Acta Numerica (2025), pp. 123–291
doi:10.1017/S0962492924000060

Ensemble Kalman methods: A mean-field perspective

Edoardo Calvello
California Institute of Technology, Pasadena, CA 91125, USA
E-mail: e.calvello@caltech.edu

Sebastian Reich
Institut für Mathematik, Universität Potsdam, D-14476 Potsdam, Germany
E-mail: sebastian.reich@uni-potsdam.de

Andrew M. Stuart
California Institute of Technology, Pasadena, CA 91125, USA
E-mail: astuart@caltech.edu

Ensemble Kalman methods, introduced in 1994 in the context of ocean state estimation, are now widely used for state estimation and parameter estimation (inverse problems) in many arenae. Their success stems from the fact that they take an underlying computational model as a black box to provide a systematic, derivative-free methodology for incorporating observations; furthermore the ensemble approach allows for sensitivities and uncertainties to be calculated. Analysis of the accuracy of ensemble Kalman methods, especially in terms of uncertainty quantification, is lagging behind empirical success; this paper provides a unifying mean-field-based framework for their analysis. Both state estimation and parameter estimation problems are considered, and formulations in both discrete and continuous time are employed. For state estimation problems, both the control and filtering approaches are considered; analogously for parameter estimation problems, the optimization and Bayesian perspectives are both studied. As well as providing an elegant framework, the mean-field perspective also allows for the derivation of a variety of methods used in practice. In addition it unifies a wide-ranging literature in the field and suggests open problems.

2020 Mathematics Subject Classification: 35, 37, 60, 62, 65, 93

CONTENTS

1. Introduction

The ensemble Kalman methodology comprises an innovative and flexible set of tools which can be used for both state estimation in dynamical systems and parameter estimation for generic inverse problems. It has primarily been developed by practitioners in the geophysical sciences, with notable impact on the fields of oceanography, oil reservoir simulation and weather forecasting. Despite its widespread adoption in the geosciences over several decades, firm theoretical foundations are only recently starting to emerge; the methodology is hard to analyse. The purpose of this article is twofold: (a) to introduce a mathematical framework for the analysis of ensemble Kalman methods, describing what is known and highlighting the many open mathematical challenges in the field, (b) to provide a literature survey which bridges the domain-specific development of the methodology with emerging mathematical analyses. In so doing we will also highlight the flexibility of the methodology for use in widespread applications, beyond its historical development in the geosciences.

The material is organized around the two separate themes of state estimation and inverse problems; within each, both discrete-time and continuous-time approaches are explained. The novel perspective which underlies all of this material is the derivation of ensemble Kalman methods as particle approximations of carefully designed mean-field models. The relationship of these mean-field models to exact transport models, for Gaussian problems, serves to motivate their form.

In Section 1.1 we give an overview of the history of ensemble Kalman methods. Section 1.2 describes the organization of the paper. In Section 1.3 we make brief remarks about the pseudo-code that we make available as a supplementary resource. The introduction concludes, in Section 1.4, with a summary of the notation that we adopt throughout.

1.1. Historical context

The Kalman filter (KF) is arguably the first setting in which the systematic integration of observational data with a dynamical system was considered, leading to both discrete-time (Kalman 1960) and continuous-time (Kalman and Bucy 1961) formulations; see Welch and Bishop (1995) for an overview. The Kalman filter applies only in the setting of linear Gaussian dynamics and observations. In this setting it computes the distribution of the state of the dynamical system, given observations, exactly; this Bayesian perspective on the filter was highlighted in Ho and Lee (1964) subsequent to the original derivation in Kalman (1960), which proceeded by computing the best linear predictor of the state, given data. The extended Kalman filter (historically denoted EKF, but the acronym ExKF is also used and is useful) was introduced in order to extend Kalman's ideas to nonlinear problems; see the books by Jazwinski (2007) and Anderson and Moore (2012) for overviews. The extended Kalman approach is based on a linearization approximation; it hence fails to exactly compute the distribution of the state of the dynamical system, given observations, in general. Furthermore, it requires propagation of covariance matrices, which can be very large for applications arising in the geosciences (Ghil *et al.* 1981).

The ensemble Kalman filter (EnKF) was introduced in the celebrated paper by Evensen (1994), which made the consequential observation that if an *ensemble* of state estimators is employed then it can also be used to make an approximation of the covariance. In geosciences applications this circumvents the computation of large covariances, replacing them instead with low-rank approximations, with rank determined by the number of ensemble members. The original paper developed the idea in the context of ocean models, but was rapidly and concurrently developed in a variety of geoscience application domains (van Leeuwen and Evensen 1996, Burgers, van Leeuwen and Evensen 1998, Houtekamer and Mitchell 1998); van Leeuwen (2020) provides a historical overview. These methods are sometimes referred to as the *stochastic EnKF*: they require simulation of random variables to implement. A different class of ensemble methods, known collectively as *ensemble square root filters*, was subsequently developed (Anderson 2001, Whitaker and Hamill 2002, Bishop, Etherton and Majumdar 2001, Hunt, Kostelich and Szunyogh 2007, Tippett *et al.* 2003, Sakov and Oke 2008). These methods are a form of *deterministic EnKF*: they do not require simulation of random variables to implement.

Central to our mathematical presentation of Kalman-based methods is the adoption of mean-field and transport perspectives on the subject. The incorporation of data within filtering constitutes (possibly approximate) application of Bayes' theorem; Daum, Huang and Noushin (2010), Reich (2011), El Moselhy and Marzouk (2012) and the survey by Cotter and Reich (2013) introduce novel approaches to Bayesian inversion, rooted in transport and mean-field models. While El Moselhy and Marzouk (2012) and Spantini, Baptista and Marzouk (2022) propose a direct numerical approximation of the underlying optimal transport problem, Daum *et al.*

(2010) and Reich (2011) pursue a homotopy approach. We note that the homotopy approach is closely related to iterative implementations of the EnKF, as first considered by Li and Reynolds (2009), Gu and Oliver (2007) and Sakov, Oliver and Bertino (2012); these homotopy approaches lead to continuous-time formulations of the EnKF in the limit of infinitely many iterations, as first considered by Bergemann and Reich (2010*a*,*b*). Directly starting from the continuous-time filtering perspective, mean-field models have been introduced independently in Crisan and Xiong (2010) and Yang, Mehta and Meyn (2013).

The key connection between mean-field models and ensemble methods is that the latter can be derived as particle approximations of the mean-field limit; this viewpoint will play a guiding role in our presentation of the subject of ensemble methods in this paper. In this context it is notable that the field of optimization, which is linked to Bayesian sampling through MAP estimation (Kaipio and Somersalo 2006), has also seen recent development using mean-field models; see Carrillo, Choi, Totzeck and Tse (2018) for an overview and unifying mathematical framework.

The methods covered in this survey provide only approximate solutions to the underlying filtering, inference and/or optimization problem, in the mean-field limit. The approximations invoked are based on assuming linear Gaussian structure, where the mean-field models are exact, but applying the resulting methodology outside this regime. In the context of the optimization and Bayesian approaches to inversion, affine-invariant algorithms (introduced in Goodman and Weare 2010) play an important conceptual role in understanding the power of ensemble Kalman methods: affine invariance can be used to show universal convergence rates for the class of all linear Gaussian problems. Empirically the affine invariance confers advantages when ensemble Kalman methods are applied beyond the linear Gaussian setting. In practice, this benefit must be weighed against the error resulting from using ensemble Kalman methods outside the linear Gaussian setting.

