Rapid Communications

Optimal nonlinear filtering using the finite-volume method

Colin Fox,[1,*] Malcolm E. K. Morrison,[1] Richard A. Norton,[2] and Timothy C. A. Molteno[1]
[1]*Department of Physics, University of Otago, Dunedin, New Zealand*
[2]*Department of Mathematics and Statistics, University of Otago, Dunedin, New Zealand*

(Received 23 July 2017; published 3 January 2018)

Optimal sequential inference, or filtering, for the state of a deterministic dynamical system requires simulation of the Frobenius-Perron operator, that can be formulated as the solution of a continuity equation. For low-dimensional, smooth systems, the finite-volume numerical method provides a solution that conserves probability and gives estimates that converge to the optimal continuous-time values, while a Courant-Friedrichs-Lewy-type condition assures that intermediate discretized solutions remain positive density functions. This method is demonstrated in an example of nonlinear filtering for the state of a simple pendulum, with comparison to results using the unscented Kalman filter, and for a case where rank-deficient observations lead to multimodal probability distributions.

DOI: 10.1103/PhysRevE.97.010201

I. INTRODUCTION

Consider a nonterminating dynamical system that evolves according to the differential equation

$$\dot{\boldsymbol{x}} = \boldsymbol{f}(\boldsymbol{x}), \tag{1}$$

where $\boldsymbol{f}$ is a known velocity field and $\boldsymbol{x}(t)$ is the state vector of the system at time t. Given an initial condition, $\boldsymbol{x}(0) = \boldsymbol{x}_0$, Eq. (1) may be simulated, at least in concept, to determine the future state $\boldsymbol{x}(t)$, $t > 0$, that we will write $\boldsymbol{x}(t;\boldsymbol{x}_0)$ to explicitly show the dependence on the initial condition. However, when the initial state of the system $\boldsymbol{x}_0$ is unknown, or uncertain, the future state $\boldsymbol{x}(t)$ is also uncertain.

At increasing discrete times t_k, $k = 1,2,3,\ldots$, the system is observed, with observation $\boldsymbol{z}_k$ at time t_k providing noisy and incomplete information about the state $\boldsymbol{x}_k = \boldsymbol{x}(t_k)$. We assume that the conditional distribution over the observed value $\boldsymbol{z}_k$, given the state vector $\boldsymbol{x}_k$,

$$\rho(\boldsymbol{z}_k|\boldsymbol{x}_k), \tag{2}$$

is known. Let $Z_t = \{\boldsymbol{z}_k : t_k \leqslant t\}$ denote the set of observations up to time t, and $\boldsymbol{x}_t = \boldsymbol{x}(t)$ for brevity.

This Rapid Communication is concerned with performing sequential Bayesian inference for the unknown true state of the continuous-time system $\boldsymbol{x}_t$ at times $t \geqslant 0$, in which knowledge of the state vector is improved upon through the multiple imperfect observations and knowledge of the dynamics. Uncertainty in the state is modeled as a probability distribution over the state and the formal solution corresponds to determining the time-varying sequence of distributions

$$\rho(\boldsymbol{x}_t|Z_t) \tag{3}$$

over the state at time t conditioned on all available measurements to time t. We assume throughout that all distributions have a probability density function (pdf), and will use the same notation for the pdf as the distribution.

*Corresponding author: fox@physics.otago.ac.nz

While the focus of this Rapid Communication is on state estimation, parameter estimation may also be performed within this framework by augmenting the system equations (1) to include uncertain parameters that do not evolve.

An approach that one often sees [1–4] is to discretize (1), and treat the discrete-time problem [5]. When uncertainty in $\boldsymbol{f}$ is included via "process noise" $\boldsymbol{v}_k$, the discrete-time system and observation equations are written [3]

$$\boldsymbol{x}_k = f_k(\boldsymbol{x}_{k-1},\boldsymbol{v}_k), \tag{4}$$

$$\boldsymbol{z}_k = h_k(\boldsymbol{x}_k,\boldsymbol{n}_k), \tag{5}$$

respectively, with functions f_k and h_k assumed known. Here, $\boldsymbol{n}_k$ is a realization from a (random) noise process that accounts for observation errors. When the random processes $\boldsymbol{v}_k$ and $\boldsymbol{n}_k$ are independently distributed from the current and previous states, the system equation (4) defines a Markov process, as does (1), while the observation equation (5) defines the conditional probability (2).

