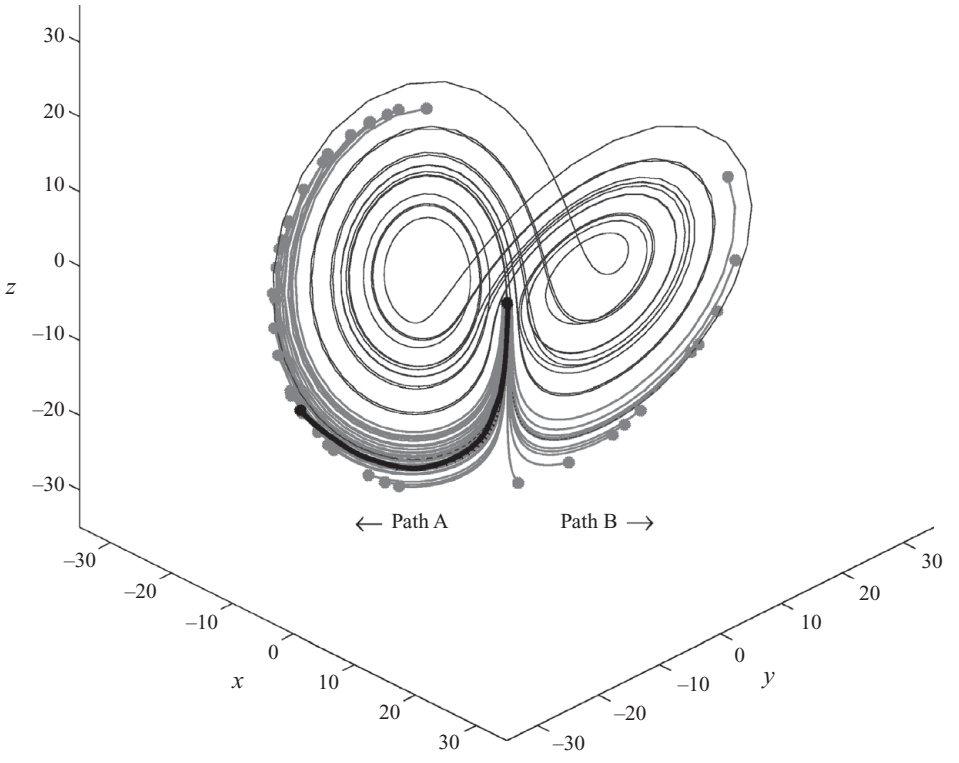

FIGURE 1. Non-Gaussian uncertainty propagation in the Lorenz system (see Lorenz 1963). The black point in the centre shows a typical point located in a sensitive area of this chaotic system's attractor in phase space, representing a current estimate of the state. The thick black line represents the evolution in time of the trajectory from this estimate. If the uncertainty of the estimate is modelled as a very small cloud of points, centred at the original estimate with an initially Gaussian distribution, then the additional grey lines show the evolution of each of these perturbed points in time. A Gaussian model of the resulting distribution of points is, clearly, completely invalid.

At its core, the linear thinking associated with the uncertainty propagation in the KF and EKF breaks down in chaotic systems. Chaotic systems are characterized by stable manifolds or 'attractors' in n-dimensional phase space. Such attractors are fractional-dimensional subsets (a.k.a. 'fractal' subsets) of the entire phase space. Trajectories of chaotic systems are stable with respect to the attractor in the sense that initial conditions off the attractor converge exponentially to the attractor, and trajectories on the attractor remain on the attractor. On the attractor, however, trajectories of chaotic systems are characterized by an *exponential divergence* of slightly perturbed trajectories. That is, two points infinitesimally close on the attractor at one time will diverge exponentially from one another as the system evolves until they are effectively uncorrelated.

Just as individual trajectories diverge along the attractor, so does the uncertainty associated with them. This uncertainty can diverge in a highly non-Gaussian fashion when such uncertainties are not infinitesimal (see figure 1). Estimation techniques that attempt to propagate probability distributions under linear, Gaussian assumptions fail to capture the true uncertainty of the estimate in such settings, and thus improved estimation techniques are required. In this case, when estimating an ordinary

differential equation (ODE) with n states, the probability density function (PDF) of the estimate in phase space must be discretized and propagated for accurate results. This converts the straightforward propagation of statistics in the KF problem (with two coupled ODEs, of order n and n^2) into a much more difficult PDE, of dimension n, governing the evolution of the PDF itself; this PDE is known as the Fokker–Plank equation (see, e.g. Jazwinski 1970, p. 164). This PDE may be approximated and evolved with a Lagrangian method, referred to in this setting as a particle filter (PF; see Arulampalam *et al.* 2002), or with a grid-based method, which may be made tractable by exploiting the sparsity of the PDF in phase space (see, Bewley & Sharma 2009); both approaches are numerically tractable only for extremely small values of n [i.e. $n \lesssim 5$].

1.2. *Prior work on the estimation of turbulent channel flow*

In high-dimensional estimation problems [e.g. $n \gtrsim O(10^6)$], Bayesian methods, based on propagating the full Fokker–Plank PDE, such as the PF method, are completely out of the question. The KF and EKF approaches are also infeasible, unless further decoupling or approximation is applied, due to their reliance on the propagation of the covariance (of order n^2) of the PDF of the estimate.

Exploiting a spectral decomposition to decouple the associated equations, Högberg *et al.* (2003) solved the full KF problem for the state estimation of near-wall flows. They did this by solving the (decoupled) estimation Riccati equations for the individual Fourier modes of the linearized system, which makes these equations tractable, then inverse-transforming the result back to physical space to obtain implementable feedback convolution kernels associated with the estimation problem. Using an improved problem formulation, Hœpffner *et al.* (2005) found effective kernels for the state estimation problem in transitional channel flow (that is, for small perturbations from the laminar state) that converged properly upon grid refinement. Chevalier *et al.* (2006) then attempted to develop a nonlinear extension of this work in order to apply it effectively to a fully developed turbulent flow using an EKF. This work meticulously calculated a numerical model of the statistics of the nonlinear terms of a fully developed channel flow, then used these statistics as the covariance of the state disturbances when computing the estimator feedback gains via the (linear) Kalman formulation.