Alternative methods, such as sequential Monte Carlo, can be designed to be consistent with the underlying nonlinear filtering problem, and do not rely on being exact only for linear Gaussian problems. Doucet, De Freitas and Gordon (2001) and Chopin and Papaspiliopoulos (2020) survey the use of sequential Monte Carlo methods for general discrete-time filtering and inference problems; Del Moral (1997) and Del Moral and Guionnet (2001) prove convergence of sequential Monte Carlo methods, including in some specific cases over long time horizons. However, sequential Monte Carlo methods suffer from the curse of dimensionality and are currently not directly applicable to high-dimensional problems as arising, for example, from geophysical applications. This issue with the curse of dimensionality provides an important motivation for the ensemble methods covered in this survey. See Snyder, Bengtsson, Bickel and Anderson (2008), Bickel, Li and Bengtsson (2008), Rebeschini and Van Handel (2015) and Agapiou, Papaspiliopoulos, Sanz-Alonso and Stuart (2017) for detailed discussion of these issues. Related issues also arise for Monte Carlo Markov chain (MCMC) when studying Bayesian inverse

problems (Kaipio and Somersalo 2006). See Hairer, Stuart and Vollmer (2014) for an analysis of the degeneration of performance of standard MCMC methods in high dimensions, as well as analysis of special MCMC methods tailored to infinite-dimensional problems; the subject is reviewed in Cotter, Roberts, Stuart and White (2013).

This completes our chronological overview of the historical context for the development of ensemble Kalman methods, and the specific mathematical context which will be our focus. Each of the four subsequent sections concludes with a bibliographical subsection in which a deeper literature review is given.

1.2. Overview

Section 2 is devoted to the problem of state estimation for discrete-time (possibly stochastic) dynamical systems, given noisy, and possibly indirect, observations. We formulate the problem from the perspectives of both control theory and probability, and we provide a unifying approach to algorithms for these problems; the approach rests on transport of measures and mean-field stochastic dynamical systems. Ensemble Kalman methods are then derived as particle approximations of the mean-field models. Section 3 adopts a perspective that parallels the previous section, but in the continuous-time setting. Ordinary differential equations (ODEs) and stochastic differential equations (SDEs) are used to describe the state and its observation process, and mean-field SDEs and ODEs, and related (stochastic) partial differential equations ((S)PDEs), are used to provide the underpinnings of algorithms; particle approximations of the mean-field systems give rise to interacting systems of SDEs which describe ensemble Kalman methods. The formulation in continuous time is useful both because in some applications state estimation problems are most naturally formulated this way, and because they provide insight into discrete-time algorithms, giving rise to cleaner analysis of phenomena present in both discrete and continuous time.

Sections 4 and 5 are devoted to the use of ensemble Kalman methods for inverse problems, including parameter estimation, demonstrating how a useful change of perspective opens up the use of state estimation in this broader setting. Sections 4 and 5 consider discrete and continuous time respectively, and each parallels the ideas developed for state estimation in Sections 2 and 3 respectively. Sections 2, 3, 4 and 5 are all organized according to the flow of ideas displayed in Figure 1.1. Indeed, starting from a probabilistic perspective, which describes the evolution of the filtering distribution, it is possible to formulate mean-field maps whose sample paths on state space are equal in law to the filtering evolution. Since it is typically not tractable to identify these maps explicitly, it is of interest to determine mean-field maps whose sample paths are only *approximately* equal in law to the filtering evolution. This leads to the idea of second-order mean-field models whose sample paths have law with first- and second-order moments which match those of a Gaussian approximation of the Bayesian data incorporation step. Particle

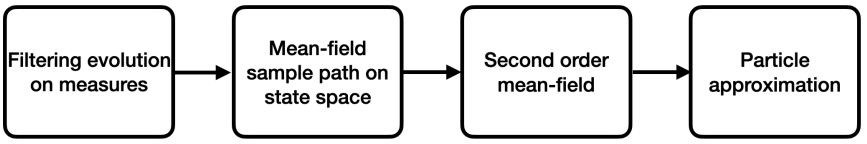


Figure 1.1. Organizational flow employed in each of Sections 2 (discrete time) and 3 (continuous time), concerning state estimation. Sections 4 and 5 apply this methodology to inverse problems, in discrete and continuous time respectively, by formulating them as state estimation problems.

approximations of these second-order mean-field maps are then used to derive implementable numerical algorithms which are the ensemble Kalman methods implemented in practice.

Section 6 concludes the article, and in particular highlights open problems in the area, of potential interest to the mathematical community. Appendix A is devoted to pseudo-code for the algorithms introduced in this paper. Appendix B contains background on the underlying Lorenz '96 model problems that we use throughout the paper in a number of illustrative numerical experiments. Appendices C and D provide foundational material on mean-field maps and on stochastic integration. Appendix E contains some observations linking different flows, in the manifold of Gaussian probability measures, that arise in the main body of the text.

1.3. Pseudo-code

Pseudo-code describing several of the algorithms that we present and deploy in this paper is given in Appendix A.[1] The reader is encouraged to consult Algorithms 1 and 2, 3DVAR and the ensemble Kalman filter (EnKF) respectively, in the context of the problem of state estimation for discrete-time dynamical systems presented in Section 2. The scheme 3DVAR is employed in Examples 2.3, 2.5 and 2.16. The ensemble Kalman filter is applied in Example 2.16. Ensemble Kalman methods for inversion, as shown in Algorithms 3, 4 and 5, are presented in Section 4 and applied in Examples 4.22 and 4.23.

1.4. Notation

Throughout we denote the positive integers and non-negative integers by $\mathbb{N} = \{1, 2, \ldots\}$ and $\mathbb{Z}^+ = \{0, 1, 2, \ldots\}$ respectively, and the notation $\mathbb{R} = (-\infty, \infty)$ and $\mathbb{R}^+ = [0, \infty)$ for the reals and the non-negative reals. We let $\langle \cdot, \cdot \rangle, |\cdot|$ denote the Euclidean inner product and norm, noting that $|v|^2 = \langle v, v \rangle$. We also use $|\cdot|$ to denote the resulting induced norm on matrices. We use : to denote the Frobenius inner product between matrices, and $|\cdot|_{\mathrm{F}}$ the induced norm on matrices. For any function $g \colon \mathbb{R}^{d_1} \mapsto \mathbb{R}^{d_2}$, we let $Dg(v) \in \mathbb{R}^{d_2 \times d_1}$ denote the Jacobian matrix of

[1] The code is available at https://github.com/EdoardoCalvello/EnsembleKalmanMethods/.

first derivatives at $v \in \mathbb{R}^{d_1}$, and let $D^2 g(v)[\cdot,\cdot]$ denote the symmetric bilinear form induced by second derivatives. We also use ∇, $\nabla\cdot$ and Δ to denote the gradient, divergence and Laplacian operations respectively.

Throughout the article we will distinguish operators acting on infinite-dimensional spaces by using a calligraphic font. For example, $\mathcal{G}$ will denote the operator acting on probability measures effecting a projection onto the nearest Gaussian; in contrast, $G\colon \mathbb{R}^{d_u} \to \mathbb{R}^{d_w}$ will denote the forward model in a generic inverse problem. We will use the mathsf font to distinguish matrices and functions between Euclidean spaces that arise in continuous time from their discrete-time counterparts. For example, $\mathsf{\Sigma}$ will denote a covariance arising in continuous time, whilst Σ will denote a covariance arising in discrete time; and h will denote a function defining the observation process in continuous time, whilst h will be the analogous function in discrete time.

We use $\mathbb{E}$ and $\mathbb{P}$ to denote expectation and probability under the prevailing probability measure; if we wish to make clear that measure π is the prevailing probability measure then we write $\mathbb{E}^\pi$ and $\mathbb{P}^\pi$. For a measure π we let $T^\sharp\pi$ denote the pushforward measure induced by the map T. We use $Law(\mathsf{rv})$ to denote the law of random variable rv. We let $\mathfrak{P} = \mathfrak{P}(\mathbb{R}^d)$ denote the set of probability measures on $\mathbb{R}^d$. Under the assumptions made in this article we are mostly able to work with measures that have density with respect to Lebesgue measure, and so will blur the distinction between measures and their densities. However, use of Dirac masses will occasionally be useful.