We take a different approach here, which is to treat the continuous-time problem in (1) and (2) directly, and to define a family of numerical approximations that converge in distribution to the continuous-time pdf's (3).

This Rapid Communication is organized as follows. Recursive Bayesian filtering is outlined in Sec. II, and then Sec. III derives a continuity equation for probability flow implied by the dynamical system. That partial differential equation (PDE) is solved using the finite-volume method, outlined in Sec. IV. Some specifics for filtering the state of a simple pendulum are presented in Sec. V, with numerical results in Secs. VI and VII for unimodal and multimodal distributions, respectively. We conclude with a brief discussion in Sec. VIII.

II. RECURSIVE BAYESIAN FILTERING

The family of time-varying pdf's (3) may be generated by iterating two operations corresponding to Z_t remaining constant between measurements times, i.e., $t \in (t_k,t_{k+1})$, requiring *prediction* of the pdf, and the change in Z_t at measurement

times t_k that requires an *update* of the pdf [6]. Practical filtering also requires an *inference* step to calculate summary statistics over the pdf, though details are application specific.

The prediction step may be derived from the (conditional) Chapman-Kolmogorov equation that, for $\Delta t \geqslant 0$,

$$\rho(\boldsymbol{x}_{t+\Delta t}|Z_t) = \int \rho(\boldsymbol{x}_{t+\Delta t}|\boldsymbol{x}_t, Z_t)\rho(\boldsymbol{x}_t|Z_t)d\boldsymbol{x}_t$$
$$= \int \delta(\boldsymbol{x}_{t+\Delta t} - \boldsymbol{x}(\Delta t; \boldsymbol{x}_t))\rho(\boldsymbol{x}_t|Z_t)d\boldsymbol{x}_t, \quad (6)$$

where δ is the Dirac delta, and we have used the deterministic solution to the autonomous system (1). Straightforward generalizations of (6) can be written for stochastic and nonautonomous system equations. Note that (6) defines a linear operator on the space of probability distributions

$$S_{\Delta t} : \rho(\boldsymbol{x}_t|Z_t) \mapsto \rho(\boldsymbol{x}_{t+\Delta t}|Z_t), \quad (7)$$

that is the element of the continuous semigroup of Frobenius-Perron (F-P) operators, associated with the system (1) [7], indexed by the time increment Δt.

The update step is simply Bayes rule written at observation time t_k [1,2,6],

$$\rho(\boldsymbol{x}_k|Z_k) = \frac{\rho(\boldsymbol{z}_k|\boldsymbol{x}_k)\rho(\boldsymbol{x}_k|Z_{k-1})}{\rho(\boldsymbol{z}_k|Z_{k-1})}, \quad (8)$$

in which $Z_k = Z_{t_k}$, and conditional independence of $\boldsymbol{z}_k$ and Z_{k-1} given $\boldsymbol{x}_k$ has been used.

The recursive Bayesian solution for *discrete* time then consists of repeatedly applying the prediction step $S_{t_{k+1}-t_k} : \rho(\boldsymbol{x}_k|Z_k) \mapsto \rho(\boldsymbol{x}_{k+1}|Z_k)$ and the update step (8) at time t_{k+1}, initialized at the distribution over $\boldsymbol{x}_0$ that we assume is known, or can be adequately modeled. F-P operators indexed by *continuous*-time increments $\Delta t \in (0, t_{k+1} - t_k)$ allow inference for the state in continuous time, i.e., continuous-time Bayesian filtering.

It interesting to note that steps (6) and (8) determine the distributions in (3) for times $t' > t$ from the distribution $\rho(\boldsymbol{x}_t|Z_t)$ and subsequent measurements $\boldsymbol{z}_l$ for $t_l > t$, only. Hence, it is only necessary to determine the $\boldsymbol{x}_t$ dependence of the distribution $\rho(\boldsymbol{x}_t|Z_t)$ at each t, and the measurements Z_t play no further role.

Sequential inference methods on dynamical systems differ in four main aspects: assumptions about the dynamical system (1), representation of the pdf's (3), simulating the F-P operator (7), and evaluating the Bayes rule update (8). Algorithms that implement these aspects essentially exactly are called "optimal," whereas appreciably approximate methods are termed "suboptimal" [3]. While the prediction and update equations for optimal Bayesian filters in both discrete time and continuous time have been known for half a century [6], implementation of optimal nonlinear filtering still presents challenges.