Alternative methods have attempted the use of a Wiener filter (Martinelli 2009) and 4DVar (Bewley & Protas (2004), a vector-based variant of the Kalman smoother reviewed) to estimate $Re_\tau = 100$ and $Re_\tau = 180$ flows, respectively. In Bewley & Protas (2004) the authors also examined the direct extrapolation of wall skin friction and pressure measurements of a turbulent flow into the flow field via Taylor series analysis; unfortunately, it was found that the domain of convergence was much smaller than 20 viscous units. Chandrasekhar's method has also been proposed for reducing the dimension of the covariance propagation equations in the KF/EKF (see, Kailath 1973). This approach involves the propagation of a reduced-order factored form of the time derivative of the covariance matrix, as well as the propagation of the feedback gain matrix itself, rather than the (numerically intractable) propagation of the full covariance matrix. This approach is promising for problems of this class and has yet to be tried for the estimation of near-wall turbulence.

The groundwork described in the previous two paragraphs, upon which the present paper is based, is reviewed further, and put into a broader context, in Kim & Bewley (2007). The broad range of existing studies on this canonical problem provides a benchmark against which new approaches may be compared. The present paper

applies the EnKF to estimate a $Re_\tau = 100$ turbulent flow based on wall skin friction and wall pressure measurements alone. The remainder of this paper summarizes briefly the EnKF and the principal heuristics, localization and covariance inflation, required in its application to large-scale systems. The results, presented in § 3, might be considered the first 'successful' estimation of this difficult benchmark problem.

2. The ensemble Kalman filter

The EnKF, first proposed by Evensen (1994), is a modern stochastic alternative to Chandrasekhar's method, described above, for state estimation in high dimensional systems and is reviewed in depth in Evensen (2003, 2009*a*,*b*). Simply stated, in standard implementations of the EnKF, a sample covariance matrix replaces the forecast of the covariance matrix in (1.6). Even though only second-order statistics of the distribution are typically used at each measurement update in the EnKF approach, the full nonlinear dynamics of the system are used to propagate each candidate realization between measurement updates.

We now review briefly the formulation of the standard EnKF, using a continuous-time representation of the system state $\boldsymbol{x}(t)$ and measurements $\boldsymbol{y}_k = \boldsymbol{y}(t_k)$ available at discrete times t_k,

$$\frac{\partial \boldsymbol{x}(t)}{\partial t} = f(\boldsymbol{x}(t), \boldsymbol{u}(t), \boldsymbol{w}(t)), \tag{2.1}$$

$$\boldsymbol{y}_k = h(\boldsymbol{x}_k) + \boldsymbol{v}_k. \tag{2.2}$$

The system dynamics $f(\cdot)$ in this formulation may be nonlinear and forced by some known function $\boldsymbol{u}(t)$, and are also assumed to be corrupted by random 'state disturbances' $\boldsymbol{w}(t)$ with known statistics. Similarly, the measurement operator $h(\cdot)$ may be nonlinear and is assumed to be corrupted by additive white 'measurement noise' $\boldsymbol{v}_k$ with covariance $\boldsymbol{R}_k$.

Recall from the Introduction that the EKF propagates the full covariance matrix and uses it to perform measurement updates according to Bayes' rule, assuming a Gaussian PDF. The EnKF is, in a sense, quite similar, but builds an estimate $\boldsymbol{P}^e$ of the covariance matrix $\boldsymbol{P}$ based on an outer product matrix quantifying the deviation of the ensemble members from their mean,

$$\boldsymbol{P}^e_{k|k-1} = \frac{(\delta\hat{\boldsymbol{X}}_{k|k-1})(\delta\hat{\boldsymbol{X}}_{k|k-1})^H}{N-1}, \tag{2.3a}$$

$$\delta\hat{\boldsymbol{X}}_{k|k-1} = \left[\delta\hat{\boldsymbol{x}}^1_{k|k-1} \quad \delta\hat{\boldsymbol{x}}^2_{k|k-1} \ldots \delta\hat{\boldsymbol{x}}^N_{k|k-1}\right], \tag{2.3b}$$

$$\delta\hat{\boldsymbol{x}}^j_{k|k-1} = \hat{\boldsymbol{x}}^j_{k|k-1} - \hat{\boldsymbol{x}}_{k|k-1}, \tag{2.3c}$$

$$\hat{\boldsymbol{x}}_{k|k-1} = \frac{1}{N}\sum_j \hat{\boldsymbol{x}}^j_{k|k-1}. \tag{2.3d}$$

Using this 'sample' (that is, approximate) covariance matrix $\boldsymbol{P}^e$, a standard KF measurement update may be performed,

$$\hat{\boldsymbol{x}}^j_{k|k} = \hat{\boldsymbol{x}}^j_{k|k-1} + \boldsymbol{P}^e_{k|k-1}\boldsymbol{H}^H\left[\boldsymbol{H}\boldsymbol{P}^e_{k|k-1}\boldsymbol{H}^H + \boldsymbol{R}_k\right]^{-1}\left(\boldsymbol{y}^j_k - \boldsymbol{H}\,\hat{\boldsymbol{x}}^j_{k|k-1}\right), \tag{2.4}$$

where $\hat{\boldsymbol{x}}^j_{k|k-1}$ denotes the jth ensemble member at time step k based on measurements up to $\boldsymbol{y}_{k-1}$, $\boldsymbol{P}^e_{k|k-1}$ denotes the sample covariance matrix, as given in (2.3*a*), based on the collection of ensemble members, $\boldsymbol{y}^j_k$ denotes a discrete-time random vector with

statistical distribution $N(\boldsymbol{y}_k, \boldsymbol{R})$ and $\boldsymbol{H}$ denotes a linearization of the operator $h(\cdot)$ about the mean state estimate $\hat{\boldsymbol{x}}_{k|k-1}$. For more information on the standard EnKF measurement update and its properties, the reader is referred to Evensen (2003).

