Filtering Dynamical Systems Using Observations of Statistics

Eviatar Bach,[1,2,a] Tim Colonius,[3] Isabel Scherl,[3] and Andrew Stuart[2]
[1] *Department of Environmental Science and Engineering, California Institute of Technology, Pasadena, California 91125, USA*
[2] *Department of Computing and Mathematical Sciences, California Institute of Technology, Pasadena, California 91125, USA*
[3] *Department of Mechanical and Civil Engineering, California Institute of Technology, Pasadena, California 91125, USA*

(Dated: 19 January 2024)

We consider the problem of filtering dynamical systems, possibly stochastic, using observations of statistics. Thus the computational task is to estimate a time-evolving density $\rho(v,t)$ given noisy observations of the true density $\rho^\dagger$; this contrasts with the standard filtering problem based on observations of the state v. The task is naturally formulated as an infinite-dimensional filtering problem in the space of densities ρ. However, for the purposes of tractability, we seek algorithms in state space; specifically we introduce a mean-field state-space model and, using interacting particle system approximations to this model, we propose an ensemble method. We refer to the resulting methodology as the ensemble Fokker–Planck filter (EnFPF).

Under certain restrictive assumptions we show that the EnFPF approximates the Kalman–Bucy filter for the Fokker–Planck equation, which is the exact solution of the infinite-dimensional filtering problem. Furthermore, our numerical experiments show that the methodology is useful beyond this restrictive setting. Specifically, the experiments show that the EnFPF is able to correct ensemble statistics, to accelerate convergence to the invariant density for autonomous systems, and to accelerate convergence to time-dependent invariant densities for non-autonomous systems. We discuss possible applications of the EnFPF to climate ensembles and to turbulence modelling.

Data assimilation (DA) is the process of estimating the state of a dynamical system using observations. Here, we modify the standard DA setting to allow for observations of *statistics* of a system with respect to its time-evolving probability density. We propose a mathematical framework, a resulting ensemble method, and present numerical experiments demonstrating accelerated convergence of a system to its attractor. We propose further applications to problems in climate and turbulence modelling.

I. INTRODUCTION

The goal of this paper is to introduce a filtering methodology that incorporates statistical information into a (possibly stochastic) dynamical system. In the next three subsections, we present, respectively, a high-level overview of the problem, discuss the motivation and previous literature, and outline the paper structure and our contributions.

A. Assimilating Statistical Observations

We start by presenting a high-level overview of the problem of incorporating statistical information into a dynamical system; a detailed problem statement follows in section II A.

Data assimilation (DA) is overviewed in a number of books, including[1–4]. The problem is to estimate the state of a dynamical system by combining noisy, partial observations with a model for the system. In the continuous-time DA problem, we have a stochastic differential equation (SDE)

$$dv^\dagger = f(v^\dagger, t)\,dt + \sqrt{\Sigma(t)}\,dW, \tag{I.1}$$

$$v^\dagger(0) = v_0^\dagger, \tag{I.2}$$

with solution $v^\dagger \in \mathbb{R}^d$, and observations given by

$$dz^\dagger = h(v^\dagger(t), t)\,dt + \sqrt{\Gamma(t)}\,dB, \tag{I.3}$$

with $z^\dagger \in \mathbb{R}^p$. The equations for $v^\dagger$ and $z^\dagger$ are driven by independent standard Wiener processes W and B. These SDEs, as with all the SDEs in the paper, are to be interpreted in the Itô sense. Filtering is then the problem of obtaining the best possible estimate of the posterior density on $v^\dagger(t)$ given the past observations $\{z^\dagger(s)\}_{s\in[0,t]}$. Throughout the paper, we use the † superscript to indicate the true quantities, and omit it for filtered quantities.

Instead of observing a specific trajectory of a dynamical system, as $\{z^\dagger(t)\}$ given by Eq. (I.3) does, one can also consider *observations of the system's statistical behavior*, that is, observations of functionals of the probability density $\rho^\dagger(v,t)$ over trajectories. This density reflects the randomness from the initial conditions for v and/or from the Brownian forcing. For a deterministic dynamical system ($\Sigma \equiv 0$), if the initial conditions are random, then

[a] Electronic mail: eviatarbach@protonmail.com

$\rho^\dagger(v,t)$ will reflect the changing density over time under the action of the system's dynamics, governed by the Liouville equation.[1] If noise is present, the changing density is also affected by the Brownian noise W and is governed by the Fokker–Planck equation, a diffusively-regularized Liouville equation. In this paper we focus on observations of $\rho^\dagger(v,t)$ defined by replacing Eq. (I.3) with

$$dz^\dagger = \left(\int \mathfrak{h}(v,t)\rho^\dagger(v,t)\,dv\right)dt + \sqrt{\Gamma(t)}\,dB. \tag{I.4}$$

Here $\mathfrak{h}(v,t)$ defines the observed statistics of v, B is a Wiener process, and $z^\dagger \in \mathbb{R}^p$. The filtering problem is to estimate a density $\rho(v,t)$ given all the past observations $\{z^\dagger(s)\}_{s\in[0,t]}$. As in the observation equation (I.3), the observations are finite-dimensional, noisy, and partial. However, since the observations are now of $\rho^\dagger(v,t)$ instead of $v^\dagger(t)$, we must specify the dynamics of $\rho^\dagger(v,t)$. This is given by the Fokker–Planck (FP) or Kolmogorov forward equation:

$$\frac{\partial\rho^\dagger}{\partial t} = \mathcal{L}^*(t)\rho^\dagger, \tag{I.5a}$$

$$\mathcal{L}^*(t)\psi = -\nabla\cdot(\psi f) + \frac{1}{2}\nabla\cdot\big(\nabla\cdot(\psi\Sigma)\big), \tag{I.5b}$$

where $\mathcal{L}^*$ is the adjoint of the generator of Eq. (I.1).[2]. For a deterministic system, with $\Sigma\equiv 0$, the Fokker–Planck equation reduces to the Liouville equation.

An important question is how one would obtain observations of a system's statistics for problems of practical relevance. We discuss this in detail in I Ba. For now we proceed on the assumption that $z^\dagger$ solving Eq. (I.4) is given.

Now, Eqs. (I.5) and (I.4) define a filtering problem for $\rho(v,t)$. This is an infinite-dimensional filtering problem, in contrast to the finite-dimensional filtering problem for $v(t)$ defined by Eqs. (I.1) and (I.3). We refer to the filtering problem defined by Eqs. (I.5) and (I.4) as the *Fokker–Planck filtering problem*. Note that both Eqs. (I.5) and (I.4) are *linear* in $\rho^\dagger$, meaning that the solution of the problem can be written using the infinite-dimensional Kalman–Bucy (KB) filter; see subsection IV A for more details.

Despite the existence of an exact solution to the filtering problem, through the infinite-dimensional Kalman–Bucy (KB) filter, approximating the Gaussian conditional density ρ is in most setting computationally intractable since the mean is a probability density function and the covariance is an operator. Thus we seek inspiration from the success of ensemble Kalman filtering[7]: we work in state space and seek an ensemble that evolves in time a number of states whose empirical density approximates the filtered ρ. We note that the particle filter similarly substitutes the problem of evolving a probability density with that of evolving a number of particles and weights[8]. Furthermore, derivation of ensemble Kalman methods via a mean-field limit provides a systematic methodology for the derivation of equal-weight approximate filters[9]. We call the resulting method the *ensemble Fokker–Planck filter* (EnFPF).

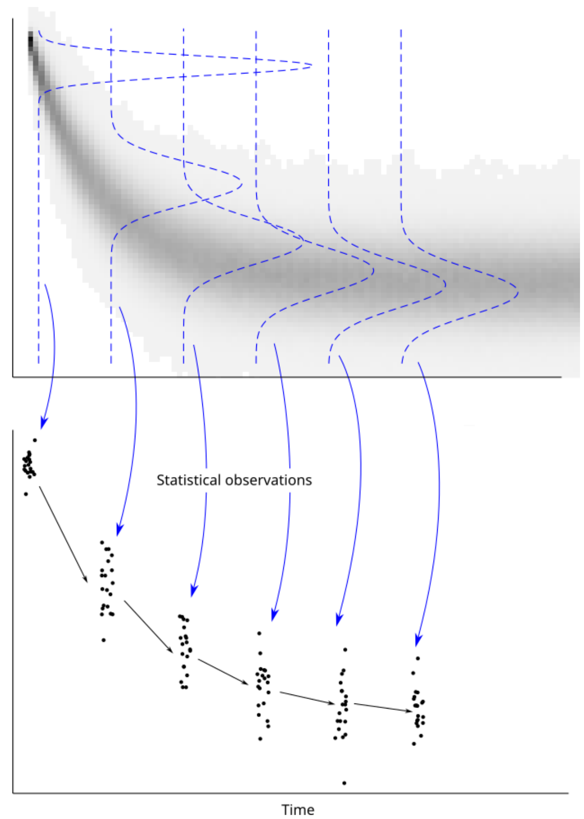

FIG. 1: The density of an Ornstein–Uhlenbeck process evolving in time (top panel). At regular intervals, we make observations of this density and use them to inform the evolution of an ensemble (bottom panel).

Figure 1 shows a schematic of such an ensemble method. In the top panel is the true time-varying probability density, in this case of an Ornstein–Uhlenbeck process. In the bottom panel is an ensemble of states. At regular intervals, we observe expectations over the density in the top panel. Using these observations and our model of the system, we evolve the ensemble over the time interval between the current and next observations.

---

[1] Here we use the term Liouville equation for the equation governing evolution of the density of any ordinary differential equation, not just in the Hamiltonian setting.

[2] We define the divergence of a matrix as is standard in continuum mechanics; see Gurtin (1981)[5] and Gonzalez and Stuart (2008)[6]. The divergence of a matrix S is defined by the identity $(\nabla\cdot S)\cdot a = \nabla\cdot(S^T a)$ holding for any vector a.

B. Motivation and Literature Review

The subject of Kalman filtering and Kalman–Bucy (KB) filtering in infinite-dimensional spaces is studied in the control theory literature[10]. We emphasize that, although we sketch out the basic mathematical foundations of the Fokker–Planck filtering problem in section IV, many interesting mathematical problems in analysis and probability remain open in this area. To the best of our knowledge, the methodology proposed here is the first general method for assimilating observations of statistics directly into a state-space formulation of dynamical systems. Our methodology is built on the conceptual approach introduced in the feedback particle filter[11,12], and earlier related work[13], seeking a mean-field model which achieves the goal of filtering and can be approximated by particle methods[14]; in particular we seek particle approximations of the mean-field model inspired by ensemble Kalman methods[9].