We let δ_u denote the Dirac mass on $\mathbb{R}^d$, centred at point $u \in \mathbb{R}^d$. The notation $\mathsf{N}(m, C)$ denotes the distribution of a Gaussian random variable with mean m and covariance C. We let $\mathfrak{G} = \mathfrak{G}(\mathbb{R}^d)$ denote the set of Gaussian probability measures on $\mathbb{R}^d$ (including indefinite covariances and hence all Dirac masses).

In the following let A, B, C be symmetric matrices. We write $A \succ B$ when $A - B$ is positive definite and $A \succeq B$ when $A - B$ is positive semi-definite. We will also write $A \prec B$ when $B - A \succ 0$ and $A \preceq B$ when $B - A \succeq 0$. For $C \succ 0$ (and therefore a covariance matrix) we define $\langle\cdot,\cdot\rangle_C$ and $|\cdot|_C$, the covariance-weighted Euclidean inner product and norm, by $\langle u, v\rangle_C = \langle u, C^{-1}v\rangle$ and $|v|_C^2 = \langle v, v\rangle_C$.

2. State estimation: discrete time

In Section 2.1 we provide the set-up for the problem of state estimation in discrete time. Section 2.2 introduces an algorithm for this problem based on a control-theoretic perspective. Section 2.3 describes the Bayesian probabilistic perspective; in this subsection algorithms are not presented but foundations for the creation of algorithms are laid through the decomposition of the filtering cycle into an iteration which alternates prediction and the assimilation of data. In Section 2.4 we introduce the important idea of Gaussian projection, and the resulting Gaussian projected filter. Section 2.5 introduces various mean-field dynamical systems which approximate the filtering cycle; a unifying transport map perspective is

adopted. This leads, in Section 2.6, to definitions of the ensemble Kalman filter, the ensemble adjustment filter and the ensemble transform filter, all derived as particle approximations of specific mean-field models. We conclude in Section 2.7 with bibliographical notes in which we review relevant literature, and include discussion of a variety of other algorithms for state estimation, relating them to the perspective we adopt here.

We emphasize that all the mean-field methods we derive in this section for the solution of filtering problems, and their continuous-time counterparts in Section 3, are exact only in the linear Gaussian setting. They may be implemented beyond this setting, however, and are empirically found to work well for many problems. Theory to justify this observation is very much needed. Our exposition provides a framework for such a theory.

2.1. *Set-up*

The objective of *sequential data assimilation* is to iteratively update the state of a (possibly stochastic) dynamical system based on (possibly noisy, nonlinear and indirect) observations and knowledge of the dynamical and observational processes; we refer to this as *state estimation*. The typical setting is one in which the initial condition is uncertain, but this uncertainty is compensated for by (typically noisy) observations of a (possibly nonlinear and indirect) function of the state. These observations often live in a space of lower dimension than the dimension of the state space, meaning that the goal of state estimation goes beyond denoising, and into the realm of control-theoretic considerations such as observability.

A useful starting point is to consider a stochastic dynamical system in which the evolution of the state, and the relationship between the observations (which we also refer to as data) and the state, are defined, respectively, by the equations

$$\boxed{\begin{aligned} v_{n+1} &= \Psi(v_n) + \xi_n, \qquad (2.1a)\\ y_{n+1} &= h(v_{n+1}) + \eta_{n+1}, \qquad (2.1b) \end{aligned}}$$

taken to hold for all $n \in \mathbb{Z}^+$. We assume that for each fixed $n \in \mathbb{Z}^+$, the *state* $v_n \in \mathbb{R}^{d_v}$ and the *observations* $y_n \in \mathbb{R}^{d_y}$. The maps $\Psi(\cdot)$ and $h(\cdot)$ describe the systematic, deterministic components of the dynamics and observation processes, and are assumed to be known measurable functions (with respect to the Borel algebra), bounded on compact sets. The initial condition for v_0 is assumed random and the systematic components of the model are subjected to mean zero noise, ξ_n and η_{n+1}. To be concrete we assume that $v_0, \{\xi_n\}_{n\in\mathbb{Z}^+}$ and $\{\eta_n\}_{n\in\mathbb{N}}$ are mutually independent Gaussians defined by

$$v_0 \sim \mathsf{N}(m_0, C_0), \quad \xi_n \sim \mathsf{N}(0, \Sigma) \text{ i.i.d.}, \quad \eta_n \sim \mathsf{N}(0, \Gamma) \text{ i.i.d.} \qquad (2.2)$$

In practice we will have available to us the observation coordinates of a specific true realization of the random dynamical system (2.1), from which we wish to recover the specific true realization of the state that gave rise to these observations.

We denote the true realizations of the state of the system by the sequence $\{v_n^\dagger\}_{n\in\mathbb{Z}^+}$, and the observed data by $\{y_n^\dagger\}_{n\in\mathbb{N}}$. These are generated by $v_0^\dagger, \{\xi_n^\dagger\}_{n\in\mathbb{Z}^+}$ and $\{\eta_n^\dagger\}_{n\in\mathbb{N}}$, specific realizations of the initial condition and state and observational noise from the distribution defined by (2.2).

We let $Y_n^\dagger = \{y_\ell^\dagger\}_{1\le\ell\le n}$. Our objective is to estimate the state $v_n^\dagger$ at time n from data $Y_n^\dagger$. Specifically, it is natural to think about this design objective in two different ways.

Objective 1. Design an algorithm producing output v_n from $Y_n^\dagger$ so that $\{v_n\}_{n\in\mathbb{Z}^+}$ estimates $\{v_n^\dagger\}_{n\in\mathbb{Z}^+}$, the true state generated by (2.1a).

Objective 2. Design an algorithm which estimates the distribution of random variable $v_n|Y_n^\dagger$, the conditional distribution defined by (2.1).

In both cases we are interested in *Markovian* formulations which update the estimate v_n, or the distribution $v_n|Y_n^\dagger$, sequentially as the data is acquired. In the next two subsections we describe control-theoretic and probabilistic approaches to this problem which, respectively, provide the basis for algorithms addressing Objectives 1 and 2. We note that fulfilling Objective 2 immediately implies resolution of Objective 1, for example by taking the mean of $v_n|Y_n^\dagger$ as state estimator; but the reverse is not typically true. However, Objective 1 is easier to address and is especially relevant when noise levels are small; furthermore, its failure modes serve to motivate approaches which are used to address Objective 2.

2.2. *Control theory perspective*

A very natural idea from control theory underlies ensemble Kalman filtering and is encapsulated in the following algorithmic approach. This way of attacking state estimation is most appropriate when $|C_0|, |\Gamma|$ and $|\Sigma|$ are small so that the state and observations are close to deterministic. To understand this setting we will first study algorithms in which the covariances Γ and Σ are set to zero, and then return to the inclusion of noise later.

The algorithmic idea works as follows: from current state estimate v_n, given $Y_n^\dagger$, *predict* the outcome of the model and data, denoted by $(\widehat{v}_{n+1}, \widehat{h}_{n+1})$, using the update equations (2.1), but ignoring the noise; then *correct* the state estimate by nudging the prediction using the mismatch between observed and predicted data $(y_{n+1}^\dagger, \widehat{h}_{n+1})$. This results in a deterministic map $v_n \mapsto v_{n+1}$, assumed to hold for all $n \in \mathbb{Z}^+$, of the following form:

$$\widehat{v}_{n+1} = \Psi(v_n), \tag{2.3a}$$

$$\widehat{h}_{n+1} = h(\widehat{v}_{n+1}), \tag{2.3b}$$

$$v_{n+1} = \widehat{v}_{n+1} + K(y_{n+1}^\dagger - \widehat{h}_{n+1}). \tag{2.3c}$$

Equations (2.3a) and (2.3b) create predicted state and data from current estimate v_n of the state. The difference between the predicted data $\widehat{h}_{n+1}$ and the true data $y^\dagger_{n+1}$, the latter found from a fixed realization of (2.1), is then used to correct the predicted state resulting in (2.3c). We can write the algorithm compactly in the form

$$\boxed{v_{n+1} = \Psi(v_n) + K\big(y^\dagger_{n+1} - h(\Psi(v_n))\big).} \tag{2.4}$$

Choice of the *gain matrix* K completes definition of an algorithm which we refer to as 3DVAR.