It is important to differentiate between the PF and the EnKF, since they are perhaps easily confused. Although both estimation methods use the governing equations to propagate sets of perturbed realizations through phase space, the measurement update in each method is fundamentally different. The PF uses a *weighted linear combination* of these perturbed candidate realizations to approximate the PDF of the estimate, with the weights being adjusted each time a measurement is taken via Bayes' rule. In contrast, the EnKF effectively uses identical weights on each realization, instead *shifting the realizations themselves*, according to a Kalman-like update formula, whenever measurements are taken.

In the computationally efficient implementation of the EnKF, the outer product formula for the sample covariance matrix (2.3*a*) is kept in its factored form when calculating the update (2.4) (or the modified form of this update, given below) in order to retain the numerical tractability of the result. That is, $\boldsymbol{P}^e$ is represented as the product of two matrices of order $n \times N$ and $N \times n$, bypassing the full computation of the $n \times n$ matrix $\boldsymbol{P}^e$, which is important because, in the implementation, $N \ll n$. We thus rewrite (2.4) as

$$\hat{\boldsymbol{x}}^j_{k|k} = \hat{\boldsymbol{x}}^j_{k|k-1} + \alpha \delta \hat{\boldsymbol{X}}_{k|k-1}\, \delta \hat{\boldsymbol{Y}}^H_{k|k-1} \left[\alpha \delta \hat{\boldsymbol{Y}}_{k|k-1} \quad \delta \hat{\boldsymbol{Y}}^H_{k|k-1} + \boldsymbol{R}_k\right]^{-1} \left(\boldsymbol{y}^j_k - \boldsymbol{H}\, \hat{\boldsymbol{x}}^j_{k|k-1}\right), \tag{2.5}$$

where $\delta \hat{\boldsymbol{Y}}_{k|k-1} = \boldsymbol{H}\, \delta \hat{\boldsymbol{X}}_{k|k-1}$ and $\alpha = 1/(N-1)$.

Butala *et al.* (2008) established that the EnKF, when formulated correctly for systems with linear dynamics, asymptotically converges to the Kalman result as the number of ensemble members becomes sufficiently large. An abbreviated proof is provided below for convenience. Note in particular that the $\boldsymbol{y}^j_k$ are perturbed in a statistically consistent fashion so that the covariance term $\boldsymbol{K}\boldsymbol{R}\boldsymbol{K}^H$ is properly recovered below in (2.14).

THEOREM 2.1. (Equivalence of the EnKF to KF) *In the limit of an infinite number of ensemble members (i.e. $N \to \infty$), the estimated ensemble mean and covariance converge to the equivalent KF equations* (1.6) *and* (1.7)*, respectively, when using the EnKF update* (2.4).

Proof. Consider the rewritten EnKF update (2.4) as the unique solution for the random variable $\hat{\boldsymbol{x}}^j_{k|k}$ conditioned on the random variables $\hat{\boldsymbol{x}}^j_{k|k-1}$ and $\boldsymbol{y}^j_k$,

$$\hat{\boldsymbol{x}}^j_{k|k} = \hat{\boldsymbol{x}}^j_{k|k-1} + \boldsymbol{K}^e \left(\boldsymbol{y}^j_k - \boldsymbol{H}\, \hat{\boldsymbol{x}}^j_{k|k-1}\right), \tag{2.6}$$

$$\boldsymbol{K}^e = \boldsymbol{P}^e_{k|k-1} \boldsymbol{H}^H \left[\boldsymbol{H}\boldsymbol{P}^e_{k|k-1} \quad \boldsymbol{H}^H + \boldsymbol{R}_k\right]^{-1}. \tag{2.7}$$

After recalling the definition of the covariance matrix

$$\boldsymbol{P}_{k|k-1} = \lim_{N\to\infty} \frac{1}{N-1} \sum_{j=1}^{N} \left(\hat{\boldsymbol{x}}^j_{k|k-1} - \hat{\boldsymbol{x}}_{k|k-1}\right)\left(\hat{\boldsymbol{x}}^j_{k|k-1} - \hat{\boldsymbol{x}}_{k|k-1}\right)^H \tag{2.8}$$

$$= \lim_{N\to\infty} \boldsymbol{P}^e_{k|k-1}, \tag{2.9}$$

as discussed in Butala *et al.*, the EnKF gain matrix $\boldsymbol{K}^e$ converges to the Kalman gain matrix $\boldsymbol{K}$ as $n \to \infty$ by Slutsky's theorem (Slutsky 1925; for explanation, see Gut

2005, Theorem 11.4, p. 249). Thus, the expected value of (2.6) can be rewritten as

$$\hat{\boldsymbol{x}}_{k|k} = \lim_{N\to\infty} \frac{1}{N}\sum_{j=1}^{N} \hat{\boldsymbol{x}}^j_{k|k} \tag{2.10}$$

$$= \lim_{N\to\infty} \frac{1}{N}\sum_{j=1}^{N} \hat{\boldsymbol{x}}^j_{k|k-1} + \boldsymbol{K}\left[\frac{1}{N}\sum_{j=1}^{N} \boldsymbol{y}^j_k - \boldsymbol{H}\,\frac{1}{N}\sum_{j=1}^{N} \hat{\boldsymbol{x}}^j_{k|k-1}\right] \tag{2.11}$$

$$= \hat{\boldsymbol{x}}_{k|k-1} + \boldsymbol{K}(\boldsymbol{y}_k - \boldsymbol{H}\,\hat{\boldsymbol{x}}_{k|k-1}), \tag{2.12}$$

which is identical to the KF state update (1.6).