The problem of recovering a probability density from a finite number of known moments is called a moment problem. When $\mathfrak{h}$ in Eq. (I.4) consists of monomials in v, the problem of reconstructing ρ is similar to a moment problem, with the major difference that ρ evolves in time according to a dynamical system. Moment problems are typically regularized by a maximum entropy approach[15]; in the Fokker–Planck filtering problem, regularization is provided by the system's dynamics.

Our motivation comes from a number of applications around which we organize the remainder of our literature review, after first discussing the general question of how to obtain observations of statistics.

a. Obtaining observations of statistics In typical applications one can only observe a single trajectory of a dynamical system, and thus the statistics of the density will not be directly available. If we are interested in the statistics of the invariant measure, as we are for several of the applications discussed below, then for ergodic systems we have that

$$\lim_{T\to\infty} \frac{1}{T}\int_0^T \mathfrak{h}(v^\dagger(t))\,dt = \int \mathfrak{h}(v)\rho^\dagger(v)\,dv, \qquad \text{(I.6)}$$

where $\rho^\dagger$ is the invariant density, and thus an approximation of the statistics of the invariant measure can be obtained from a long observed or simulated trajectory of the dynamical system.

For nonautonomous systems, due to lack of ergodicity, observations of the statistics cannot be made using long time averages. If the nonstationary forcing is slow enough, however, an adiabatic approximation, in which the fast scales are considered to be ergodic with an invariant measure parameterized by the value of the slow forcing, may be justified[16,17]. If the forcing is periodic, then observations of the phase-dependent statistics could be obtained by averaging the observables at a given phase over multiple periods.

For certain systems, invariant statistics may be acquired analytically, or by numerically solving a different set of equations. For example, for the Navier–Stokes equations, the Reynolds-averaged Navier–Stokes (RANS) equations can be used to approximate the stationary statistics.

It may be possible to instead formulate a filtering problem using an observation operator that involves averaging over a finite time window; we leave this for future work. This problem was considered in[18], but only a heuristic solution was proposed. We note that other works have made use of observation operators with time-delayed observations[19,20], albeit for different purposes.

In the next four subsections we review the possible applications of the ensemble Fokker–Planck filter.

b. Acceleration of convergence to a (possibly time-dependent) invariant measure Acceleration of the time to convergence of dynamical models to their invariant measure (often referred to as the "spin-up" period, or the transient) is of importance in many fields, including climate[21–24] and other fluids problems[25], Langevin sampling[26,27], and turbulence simulation[28].

For a stochastic differential equation with an invariant measure, under conditions described in Goldys and Maslowski (2005)[29], the convergence to this invariant measure is exponential with an exponent related to the spectral gap of the corresponding generator.

In this paper we show that this convergence can be accelerated using the ensemble Fokker–Planck filter, and this is the primary application we test in the numerical experiments. In particular, if some statistics of the invariant measure are known, these statistics can be assimilated into the ensemble, obtaining an ensemble whose empirical density is closer to the invariant measure.

To our knowledge, existing methods of accelerating convergence of model trajectories to the invariant measure have been problem-dependent, as in Bryan (1984)[21]. Isik (2013)[30] and Isik, Takhirov, and Zheng (2017)[31] studied a relaxation-based method of accelerating the convergence to equilibrium of the Navier–Stokes equations, which bears some resemblance to our approach.

Non-autonomous (also referred to as non-stationary) and random dynamical systems can have time-dependent attractors, known as pullback attractors, to which the evolution converges[32]. A pullback attractor is the set that the dynamical system approaches when evolved in time from the infinite past to a fixed time (say time 0 without loss of generality). We refer to the probability measure associated with these attractors as time-dependent invariant measures, following Chekroun, Simonnet, and Ghil (2011)[33]. These objects are of considerable interest for climate[22,33,34]. The EnFPF can also accelerate convergence to these invariant measures.

The problem of accelerating convergence to the invariant measure is related to the problem of controlling the Fokker–Planck equation, where a density is controlled in order to reach to a specified target distribution[35], and to statistical control, wherein one aims to return a perturbed system to its equilibrium statistics[36].

Furthermore, the EnFPF could be tested for accelerating the convergence of sampling algorithms such as Langevin sampling and Markov chain Monte Carlo, when some statistics of the target density are known *a priori*.

Finally we note that, when estimating Koopman or Perron–Frobenius operators, it is often necessary to have a large number of trajectories from initial conditions sampled from the invariant measure.

c. Parameter estimation The EnFPF could be used for jointly updating states and parameters using statistical observations, by adopting a state augmentation approach. Other work has adapted methods from data assimilation for parameter estimation using time-averaged statistics, assumed to be close to the statistics on the invariant measure by ergodicity[37–40].

d. Correcting for model error Generally, methods that correct for model error are formulated in terms of forecast performance at some lead time[41,42]. If one is instead interested in correcting statistical properties, one can postulate a parametric form for the model error and use time-averaged observations to estimate the parameters, as discussed in the preceding paragraph. Alternatively, the EnFPF could be tested for directly correcting model error using statistical observations, in a similar manner to the use of classical DA in reducing the impact of model error for forecast applications[43,44]. The analysis increments could then be taken to approximate model error corrections, and training a machine learning model to predict these corrections could be tested, as has been done for classical DA[43–45].

Statistical properties have previously been used to learn closure models for the Navier–Stokes equation using a 3DVar-like scheme[46].

e. Assimilation of time-averaged observations In paleoclimate, proxy records often represent time averages instead of instantaneous measurements. Methods have been developed for making use of time-averaged observations for state estimation in the paleoclimate data assimilation literature[18]. As discussed above, in the case of slow forcing, time-averaged observations can be used to approximately track the system's time-varying statistics, enabling their use in the EnFPF.

C. Contributions and Paper Outline

The primary contributions of this work are: (i) to establish a framework for the filtering of stochastic dynamical systems, or dynamical systems with random initial data, given only observations of statistics; (ii) to introduce ensemble-based state-space methods for this filtering problem via a mean field perspective; and (iii) to demonstrate numerically that the proposed methods are effective at guiding dynamical systems towards observed statistics. (i) is covered in section II A and section IV; (ii) is covered in sections II B–II F; and (iii) is covered in section III.

In section II A we outline the Fokker–Planck filtering problem and distinguish it from the standard filtering problem. In sections II B–II D we introduce a mean-field algorithm and its particle and discrete-time approximations, culminating in the ensemble Fokker–Planck filter (EnFPF). In section II F we discuss implementation details, including the approximation of the score function and a square-root ensemble formulation with reduced computational effort.

In section III we carry out numerical experiments with several chaotic dynamical systems, both autonomous and non-autonomous, and based on the Lorenz63, Lorenz96, and Kuramoto–Sivashinsky models. In particular, we demonstrate that the EnFPF can accelerate the convergence of these systems to their invariant densities, using information about the moments of these densities.

In section IV we provide a justification of our algorithm. We first formulate the KB filter for densities (section IV A), which provides a solution to the Fokker–Planck filtering problem in function space, and analyze some of its properties in Appendix A. We then propose an ansatz amenable to a mean-field model (section IV B), and show its equivalence to the KB filter for densities under some assumptions (Theorem 1 in Appendix B). We then show how this ansatz can be approximated by a mean-field model (section IV C, providing further details in Appendix C).

Finally, in section V we give conclusions and outlook for future work.

II. PROBLEM AND ALGORITHM

In subsection II A we introduce the probabilistic formulation of the standard filtering problem, and then contrast it with the Fokker–Planck filtering problem, where data is in the form of statistics. Subsection II B demonstrates an approach to this problem using a mean-field model. In subsection II C we introduce a particle approximation of the mean-field algorithm, which forms the basis of the proposed EnFPF.

A. Problem Statement

1. The Standard Filtering Problem

In the standard filtering problem, we are given state observations $z^\dagger(t)$ of $v^\dagger(t)$, defined by Eq. (I.3), and the dynamics of $v^\dagger(t)$ are given by Eq. (I.1). The problem is then to find an equation for the conditional distribution of $v|Z^\dagger(t)$, where $Z^\dagger(t) = \{z^\dagger(s)\}_{s\in[0,t]}$ are the observations accumulated up to time t under a fixed realization of B. The solution to the filtering problem is given by the Kushner–Stratonovich equation:

$$\frac{\partial \rho}{\partial t} = \mathcal{L}^*(t)\rho + \left\langle h(v,t) - \mathbb{E}h, \frac{dz^\dagger}{dt} - \mathbb{E}h \right\rangle_{\Gamma(t)} \rho, \quad \text{(II.1)}$$

where $\langle \cdot,\cdot \rangle_A \equiv \langle A^{-1/2}\cdot, A^{-1/2}\cdot \rangle$ is the weighted Euclidean inner product. Treatments of the standard filtering problem can be found in, e.g., Jazwinski (1970)[1] and Bain and Crisan (2009)[47].

2. The Fokker–Planck Filtering Problem

In this paper we consider instead noisy observations of $\rho^\dagger(v,t)$: the observation process $z^\dagger(\cdot)$ is given by

$$dz^\dagger = H(t)\big(\rho^\dagger(\cdot,t)\big)\,dt + \sqrt{\Gamma(t)}\,dB. \tag{II.2}$$

Here $H(t)$ is a linear operator mapping the space of probability densities into a finite-dimensional Euclidean space, and the dynamics of $\rho^\dagger$ are given by the Fokker–Planck equation (I.5). That is, we make observations of *statistics* of the dynamical system. We refer to the problem of finding the conditional density of $v|Z^\dagger(t)$, where $Z^\dagger(t) = \{z^\dagger(s)\}_{s\in[0,t]}$ is given by Eq. (II.2), as the *Fokker–Planck filtering problem.* In the following subsection, we propose an approximation to the solution of this problem in state space.

B. Mean-Field Equation

Although in section IV A we treat the Fokker–Planck filtering problem for more general H, in the rest of what follows we focus on the setting where

$$H(t)\rho = \mathbb{E}[\mathfrak{h}(v,t)] = \int \mathfrak{h}(v,t)\rho(v,t)\,dv, \tag{II.3}$$

for some $\mathfrak{h}$. With this assumption on H, Eq. (II.2) reduces to Eq. (I.4). In particular, if $\mathfrak{h}$ is a monomial in v, e.g., $\mathfrak{h}(v) = v$ or $\mathfrak{h}(v) = \mathrm{vec}(v\otimes v)$, then $H\rho$ will correspond to moments of ρ. We will henceforth use $\mathbb{E}$ to denote expectation under ρ, unless otherwise indicated.