Remark 2.1. The nomenclature '3DVAR' was coined in the geophysics community, and stands for three-dimensional variational data assimilation. This is natural for algorithms which incorporate spatially distributed data sequentially in time, in the context where the state variable v varies across three spatial coordinates. In our setting the state variable v is not required to have any spatial structure, but the control formulation (2.4) reproduces the 3DVAR algorithm from the geophysics community when it does. Hence we still refer to it as 3DVAR. We also note that the method is perhaps more properly termed cycled 3DVAR: '3DVAR' refers to the assimilation of data at each observation time, and 'cycled' refers to doing this sequentially in time as successive observations are acquired. Pseudo-code for 3DVAR may be found as Algorithm 1 in Appendix A. The variant on 3DVAR which uses data distributed over several time steps is known as 4DVAR; see Section 2.7.

We note that the difference between the observed value $y^\dagger_{n+1}$ and its estimator $\widehat{h}_{n+1} = h(\Psi(v_n))$ is often referred to as the *innovation*, and for this we introduce the notation

$$\mathfrak{I}_n = y^\dagger_{n+1} - \widehat{h}_{n+1}. \tag{2.5}$$

□

To illustrate 3DVAR we consider the linear setting.

Example 2.2. Assume that for matrices M, H of appropriate dimensions,

$$\Psi(\cdot) := M\cdot, \quad h(\cdot) = H\cdot \tag{2.6}$$

and consider the setting in which there is no noise present. The 3DVAR algorithm (2.4) is appropriate in this setting and reduces to

$$v_{n+1} = Mv_n + K\big(y^\dagger_{n+1} - HMv_n\big). \tag{2.7}$$

Since there is no noise present, it follows that $y^\dagger_{n+1} = Hv^\dagger_{n+1} = HMv^\dagger_n$, and hence that

$$v^\dagger_{n+1} = Mv^\dagger_n + K\big(y^\dagger_{n+1} - HMv^\dagger_n\big). \tag{2.8}$$

Subtracting (2.8) from (2.7) and defining $e_n = v_n - v^\dagger_n$, we find that

$$e_{n+1} = (I - KH)Me_n.$$

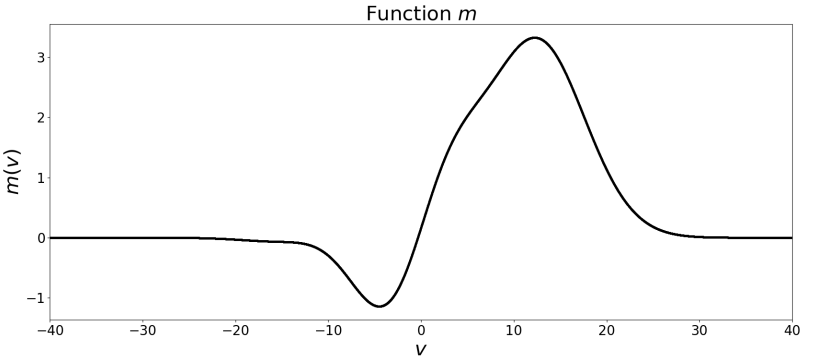


Figure 2.1. Graph of function m appearing in Lorenz '96 model (2.9).

Since the goal of data assimilation is to recover the true state from partial observations, we wish to drive e_n to zero as $n \to \infty$. Thus a key question, for given forward dynamical model M and observation operator H, is whether it is possible to design K to ensure that the spectrum of $(I - KH)M$ is inside the unit circle. Such questions are at the heart of the theory of linear control. Thus the subject of control theory is fundamentally aligned with Objective 1. □

In the following example we illustrate ideas similar to those from the preceding, linear, example but within the nonlinear context. In subsequent subsections we will show how the ideas can be generalized to an adaptive gain matrix K_n, leading us to the ensemble Kalman methodology and to addressing both Objectives 1 and 2.

Example 2.3. To illustrate the 3DVAR algorithm (2.4) for state estimation, we consider the Lorenz '96 (single-scale) model from Appendix B. The unknown $v \in C(\mathbb{R}^+, \mathbb{R}^L)$ satisfies the equations

$$\dot{v}_\ell = -v_{\ell-1}(v_{\ell-2} - v_{\ell+1}) - v_\ell + F + h_v m(v_\ell), \quad \ell = 1 \dots L, \tag{2.9a}$$

$$v_{\ell+L} = v_\ell, \quad \ell = 1 \dots L. \tag{2.9b}$$

Here we set $L = 9$, $h_v = -0.8$ and $F = 10$. Function m is shown in Figure 2.1.[2]

We let Ψ denote the solution operator for (2.9) over the observation time interval τ, omitting the explicit dependence on τ for notational convenience. We emphasize that at the parameter values we have chosen, the solution to (2.9) is chaotic and exhibits sensitivity to perturbations. Prediction is thus challenging. But we will show that use of data enables accurate prediction.

[2] We note that setting $h_v = 0$ in (2.9) leads to the standard single-scale Lorenz '96 model. We have $h_v \neq 0$ leading to a non-standard version of the model. However, the specific choice of function $m(\cdot)$ does not make any material difference to what is presented in this example; many functions $m(\cdot)$ for which the equation is well-posed will lead to similar conclusions. However, the specific choice of $m(\cdot)$ shown in Figure 2.1 is relevant within Example B.1.

We consider observations $\{y_n^\dagger\}_{n\in\mathbb{Z}^+}$ arising from the model

$$v_{n+1}^\dagger = \Psi(v_n^\dagger) + \xi_n^\dagger,$$
$$y_{n+1}^\dagger = h(v_{n+1}^\dagger) + \eta_{n+1}^\dagger,$$

where $\{\xi_n^\dagger\}_{n\in\mathbb{Z}^+}$, $\{\eta_n^\dagger\}_{n\in\mathbb{N}}$ are mutually independent Gaussian sequences defined by

$$\xi_n^\dagger \sim \mathsf{N}(0,\sigma^2 I) \text{ i.i.d.}, \quad \eta_n^\dagger \sim \mathsf{N}(0,\gamma^2 I) \text{ i.i.d.}$$

Because of the chaotic nature of the dynamical system defined by iteration of Ψ, a key question concerning the problem of determining the state $v_n^\dagger$ from $Y_n^\dagger$ is whether the observations compensate for the sensitive dependence of the state evolution, enabling accurate recovery of the state; and whether there is then a choice of K in 3DVAR which enables this data to be used to accurately recover the state.