By performing an equivalent analysis for the covariance of the random variable $\hat{\boldsymbol{x}}^j_{k|k}$, the covariance update equation is recovered similarly

$$\boldsymbol{P}_{k|k} = \lim_{N\to\infty} \frac{1}{N-1}\sum_{j=1}^{N} \left(\hat{\boldsymbol{x}}^j_{k|k} - \hat{\boldsymbol{x}}_{k|k}\right)\left(\hat{\boldsymbol{x}}^j_{k|k} - \hat{\boldsymbol{x}}_{k|k}\right)^H \tag{2.13}$$

$$= (\boldsymbol{I} - \boldsymbol{K}\boldsymbol{H})\boldsymbol{P}_{k|k-1}(\boldsymbol{I} - \boldsymbol{K}\boldsymbol{H})^H + \boldsymbol{K}\boldsymbol{R}\boldsymbol{K}^H + \boldsymbol{\Phi} + \boldsymbol{\Phi}^H, \tag{2.14}$$

where

$$\boldsymbol{\Phi} = \lim_{N\to\infty} \frac{1}{N-1}\sum_{j=1}^{N} \left[(\boldsymbol{I} - \boldsymbol{K}\boldsymbol{H})\left(\hat{\boldsymbol{x}}^j_{k|k-1} - \hat{\boldsymbol{x}}_{k|k-1}\right)\left(\boldsymbol{y}^j_k - \boldsymbol{y}_k\right)^H \boldsymbol{K}^H\right] = \boldsymbol{0}, \tag{2.15}$$

which implies that (2.14) and (1.7) are equivalent and the proof is complete.

Although this theoretical result justifies applying the EnKF to many problems, it does not provide practical guidelines for choosing ensemble size for more general applications, which is necessary during implementation. For linear problems of very high dimension ($n \geqslant O(10^6)$) Furrer & Bengtsson (2007) show that convergence of the trace of the covariance matrix is possible when the number of ensemble members scales like the square of the order of the state (i.e. $N \sim O(n^2)$). Furrer & Bengtsson also show that there can sometimes be considerable bias in the estimator even when the number of ensemble members is of the same order as the order of the state dimension (i.e. $N \sim O(n)$, also discussed in Evensen 2003, 2009*a*), a much less strict requirement.

Based on this analysis alone, the requirements for convergence, which scale as poorly as the storage requirements for the KF and EKF, restrict applications of the EnKF to low-dimensional systems. As a result, two heuristics must be implemented in practice to reduce the negative side effects associated with a reduced ensemble size: *localization* and *covariance inflation*. As mentioned in the Introduction, it is often found in practice that when these two heuristics are used with an ensemble-based approach, the dominant directions of uncertainty of the estimation error (in phase space) are captured accurately even when a relatively small number of ensemble members is used.

2.1. *Localization*

Localization is an artificial distance-based suppression of the off-diagonal components of the sample covariance matrix $\boldsymbol{P}^e_{k|k-1}$ as represented by (2.3*a*). It was first proposed by Houtekamer & Mitchell (2001) and is an essential ingredient to the success of the EnKF in practice. It is introduced to eliminate spurious correlations in the covariance matrix that arise from the fact that it is usually grossly under-sampled (that is,

$N \ll n$). Note in (2.3a) that the off-diagonal components of the covariance matrix $\boldsymbol{P}^e_{k|k-1}$ are obtained by averaging the product of a flow field perturbation at one point in the physical domain with a flow field perturbation at another point in the physical domain. If these two points are separated by a large distance, it may be argued on physical grounds that this averaged product should be small; localization thus imposes this decay of correlation with distance, even if the system is so grossly under-sampled that (2.3a) does not capture this decay (which is usually the case).

The sample covariance matrix $\boldsymbol{P}^e_{k|k-1}$ in (2.4) may thus be replaced by

$$\boldsymbol{P}^e_{k|k-1} = \frac{\rho \bullet (\delta\hat{\boldsymbol{X}}_{k|k-1})(\delta\hat{\boldsymbol{X}}_{k|k-1})^H}{N-1}, \tag{2.16}$$

where ρ is a distanced-based localization function and $\bullet$ denotes the element-wise product. Using this modified sample covariance formula, (2.5) may be rewritten as

$$\hat{\boldsymbol{x}}^j_{k|k} = \hat{\boldsymbol{x}}^j_{k|k-1} + \alpha\rho_1 \bullet (\delta\hat{\boldsymbol{X}}_{k|k-1} \quad \delta\hat{\boldsymbol{Y}}^H_{k|k-1}) \times \left[\alpha\rho_2 \bullet \quad (\delta\hat{\boldsymbol{Y}}_{k|k-1} \quad \delta\hat{\boldsymbol{Y}}^H_{k|k-1}) + \boldsymbol{R}_k\right]^{-1} \left(\boldsymbol{y}^j_k - \boldsymbol{H}\,\hat{\boldsymbol{x}}^j_{k|k-1}\right) \tag{2.17}$$

where $\alpha = 1/(N-1)$ is the constant defined in (2.5), $\rho_1^{i,m}$ is a distance-based localization function relating the ith state and the mth measurement and $\rho_2^{m_1,m_2}$ is a localization function relating m_1th measurement and m_2th measurement; both functions approach unity as the distance between the corresponding flow quantities approaches zero, and both approach zero as the distance between the corresponding flow quantities becomes large.

2.2. *Covariance inflation*

Another challenge when using under-sampled representations of probability distributions in high-dimensional state-space systems is 'covariance collapse'. This occurs when the majority of ensemble members are distributed on a fraction of the attractor, and thus the computed statistics do not capture all of the principal directions of uncertainty. Anderson & Anderson (1999) review this phenomena and the effect it has in weather forecasting applications and discuss its simple practical solution *covariance inflation*. Much earlier, Anderson & Moore (1979, pp. 131–134) demonstrated that, even in linear KF applications, insensitivity to measurements (resulting from accounting improperly for model uncertainty) can lead to filter divergence. Furrer & Bengtsson (2007) further argue that most sources of error in ensemble filters result in underestimation of the ensemble variance; thus, covariance inflation is a natural mechanism for correcting the unknown deficiencies that lead to an underestimated prior variance. Covariance inflation, as proposed by Anderson & Anderson (1999), simply pushes the ensemble members away from the mean by some arbitrary growth factor β at each time step,

$$\hat{\boldsymbol{x}}^j = \beta\,(\hat{\boldsymbol{x}}^j - \hat{\boldsymbol{x}}) + \hat{\boldsymbol{x}}. \tag{2.18}$$

In weather forecasting applications, inflation parameters in the range of $\beta \in [1.005, 1.05]$ demonstrate significantly improved estimator performance. Attempts to develop adaptive methods for tuning β have been made in Anderson (2007, 2009) and Wang & Bishop (2003), but unfortunately the results so far do not appear to justify their complexity.