Remark 1. *Note that if $\rho^\dagger(v,0) = \delta(v - v_0^\dagger)$ for some $v_0^\dagger$, and $\Sigma = 0$, then the Fokker–Planck filtering problem is equivalent to the standard filtering problem with $v^\dagger(0) = v_0^\dagger$, observation operator $\mathfrak{h}$, and $\Sigma = 0$.*

Our proposed methodology is to introduce a mean-field model for variable v, depending on its own probability density function $\rho(v,t)$. The mean-field model is chosen to drive the system towards the observed statistical information. Algorithms are then based on particle approximation of this model, leading to ensemble Kalman–type methods. The mean-field model considered is

$$dv = f(v,t)\,dt + \sqrt{\Sigma(t)}\,dW + K(t)\big(dz^\dagger - d\hat{z}\big), \tag{II.4a}$$

$$d\hat{z} = (\mathbb{E}\mathfrak{h})(t)\,dt + \sqrt{\Gamma(t)}dB, \tag{II.4b}$$

$$K(t) = C^{v\mathfrak{h}}(t)\Gamma(t)^{-1}, \tag{II.4c}$$

$$C^{v\mathfrak{h}}(t) = \mathbb{E}\big[\big(v(t) - \mathbb{E}v(t)\big)\big(\mathfrak{h}(v,t) - (\mathbb{E}\mathfrak{h})(t)\big)^T\big]. \tag{II.4d}$$

The terms in the mean-field model can be understood intuitively as follows. The first two terms on the right-hand side of Eq. (II.4a) are simply the dynamics of the system (I.1). The third term resembles the standard nudging observer term from control theory, with an ensemble Kalman–inspired gain, and the use of noisy simulated data, as in the stochastic ensemble Kalman filter.

In some problems we find that it is beneficial to include an additional score-based term in the model, replacing Eq. (II.4a) by

$$dv = f(v,t)\,dt + \sqrt{\Sigma(t)}\,dW + K(t)\Big(dz^\dagger - d\hat{z}\Big) + K(t)\Gamma(t)K(t)^T\nabla\log\rho(v,t)\,dt. \tag{II.5}$$

The additional term induces negative diffusion in the equation for the density of v, exactly balancing the diffusion introduced through $z^\dagger$ and $\hat{z}$. We justify equations (II.4) and (II.5) in detail in section IV by building on the Fokker–Planck picture in density space.

C. Particle Approximation of Mean-Field Equation

In order to tractably implement the mean-field equations (II.4), we use a particle (or ensemble) approximation. That is, given J particles, we consider the following interacting particle system for $\{v^{(j)}\}_{j=1}^{\mathsf{J}}$:

$$dv^{(j)} = f(v^{(j)},t)\,dt + \sqrt{\Sigma(t)}\,dW^{(j)} + K(t)\big(dz^\dagger - d\hat{z}^{(j)}\big), \tag{II.6a}$$

$$d\hat{z}^{(j)} = (\mathbb{E}^{\mathsf{J}}\mathfrak{h})(t)\,dt + \sqrt{\Gamma(t)}\,dB^{(j)}, \tag{II.6b}$$

$$K(t) = (C^{v\mathfrak{h}}(t))^{\mathsf{J}}\,\Gamma(t)^{-1}. \tag{II.6c}$$

Here $\mathbb{E}^{\mathsf{J}}$ denotes expectation with respect to the empirical measure formed by equally weighting Dirac measures at the particles $\{v^{(j)}\}_{j=1}^{\mathsf{J}}$; $(C^{v\mathfrak{h}})^{\mathsf{J}}$ denotes the sample cross-covariance computed using this empirical measure:

$$C^{v\mathfrak{h}}(t) = \mathbb{E}^{\mathsf{J}}\big[\big(v(t) - \mathbb{E}v(t)\big)\big(\mathfrak{h}(v,t) - (\mathbb{E}\mathfrak{h})(t)\big)^T\big].$$

Note that, unlike the ensemble Kalman filter, the predicted observation for each ensemble member, Eq. (II.6b), involves the expectation of $\mathfrak{h}$ over the ensemble, instead of the observation operator applied to that ensemble member.

D. Discrete-Time Approximation of Mean-Field Equation

A discrete-time analogue of Eqs. (II.6) is given by

$$\hat{v}_{i+1}^{(j)} = \Psi_i(v_i^{(j)}) + \xi_i^{(j)}, \tag{II.7a}$$

$$v_{i+1}^{(j)} = \hat{v}_{i+1}^{(j)} + K_{i+1}(y_{i+1}^\dagger - \hat{y}_{i+1}^{(j)}), \tag{II.7b}$$

$$\hat{y}_{i+1}^{(j)} = \mathbb{E}^{\mathsf{J}}[\mathfrak{h}_{i+1}(\hat{v}_{i+1})] + \eta_{i+1}^{(j)}, \tag{II.7c}$$

$$K_{i+1} = (\hat{C}_{i+1}^{v\mathfrak{h}})^{\mathsf{J}}((\hat{C}_{i+1}^{\mathfrak{h}\mathfrak{h}})^{\mathsf{J}} + (\Gamma_d)_{i+1})^{-1}, \tag{II.7d}$$

where $\xi_i^{(j)} \sim \mathcal{N}(0, (\Sigma_d)_i)$, $\eta_i^{(j)} \sim \mathcal{N}(0, (\Gamma_d)_i)$, $\mathfrak{h}_i(v) = \mathfrak{h}(v, t)$, and

$$(\hat{C}^{v\mathfrak{h}}_{i+1})^{\mathsf{J}} = \mathbb{E}^{\mathsf{J}}[(\hat{v}_{i+1} - \mathbb{E}^{\mathsf{J}}\hat{v}_{i+1}) \otimes (\mathfrak{h}_{i+1}(\hat{v}_{i+1}) - \mathbb{E}^{\mathsf{J}}[\mathfrak{h}_{i+1}(\hat{v}_{i+1})])], \tag{II.8}$$

$$(\hat{C}^{\mathfrak{h}\mathfrak{h}}_{i+1})^{\mathsf{J}} = \mathbb{E}^{\mathsf{J}}[(\mathfrak{h}_{i+1}(\hat{v}_{i+1}) - \mathbb{E}^{\mathsf{J}}[\mathfrak{h}_{i+1}(\hat{v}_{i+1})]) \otimes (\mathfrak{h}_{i+1}(\hat{v}_{i+1}) - \mathbb{E}^{\mathsf{J}}[\mathfrak{h}_{i+1}(\hat{v}_{i+1})])]. \tag{II.9}$$

Furthermore, we introduce the following rescalings adopted in Law, Stuart, and Zygalakis (2015)[3]:

$$\begin{aligned} f(\cdot, t) &= (\Psi_i(\cdot) - I\cdot)/\tau, \quad z^\dagger_{i+1} - z^\dagger_i = \tau y^\dagger_{i+1}, \\ \Sigma(i\tau) &= (\Sigma_d)_i/\tau \quad \Gamma(i\tau) = \tau(\Gamma_d)_i, \quad i = t/\tau, \end{aligned} \tag{II.10}$$

Then Eqs. (II.7) can be seen to be a discretization of Eqs. (II.6) with time-step τ. More justification is given for these rescalings in Salgado, Middleton, and Goodwin (1988)[48] and Simon (2006)[49]. Note that both $K_{i+1} = (\hat{C}^{v\mathfrak{h}}_{i+1})^{\mathsf{J}}(\Gamma_d)^{-1}_{i+1}$ and Eq. (II.7d) are consistent with the continuous-time gain as $\tau \to 0$. We use the latter, similar to the discrete-time Kalman filter.

E. Score Function Term

We now discuss further computational issues that arise when Eq. (II.4a) is replaced by (II.5). This term involves the score function, defined as $\nabla \log \rho$, but with an additional preconditioning. If this term is added to the discrete-time particle version of the filter, Eq. (II.7b) becomes

$$v^{(j)}_{i+1} = \hat{v}^{(j)}_{i+1} + K_{i+1}(y^\dagger_{i+1} - \hat{y}^{(j)}_{i+1}) + K_{i+1}(\Gamma_d)_{i+1}K^T_{i+1}(\nabla \log \rho_{i+1})^{\mathsf{J}}(\hat{v}^{(j)}_{i+1}). \tag{II.11}$$

where $(\nabla \log \rho_{i+1})^{\mathsf{J}}$ denotes particle-based approximation of the score function using the $\{\hat{v}^{(j)}_{i+1}\}^{\mathsf{J}}_{j=1}$. If we make the assumption that the density is Gaussian with mean $\mathbb{E}v$ and covariance C^{vv}, the score function takes on a simple form,

$$\nabla \log \rho = -(C^{vv})^{-1}(v - \mathbb{E}v). \tag{II.12}$$

A natural particle approximation $(\nabla \log \rho)^{\mathsf{J}}$ follows by replacing the mean and covariance with the corresponding quantities computed under the empirical measure of the set of particles.

More general kernel-based nonparametric estimators for the score function have been developed, such as those defined in Zhou, Shi, and Zhu (2020)[50] and implemented in the `kscore` package. In the numerical experiments reported in this paper, we either omit the score term completely, or use it and employ only the Gaussian approximation.

F. Implementation

1. Ensemble Square-Root Formulation

In order to make the method scale well to high dimensions, an ensemble square-root formulation[51] of Eq. (II.7) can be used, although we do not use it in the numerical experiments reported here. The advantage of this formulation is that the most expensive linear algebra operations are rewritten in the ensemble space, resulting in favorable computational complexity when J is much smaller than the state-space dimension d or observation-space dimension p.[3]

To implement this method we write $(C^{vv})^{\mathsf{J}} = VV^T$, $(C^{v\mathfrak{h}})^{\mathsf{J}} = VY^T$, and $(C^{\mathfrak{h}\mathfrak{h}})^{\mathsf{J}} = YY^T$, where the jth column of V and the jth column of Y are given by

$$\begin{aligned} V^{(j)} &= (v^{(j)} - \mathbb{E}^{\mathsf{J}}v)/\sqrt{J-1}, \\ Y^{(j)} &= (\mathfrak{h}(v^{(j)}) - \mathbb{E}^{\mathsf{J}}\mathfrak{h})/\sqrt{J-1}, \end{aligned} \tag{II.13}$$

respectively. Then, K can be written as

$$K = VY^TW, \tag{II.14}$$

where $W = (\Gamma_d^{-1} - \Gamma_d^{-1}Y(I + Y^T\Gamma_d^{-1}Y)^{-1}Y^T\Gamma_d^{-1})$ by the Woodbury identity.