We assume that the observation function is linear: $h(v) = Hv$ for matrix $H\colon \mathbb{R}^9 \to \mathbb{R}^6$ defined by

$$Hv = (v_1, v_2, v_4, v_5, v_7, v_8)^\top. \tag{2.10}$$

The 3DVAR algorithm (2.3) reduces, in the setting of this example, to the mapping

$$v_{n+1} = (I - KH)\Psi(v_n) + Ky_{n+1}^\dagger. \tag{2.11}$$

We define the filter by choosing gain $K\colon \mathbb{R}^6 \to \mathbb{R}^9$ to be

$$Kw = (w_1, w_2, 0, w_3, w_4, 0, w_5, w_6, 0)^\top. \tag{2.12}$$

To interpret the algorithm, and motivate the choice of K, given H, notice that

$$KHv = (v_1, v_2, 0, v_4, v_5, 0, v_7, v_8, 0)^\top, \tag{2.13a}$$
$$(I - KH)v = (0, 0, v_3, 0, 0, v_6, 0, 0, v_9)^\top, \tag{2.13b}$$
$$HK = I. \tag{2.13c}$$

Applying the observation map H to the recursion (2.11) and using (2.13c), we find that

$$Hv_{n+1} = y_{n+1}^\dagger = H\big(\Psi(v_n^\dagger) + \xi_n^\dagger\big) + \eta_{n+1}^\dagger.$$

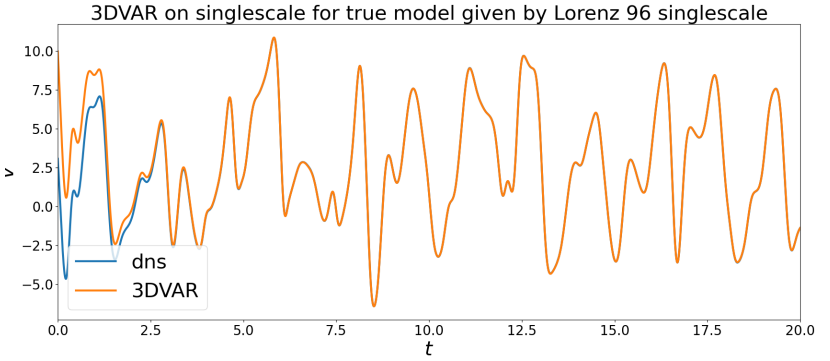
Assuming that σ^2 and γ^2 are small and neglecting the noise contributions shows that

$$y_{n+1}^\dagger \approx H\Psi(v_n^\dagger) \approx Hv_{n+1}^\dagger.$$

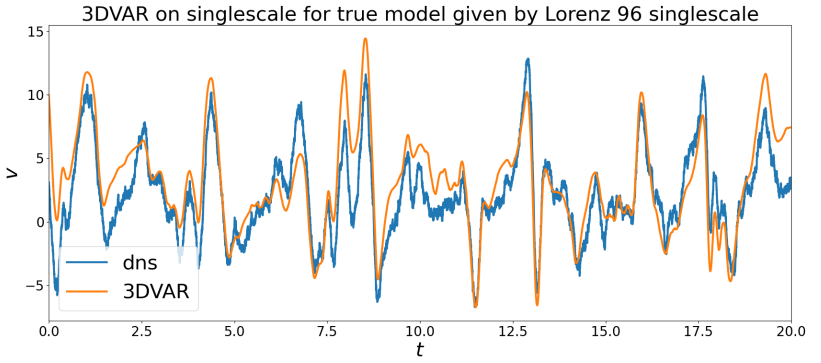
Thus, ignoring small noise perturbations, $y_{n+1}^\dagger \approx Hv_{n+1}^\dagger$. Thus (2.11) gives

$$v_{n+1} \approx (I - KH)\Psi(v_n) + KHv_{n+1}^\dagger. \tag{2.14}$$

Then, using (2.13a), (2.13b) and (2.14),we see that the algorithm (2.11) has the very natural (approximate) interpretation of iterating using the model Ψ to update the unobserved components and using the observed true state to update the observed components. This explains why the specific choice of K is reasonable.



(a)

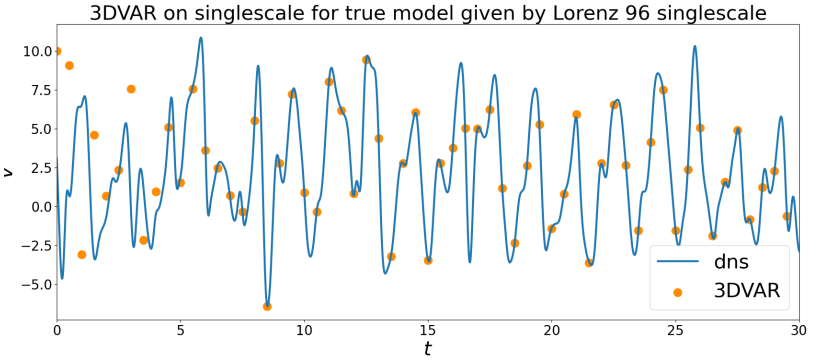


(b)

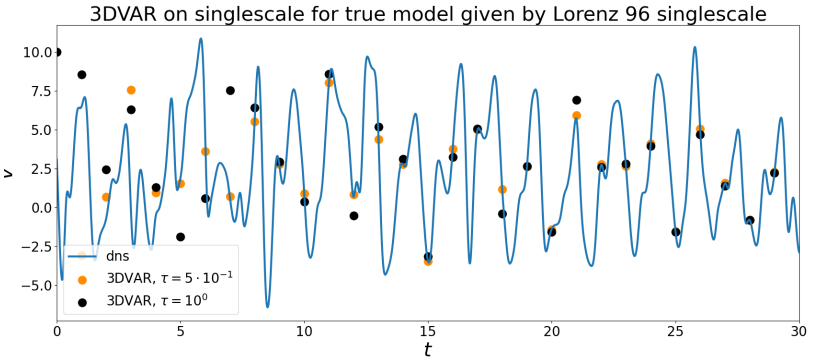
Figure 2.2. In both (a) and (b) the estimates of $v := v_3$ in time produced by 3DVAR using observation time interval $\tau = 10^{-3}$ are displayed and compared with dns. In (a) σ and γ are set to 10^{-3}, while in (b) they are set to 10^{-1}. The acronym 'dns' refers to direct numerical simulation of a true trajectory of the chaotic dynamical system. In both cases, it is noteworthy that 3DVAR is able to synchronize with the dns even though it is initialized far from the true initial condition. This is an example of data assimilation overcoming sensitive dependence in a chaotic system.

Because of this interpretation it is natural to study the 3DVAR algorithm, for this example, by displaying the output of 3DVAR on one of the unobserved components, and comparing with the truth; the key question is whether the observed components induce synchronization of 3DVAR with the truth in the unobserved components. Thus, in the following numerical experiments, we display component $v := v_3$.

Figure 2.2 illustrates the foregoing intuition about the behaviour of 3DVAR in experiments conducted with the choice $\tau = 10^{-3}$. For small noise with standard deviations of size 10^{-3}, we observe in Figure 2.2(a) the phenomenon of near-perfect synchronization of the 3DVAR output with the truth. Although not shown



(a)



(b)

Figure 2.3. In both (a) and (b) the noise standard deviations σ and γ are set to 10^{-3}. In (a) we display the estimates of $v := v_3$ in time produced by 3DVAR using observation time interval $\tau = 5 \times 10^{-1}$, compared with dns; (b) displays the estimates obtained at unit time using assimilation at $\tau = 5 \times 10^{-1}$ and $\tau = 10^0$. Again the acronym 'dns' refers to direct numerical simulation of the true chaotic dynamics. 3DVAR successfully synchronizes with the dns at the smaller value of τ but fails to do as well when the observation time interval τ is larger.

here, this synchronization occurs in all components of the solution, observed and unobserved, and thus the entire state of 3DVAR synchronizes with the truth, up to a small error on the scale of the noise. The algorithm thus produces an accurate estimate of the true state. In Figure 2.2(b) larger state and observational noise, of standard deviation 10^{-1}, is considered. In this scenario 3DVAR still captures the correct trend of the true dynamics, but there are clear overshoots and undershoots in the estimates; this occurs because the noise is larger than in Figure 2.2(a) and because noise is not accounted for in the 3DVAR algorithm.