3. Numerical results

We now characterize the ability of the EnKF, as described above, to estimate a 3D incompressible turbulent channel flow, given measurements of the skin friction and pressure on uniformly spaced 16×16 array on each wall. The numerical computations presented use the standard spectral-spectral-second-order-finite-difference approach of Bewley, Moin & Temam (2001) to simulate the uncontrolled, constant mass-flux turbulent channel flow on a 64^3 grid with $L_x = 2\pi$, $L_z = \pi$ and $L_y = 2$. (Note that this code-base uses L_x, L_z and L_y to denote the streamwise, spanwise and wall-normal directions, respectively. The reader is referred to Bewley *et al.* (2001) for more details.) The flow is governed by the incompressible Navier–Stokes equation with uniform density and viscosity. By defining the half-channel height δ, mean skin friction $\bar{\tau}_w = -\nu\, \partial \bar{u}_1/\partial n$ and mean friction velocity $u_\tau = (\bar{\tau}_w/\rho)^{0.5}$, this equation can be conveniently non-dimensionalized where time, space and velocity are normalized by ν/u_τ^2, ν/u_τ and u_τ, respectively. As a result, the Reynolds number becomes simply a function of viscosity $Re_\tau = 1/\nu$. By choosing $\nu = 0.01$, the domain is re-expressed, in non-dimensional form, with $L_x^+ = 628$, $L_z^+ = 314$ and $L_y^+ = 200$. Although the domain size is not identical to our previous work in flow estimation, it is larger than the minimal flow unit required for the onset of turbulence (Jimenez & Moin 1991). This, along with statistics generated for validation purposes, ensures that the results are in fact representative of an estimator tracking a turbulent flow at low-Reynolds number. The 'truth' model is calculated as an identical simulation running in parallel with the EnKF-based estimator.

Since the simulation relies on Fourier transforms for computational efficiency, all data are stored in frequency space as the simulation advances. This conveniently allows for measurements to be extracted through high-order spectral interpolation schemes. Wall shear stress in the streamwise and spanwise directions were calculated via spectral interpolation in the wall-parallel directions and second-order interpolation in the wall-normal direction. When pressure measurements were required, a Poisson equation was solved.

As mentioned in §§ 2.1 and 2.2, localization and inflation are *ad hoc* yet essential ingredients to the success of any large-scale EnKF implementation. The distance-dependent localization functions ρ used in the present work were chosen to be exponential in shape,

$$\rho^{i,j} = \rho(\boldsymbol{x}(i), \boldsymbol{x}(j)) = e^{|\boldsymbol{x}^i - \boldsymbol{x}^j|^2_{\boldsymbol{Q}}}, \tag{3.1}$$

where $\boldsymbol{Q} = diag\{[q_1,\ q_2,\ q_3]\} > 0$ is a diagonal weighting matrix related to three length scales. The appropriate selection of each q_i reflects the 'trust' associated with the ability of the sample covariance matrix to construct accurate correlations in the streamwise, wall-normal and spanwise directions. These diagonal elements of $\boldsymbol{Q}$ correspond to the half-height of the exponential function decay and are generally selected to correspond to known flow statistics. These length scales were determined from correlation studies of uncontrolled turbulence, analogous to those reported in Kim, Moin & Moser (1987) and Bewley *et al.* (2001) and are subjected to a minor amount of additional variation. The inflation parameter $\beta = 1.01$ was selected based on reported results from the weather forecasting community (Anderson & Anderson 1999).

The simulation was performed using the Triton cluster at San Diego's Super Computing Center (SDSC), where each simulation required 70 h of compute time on 66 parallel cores (the details of the cluster can be found at the SDSC website), with each core corresponding to an ensemble member. This algorithm design choice was determined by two unrelated but important constraints: the discretization of

the domain and the size of the cluster on which the simulation was performed. The computational expense of the EnKF simulation performed and the desire to drive it all the way to statistical steady state (2200 ensemble updates were performed during each simulation) prevented us from performing more extensive parametric studies on these three length parameters at the present time, or repeating the study at higher Reynolds numbers, both of which are left for future work.

The quality of the reconstruction is determined by comparing the perturbation component of the true velocity field with the same perturbation component of the estimated velocity field. Normalized error and correlation measures, as defined by Bewley & Protas (2004), are used for comparison with previous work,

$$\mathrm{Errn}(q'_{est},\, q'_{tru}) = \frac{\displaystyle\int_0^{L_x}\int_0^{L_z} (q'_{est} - q'_{tru})^2 \,\mathrm{d}z\,\mathrm{d}x}{\displaystyle\int_0^{L_x}\int_0^{L_z} (q'_{tru})^2 \,\mathrm{d}z\,\mathrm{d}x}, \tag{3.2}$$

$$\mathrm{Corr}(q'_{est},\, q'_{tru}) = \frac{\displaystyle\int_0^{L_x}\int_0^{L_z} q'_{est} q'_{tru} \,\mathrm{d}z\,\mathrm{d}x}{\sqrt{\displaystyle\int_0^{L_x}\int_0^{L_z} (q'_{est})^2 \,\mathrm{d}z\,\mathrm{d}x}\sqrt{\displaystyle\int_0^{L_x}\int_0^{L_z} (q'_{tru})^2 \,\mathrm{d}z\,\mathrm{d}x}}. \tag{3.3}$$

Note that quantities are primed to emphasize that the perturbation component of the velocity field (that is, the instantaneous velocity component minus its planar average) is being used in the comparison. The subscripts $()_{est}$ and $()_{tru}$ correspond to the 'estimated' and 'truth' values, respectively. These two normalized measures account for the (x, z)-plane averaged statistics as a function of time and distance from the wall. The long-time average of these measures provides a rigorous quantification of the quality of the state estimate as a function of distance from the wall, approximating their corresponding expected values, $E[\mathrm{Errn}(y, t)]$ and $E[\mathrm{Corr}(y, t)]$, at statistical steady state.