We assume that Γ_d^{-1} is provided and can be applied cheaply, for example if it is diagonal. This is a standard assumption[51]. With this expression, K can be computed in $\mathcal{O}(J^3 + J^2p + Jp^2 + dJp)$.

Note that the Gaussian score function approximation Eq. (II.12) cannot be applied in cases when $J < d$, since $(C^{vv})^{\mathsf{J}}$ will be singular. We do not consider the score function term in the complexity analysis.

The complexity is thus a quadratic polynomial in d and p, whereas various ensemble square-root filters can be implemented to be linear in p and d. The latter rely on the fact that the in the standard Kalman filter the updated covariance can be written as $(I - KH)C^{vv}$, where H is the observation operator. The EnFPF cannot be written in this way. Whether the EnFPF can be reformulated to be linear in p and d by another approach is a topic for future research.

2. Code

The open-source Julia code for the EnFPF is available at `https://github.com/eviatarbach/EnFPF`. In the numerical experiments that follow, we compute the Wasserstein distance (explained in section III) using the Python Optimal Transport library[52]. We used the parasweep library to facilitate parallel experiments[53].

---

[3] Note, however, that in many applications with a high-dimensional state space, the statistics of interest may be relatively low-dimensional, such that the regular version of the algorithm (II.7) will be feasible.

3. Numerical Methods for the Test Models

In section III, we will present numerical experiments with the Lorenz63, Lorenz96, and Kuramoto–Sivashinsky models. We integrate the Lorenz63 and Lorenz96 models using the fourth-order Runge–Kutta method, with a time step of 0.05 for both. We integrate the Kuramoto–Sivashinsky equation in Fourier space using the exponential time differencing fourth-order Runge–Kutta method[54] with 64 Fourier modes and a time step of 0.25.

III. NUMERICAL EXPERIMENTS

In this section we present the results of numerical experiments applying the discrete-time EnFPF of section II D to the Lorenz63, Lorenz96, and Kuramoto–Sivashinsky systems. The first three subsections are devoted, respectively, to these three models; a final fourth subsection returns to the Lorenz63 model, with quasiperiodic forcing.

We found in the experiments that assimilating too often can cause degraded results for some systems, as opposed to the situation in standard filtering, where increased assimilation frequency is typically preferred. In the standard filtering problem, there is a single true trajectory, and under certain conditions the filtering distribution will converge to this trajectory in the limit of zero observational noise[3,55]. In the non-zero noise case, however, the filtered time-series (e.g., the maximum *a posteriori* estimate) will not even generally be a trajectory of the dynamical system, except in methods such as strong-constraint 4DVar. Here, we expect that the problem of ensemble members deviating from being trajectories can be amplified, since the method only aims to match statistical features of the entire ensemble. Thus, if assimilation is done too frequently, then ensemble members may be pushed too far from being trajectories into unphysical or unstable parts of the phase space. In fact, we found the assimilation frequency to be a key tuning parameter. We refer to a single forecast–assimilation step (Eqs. (II.7)) as a *cycle*, as is common in the DA literature, and each cycle lasts for τ time units.

We found, furthermore, that the score term did not consistently improve filtering performance. In the experiments that follow, we omit the score term except in the experiments with the Kuramoto–Sivashinsky system in section III C, where it leads to clear improvements when used, together with the Gaussian approximation, in the form Eq. (II.12). For both the Lorenz models we found that the inclusion of the Gaussian approximation of the score degraded performance and that use of the kernel-based score approximations, based on the paper of Zhou, Shi, and Zhu (2020)[50], was no better than simply omitting the term altogether.

In the experiments below, we use a Wasserstein metric to quantify the distance between the ensemble distribution and the invariant density. We estimate the invariant density using an ensemble integrated for a sufficiently long time. We employ the W_1 Wasserstein metric which allows us to compute distances between empirical distributions. The code for computing this distance is readily available (see II F 2).

A. Lorenz63 Model

For the experiments in this subsection, we use the Lorenz (1963)[56] model

$$\begin{aligned} \frac{dx}{dt} &= \sigma(y-x), \\ \frac{dy}{dt} &= x(r-z)-y, \\ \frac{dz}{dt} &= xy-\beta z, \end{aligned} \tag{III.1}$$

with the standard parameter values $\sigma = 10$, $r = 28$, and $\beta = 8/3$.

1. Assimilating Time-Varying Means and Second Moments

We first verify the ability of the EnFPF to force an ensemble to adopt time-varying statistics. We do this by applying the EnFPF to a 10-member ensemble, with noisy statistical observations of the means and uncentered second moments of the three variables coming from a 100-member ensemble being evolved concurrently. The difference between the statistics computed over the 10- and 100-member ensembles arise due to both sampling errors and different initial conditions. The 100-member ensemble (despite having its own sampling error) better approximates the true statistics of the system, and we view these 100-member ensemble statistics as the truth, based on which we may compute errors in the statistics of 10-member ensembles. We assimilate observations every 0.2 time units, with an observation error covariance set to 20% of the time variability of each statistic computed over the 100-member ensemble.

Figure 2 shows the resulting error in the means and second moments of the 10-member ensemble, compared with the errors arising from an unfiltered run of the 10-member ensemble; in both cases the errors are computed by comparison with the 100-member ensemble. After several cycles, the filter appears to reach an asymptotic error on the order of the observation error, and this error is significantly lower than that arising in the unfiltered case.

Table I shows the impact of the observation error covariance magnitude on the filtering performance. The set-up is otherwise the same as that described above. As expected, the error increases as Γ is increased, although still outperforming the unfiltered ensemble.

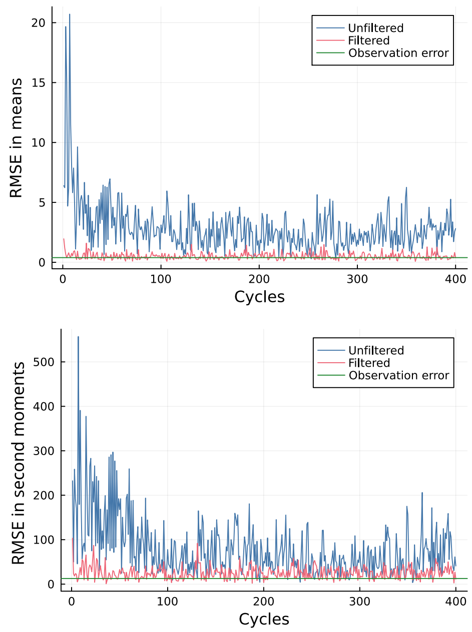

FIG. 2: The impact of filtering on the root-mean-square error (RMSE) in the mean and second moment in the Lorenz63 model.

Means

Observation error	*Filtered RMSE*
10% (0.088)	0.11
35% (0.31)	0.40
60% (0.53)	0.69
85% (0.75)	0.97

Second moments

Observation error	*Filtered RMSE*
10% (2.8)	20
35% (9.9)	23
60% (17)	29
85% (24)	35

TABLE I: The impact of the observation error covariance on filtering performance. In the first column are the percentages of the standard deviation of the time variability of each statistic taken to be the observation error, and in parentheses the square root of the total variance of the observation error in the statistic. With no filtering, the RMSE is 2.5 in the unfiltered means and 73 in the second moments. The RMSE is averaged over 1400 cycles after 100 transient cycles.

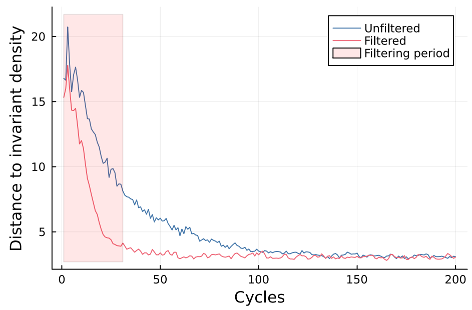

FIG. 3: The estimated Wasserstein distance to the invariant density in Lorenz63, in unfiltered and filtered cases. For the filtered case, the first and second moments are assimilated. Each curve is averaged over 10 different initializations.

2. Accelerating Convergence to the Invariant Density

We now test the ability of the EnFPF to accelerate convergence to the invariant density. We assimilate observations of fixed statistics of the invariant density, the means and second moments of the three variables, into a 100-member ensemble. We use the same assimilation frequency and observation error as in subsubsection III A 1.

Figure 3 shows the impact of the EnFPF on the convergence to the invariant density. In this case, we only apply the EnFPF for the first 30 cycles (indicated by the pink rectangle), and then let the ensemble evolve under the regular Lorenz63 dynamics. We see that the EnFPF leads to a more rapid convergence: by the end of the filtering period, the distance is close to the asymptotic one, while it takes at least 100 cycles for the unfiltered case to reach the same. Figure 4 visualizes in state space this rapid convergence toward the invariant density via the EnFPF.

3. Impact of Higher-Order Moments

Figure 5 shows the convergence to the invariant measure of Lorenz63 with different assimilated moments of x and y, namely the first, first and second, and first, second, and third marginal moments. Assimilating the first-order moments accelerates the convergence to the invariant measure compared to the unfiltered case. Adding the second and third order moments appears to result in the most rapid initial rate of convergence, and after about 50 cycles assimilating the first two and the first three moments leads to a similar asymptotic distance to the invariant measure.

B. Lorenz96 Model

We now test the convergence to the invariant density of the Lorenz (1996)[57] model

$$\frac{\mathrm{d}x_i}{\mathrm{d}t} = -x_{i-1}(x_{i-2} + x_{i+1}) - x_i + F, \qquad \text{(III.2)}$$

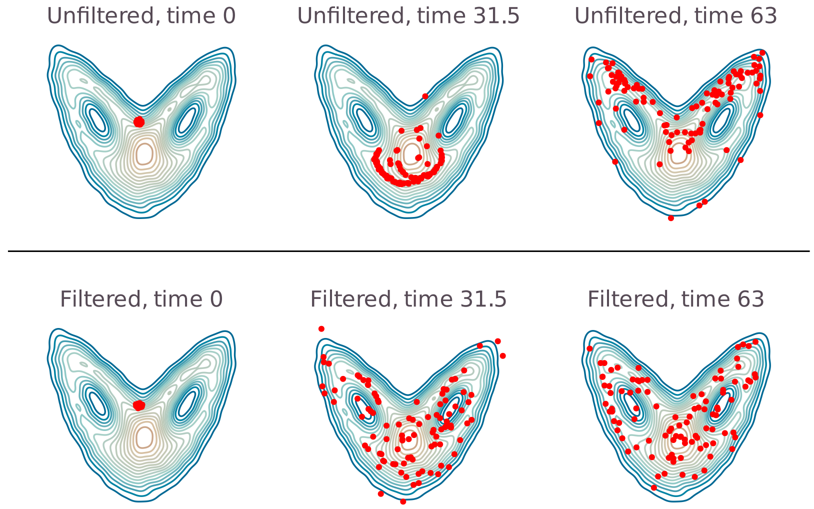

FIG. 4: Top panel: an ensemble evolving in time from left to right, superimposed on the invariant density of Lorenz63 in the $x - z$ plane. Orange corresponds to higher probability density and blue to lower. Bottom panel: the same, but with the EnFPF applied.