Perhaps the best known and most successful filter is the Kalman filter that is optimal for the discrete-time linear-Gaussian case, i.e., when functions f_k and h_k in (4) and (5) are linear and all distributions are Gaussian. The need for filters that can operate with nonlinear dynamics and more general distributions has led to many variants of the Kalman filter, such as the many varieties of the extended Kalman filter (EKF) that linearize the dynamics at each discrete time [8]. The unscented Kalman filter (UKF), based on the unscented transform [8,9], requires only that the pdf's remain close to Gaussian.

The filter presented here is an example of a grid-based method [3], that represents the pdf (3) over a grid of points, or mesh, and implements numerical prediction and update steps using that representation. The present filter approximates the continuous-space (and continuous-time) problem, to exploit the smoothness of functions and dynamics. As we show here, fixed-mesh filters can provide practical optimal filtering in low-dimensional settings.

III. CONTINUITY EQUATION

A partial differential equation (PDE) form of the F-P operator (7) may be derived directly from (6) (see, e.g., Ref. [6]), though it is instructive to derive this PDE using just elementary principles, as follows. Consider a time-varying probability density $\rho(\boldsymbol{x}; t)$ over the state of a system that evolves as (1), with an initial value

$$\rho(\boldsymbol{x}; 0) = \rho_0(\boldsymbol{x}), \quad (9)$$

and consider an infinitesimal volume in state space $d\boldsymbol{x}$ at location $\boldsymbol{x}$ and time t. Probability mass is conserved during evolution of the system, so it is natural to define a *flux* of probability mass equal to $\rho(\boldsymbol{x}; t)\boldsymbol{f}(\boldsymbol{x})$. For smooth ρ and $\boldsymbol{f}$, the rate of probability mass entering the element is then $-\nabla \cdot [\rho(\boldsymbol{x}; t)\boldsymbol{f}(\boldsymbol{x})]|d\boldsymbol{x}|$. Since the volume of the element $|d\boldsymbol{x}|$ is not changing, the probability density function evolves as

$$\dot{\rho} = -\nabla \cdot (\rho \boldsymbol{f}), \quad (10)$$

which is the continuity equation, well known in many branches of physics. A derivation of (10) via the adjoint (Koopman) operator is given in Ref. [7].

The F-P operator (7) then corresponds to solving the initial value problem (IVP) (9) and (10) with $\rho_0(\boldsymbol{x}) = \rho(\boldsymbol{x}_t|Z_t)$ to evaluate $\rho(\boldsymbol{x}; \Delta t) = \rho(\boldsymbol{x}_{t+\Delta t}|Z_t)$.

The continuity equation (10) is a linear advection equation. When the state equation has additive stochastic forcing, evolution of the pdf is governed by a linear advection-diffusion (Fokker-Planck) equation [7].

The filter we develop here uses an efficient numerical solver, developed precisely for conservation laws, such as (10), that we describe next.

IV. FINITE-VOLUME METHOD

The finite-volume method (FVM) [10] discretizes the continuity equation (10) in its integral form,

$$\frac{\partial}{\partial t}\int_K \rho\, d\boldsymbol{x} + \oint_{\partial K} \rho(\boldsymbol{f} \cdot \hat{\boldsymbol{n}})dS = 0, \quad (11)$$

that holds for each volume K in state space. Here, $\hat{\boldsymbol{n}}$ is the unit outwards facing normal vector on the boundary ∂K, and dS is the surface area element.

Denote state space by X, and define a mesh T of X, that is a family of disjoint regions such that $X = \cup_{K \in T} \overline{K}$. Each such region K is called a "cell," or control volume, in the FVM. We write $L \sim K$ if cells L and K share a common interface, denoted E_{KL}, that we assume is a subset of a hyperplane. Denote by $\hat{\boldsymbol{n}}_{KL}$ the unit normal on E_{KL} directed from K to L.