It is intuitive that the synchronization phenomena studied above will depend not only on the size of the noise, but also on the observation time intervals τ. Figure 2.3 illustrates the effect of varying τ, in the setting where the noise standard deviation is 10^{-3}. The simulations show that for the larger value of τ, 3DVAR does not estimate the true state as accurately as for the smaller value. □

Remark 2.4. A common source of error in application of data assimilation algorithms in practice arises from the fact that the data is not produced by the mathematical model used for assimilation. This is known as *model misspecification*. For the majority of this paper we will make the *perfect model* assumption, avoiding the model misspecification issue. However, we do illustrate model misspecification in Example B.1. In that example we assimilate data produced from the Lorenz '96 multiscale model, but we use the Lorenz '96 single-scale model (2.9) as the basis for 3DVAR; thus the model used for assimilation differs from (but is close to) the model generating the data. The relationship between the multiscale and single-scale models is detailed in Appendix B, where Example B.1 may also be found. □

The control-theoretic approach of 3DVAR addresses Objective 1. However, when $|C_0|, |\Gamma|$ and $|\Sigma|$ are no longer necessarily small, so that the state and observations are subject to noise, it is natural to try and generalize the approach to address Objective 2, taking account of non-zero covariances; Figure 2.2(b) from Example 2.3 shows that this may indeed be needed when the noise is larger. A natural stochastic generalization of the observer approach is as follows: from current state estimate v_n, given $Y_n^\dagger$, predict the outcome of the model and data from the update equations (2.1), which we denote by $(\widehat{v}_{n+1}, \widehat{y}_{n+1})$; then correct the state estimate by nudging the mean of the prediction using the mismatch between observed and predicted data $(y_{n+1}^\dagger, \widehat{y}_{n+1})$. Given v_n computed from $Y_n^\dagger$, this results in state estimate v_{n+1} from $Y_{n+1}^\dagger$ defined through the following stochastic dynamical system, assumed to hold for all $n \in \mathbb{Z}^+$:

$$\boxed{\begin{aligned} \widehat{v}_{n+1} &= \Psi(v_n) + \xi_n, && \text{(2.15a)} \\ \widehat{y}_{n+1} &= h(\widehat{v}_{n+1}) + \eta_{n+1}, && \text{(2.15b)} \\ v_{n+1} &= \widehat{v}_{n+1} + K_n\left(y_{n+1}^\dagger - \widehat{y}_{n+1}\right). && \text{(2.15c)} \end{aligned}}$$

Here v_0, ξ_n and η_{n+1} are random variables given by the known distributions in (2.2). Note that the innovation $\mathfrak{I}_n$ is now modified from (2.5) to read

$$\mathfrak{I}_n = y_{n+1}^\dagger - \widehat{y}_{n+1}. \tag{2.16}$$

The data $\{y_{n+1}^\dagger\}_{n\in\mathbb{Z}^+}$ is a fixed realization of (2.1). Given this data, equations (2.15) then define a random map $v_n \mapsto v_{n+1}$ which uses knowledge of the model and the observed data to update our state estimate. The predicted state and data $(\widehat{v}_{n+1}, \widehat{y}_{n+1})$ are also referred to as the *simulated state and data*. Choice of the *gain matrix* K_n will complete definition of an algorithm. The key question we proceed

to study in subsequent sections concerns the choice of $\{K_n\}_{n\in\mathbb{Z}^+}$ when addressing Objective 2. Before doing so we build on Example 2.3, studying the effect of noise on the 3DVAR algorithm, which is aimed at addressing Objective 1, and for which $K_n = K$ is constant. The resulting Example 2.5 demonstrates the need for an adaptive choice of K_n when noise is significant; it serves to motivate algorithms which address Objective 2.

Example 2.5. We return to the setting of Example 2.3 and consider the effect of including noise in 3DVAR as in (2.15). We again study the Lorenz '96 single-scale model (2.9) for unknown $v \in C(\mathbb{R}^+, \mathbb{R}^L)$, with $L = 9, h_v = -0.8$ and $F = 10$. We let Ψ denote the solution operator for (2.9) over the observation time interval τ and consider observations $\{y_n^\dagger\}_{n\in\mathbb{Z}^+}$ defined by

$$\begin{aligned} v_{n+1}^\dagger &= \Psi(v_n^\dagger) + \xi_n^\dagger, \\ y_{n+1}^\dagger &= h(v_{n+1}^\dagger) + \eta_{n+1}^\dagger, \end{aligned}$$

where $\{\xi_n^\dagger\}_{n\in\mathbb{Z}^+}$, $\{\eta_n^\dagger\}_{n\in\mathbb{N}}$ are mutually independent Gaussian sequences

$$\xi_n^\dagger \sim \mathsf{N}(0, \sigma^2 I) \text{ i.i.d.}, \quad \eta_n^\dagger \sim \mathsf{N}(0, \gamma^2 I) \text{ i.i.d.},$$

with $\sigma = 0.1$ and $\gamma = 0.1$. We again assume that the observation function is linear: $h(v) = Hv$ for matrix $H\colon \mathbb{R}^9 \to \mathbb{R}^6$ defined by (2.10). As in Example 2.3 we choose fixed gain $K_n \equiv K$ with $K\colon \mathbb{R}^6 \to \mathbb{R}^9$ defined by (2.12).

Figure 2.4 illustrates that this version of noisy 3DVAR produces trajectories that resemble the true signal better than noise-free 3DVAR (2.11). However, the noisy 3DVAR does not perform better than the noise-free 3DVAR, in this setting where true state and true observation noise levels are high, in a quantitative sense. To demonstrate this, we compute the mean squared error between the estimates arising from the 3DVAR algorithm (2.11) and the true states, and the mean squared error between the estimates obtained using noisy 3DVAR algorithm (2.15) and the true states. Given either method the error e is computed using the following formula, in which, recall, $\{v_n^\dagger\}$ is the truth and $\{v_n\}$ is the output of the 3DVAR or noisy 3DVAR algorithm:

$$e = \frac{1}{N \cdot d_v} \sum_{n=1}^{N} \left| v_{n^*+n}^\dagger - v_{n^*+n} \right|^2. \tag{2.17}$$

Time $t^* = n^*\tau$ is chosen to remove error from the incorrect initialization, and focus on quantifying error in the statistical steady state, after synchronization. Here N is such that $(n^* + N)\tau = T$ and $T - t^*$ is chosen large enough to allow time-averaging over a long enough window to capture the statistical steady state. In this case, we compute this average for the estimates obtained after time $t^* = 3$, up to $T = 20$, and find errors $e_{\text{3DVAR}} = 1.65 \times 10^0$ and $e_{\text{noisy3DVAR}} = 3.30 \times 10^0$. Clearly use of noisy 3DVAR does not improve the error, in comparison with noise-free 3DVAR. ∎

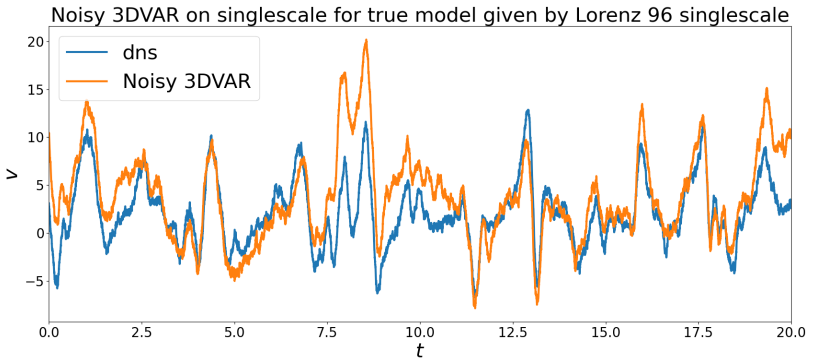


Figure 2.4. In this experiment we set the noise levels $\sigma = 10^{-1}, \gamma = 10^{-1}$. Again the acronym 'dns' refers to direct numerical simulation. We display the estimates of $v := v_3$ in time produced by noisy 3DVAR against the true dynamics using observation time interval $\tau = 10^{-3}$. This should be compared with Figure 2.2(b), in which the noise-free 3DVAR is deployed to solve the same problem. Notice that adding noise to 3DVAR has not improved the recovery of the true trajectory. However, qualitatively the output of 3DVAR now resembles the true signal more closely.