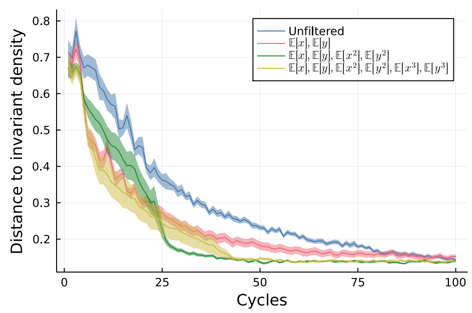

FIG. 5: The estimated Wasserstein distance to the invariant density in Lorenz63, in unfiltered and filtered cases when different moments are assimilated. The curves are averaged over 25 initial conditions, and the shaded areas correspond to ± the standard error over the initializations. Here, for the filtered cases, the EnFPF is applied at every cycle.

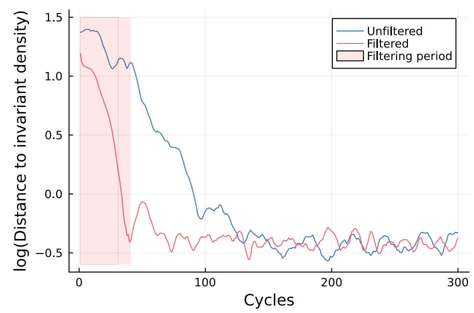

FIG. 6: The estimated Wasserstein distance to the invariant density in Lorenz96, in unfiltered and filtered cases. For the filtered case, the first and second moments are assimilated. Here, we show the mean of the Wasserstein distances corresponding to the marginal density for each variable.

where the indices i range from 1 to D and are cyclical. We use $F = 8$ and $D = 40$ variables. This is a model of an atmospheric latitude circle that is commonly used in data assimilation experiments.

We assimilate the means and second moments of the 40 variables on the invariant density, with an observation error covariance of 20% of the temporal variability of the statistics computed over a 100-member ensemble. We assimilate every 0.05 time units into a 100-member ensemble for 40 cycles. Figure 6 shows that the convergence towards the invariant density is thereby significantly accelerated.

C. Kuramoto–Sivashinsky Model

We now carry out experiments with the Kuramoto–Sivashinsky model, a chaotic partial differential equation in one spatial dimension:

$$u_t + u_{xxxx} + u_{xx} + uu_x = 0, \quad x \in [0, L]. \qquad \text{(III.3)}$$

We use $L = 22$ and periodic boundary conditions, discretized using 64 Fourier modes (see II F 3 for details on the numerical method).

We assimilate the means and second moments of the invariant density of the 64 variables in physical space, every 2.0 time units. We assimilate for 30 cycles using a 100-member ensemble, and again use an observational

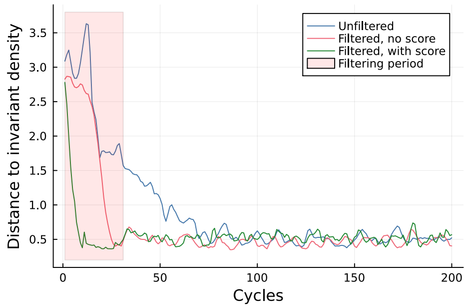

FIG. 7: The estimated Wasserstein distance to the invariant density in the Kuramoto–Sivashinsky equation, in unfiltered and filtered cases. For the filtered case, the first and second moments are assimilated. Here, we show the mean of the Wasserstein distances corresponding to the marginal density for each variable.

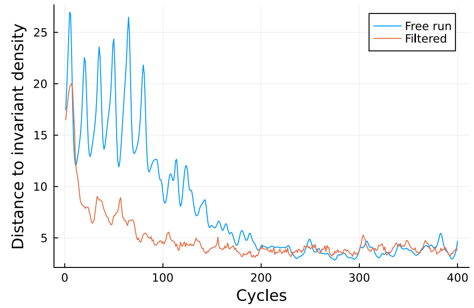

FIG. 8: The estimated Wasserstein distance to the invariant density in the non-autonomous Lorenz63 model, in unfiltered and filtered cases. For the filtered case, $\mathbb{E}[x^i]$, $\mathbb{E}[y^i]$, and $\mathbb{E}[z^i]$ for $i = 1, 2, 3$ are assimilated.

error covariance of 20% of the temporal variability. Figure 7 shows the results with and without the score term included. In both cases, there is an acceleration compared to the unfiltered case; inclusion of the score term considerably accelerates convergence.

D. Time-Dependent Invariant Measures

We now use the Lorenz63 model (Eq. (III.1)), but with the r parameter subject to quasiperiodic forcing, as in Daron and Stainforth (2015)[22]:

$$r(t) = 28 + \sin(2\pi t) + \sin\left(\sqrt{3}t\right) + \sin\left(\sqrt{17}t\right). \quad \text{(III.4)}$$

Since this system is non-autonomous, it possesses for each time t a pullback attractor with a corresponding time-dependent invariant measure, as discussed in section I B. The measure at time t can be approximated by the empirical density at time t of an ensemble initialized sufficiently far back in time, at $t - T$ for some large T. Here, we evolve a 100-member ensemble using $T = 500$ time units to approximate the invariant measures at time t. Then, we evolve the ensemble for the additional time period of t to $t + 20$ to obtain approximations to the invariant measures in this period.

We evolve two separate 100-member ensembles for the same time period t to $t+20$, but with $T = 0$ (no spin-up). We apply the EnFPF to one of these ensembles and not the other. For the EnFPF, we assimilate every 0.05 time units with an observation error covariance of 20% of the temporal variability. We then measure the distance between the empirical densities of these two ensembles and the one approximating the invariant measure described in the previous paragraph.

Figure 8 shows that the convergence to the time-dependent invariant measures is indeed accelerated by the EnFPF, reaching a comparable asymptotic distance to the invariant measure in less than half the time.

IV. JUSTIFICATION OF ALGORITHM

A. Kalman–Bucy (KB) Filter for Densities

Since both the Fokker–Planck equation (I.5) and the observation equation (II.2) are linear, and since all noise is additive Gaussian, the conditional probability measure over densities, $\rho|Z^\dagger(t)$, is a Gaussian. This filtering problem can be solved using a Kalman–Bucy filter in Hilbert space, posing significant challenges because it involves finding a sequence of probability measures on an infinite-dimensional space of functions (densities).

We start by defining the Hilbert space $\mathcal{H} = L^2(\mathbb{R}^d, \mathbb{R})$ with inner product

$$\langle a, b\rangle_{\mathcal{H}} \equiv \int ab\, dv. \quad \text{(IV.1)}$$

We consider density functions $\rho \in \mathcal{H}$, and we require that $\rho(v, t) \to 0$ as $v \to \infty$. Note that we will sometimes use this inner product in situations where one of the arguments is only locally square integrable; in particular we will need to use the constant function $\mathbb{1}(v) = 1$. To distinguish them from the Hilbert space inner product, we denote the standard Euclidean inner product in $\mathbb{R}^p$ as $\langle\cdot,\cdot\rangle_{\mathbb{R}^p}$ and the weighted Euclidean inner product, defined for any strictly positive-definite and symmetric $A \in \mathbb{R}^{p\times p}$, as $\langle\cdot,\cdot\rangle_A \equiv \langle A^{-1/2}\cdot, A^{-1/2}\cdot\rangle_{\mathbb{R}^p}$.

Recall definition Eq. ((I.5)b) of the adjoint of the generator $\mathcal{L}$. We are given the dynamics and observation equations (I.5) and (II.2):

$$d\rho^\dagger(v, t) = \mathcal{L}^*(t)\rho^\dagger(v, t)\, dt, \quad \text{(IV.2)}$$

$$dz^\dagger(t) = H(t)\rho^\dagger(v, t)\, dt + \sqrt{\Gamma(t)}dB. \quad \text{(IV.3)}$$

Then, given all observations up to time t, $Z^\dagger(t) = \{z^\dagger(s)\}_{s\in[0,t]}$, the filtering distribution is given by

$$\rho(\cdot, t)|Z^\dagger(t) \sim \mu(t) \equiv \mathcal{N}\big(m(t), C(t)\big), \quad \text{(IV.4)}$$

where $\mathcal{N}$ is a Gaussian measure on $\mathcal{H}$ with mean $m(t)$ and covariance operator $C(t)$. For notational simplicity, we have dropped the explicit dependence of $m(t)$, $C(t)$, and $\rho(t)$ on v. Here $C \in L(\mathcal{H},\mathcal{H})$ is necessarily self-adjoint and trace class[58]; that is, $\mathrm{tr}(C) < \infty$. In what follows the expectation $\mathbb{E}_\mu$ is defined with respect to the measure μ on the space of L^2 densities ρ.

Using Theorem 7.10 in Falb (1967)[59], the KB filter for this system can be written as

$$\begin{aligned} dm(t) &= \mathcal{L}^*(t)m(t)\,dt \qquad\qquad \text{(IV.5a)}\\ &\quad + C(t)H^*(t)\Gamma(t)^{-1}\big(dz^\dagger(t) - H(t)m(t)\big)\,dt,\\ dC(t) &= \mathcal{L}^*(t)C(t)\,dt + C(t)\mathcal{L}(t)\,dt \qquad \text{(IV.5b)}\\ &\quad - C(t)H^*(t)\Gamma(t)^{-1}H(t)C(t)\,dt,\\ m(0) &= m_0, C(0) = C_0, \qquad\qquad \text{(IV.5c)}\end{aligned}$$

where

$$C(t) = \mathrm{cov}(\rho(t) - m(t), \rho(t) - m(t)), \tag{IV.6}$$

and

$$\mathrm{cov}(x,y) \equiv \mathbb{E}_\mu[x\otimes y] - \mathbb{E}_\mu[x]\otimes\mathbb{E}_\mu[y]. \tag{IV.7}$$

The outer-product $x_1 \otimes y_1$ is defined by the identity

$$(x_1\otimes y_1)x = x_1\langle y_1, x\rangle_{\mathcal{H}} \tag{IV.8}$$

holding for all $x \in \mathcal{H}$. Note that Falb (1967)[59] requires boundedness of $\mathcal{L}^*$, but the results have been extended to unbounded operators[60]. However, we still require boundedness of H. For the rest of the paper, we will assume that H takes the form in Eq. (II.3).