The FVM we use evaluates a numerical approximation to the solution of the IVP (9) and (10) in the following way. At each time t the pdf's (3) are represented by piecewise constant functions, constant on each cell K. Define the initial vector $\boldsymbol{P}^0$ of cell values by

$$\boldsymbol{P}_K^0 = \frac{1}{|K|}\int_K \rho_0(\boldsymbol{x})d\boldsymbol{x}, \tag{12}$$

then for $m = 0,1,\ldots,r$ compute $\boldsymbol{P}^{m+1}$ as [11]

$$\frac{\boldsymbol{P}_K^{m+1} - \boldsymbol{P}_K^m}{\Delta t} + \frac{1}{|K|}\sum_{L\sim K} f_{KL}\boldsymbol{P}_{KL}^m = \boldsymbol{0}, \tag{13}$$

where

$$f_{KL} = \int_{E_{KL}} \boldsymbol{f}\cdot\hat{\boldsymbol{n}}_{KL}\,dS \tag{14}$$

gives the normal velocity on the interface between cells K and L, and $\boldsymbol{P}_{KL}^m$ denotes the first-order *upwinding* scheme

$$\boldsymbol{P}_{KL}^m = \begin{cases} \boldsymbol{P}_K^m, & \text{if } f_{KL} \geqslant 0, \\ \boldsymbol{P}_L^m, & \text{if } f_{KL} < 0. \end{cases} \tag{15}$$

In matrix form, this is

$$\boldsymbol{P}^{m+1} = (\boldsymbol{I} - \Delta t\boldsymbol{A})\boldsymbol{P}^m, \tag{16}$$

where $\boldsymbol{I}$ is the identity matrix and $\boldsymbol{A}$ is a sparse matrix defined by (13)–(15). Since $f_{KL} = -f_{LK}$, the FVM conserves probability at each step, i.e., $\sum_K |K|\boldsymbol{P}_K^{m+1} = \sum_K |K|\boldsymbol{P}_K^m$. The FVM step (13) also preserves positivity when the time step Δt is small enough that the matrix $\boldsymbol{I} - \Delta t\boldsymbol{A}$ has all non-negative entries. For regular cubic meshes [12] with edge length h in d dimensions that requirement may be written as the Courant-Friedrichs-Lewy (CFL)-type condition

$$\Delta t \leqslant \min_K \frac{h^d}{\sum_{L\sim K}\max(0, f_{KL})}. \tag{17}$$

The F-P operator (7) is simulated by discretizing $\rho(\boldsymbol{x}_t|Z_t)$ using (12), then iterating (13). Thus, our numerical discrete-time F-P operator is the matrix

$$S_{t_{k+1}-t_k} = (\boldsymbol{I} - \Delta t\boldsymbol{A})^r, \tag{18}$$

where r is the largest integer with $r\Delta t = t_k - t_{k-1}$ that satisfies (17). Weak convergence of the FVM means that the piecewise-constant pdf defined by cell values $\boldsymbol{P}^m$ converges in distribution to the true pdf (3) as the mesh is refined and the time step decreases [13], and expectations over the discrete pdf converge to the correct continuous-time continuous-space values.

V. FILTERING THE STATE OF A PENDULUM

Let θ and ω be the angular displacement (from the vertical, downwards) and angular velocity, respectively, of a simple pendulum. The dynamics is described by (1) with the state $\boldsymbol{x} = (\theta,\omega)$ and the velocity field

$$\boldsymbol{f}(\theta,\omega) = \left(\omega, -\frac{g}{l}\sin(\theta)\right), \tag{19}$$

where g is the acceleration due to gravity, l the length of the pendulum.

Phase space is the region $(\theta,\omega) \in S^1 \times \mathbb{R}$. For computational purposes, the range of ω is restricted to $(-\omega_{\max},\omega_{\max})$ chosen so that the pdf's are always numerically zero at these boundaries. We set $\boldsymbol{f}\cdot\hat{\boldsymbol{n}} = \boldsymbol{0}$ on the ω boundaries to ensure conservation of probability.

We discretize the region $(-\pi,\pi]\times(-\omega_{\max},\omega_{\max}) \subset \mathbb{R}^2$ with a uniform square mesh of $n \times n$ cells of width $\Delta\theta = \Delta\omega$ (t scaled accordingly), with the (i,j) cell having center (θ_i,ω_j). The normal velocity (14) between neighboring cells with different θ values is

$$f_{(i,j)(i+1,j)} = \int_{\omega_j-\frac{1}{2}\Delta\omega}^{\omega_j+\frac{1}{2}\Delta\omega} \omega\,d\omega$$

(index i is evaluated modulo n), while for neighboring cells with different ω values,

$$f_{(i,j)(i,j+1)} = -\frac{g}{l}\int_{\theta_i-\frac{1}{2}\Delta\theta}^{\theta_i+\frac{1}{2}\Delta\theta} \sin\theta\,d\theta.$$

Together with the Bayes update (8), this defines the recursive Bayesian filter that we call the finite-volume filter (FVF). In the following numerical examples, we set the constants g, m, and l to be unity, for simplicity.