The error norm defined above is perhaps the more sensitive of the two criteria. It is normalized by the planar-averaged mean-squared energy of the truth simulation, which makes it a particularly sensitive measure near the wall, where this quantity approaches zero. (Note that an error norm near unity indicates that the estimate is completely decoupled from the truth, whereas an error norm near zero indicates that the estimate is in perfect agreement with the truth.) When significant error is present, the correlation is useful to quantify the planar-averaged phase error, as distinct from the planar-averaged amplitude error; an error in the amplitude of the estimate (but not its phase) will adversely affect the error norm, but not the correlation. Note also that a correlation near unity indicates perfect phase alignment of the estimate with the truth.

Figure 2 shows the L_2 energy of the difference between the estimate and the truth as a function of t^+. Note that the error norm and correlation were calculated, and averaged, from $t^+ = 1000$ to 2200 (1200 viscous time units after apparently converging to statistical steady state). Through observation of the variations in averaged skin-friction (skin friction averaged over the top/bottom walls), it is clear that temporal averages are taken over multiple flow-throughs with substantial statistical averaging.

The most significant test of the estimator in this problem, of course, is to quantify its convergence starting from arbitrary initial conditions. Figures 3 and 4 thus report such a test, using the most suitable values of the localization parameters identified

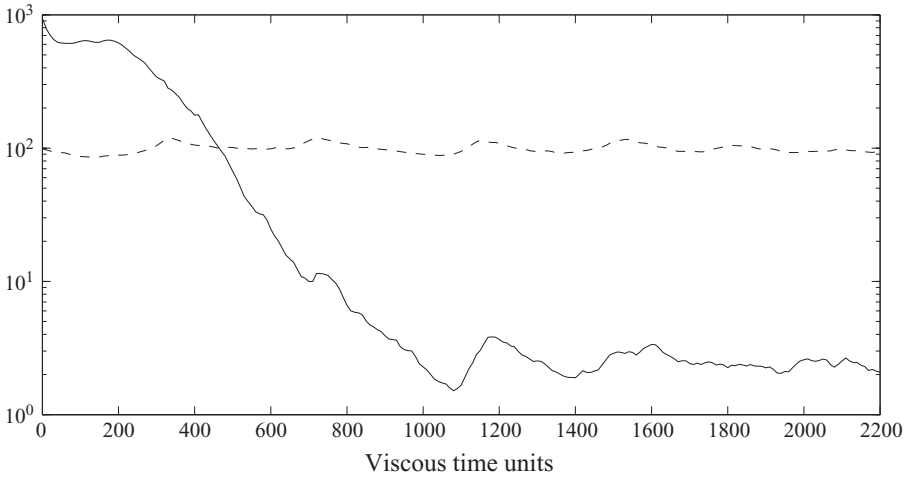

FIGURE 2. Convergence statistics versus viscous time units. L_2 energy of the error of the estimation, solid, and total wall drag, dashed, as a function of t^+. Time averaged results in figures 3 and 4 are averaged from time $t^+ = 1000$ to 2200. The magnitude and variations in total wall drag are consistent with those of an $Re_\tau = 100$ turbulent flow.

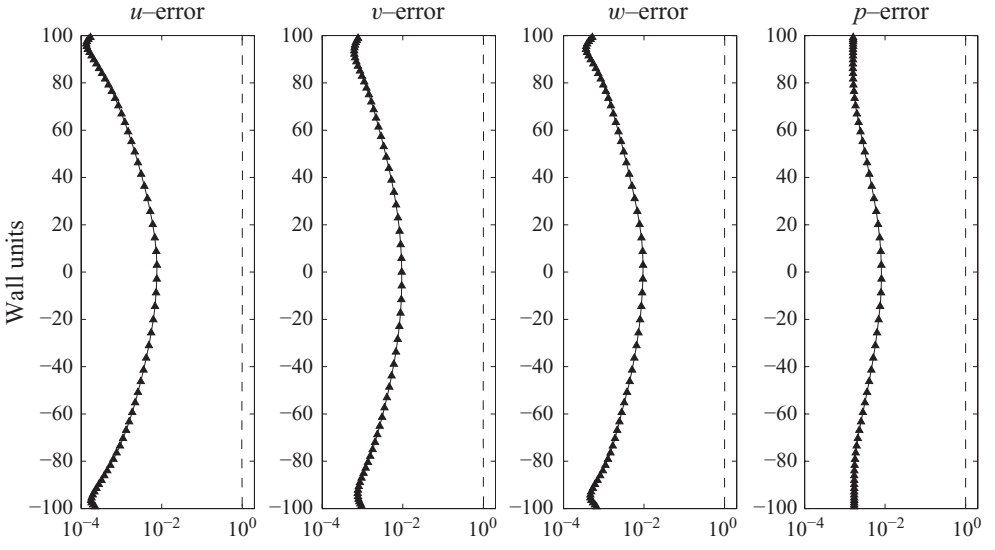

FIGURE 3. Error norm in the estimation of an $Re_\tau = 100$ flow using the EnKF, starting from bad initial conditions. The localization used is given in (3.1), with localization length scales of $\boldsymbol{Q} = \mathrm{diag}([50, 50, 25])$ and localization constant $\beta = 1.01$. The dashed line represents the statistics of the error before estimation was attempted (averaged over 2400 viscous time units), and the line denoted ▲ represents the error norm of the EnKF at statistical steady state.

thus far, $\boldsymbol{Q} = diag\{[50, 50, 25]\}$. An initial condition for the estimator was generated from a fully developed turbulent flow simulation (independent of the truth model) and snapshots of this flow were taken every $\Delta t^+ = 200$ viscous time units to initialize each ensemble member. Symmetry across the channel centreline in figures 3 and 4 reflect the approximate statistical convergence in the simulation. This symmetry would be perfect if statistical steady state were in fact reached.