The adjoint operator H^* is then given by

$$H^*(t)u = \langle \mathfrak{h}(v,t), u\rangle_{\mathbb{R}^p}, \tag{IV.9}$$

for $u \in \mathbb{R}^p$. Note that, formally, $H^*(t)u$ is to be viewed as a function of v, in the space $\mathcal{H}$.

In general the solution of Eq. (IV.5a), $m(t)$, will not be normalized. However, in Appendix A we show that normalization is preserved under certain conditions on the initializations m_0 and C_0 from Eq. (IV.5c). However, $m(t)$ is not guaranteed to be non-negative for all v and t, and thus cannot be a proper probability density. Nonetheless, we can still consider integrals against it.

B. Ansatz and Relation to KB Filter for Densities

Solving the KB filter equations directly is intractable. We therefore seek an equation which is amenable to a mean-field model, which in turn can be approximated by a particle system. We propose the following ansatz for the density of $v|Z^\dagger(t)$:

$$\frac{\partial\rho}{\partial t} = \mathcal{L}^*(t)\rho + \left\langle \mathfrak{h}(v,t) - H(t)\rho, \frac{dz^\dagger}{dt} - H(t)\rho\right\rangle_{\Gamma(t)}\rho. \tag{IV.10}$$

Note the similarity to the Kushner–Stratonovich (KS) equation (II.1). Although the solutions of this equation do not match the KB filter for densities in general, we show in Theorem 1 that they coincide in observation space for linear f and $\mathfrak{h}$, under additional assumptions detailed there. The proof sketch is provided in Appendix B.

C. Mean-Field Approximation

We would now like to find a mean-field model which has, as its FP equation, Eq. (IV.10). We postulate the following form:

$$\begin{aligned} dv &= f(v,t)\,dt + \sqrt{\Sigma(t)}\,dW + a(v,\rho,t)\,dt \qquad \text{(IV.11)}\\ &\quad + K(v,\rho,t)\Big(dz^\dagger - H(t)\rho(v,t)\,dt - \sqrt{\Gamma(t)}dB\Big).\end{aligned}$$

Specifically, we aim to choose the pair of functions (a, K) so that the Fokker–Planck equation for v governed by this mean-field model coincides with Eq. (IV.10). In Appendix C we detail the choices which achieve this and, after making a further approximation of K, we obtain equations (II.4) with (II.4a) replaced by (II.5). However, as explained there, in many cases use of Eq. (II.4a), which corresponds to setting $a \equiv 0$ and using a simple approximation of K, leads to algorithms which empirically perform well.

V. CONCLUSIONS

In this paper we introduce the Fokker–Planck filtering problem, which consists of estimating the evolving probability density of a (possibly stochastic) dynamical system given noisy observations of expectations evaluated with respect to it. We provide a solution for this problem using the KB filter in Hilbert space, and introduce an ensemble algorithm, the ensemble Fokker–Planck filter (EnFPF), that approximates it under conditions on the dynamics and observables. We also show, through numerical experiments, that this method can be used to accelerate convergence to the invariant measure of dynamical systems, and that this acceleration phenomenon applies beyond the conditions on the dynamics and observables required to provably link the KB filter and the mean-field model underlying our proposed ensemble method.

Future work will test this method on higher-dimensional models, such as turbulent channel flows and ocean models. Other future directions, as described in section I B, include: (i) the testing of this method as an approach to counteract model error; (ii) use in parameter estimation; and (iii) use in the acceleration of sampling methods such as Langevin dynamics and Markov chain Monte Carlo when some statistics of the target density are known. Furthermore, many of the numerical results require deeper understanding; these include the impact of the assimilation frequency, the score term, and the incorporation of higher-order moments, or other observables, on the filter performance. Finally, on the theoretical side, there is a considerable need for deeper analysis.

ACKNOWLEDGMENTS

EB is supported by the the Foster and Coco Stanback Postdoctoral Fellowship. AS is supported by the Office of Naval Research (ONR) through grant N00014-17-1-2079. TC and AS acknowledge recent support through ONR grant N00014-23-1-2654. EB and AS are also grateful for support from the Department of Defense Vannevar Bush Faculty Fellowship held by AS. We thank Tapio Schneider, Dimitris Giannakis, and two anonymous referees for helpful comments.

[1] A. H. Jazwinski, *Stochastic Processes and Filtering Theory*, Mathematics in Science and Engineering No. 64 (Academic Press, Inc., New York, 1970).
[2] E. Kalnay, *Atmospheric Modeling, Data Assimilation and Predictability* (Cambridge University Press, New York, 2002).
[3] K. Law, A. Stuart, and K. Zygalakis, *Data Assimilation: A Mathematical Introduction*, Texts in Applied Mathematics No. 62 (Springer International Publishing, 2015).
[4] S. Reich and C. Cotter, *Probabilistic Forecasting and Bayesian Data Assimilation* (Cambridge University Press, Cambridge, 2015).
[5] M. E. Gurtin, *An Introduction to Continuum Mechanics*, Mathematics in Science and Engineering, Vol. 158 (Academic Press, 1981).
[6] O. Gonzalez and A. M. Stuart, *A First Course in Continuum Mechanics*, Cambridge Texts in Applied Mathematics (Cambridge University Press, Cambridge, 2008).
[7] G. Evensen, "Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics," Journal of Geophysical Research: Oceans **99**, 10143–10162 (1994).
[8] D. Crişan, P. Del Moral, and T. J. Lyons, "Interacting Particle Systems Approximations of the Kushner-Stratonovitch Equation," Advances in Applied Probability **31**, 819–838 (1999), 1428398.
[9] E. Calvello, S. Reich, and A. M. Stuart, "Ensemble Kalman Methods: A Mean Field Perspective," (2022), arxiv:2209.11371 [cs, math].
[10] R. F. Curtain and A. J. Pritchard, eds., *Infinite Dimensional Linear Systems Theory*, Lecture Notes in Control and Information Sciences, Vol. 8 (Springer-Verlag, Berlin/Heidelberg, 1978).
[11] T. Yang, P. G. Mehta, and S. P. Meyn, "Feedback Particle Filter," IEEE Transactions on Automatic Control **58**, 2465–2480 (2013).
[12] S. Reich, "Data assimilation: The Schrödinger perspective," Acta Numerica **28**, 635–711 (2019).
[13] D. Crisan and J. Xiong, "Approximate McKean–Vlasov representations for a class of SPDEs," Stochastics **82**, 53–68 (2010).
[14] S. Pathiraja, S. Reich, and W. Stannat, "McKean–Vlasov SDEs in Nonlinear Filtering," SIAM Journal on Control and Optimization **59**, 4188–4215 (2021).
[15] R. V. Abramov, "The multidimensional maximum entropy moment problem: A review of numerical methods," Communications in Mathematical Sciences **8**, 377–392 (2010).
[16] G. A. Pavliotis and A. M. Stuart, *Multiscale Methods: Averaging and Homogenization*, Texts in Applied Mathematics, Vol. 53 (Springer, New York, NY, 2008).
[17] G. Drótos, T. Bódai, and T. Tél, "Quantifying nonergodicity in nonautonomous dissipative dynamical systems: An application to climate change," Physical Review E **94**, 022214 (2016).
[18] S. Dirren and G. J. Hakim, "Toward the assimilation of time-averaged observations," Geophysical Research Letters **32** (2005), 10.1029/2004GL021444.
[19] D. Rey, M. Eldridge, U. Morone, H. D. I. Abarbanel, U. Parlitz, and J. Schumann-Bischoff, "Using waveform information in nonlinear data assimilation," Physical Review E **90**, 062916 (2014).
[20] E. Bach, S. Mote, V. Krishnamurthy, A. S. Sharma, M. Ghil, and E. Kalnay, "Ensemble Oscillation Correction (EnOC): Leveraging Oscillatory Modes to Improve Forecasts of Chaotic Systems," Journal of Climate **34**, 5673–5686 (2021).
[21] K. Bryan, "Accelerating the Convergence to Equilibrium of Ocean-Climate Models," Journal of Physical Oceanography **14**, 666–673 (1984).
[22] J. D. Daron and D. A. Stainforth, "On quantifying the climate of the nonautonomous Lorenz-63 model," Chaos: An Interdisciplinary Journal of Nonlinear Science **25**, 043103 (2015).
[23] G. Drótos, T. Bódai, and T. Tél, "On the importance of the convergence to climate attractors," The European Physical Journal Special Topics **226**, 2031–2038 (2017).
[24] C. Deser, F. Lehner, K. B. Rodgers, T. Ault, T. L. Delworth, P. N. DiNezio, A. Fiore, C. Frankignoul, J. C. Fyfe, D. E. Horton, J. E. Kay, R. Knutti, N. S. Lovenduski, J. Marotzke, K. A. McKinnon, S. Minobe, J. Randerson, J. A. Screen, I. R. Simpson, and M. Ting, "Insights from Earth system model initial-condition large ensembles and future prospects," Nature Climate Change **10**, 277–286 (2020).
[25] Z. Wang, J. Xin, and Z. Zhang, "DeepParticle: Learning invariant measure by a deep neural network minimizing Wasserstein distance on data generated from an interacting particle method," Journal of Computational Physics **464**, 111309 (2022).
[26] C.-R. Hwang, S.-Y. Hwang-Ma, and S.-J. Sheu, "Accelerating diffusions," The Annals of Applied Probability **15**, 1433–1444 (2005).
[27] A. Abdulle, G. A. Pavliotis, and G. Vilmart, "Accelerated convergence to equilibrium and reduced asymptotic variance for Langevin dynamics using Stratonovich perturbations," Comptes Rendus Mathematique **357**, 349–354 (2019).
[28] K. S. Nelson and O. B. Fringer, "Reducing spin-up time for simulations of turbulent channel flow," Physics of Fluids **29**, 105101 (2017).
[29] B. Goldys and B. Maslowski, "Exponential ergodicity for stochastic reaction–diffusion equations," in *Stochastic Partial Differential Equations and Applications - VII*, edited by G. D. Prato and L. Tubaro (Chapman and Hall/CRC, 2005).
[30] O. R. Isik, "Spin up problem and accelerating convergence to steady state," Applied Mathematical Modelling **37**, 3242–3253 (2013).
[31] O. R. Isik, A. Takhirov, and H. Zheng, "Second order time relaxation model for accelerating convergence to steady-state equilibrium for Navier–Stokes equations," Applied Numerical Mathematics **119**, 67–78 (2017).
[32] L. Arnold, *Random Dynamical Systems*, Springer Monographs in Mathematics (Springer, Berlin, Heidelberg, 1998).
[33] M. D. Chekroun, E. Simonnet, and M. Ghil, "Stochastic climate dynamics: Random attractors and time-dependent invariant measures," Physica D: Nonlinear Phenomena **240**, 1685–1700 (2011).
[34] M. Ghil, M. D. Chekroun, and E. Simonnet, "Climate dynamics and fluid mechanics: Natural variability and related uncertainties," Physica D: Nonlinear Phenomena Euler Equations: 250 Years On, **237**, 2111–2126 (2008).
[35] M. Annunziato and A. Borzì, "A Fokker–Planck control framework for multidimensional stochastic processes," Journal of Computational and Applied Mathematics **237**, 487–507 (2013).
[36] J. Covington, D. Qi, and N. Chen, "Effective Statistical Control Strategies for Complex Turbulent Dynamical Systems," (2023), arxiv:2307.15637 [math-ph].
[37] A. Köhl and J. Willebrand, "An adjoint method for the assimilation of statistical characteristics into eddy-resolving ocean mod-

els," Tellus A: Dynamic Meteorology and Oceanography **54**, 406–425 (2002).