VI. RESULTS FOR FVF AND UKF

In this section we present filtering for the state of the simple pendulum from noisy observations of the state, comparing results for the FVF with the UKF as presented in Ref. [9]. Initial and noise distributions are Gaussian and pdf's over state remain unimodal, hence the UKF is applicable. We compare the filters for various initial positions, observation frequencies, and noise variance.

The "true" path of the pendulum was computed with initial position and velocity $(\theta_0,0)$ using a Runge-Kutta 4-5 to solve Eq. (1). Noisy observations were simulated by adding Gaussian noise to the true state at evenly spaced discrete observation times giving the observation model $\boldsymbol{z}_k = \boldsymbol{I}\boldsymbol{x}_k + \boldsymbol{n}_k$ at times t_k, where $\boldsymbol{n}_k \overset{\text{iid}}{\sim} N(0,\sigma_z^2\boldsymbol{I})$, and likelihood function

$$\rho(z_k|\boldsymbol{x}_k) \propto \exp\left\{-\frac{\|\boldsymbol{z}_k - \boldsymbol{x}_k\|^2}{2\sigma_z^2}\right\}.$$

The FVF was implemented with $n = 200$ intervals in each angular displacement and angular velocity, chosen so that run times of the FVF and UKF are approximately the same [14]. Both UKF and FVF used the same observations and initial Gaussian $N((\theta_0,0),\sigma_0^2\boldsymbol{I})$ with $\sigma_0 = 0.4$. We found it necessary to introduce added Gaussian process noise in the UKF to prevent prediction drifting from the true solution, while not overweighting observations. This process noise is required to compensate for the approximations to pdf's used by the UKF. We found a standard deviation of 0.01 provided

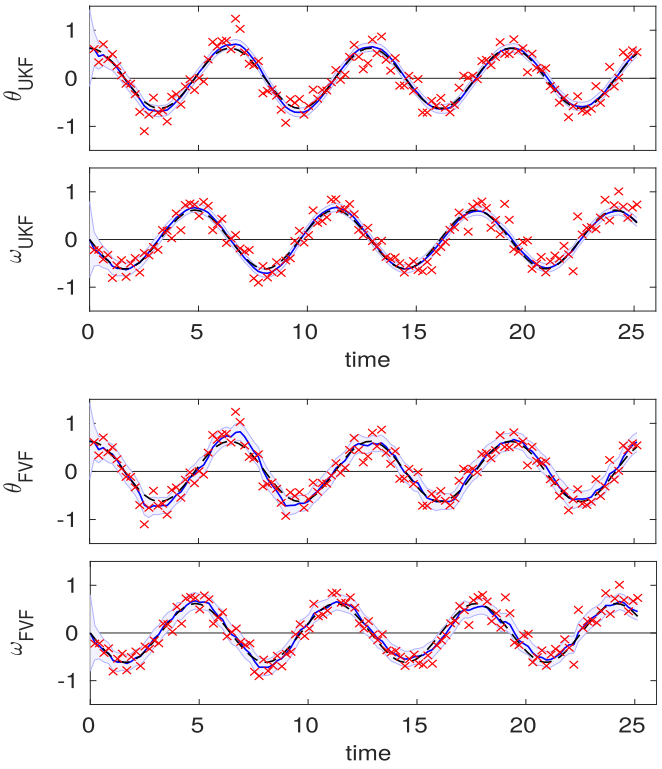

FIG. 1. Filter results for $\theta_0 = 0.2\pi$ with 30 observations per 2π time, with $\sigma_z = 0.2$. Each UKF (upper pair) and FVF (lower pair) shows a true and filtered angular displacement θ and angular velocity ω. Predicted mean (blue solid), 95% region (shaded), observations (red crosses), and true solution (black dashed).