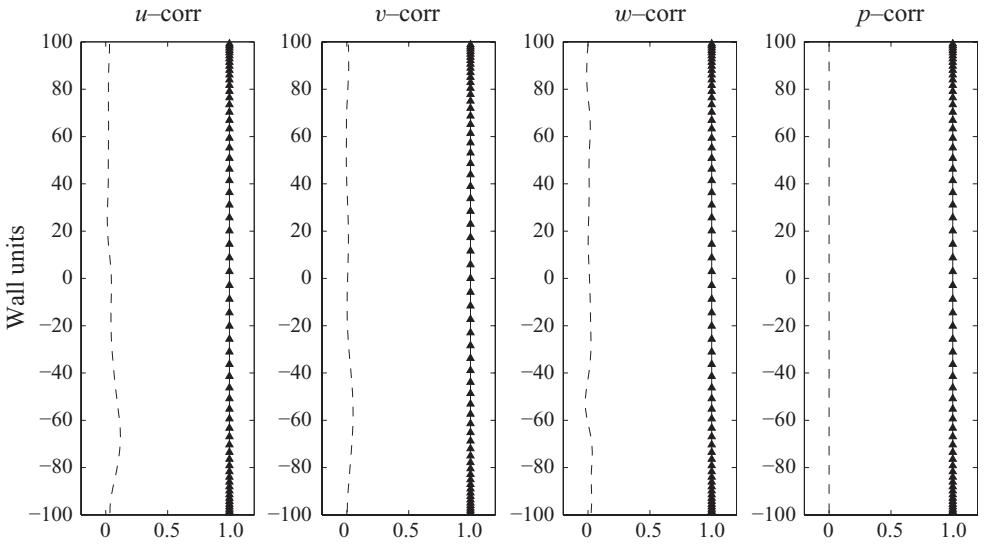

FIGURE 4. As in figure 3, but reporting the correlation.

In this case, near perfect synchronization of the EnKF with the truth is observed. The EnKF estimation was found to perform with at least an *order of magnitude less error* at 20 viscous units from the wall than the previous best estimation results reported in the literature on this problem (see § 1.2). In particular, we observe error measurements of 0.001, 0.001 and 0.001 and correlation measurements of 0.99, 0.99 and 0.99 (in the u, v and w velocity components, respectively). As a point of comparison Chevalier *et al.* (2006, when using the EKF) reported error measurements of 0.5, 0.88 and 0.9 and correlation measurements of 0.87, 0.59 and 0.59 at the same location.

4. Conclusions and future work

The EnKF depends on the sample covariance matrix $\boldsymbol{P}^e$, which is a low-rank approximation of the estimated covariance matrix $\boldsymbol{P}$. It is well known by those who use the EnKF in weather forecasting applications that the finite size of the ensemble in the EnKF causes spurious correlations and covariance collapse. In practice, these phenomena must be compensated for through a distance-dependent localization function and covariance inflation in order to ensure adequate convergence of the estimator.

This paper has presented the first near-perfect state estimation of $Re_\tau = 100$ turbulent flow using wall information only; that is, we have demonstrated a sustained synchronization of the state estimate with the truth when a random initial condition is used in the estimator, and the localization function is tuned appropriately. When comparing with previous results in the published literature, at least an order of magnitude less error was observed at 20 viscous units from the wall.

Previous unreported results found that the ability to achieve such estimator convergence is apparently fairly sensitive to the localization parameter used in the wall-normal direction, and apparently fairly insensitive to the localization parameters used in the streamwise and spanwise directions. To a certain extent, these results reflect the law of the wall, the fundamental idea that any turbulence near-wall flow varies in a predictable statistical fashion as a function of the distance from the wall. Exploring

the estimator performance as a function of these localization parameters may reveal inherent properties of the flow not yet known and warrants further investigation.

Besides further tuning of these heuristics, one interesting possibility for improving the present estimation strategy is to implement a 'Rogallo transform' (Rogallo 1981; Rogallo & Moin 1984) for the quantities being estimated. In his pioneering work, Rogallo showed that, in regions of high shear (in this case, near a wall), a convenient transformation on the domain may be defined that moves something like a windshield wiper. Such a transformation on the domain might in the present problem provide a slower evolution of the individual discretized flow perturbation quantities being estimated, thus creating an easier problem for the EnKF to estimate. Another promising idea is to explore recent hybrid methods for state estimation that consistently combine the strengths (and numerical tractability) of the EnKF and 4DVar approaches (see Cessna 2010).

Once the best estimator possible for this problem has been developed, of course, the problem of controlling a turbulent flow based on this estimate must be revisited, as well as the extension of this approach to higher Reynolds numbers. The present investigation, which represents the first reasonable high-fidelity estimate of a turbulent flow based on wall information only, represents a significant step in this direction.

The authors gratefully acknowledge the generous financial support of the National Security Education Center (NSEC) at Los Alamos National Laboratory (LANL), the many helpful conversations with Frank Alexander (LANL) and the Triton Cluster support team at the San Diego Supercomputing Center at University of California, San Diego.

REFERENCES

Anderson, B. D. O. & Moore, J. B. 1979 *Optimal Filtering*. Prentice-Hall.

Anderson, J. L. 2007 An adaptive covariance inflation error correction algorithm for ensemble filters. *Tellus* **59** (2), 210–224.

Anderson, J. L. 2009 Spatially and temporally varying adaptive covariance inflation for ensemble filters. *Tellus* A **61** (1), 72–83.

Anderson, J. L. & Anderson, S. L. 1999 A Monte Carlo implementation of the nonlinear filtering problem to produce ensemble assimilations and forecasts. *Monthly Weather Rev.* **127** (12), 2741–2758.