[38] T. Schneider, A. M. Stuart, and J.-L. Wu, "Learning stochastic closures using ensemble Kalman inversion," Transactions of Mathematics and Its Applications **5**, tnab003 (2021).

[39] V. Mons, Y. Du, and T. A. Zaki, "Ensemble-variational assimilation of statistical data in large-eddy simulation," Physical Review Fluids **6**, 104607 (2021).

[40] T. Schneider, A. M. Stuart, and J.-L. Wu, "Ensemble Kalman inversion for sparse learning of dynamical systems from time-averaged data," Journal of Computational Physics **470**, 111559 (2022).

[41] M. Levine and A. Stuart, "A framework for machine learning of model error in dynamical systems," Communications of the American Mathematical Society **2**, 283–344 (2022).

[42] E. Bach and M. Ghil, "A Multi-Model Ensemble Kalman Filter for Data Assimilation and Forecasting," Journal of Advances in Modeling Earth Systems **15**, e2022MS003123 (2023).

[43] C. M. Danforth and E. Kalnay, "Using Singular Value Decomposition to Parameterize State-Dependent Model Errors," Journal of the Atmospheric Sciences **65**, 1467–1478 (2008).

[44] T.-C. Chen, S. G. Penny, J. S. Whitaker, S. Frolov, R. Pincus, and S. Tulich, "Correcting Systematic and State-Dependent Errors in the NOAA FV3-GFS Using Neural Networks," Journal of Advances in Modeling Earth Systems **14**, e2022MS003309 (2022).

[45] A. Farchi, P. Laloyaux, M. Bonavita, and M. Bocquet, "Using machine learning to correct model error in data assimilation and forecast applications," Quarterly Journal of the Royal Meteorological Society **147**, 3067–3084 (2021).

[46] S. Ephrati, A. Franken, E. Luesink, and B. Geurts, "Data-assimilation closure for large-eddy simulation of quasi-geostrophic flow on the sphere," (2023), arxiv:2312.12858 [physics].

[47] A. Bain and D. Crisan, *Fundamentals of Stochastic Filtering*, Stochastic Modelling and Applied Probability (Springer, New York, NY, 2009).

[48] M. Salgado, R. Middleton, and G. C. Goodwin, "Connection between continuous and discrete Riccati equations with applications to Kalman filtering," IEE Proceedings D (Control Theory and Applications) **135**, 28–34 (1988).

[49] D. Simon, *Optimal State Estimation: Kalman, H_∞, and Nonlinear Approaches* (John Wiley & Sons, Ltd, 2006).

[50] Y. Zhou, J. Shi, and J. Zhu, "Nonparametric Score Estimators," in *Proceedings of the 37th International Conference on Machine Learning* (PMLR, 2020) pp. 11513–11522.

[51] M. K. Tippett, J. L. Anderson, C. H. Bishop, T. M. Hamill, and J. S. Whitaker, "Ensemble Square Root Filters," Monthly Weather Review **131**, 1485–1490 (2003).

[52] R. Flamary, N. Courty, A. Gramfort, M. Z. Alaya, A. Boisbunon, S. Chambon, L. Chapel, A. Corenflos, K. Fatras, N. Fournier, L. Gautheron, N. T. H. Gayraud, H. Janati, A. Rakotomamonjy, I. Redko, A. Rolet, A. Schutz, V. Seguy, D. J. Sutherland, R. Tavenard, A. Tong, and T. Vayer, "POT: Python Optimal Transport," Journal of Machine Learning Research **22**, 1–8 (2021).

[53] E. Bach, "parasweep: A template-based utility for generating, dispatching, and post-processing of parameter sweeps," SoftwareX **13**, 100631 (2021).

[54] A.-K. Kassam and L. N. Trefethen, "Fourth-Order Time-Stepping for Stiff PDEs," SIAM Journal on Scientific Computing **26**, 1214–1233 (2005).

[55] M. Dashti, K. J. H. Law, A. M. Stuart, and J. Voss, "MAP estimators and their consistency in Bayesian nonparametric inverse problems," Inverse Problems **29**, 095017 (2013).

[56] E. N. Lorenz, "Deterministic Nonperiodic Flow," Journal of the Atmospheric Sciences **20**, 130–141 (1963).

[57] E. N. Lorenz, "Predictability: A problem partly solved," in *Proceedings of a Seminar Held at ECMWF on Predictability, 4-8 September 1995*, Vol. 1 (ECMWF, Shinfield Park, Reading, 1996) pp. 1–18.

[58] V. I. Bogachev, *Gaussian Measures*, Mathematical Surveys and Monographs No. 62 (American Mathematical Society, 1998).

[59] P. L. Falb, "Infinite-dimensional filtering: The Kalman–Bucy filter in Hilbert space," Information and Control **11**, 102–137 (1967).

[60] R. Curtain, "A Survey of Infinite-Dimensional Filtering," SIAM Review **17**, 395–411 (1975).

Appendix A: Properties of the KB Filter for Densities

Lemma 1 and Remark 2 below give the conditions under which $m(t)$ and $\rho(t) \sim \mathcal{N}(m(t), C(t))$ will be normalized. The function $\mathbb{1}$ is defined as $\mathbb{1}(v) \equiv 1$ for all v.

Lemma 1. *Assume that* $\rho(0) \sim \mu(0) = \mathcal{N}(m_0, C_0)$ *with*

$$\begin{cases} \langle m_0, \mathbb{1} \rangle_{\mathcal{H}} = 1, \\ C_0 \mathbb{1} = 0. \end{cases} \tag{A.1}$$

Then, for $m(t)$ *and* $C(t)$ *satisfying equations* (IV.5a)–(IV.5c),

(a) $C(t)\mathbb{1} = 0$ *for all* $t \geq 0$, *and*

(b) $\langle m(t), \mathbb{1} \rangle_{\mathcal{H}} = 1$ *for all* $t \geq 0$.

Proof. (Sketch)

(a) Since $\mathcal{L}\mathbb{1} = 0$, we have

$$\frac{d}{dt}(C\mathbb{1}) = \mathcal{L}^* C\mathbb{1} - CH^*\Gamma^{-1}HC\mathbb{1}. \tag{A.2}$$

Assuming uniqueness of the solution to the equation (IV.5b) for the evolution of $C(t)$, we deduce that $C(t)\mathbb{1} = 0$ solves Eq. (A.2).

(b) Applying Itô's lemma to $\langle m, \mathbb{1} \rangle_{\mathcal{H}}$ (the Itô correction does not appear due to linearity of the inner product),

$$\begin{aligned} \frac{d}{dt}\langle m, \mathbb{1} \rangle_{\mathcal{H}} &= \langle \mathcal{L}^* m, \mathbb{1} \rangle_{\mathcal{H}} + \left\langle CH^*\Gamma^{-1}(dz^\dagger - Hm), \mathbb{1} \right\rangle_{\mathcal{H}}, \\ &= \langle m, \mathcal{L}\mathbb{1} \rangle_{\mathcal{H}} + \left\langle H^*\Gamma^{-1}(dz^\dagger - Hm), C\mathbb{1} \right\rangle_{\mathcal{H}}, \\ &= 0, \end{aligned} \tag{A.3}$$

since $\mathcal{L}\mathbb{1} = 0$, C is self-adjoint by construction, and $C\mathbb{1} = 0$ by (a). Now assuming uniqueness of the equation (IV.5a) for $m(t)$ we find that, $\langle m(t), \mathbb{1} \rangle_{\mathcal{H}} = 1$ solves Eq. (A.3).

□

Remark 2. *If the conditions in Eq.* (A.1) *hold then* $\langle \rho(t), \mathbb{1} \rangle = 1$ *for* $t \geq 0$ *almost surely, where* $\rho(t) \sim \mu(t) = \mathcal{N}(m(t), C(t))$. *This is because* $\mathbb{1}$ *is in the null-space of both the symmetric operator square-root of* $C(t)$, $\sqrt{C(t)}$, *and*

$$\rho(t) = m(t) + \sqrt{C(t)}\xi, \tag{A.4}$$

where $\xi \sim \mathcal{N}(0, \mathbb{I})$. Thus

$$\begin{aligned} \langle \rho(t), \mathbb{1} \rangle_{\mathcal{H}} &= \langle m(t), \mathbb{1} \rangle_{\mathcal{H}} + \langle \sqrt{C(t)}\xi, \mathbb{1} \rangle_{\mathcal{H}}, \\ &= 1 + \langle \xi, \sqrt{C(t)}\mathbb{1} \rangle_{\mathcal{H}}, \\ &= 1. \end{aligned} \tag{A.5}$$

This explains the importance of the conditions in Eq. (A.1)*: they ensure that $\rho(t)$ is normalized.*

Appendix B: Theorem 1

Theorem 1. *Assume that:*

1. *The system dynamics f and $\mathfrak{h}$ are linear in state space: $f(v,t) = \mathsf{L}^T v$ and $\mathfrak{h}(v,t) = \mathsf{H}v$, with injective H.*
2. *$\Sigma = 0$.*
3. *$\rho(0)$ is chosen such that its mean $\mathsf{m}(0)$ and covariance $\mathsf{C}(0)$ satisfy*

$$\begin{cases} \mathsf{Hm}(0) = Hm_0, \\ \mathsf{HC}(0)\mathsf{H}^T = HC_0H^*. \end{cases} \tag{B.1}$$

4. *$m(t)$ stays in the subspace*

$$\mathcal{S} \equiv \left\{ u \in \mathcal{H} \,\middle|\, \int |u(v)| v_i v_j dv < \infty \; \forall i,j \in \{1,\ldots,d\} \right\}$$

and $C(t)$ stays in $L(\mathcal{S},\mathcal{S})$, the space of bounded linear operators from $\mathcal{S}$ into itself.