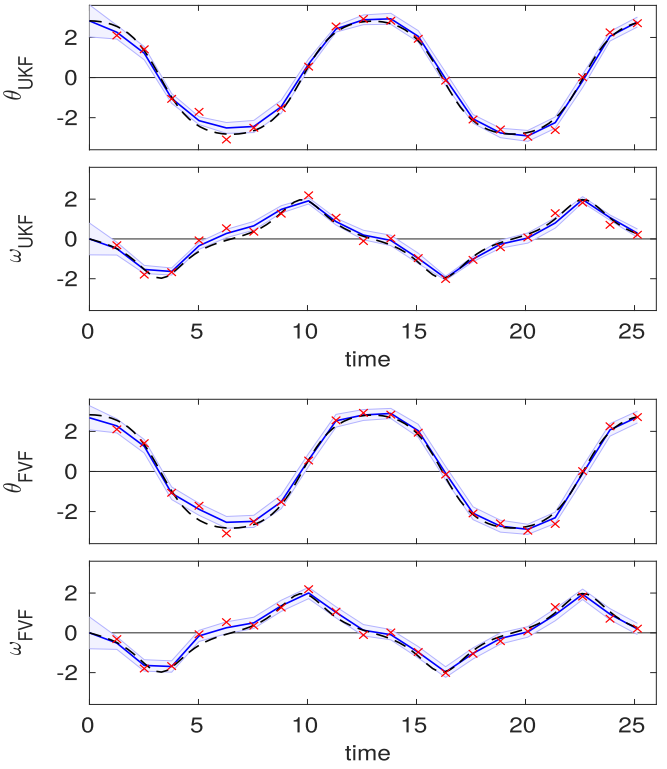

FIG. 2. Filter results for $\theta_0 = 0.9\pi$ with five observations per 2π time, having $\sigma_z = 0.2$. Equivalent plots to Fig. 1.

the best qualitative results, and was used in all examples here.

Figure 1 shows summary results for the case $\theta_0 = 0.2\pi$ with 30 observations per 2π. This small-angle and frequent observations setting has dynamics that is close to linear, so pdf's remain close to Gaussian and the UKF is close to optimal. As can be seen in Fig. 1, both UKF and FVF perform well in this example, with both tracking the state of the pendulum.

We also computed the filtering for a very long time series (not shown), and computed the root-mean-square error (RMSE) in each of the state variables, asymptotic in time. These values are presented in Table I, and provide a quantitative comparison of the UKF and FVF ($n = 200$) filters. The first column of Table I shows that the UKF achieves about half the RMSE of the FVF with $n = 200$, and therefore the UKF is preferable in this setting.

TABLE I. Asymptotic RMSE in (θ,ω) for the filtering examples shown in Figs. 1– 3.

	As Fig. 1	As Fig. 2	As Fig. 3
UKF	(0.030,0.027)	(0.13,0.088)	(0.051,0.038)
FVF ($n = 200$)	(0.069,0.064)	(0.15,0.11)	(0.035,0.028)

The mesh size used in the FVM is $2\pi/200 = 0.0314$ and this "pixel" resolution gives a lower bound on error when the support of the pdf does not cover many pixels. Since the FVM is convergent as $n \to \infty$, the FVF will approach optimal with increasing n [13], with RMSE that will decrease to that of the UKF, or smaller, though at greatly increased computational cost.

Figure 2 shows filtering for larger angles with $\theta_0 = 0.9\pi$, and only five observations per 2π time. Again, both UKF and FVF perform well. The second column of Table I shows that the UKF has the smaller RMSE. Even though the dynamics is more nonlinear than the previous example, the relatively small noise level ensures that pdf's remain compact and the Gaussian assumption of the UKF performs a little better than the piecewise-constant approximation of the FVF with $n = 200$.

Figure 3 shows both filters running with a stationary pendulum, vertically down, with 50 observations having $\sigma_z = 0.4$. This is a case with larger observation noise, and we see that the UKF does not perform as well as the FVF ($n = 200$). Figure 3 (upper) shows that the predicted mean provided by the UKF continues to jump around the true value in both state variables, and the 95% bounds repeatedly do not include the true value. In contrast, Fig. 3 (lower) shows the FVF consistently predicting the correct value, with 95% bounds that include the true value. The third column of Table I confirms that the FVF has smaller RMSE than the UKF, and so it is to be preferred in this larger-noise setting.