Arulampalam, S., Maskell, S., Gordon, N. & Clapp, T. 2002 A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking. *IEEE Trans. Signal Process.* **50**, 174–188.

Bertsekas, D. P. 2001 *Dynamic Programming and Optimal Control*, 2nd edn. Athena Scientific.

Bewley, T. R. 2001 Flow control: new challenges for a new Renaissance. *Prog. Aerosp. Sci.* **37** (1), 21–58.

Bewley, T. R. & Liu, S. 1998 Optimal and robust control and estimation of linear paths to transition. *J. Fluid Mech.* **365**, 305–349.

Bewley, T. R., Moin, P. & Temam, R. 2001 DNS-based predictive control of turbulence: an optimal benchmark for feedback algorithms. *J. Fluid Mech.* **447**, 179–225.

Bewley, T. R. & Protas, B. 2004 Skin friction and pressure: the 'footprints' of turbulence. *Physica D: Nonlinear Phenom.* **196** (1–2), 28–44.

Bewley, T. & Sharma, A. 2009 Grid-based Bayesian estimation exploiting sparsity for systems with non-Gaussian uncertainty. In *Bulletin of the American Physical Society*, vol. 54.

Bouttier, F. & Courtier, P. 1999 Data assimilation concepts and methods. *ECMWF Meteorol. Train. Course Lecture Ser.* 1–58.

Boyd, S., El Ghaoui, L., Feron, E. & Balakrishnan, V. 1994 *Linear Matrix Inequalities in System and Control Theory*. Society for Industrial Mathematics.

Butala, M. D., Yun, J., Chen, Y., Frazin, R. A. & Kamalabadi, F. 2008 Asymptotic convergence of the ensemble Kalman filter. In *15th IEEE International Conference on Image Processing*, pp. 825–828.

Cessna, J. B. 2010 The hybrid ensemble smoother (HEnS) & noncartesian computational interconnects. PhD thesis, University of California, San Diego.

Chevalier, M., Hœpffner, J., Bewley, T. R. & Henningson, D. S. 2006 State estimation in wall-bounded flow systems. Part 2. Turbulent flows. *J. Fluid Mech.* **552**, 167–187.

Evensen, G. 1994 Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *J. Geophys. Res.* **99** (C5), 143–162.

Evensen, G. 2003 The ensemble Kalman filter: Theoretical formulation and practical implementation. *Ocean Dyn.* **53** (4), 343–367.

Evensen, G. 2009*a* *Data Assimilation: The Ensemble Kalman Filter*. Springer Verlag.

Evensen, G. 2009*b* The ensemble Kalman filter for combined state and parameter estimation. *IEEE Control Syst. Magazine* **29** (3), 83–104.

Furrer, R. & Bengtsson, T. 2007 Estimation of high-dimensional prior and posterior covariance matrices in Kalman filter variants. *J. Multivariate Anal.* **98** (2), 227–255.

Gut, A. 2005 *Probability: A Graduate Course*. Springer Verlag.

Hœpffner, J., Chevalier, M., Bewley, T. R. & Henningson, D. S. 2005 State estimation in wall-bounded flow systems. Part 1. Perturbed laminar flows. *J. Fluid Mech.* **534**, 263–294.

Högberg, M., Bewley, T. R. & Henningson, D. S. 2003 Linear feedback control and estimation of transition in plane channel flow. *J. Fluid Mech.* **481**, 149–175.

Houtekamer, P. L. & Mitchell, H. L. 2001 A sequential ensemble Kalman filter for atmospheric data assimilation. *Monthly Weather Rev.* **129** (1), 123–137.

Jazwinski, A. H. 1970 *Stochastic Processes and Filtering Theory*. Academic Press.

Jimenez, J. & Moin, P. 1991 The minimal flow unit in near-wall turbulence. *J. Fluid Mech.* **225**, 213–240.

Kailath, T. 1973 Some new algorithms for recursive estimation in constant linear systems. *IEEE Trans. Inf. Theory* **19**, 750–760.

Kalman, R. E. 1960 A new approach to linear filtering and prediction problems. *J. Basic Engng* **82** (1), 35–45.

Kalman, R. E. & Bucy, R. S. 1961 New results in linear filtering and prediction theory. *J. Basic Engng* **83** (3), 95–108.

Kim, J. & Bewley, T. R. 2007 A linear systems approach to flow control. *Ann. Rev. Fluid Mech.* **39**, 383–417.

Kim, J., Moin, P. & Moser, R. 1987 Turbulence statistics in fully developed channel flow at low Reynolds number. *J. Fluid Mech.* **177** (1), 133–166.

Lorenz, E. N. 1963 Deterministic nonperiodic flow. *J. Atmos. Sci.* **20**, 130–141.

Martinelli, F. 2009 Feedback control of turbulent wall flows. PhD thesis, Politecnico di Milano.

Rauch, H. E., Tung, F. & Striebel, C. T. 1965 Maximum likelihood estimates of linear dynamic systems. *AIAA J.* **3** (8), 1445–1450.

Rogallo, R. S. 1981 Numerical experiments in homogeneous turbulence. *NASA Tech Rep.* 81315, 1–91.

Rogallo, R. S. & Moin, P. 1984 Numerical simulation of turbulent flows. *Ann. Rev. Fluid Mech.* **16** (1), 99–137.

Scherer, C., Gahinet, P. & Chilali, M. 2002 Multiobjective Output-feedback Control via LMI Optimization. *IEEE Trans. Autom. Control* **42** (7), 896–911.

Slutsky, E. 1925 Über stochastische Asymptoten und Grenzwerte. *Metron* **5** (3), 3–90.

Wang, X. & Bishop, C. H. 2003 A comparison of breeding and ensemble transform Kalman filter ensemble forecast schemes. *J. Atmos. Sci.* **60** (9), 1140–1158.

Zhou, K., Doyle, J. C. & Glover, K. 1996 *Robust and Optimal Control*. Prentice Hall.