Then, under the same noise realization for $Z^\dagger$, $\mathsf{Hm}(t) = Hm(t)$ and $\mathsf{HC}(t)\mathsf{H}^T = HC(t)H^*$ will hold for $t \geq 0$, where $\mathsf{m}(t)$ and $\mathsf{C}(t)$ are the mean and covariance of $\rho(t)$ obtained from Eq. (IV.10)*, and $m(t)$ and $C(t)$ are given by the KB filter for densities* (IV.5a)–(IV.5c).

Proof. (Sketch)

We give here the outlines of a proof, but a rigorous proof, as well as analysis of whether the equivalence holds in any setting more general than the above restrictive conditions, will require considerably more work.

We consider the evolution of the mean and covariance of the KB filter for densities (Eqs. (IV.5a) and (IV.5b)) projected into observation space:

$$d(Hm) = H\mathcal{L}^* m\, dt + HCH^*\Gamma^{-1}(dz^\dagger - Hm\, dt), \tag{B.2a}$$

$$d(HCH^*) = H\mathcal{L}^* CH^*\, dt + HC\mathcal{L}H^*\, dt - HCH^*\Gamma^{-1}HCH^*\, dt, \tag{B.2b}$$

where $H(t) = H$ is not time-dependent because $\mathfrak{h}(v,t) = \mathfrak{h}(v) = \mathsf{H}v$. These equations now describe the time evolution of the finite-dimensional quantities Hm and HCH^*.

Now, imposing $f(v,t) = \mathsf{L}^T v$ and $\mathfrak{h}(v,t) = \mathsf{H}v$ on the ansatz (Eq. (IV.10)), the time evolution of ρ can be entirely characterized by its mean and covariance, and we obtain the following equations for them:

$$d\mathsf{m} = \mathsf{L}^T\mathsf{m}\, dt + \mathsf{CH}^T\Gamma^{-1}(dz^\dagger - \mathsf{Hm}\, dt), \tag{B.3a}$$

$$d\mathsf{C} = \mathsf{L}^T\mathsf{C}\, dt + \mathsf{CL}\, dt - \mathsf{CH}^T\Gamma^{-1}\mathsf{HC}\, dt, \tag{B.3b}$$

where $\mathsf{m} \equiv \mathbb{E}[v]$ and $\mathsf{C} \equiv \mathbb{E}[(v-\mathsf{m})(v-\mathsf{m})^T]$. A similar calculation is made in, e.g., section 7.4 of Jazwinski (1970)[1]. In observation space, we have that

$$d(\mathsf{Hm}) = \mathsf{HL}^T\mathsf{m}\, dt + \mathsf{HCH}^T\Gamma^{-1}(dz^\dagger - \mathsf{Hm}\, dt), \tag{B.4a}$$

$$d(\mathsf{HCH}^T) = \mathsf{HL}^T\mathsf{CH}^T\, dt + \mathsf{HCLH}^T\, dt - \mathsf{HCH}^T\Gamma^{-1}\mathsf{HCH}^T\, dt. \tag{B.4b}$$

We would now like to show that $\mathsf{Hm}(t) = Hm(t)$ and $\mathsf{HC}(t)\mathsf{H}^T = HC(t)H^*$ for all $t \geq 0$. We do this by showing that the RHS of Eqs. (B.2a) and (B.2b) are equal to the RHS of Eqs. (B.4a) and (B.4b) at time t^* if $\mathsf{Hm}(t^*) = Hm(t^*)$ and $\mathsf{HC}(t^*)\mathsf{H}^T = HC(t^*)H^*$. Together with the initial conditions (B.1) and uniqueness, this proves the theorem.

It follows immediately that

$$\begin{aligned} &HC(t^*)H^*\Gamma^{-1}\left[\frac{dz^\dagger}{dt} - Hm(t^*)\right] \\ &\qquad = \mathsf{HC}(t^*)\mathsf{H}^T\Gamma^{-1}\left[\frac{dz^\dagger}{dt} - \mathsf{Hm}(t^*)\right], \end{aligned} \tag{B.5}$$

and that

$$HC(t^*)H^*\Gamma^{-1}HC(t^*)H^* = \mathsf{HC}(t^*)\mathsf{H}^T\Gamma^{-1}\mathsf{HC}(t^*)\mathsf{H}^T. \tag{B.6}$$

Note that

$$\begin{aligned} \mathsf{Hm}(t^*) &= Hm(t^*), \\ &= \mathsf{H}\int v m(t^*)\, dv, \end{aligned} \tag{B.7}$$

which implies that

$$\mathsf{m}(t^*) = \int v m(t^*)\, dv, \tag{B.8}$$

because H was assumed to be injective.

We proceed with the rest of the terms. For the first term of the RHS of Eq. (B.2a),

$$\begin{aligned} H\mathcal{L}^* m &= \mathsf{H}\int v\mathcal{L}^* m\, dv, \\ &= -\mathsf{H}\int v\nabla\cdot(mf)\, dv, \\ &= -\mathsf{HL}^T\int v\nabla\cdot(mv)\, dv, \\ &= \mathsf{HL}^T\int vm\, dv, \\ &= \mathsf{HL}^T\mathsf{m}, \end{aligned} \tag{B.9}$$

where the fourth line follows from integration by parts, and the last from Eq. (B.8). Note that the boundary term in the integration by parts vanishes from assumption 4. Thus,

$$H\mathcal{L}^* m = \mathsf{H}\mathsf{L}^T \mathsf{m}. \tag{B.10}$$

It remains to show that $H\mathcal{L}^* C(t^*)H^* = \mathsf{H}\mathsf{L}^T \mathsf{C}(t^*)\mathsf{H}^T$. We have that for any u,

$$HC(t^*)H^* u = \mathsf{H}\int vC(t^*)H^* u\, dv = \mathsf{H}\mathsf{C}(t^*)\mathsf{H}^T u. \tag{B.11}$$

Since H was assumed to be injective,

$$\int vC(t^*)H^* u\, dv = \mathsf{C}(t^*)\mathsf{H}^T u. \tag{B.12}$$

Then, for any w,

$$\begin{aligned} H\mathcal{L}^* C(t^*)H^* w &= \mathsf{H}\int v\mathcal{L}^* C(t^*)H^* w\, dv \\ &= -\mathsf{H}\int v\nabla\cdot\left(C(t^*)H^* w \mathsf{L}^T v\right) dv \\ &= \mathsf{H}\mathsf{L}^T \int vC(t^*)H^* w\, dv \\ &= \mathsf{H}\mathsf{L}^T \mathsf{C}(t^*)\mathsf{H}^T w \end{aligned} \tag{B.13}$$

where the third line follows from integration by parts (with the boundary term vanishing by the same argument as above), and the last line from Eq. (B.12). Taking the adjoint demonstrates that $HC(v,t^*)\mathcal{L}H^* = \mathsf{H}\mathsf{C}(t^*)\mathsf{L}\mathsf{H}^T$, completing the proof. □

Appendix C: Mean-Field Approximation

We omit the function arguments until the end of the subsection, for brevity. Using Eq. 3.30 from Calvello, Reich, and Stuart (2022)[9], we know that the FP equation of Eq. (IV.11) when $f(v,t) = 0$ and $\Sigma = 0$ is

$$\begin{aligned} \frac{\partial \rho}{\partial t} &= -\nabla\cdot(\rho(a - KH\rho)) - \left\langle \nabla\cdot(\rho K^T), \frac{dz^\dagger}{dt} \right\rangle \\ &\quad + \nabla\cdot\left(\nabla\cdot(\rho K\Gamma K^T)\right). \end{aligned} \tag{C.1}$$

We now match the terms of Eqs. (C.1) and (IV.10) to make them equal. By matching the terms involving $dz^\dagger/dt$, we obtain that

$$\Gamma^{-1}(\mathfrak{h} - H\rho)\rho = -\nabla\cdot(\rho K^T), \tag{C.2}$$

and matching the rest of the terms,

$$-\rho\langle \mathfrak{h} - H\rho, H\rho\rangle_\Gamma = -\nabla\cdot(\rho(a - KH\rho)) + \nabla\cdot(\nabla\cdot(\rho K\Gamma K^T)). \tag{C.3}$$

Substituting Eq. (C.2) into Eq. (C.3), we obtain

$$\begin{aligned} \langle \nabla\cdot(\rho K^T), H\rho\rangle &= \nabla\cdot(\rho K H\rho) \\ &= -\nabla\cdot(\rho(a - KH\rho)) \\ &\quad + \nabla\cdot(\nabla\cdot(\rho K\Gamma K^T)). \end{aligned} \tag{C.4}$$

Setting the term in the divergence to 0, we obtain

$$a = K\Gamma K^T \nabla \log \rho. \tag{C.5}$$

This is the origin of the score function term discussed in subsection II E.

We propose a test function $\psi(v) = v - \mathbb{E}v$, take the outer product of it with both sides of Eq. (C.2), and integrate by parts, obtaining the identity

$$\mathbb{E}K = \mathbb{E}[\psi(\mathfrak{h} - H\rho)^T]\Gamma^{-1} = C^{v\mathfrak{h}}\Gamma^{-1}, \tag{C.6}$$

where $C^{v\mathfrak{h}}(t) \equiv \mathbb{E}[(\mathfrak{h}(v,t) - H\rho)(\mathfrak{h}(v,t) - H\rho)^T]$.

Fixing the value of the gain K to its expectation (the constant gain approximation discussed in Calvello, Reich, and Stuart (2022)[9]), we then obtain

$$K(t) = C^{v\mathfrak{h}}(t)\Gamma(t)^{-1}. \tag{C.7}$$

Thus, the mean-field model is

$$\begin{aligned} dv &= f(v,t)\, dt + \sqrt{\Sigma(t)}\, dW + K(t)\left(dz^\dagger - d\hat{z}\right) \\ &\quad + K(t)\Gamma(t)K(t)^T \nabla \log \rho(v,t)\, dt, \\ d\hat{z} &= (\mathbb{E}\mathfrak{h})(t)\, dt + \sqrt{\Gamma(t)} dB, \end{aligned}$$

which gives Eqs. (II.4), with (II.4a) replaced by (II.5).