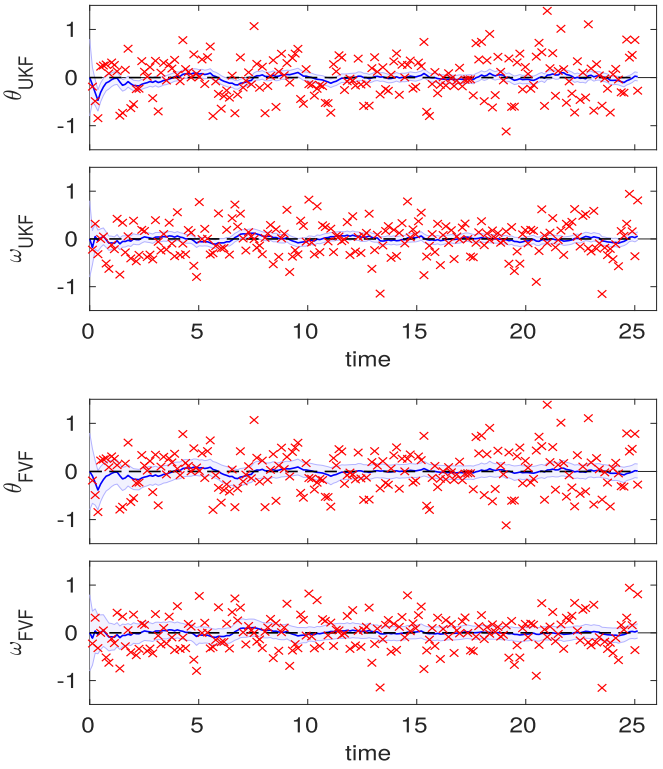

FIG. 3. Filter results for $\theta_0 = 0$ with 50 observations per 2π time, having $\sigma_z = 0.4$. Equivalent plots to Fig. 1.

VII. FILTERING WITH MULTIMODAL pdf's

We now give an example of filtering for the state of the simple pendulum where observations are made of the tension force in the string. Since tension is uninformative about the sign of angular displacement and velocity, these observations lead to multimodal filtering distributions. The UKF, or any extension to the Kalman filter that assumes Gaussian pdf's, is unable to accurately represent these multimodal distributions. However, the FVF remains applicable, and optimal for large enough n.

The tension force on a pendulum is

$$F(\boldsymbol{x}) = ml\omega^2 + mg\cos(\theta). \tag{20}$$

We simulated measurements by simulating the motion of the pendulum with initial position $(\theta_0,\omega_0) = (0.2\pi,0)$, then recording eight values of the tension, per 2π time, with added Gaussian noise of standard deviation $\sigma = 0.2$.

The FVF with $n = 200$ was initialized with the Gaussian distribution $N((0,0),\sigma_0^2\boldsymbol{I})$ with $\sigma_0 = 0.8$. Figure 4 shows four snapshots of the filtered probability density over phase space produced by the FVF.

Since the filtering pdf's are symmetric about the origin, the UKF will approximate all these distributions by zero-mean Gaussians and estimate the state as identically zero, for all time. Clearly, this is uninformative.

In contrast, the FVF has localized the true state after 3π time (about 1.5 periods), albeit with ambiguity in sign. Properties of the system that do not depend on the sign of the state, such as the period, can then be accurately estimated.

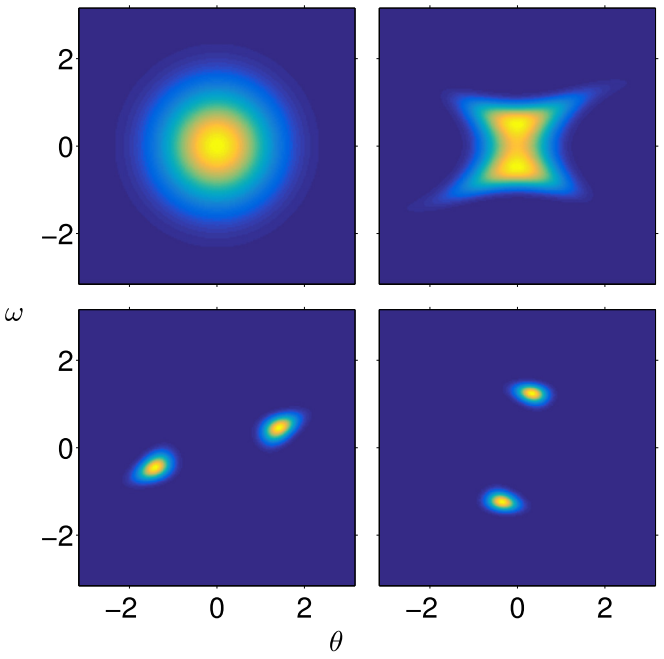

FIG. 4. Initial ($t = 0$) and filtered pdf's in phase space after measurements at times $t = \pi/4$, π, and 3π (left to bottom).