Example 2.5 highlights the need to quantify uncertainty and pass to a probabilistic interpretation (Objective 2) of the filtering problem; and, in particular, to make an informed choice of adaptive gain matrices K_n. We turn to the probabilistic interpretation in the next subsection; in later subsections we derive algorithms, leading in particular to a specific choice of adaptive gain matrices.

2.3. *Probabilistic perspective*

We have shown that the 3DVAR methodology can recover the state of a (possibly chaotic) dynamical system, even though the initial condition is not known, by exploiting the observations. However, 3DVAR does not quantify uncertainty in the state estimate; it is derived from a purely control-theoretic perspective. We now introduce a probabilistic perspective which enables us to address the issue of uncertainty quantification. Section 2.3.1 discusses the unconditioned dynamics from the perspective of evolution of probability densities. In Section 2.3.2 we define the filtering distribution and describe this from the perspective of evolution of probability densities. Section 2.3.3 introduces the *sample-path perspective* on algorithms for filtering, a central idea in this paper. In Section 2.3.4 we establish some notation, used henceforth, that is important for the reader to internalize.

2.3.1. *Unconditioned dynamics*

To open our development of the probabilistic perspective we first consider the unconditioned dynamics on state $\{v_n\}_{n\in\mathbb{Z}^+}$ defined by (2.1a). We refer to $\{v_n\}_{n\in\mathbb{Z}^+}$

defined by (2.1a) as a *Markov process*. We let r_n denote the probability density of random variable v_n and derive an evolution equation for r_n. Irrespective of whether the evolution of the state $\{v_n\}_{n\in\mathbb{Z}^+}$ defined by (2.1a) is linear, the evolution of $\{r_n\}_{n\in\mathbb{Z}^+}$ is linear. Furthermore, the evolution of the state and its probability density are uncoupled from one another. The evolution of the probability density implied by (2.1a) is given by

$$r_{n+1} = \mathcal{P} r_n, \tag{2.18a}$$

$$(\mathcal{P}r)(\mathrm{d}v) = \left(\int_{u\in\mathbb{R}^{d_v}} p(u,v) r(\mathrm{d}u) \right) \mathrm{d}v, \tag{2.18b}$$

$$p(u,v) = \frac{1}{(2\pi)^{d_v/2}\sqrt{\det(\Sigma)}} \exp\left(-\frac{1}{2} |v - \Psi(u)|_\Sigma^2 \right). \tag{2.18c}$$

Thus, in particular, r_n evolves in time n through application of a linear integral operator. The situation when we condition the state on observations is different, leading to nonlinear evolution of densities.

2.3.2. *The filtering distribution*

Here we introduce the *filtering distribution* with density μ_n: the distribution of the conditioned random variable $v_n|Y_n^\dagger$. This captures the knowledge of the state of the system, and uncertainties in the state, given the observations. To understand how uncertainty in estimates of the state evolves, it is thus important to understand how μ_n evolves with n. Unlike the unconditioned dynamics, this conditioned dynamics has a *nonlinear* structure. Nonlinear evolution equations arise in filtering through the incorporation of data. This nonlinearity renders filtering a challenging mathematical and computational problem. To define this evolution, we first define the linear operator

$$(\mathcal{Q}\mu)(\mathrm{d}v, \mathrm{d}y) = q(v,y)\mu(\mathrm{d}v)\,\mathrm{d}y, \tag{2.19a}$$

$$q(v,y) = \frac{1}{(2\pi)^{d_y/2}\sqrt{\det(\Gamma)}} \exp\left(-\frac{1}{2} |y - h(v)|_\Gamma^2 \right). \tag{2.19b}$$

Then we define the n-dependent family of two nonlinear operators

$$\mathcal{B}_n(\pi)(\mathrm{d}v) = \int_{y\in\mathbb{R}^{d_y}} \delta_{y_{n+1}^\dagger}(y)\pi(\mathrm{d}v,\mathrm{d}y) \Big/ \left(\int_{(v,y)\in\mathbb{R}^{d_v}\times\mathbb{R}^{d_y}} \delta_{y_{n+1}^\dagger}(y)\pi(\mathrm{d}v,\mathrm{d}y) \right), \tag{2.20}$$

$$\mathcal{L}_n(\mu)(\mathrm{d}v) = q\left(v, y_{n+1}^\dagger\right)\mu(\mathrm{d}v) \Big/ \left(\int_{v\in\mathbb{R}^{d_v}} q\left(v, y_{n+1}^\dagger\right)\mu(\mathrm{d}v) \right). \tag{2.21}$$

The nonlinear map $\mu_n \mapsto \mu_{n+1}$ is most easily described by first introducing $\widehat{\mu}_{n+1}$, the distribution of $v_{n+1}|Y_n^\dagger$, and π_{n+1}, the distribution of $(v_{n+1}, y_{n+1})|Y_n^\dagger$. The map from μ_n to $\widehat{\mu}_{n+1}$ is determined by equation (2.1a) and is *linear*, as a map from

the space of probability measures defined on $\mathbb{R}^{d_v}$ into itself, as discussed in the preceding subsection; the map from $\widehat{\mu}_{n+1}$ to π_{n+1} is defined by (2.1b) and is also *linear*, now as a map from the space of probability measures defined on $\mathbb{R}^{d_v}$ into the space of probability measures defined on $\mathbb{R}^{d_v+d_y}$; the map from π_{n+1} to μ_{n+1} is defined by conditioning π_{n+1} on $y^\dagger_{n+1}$ and is *nonlinear*, as a map from the space of probability measures defined on $\mathbb{R}^{d_v+d_y}$ into the space of probability measures defined on $\mathbb{R}^{d_v}$. Using the preceding definitions of linear and nonlinear operators, we have

$$\boxed{\begin{aligned} \widehat{\mu}_{n+1} &= \mathcal{P}\mu_n, && (2.22a)\\ \pi_{n+1} &= \mathcal{Q}\widehat{\mu}_{n+1}, && (2.22b)\\ \mu_{n+1} &= \mathcal{B}_n(\pi_{n+1}). && (2.22c) \end{aligned}}$$

Concatenating, we find that

$$\mu_{n+1} = \mathcal{B}_n(\mathcal{Q}\mathcal{P}\mu_n), \quad \mu_0 = \mathsf{N}(m_0, C_0). \tag{2.23}$$

This map defines an inhomogeneous nonlinear map on the space of probability measures on $\mathbb{R}^{d_v}$. The map $\mathcal{B}_n(\mathcal{Q}\mathcal{P}\cdot)$ may be decomposed into two maps: (i) the *prediction* $\mathcal{P}$, which represents application of the dynamical model (2.1a); and (ii) application of *Bayes' theorem*[3] through operator $\mathcal{L}_n\cdot := \mathcal{B}_n(\mathcal{Q}\cdot)$, which corresponds to use of likelihood defined by the observation model (2.1b). With this notation we thus obtain

$$\boxed{\begin{aligned} \widehat{\mu}_{n+1} &= \mathcal{P}\mu_n, && (2.24a)\\ \mu_{n+1} &= \mathcal{L}_n(\widehat{\mu}_{n+1}). && (2.24b) \end{aligned}}$$

We refer to iteration of (2.24) as the *filtering cycle*. The cycle involves iterative interleaving of prediction, using the dynamical model, a linear operation on measures, and Bayes' theorem, using the observation model, a nonlinear operation on measures.