VIII. CONCLUDING REMARKS

This Rapid Communication presents a method for performing sequential inference on dynamical systems that is both optimal and practical. We have shown that the Frobenius-Perron operator can be simulated by solving the continuity equation for probability using a finite-volume method. Computed examples showed that the FVF is preferable to the UKF in a nonlinear system with large noise or rank-deficient observations.

A few earlier works have implemented filters by also forming and integrating the continuity equation. An alternating direction iterative solver was used in Ref. [15], Ref. [16] solved along characteristics, while Ref. [17] estimated the matrix form of the discrete integrator using Monte Carlo simulations. The FVF developed here has a clear advantage over the methods in Refs. [15,16] by guaranteeing the Markov property, hence intermediate results are valid probability distributions, and is provably convergent [13], implying that estimates over discrete filtering pdf's converge to the correct continuous-time values, as discretization is refined. In contrast, computed examples in Refs. [15,16] show resulting numerical problems in examples that were not carefully chosen. The method in Ref. [17] guarantees a Markov operator, however, the FVF provides a much more efficient, deterministic route to evaluating the required transition matrix.

We developed the FVF as a "reference" filter for our ongoing work in dynamic weighing, for which we wish to accurately determine the weight of an object from force measurements on the moving object, and motion modeling. The example in Sec. VII is a stylized version of that problem. We were pleasantly surprised to find that the FVF easily fits into the computational capability of current desktop or embedded computers, for systems with few parameters.

We agree with Daum [18] that recent adaptive and adjoint meshing methods [19] could alleviate the growth of cost with dimension of the FVF, while maintaining the computational benefits of exploiting smoothness.

ACKNOWLEDGMENTS

This work was supported by the Royal Society of New Zealand's Marsden Fund Contract No. UOO1015, and MBIE Contract No. UOOX1208.

[1] R. S. Bucy, J. Astronaut. Sci. **17**, 80 (1969).
[2] N. Gordon, D. Salmond, and A. Smith, IEEE Proc. F **140**, 107 (1993).
[3] M. S. Arulampalam, S. Maskell, N. Gordon, and T. Clapp, IEEE Trans. Signal Process. **50**, 174 (2002).
[4] K. Judd and T. Stemler, Phys. Rev. E **79**, 066206 (2009).
[5] O. Cappé, E. Moulines, and T. Ryden, *Inference in Hidden Markov Models* (Springer, New York, 2005).
[6] A. H. Jazwinski, *Stochastic Processes and Filtering Theory* (Academic, New York, 1970).
[7] A. Lasota and M. C. Mackey, *Chaos, Fractals, and Noise*, 2nd ed. (Springer, New York, 1994).
[8] F. Daum, IEEE Aerospace Electron. Syst. Mag. **20**, 57 (2005).
[9] S. J. Julier and J. K. Uhlmann, Proc. SPIE **3068**, 182 (1997).
[10] H. Versteeg and W. Malalasekera, *An Introduction to Computational Fluid Dynamics: The Finite Volume Method* (Pearson Education, Boston, 2007).
[11] An explicit Euler step is implemented in (13). Other implementations of the FVM, including higher-order time steps, are presented in Ref. [10].
[12] Unstructured meshes are treated in Ref. [13].
[13] R. A. Norton, C. Fox, and M. E. Morrison, arXiv:1610.02106.
[14] Run times were within 10% when implemented in MATLAB with vectorized code.
[15] K. Kastella and C. Kreucher, IEEE Trans. Aerospace Electron. Syst. **41**, 549 (2005).
[16] P. Dutta and R. Bhattacharya, J. Guid. Control Dyn. **34**, 325 (2011).
[17] S. Ungarala, Z. Chen, and K. Li, Ind. Eng. Chem. Res. **45**, 4208 (2006).
[18] F. Daum, Proc. SPIE **5913**, 59130D (2005).
[19] W. Huang and R. D. Russell, *Adaptive Moving Mesh Methods* (Springer, Berlin, 2011).