It is important to appreciate that there is, in general, no closed-form expression for μ_n defined by the iteration (2.24); thus (2.24) does not constitute an algorithm. However, if Ψ and h are linear then, since $\mu_0 = \mathsf{N}(m_0, C_0)$ is Gaussian, it follows that $\widehat{\mu}_{n+1}, \pi_{n+1}, \mu_{n+1}$ are all Gaussian for all $n \in \mathbb{Z}^+$, and closed-form expressions, based on dynamical updates of means and covariances, are available. This linear Gaussian setting is discussed in the following example, and linear Gaussian examples will be used throughout the paper. However, the main thrust of the paper concerns nonlinear and non-Gaussian problems; for these problems further ideas, which we will explain in subsections below, are required to make actionable algorithms from the iteration (2.24).

[3] Often referred to as the *analysis* step in the geophysical data assimilation community. Bayes' theorem is discussed in more detail in Section 4.

Example 2.6. To define the *Kalman filter* we consider the setting (2.6), where $\Psi(\cdot)$ and $h(\cdot)$ are both linear. Then (2.1) becomes

$$v_{n+1} = Mv_n + \xi_n, \tag{2.25a}$$

$$y_{n+1} = Hv_{n+1} + \eta_{n+1}. \tag{2.25b}$$

For this problem the mapping (2.24) may be solved explicitly. In fact μ_n and $\widehat{\mu}_n$ are both Gaussian, and we write their mean–covariance pairs as (m_n, C_n) and $(\widehat{m}_n, \widehat{C}_n)$ respectively. Then $\widehat{\mu}_{n+1}$ is determined from μ_n by the formulae

$$\widehat{m}_{n+1} = Mm_n, \tag{2.26a}$$

$$\widehat{C}_{n+1} = MC_nM^\top + \Sigma, \tag{2.26b}$$

the prediction step. Measure μ_{n+1} is determined from $\widehat{\mu}_n$ by the application of a Bayesian update, solving the inverse problem defined by (2.25b); this inverse problem is for v_{n+1} given fixed realization of data $y_{n+1} = y^\dagger_{n+1}$ generated by (2.25). Since the prior and posterior are Gaussian, we may complete the square to solve the Bayesian inverse problem to give the following update formulae for the mean and precision (inverse covariance):

$$C^{-1}_{n+1}m_{n+1} = \widehat{C}^{-1}_{n+1}\widehat{m}_n + H^\top\Gamma^{-1}y^\dagger_{n+1}, \tag{2.27a}$$

$$C^{-1}_{n+1} = \widehat{C}^{-1}_{n+1} + H^\top\Gamma^{-1}H. \tag{2.27b}$$

By use of the Woodbury matrix identity, we obtain the following formulae, expressed in terms of covariances rather than precisions:

$$m_{n+1} = \widehat{m}_{n+1} + \widehat{C}_{n+1}H^\top(H\widehat{C}_{n+1}H^\top + \Gamma)^{-1}\left(y^\dagger_{n+1} - H\widehat{m}_{n+1}\right), \tag{2.28a}$$

$$C_{n+1} = \widehat{C}_{n+1} - \widehat{C}_{n+1}H^\top(H\widehat{C}_{n+1}H^\top + \Gamma)^{-1}H\widehat{C}_{n+1}. \tag{2.28b}$$

The update equations (2.26) simply represent propagation of a Gaussian under the linear dynamics defined by (2.25a); (2.28) corresponds to application of Bayes' theorem, and in this particular linear setting to conditioning a Gaussian, using the observations defined by (2.25b) with $y_{n+1} = y^\dagger_{n+1}$. The Kalman filter update equations (2.26) and (2.28) are well-defined if $\Gamma \succ 0$. □

The Gaussian setting of the preceding example is very special. But the idea of making a *Gaussian approximation* will play a central role in ensemble Kalman methods, and as a consequence the explicit calculations in the example will be generally useful. The perspective of invoking Gaussian approximations is introduced in Section 2.4; it is subsequently developed to form the backbone of the methodology highlighted in this paper.

2.3.3. The sample-path perspective

In Section 2.3.1 we demonstrated that the (typically nonlinear) state space evolution of $\{v_n\}_{n\in\mathbb{Z}^+}$ defined by (2.1a) may be alternatively viewed in terms of the linear evolution of probability density functions r_n defined by (2.18). In Section 2.3.2 we

showed that, when conditioned on observations, the evolution of the probability density function becomes nonlinear and is defined by (2.22) or (2.24). In this section we address the question of finding an evolution in state space that is consistent with this nonlinear evolution of densities defined by (2.22) or (2.24). Our aim is to generalize the control-theoretic data assimilation algorithm given by (2.15) in order to find such an evolution.

Throughout this subsection let v_n be a random variable distributed according to $Law(v_n)$, let $\widehat{v}_{n+1}$ be a random variable with $Law(\widehat{v}_{n+1}) = \mathcal{P}\, Law(v_n)$ and let $(\widehat{v}_{n+1}, \widehat{y}_{n+1})$ be a random variable with law $\mathcal{Q}\, Law(\widehat{v}_{n+1})$. Using *Law* avoids a proliferation of notation for the different measures arising from the use of various different algorithms to approximate the filtering cycle.

All of the algorithms that we will introduce in what follows are based on a prediction of state, and possibly observation, from $Law(v_n)$. To this end, recall (2.15), and for all $n \in \mathbb{Z}^+$,

$$\widehat{v}_{n+1} = \Psi(v_n) + \xi_n, \tag{2.29a}$$

$$\widehat{y}_{n+1} = h(\widehat{v}_{n+1}) + \eta_{n+1}. \tag{2.29b}$$

The distributions of $\widehat{v}_{n+1}$ and $(\widehat{v}_{n+1}, \widehat{y}_{n+1})$ given by these equations define $\mathcal{P}\, Law(v_n)$ and $\mathcal{Q}\, Law(\widehat{v}_{n+1})$, respectively. A common theme in this paper is to augment these equations for the predicted state and data with a final map, to generalize (2.15c), of the form

$$(\widehat{v}_{n+1}, \widehat{y}_{n+1}) \mapsto v_{n+1} \tag{2.30}$$

or

$$\widehat{v}_{n+1} \mapsto v_{n+1}. \tag{2.31}$$

These maps are chosen, respectively, to mimic (2.22) or (2.24). To be precise, if $v_n \sim \mu_n$, the maps are designed so that $v_{n+1} \sim \mu_{n+1}$, where μ_n and μ_{n+1} are related by the filtering update (2.22) or, equivalently, (2.24). A key observation is that these maps will necessarily depend on the distributions of $\widehat{v}_{n+1}$ and $\widehat{y}_{n+1}$ rendering the associated Markov processes of mean-field type. Thus, in contrast to the unconditioned dynamics, the state space evolution does not decouple from the evolution of the associated probability density function; rather, it depends on it. The resulting state space evolution is said to define a *nonlinear Markov process.*

Note that once (2.29) is augmented with either (2.30) or (2.31), we have a state space evolution that describes the filtering process via the probability distribution of v_n; the state space evolution can then be used as the basis for algorithms. Thus the key question for such a program is the identification of either (2.30) or (2.31). The existence of *transport maps* or, more generally, *couplings*, which define the steps (2.30) or (2.31), follows under quite general conditions. But finding them explicitly is generally difficult. Furthermore, the maps (2.30) and (2.31) are not uniquely defined, in general. As a consequence of the difficulty in identifying mean-field transport maps, the algorithms we study will be based on identifying

maps (2.30) or (2.31), which only *approximately* achieve the filtering update (2.22). However, all formulations considered in this survey will be of mean-field type.

In summary, we refer to equations (2.29), (2.30) or (2.29a), (2.31) as providing a *sample-path perspective*. The resulting algorithms update state $v_n \mapsto v_{n+1}$ and are designed to exactly (in theory), and approximately (in practice), reproduce the *probabilistic perspective* encapsulated in the three steps of (2.22), or the two steps of (2.24). These algorithms result in a sample path $\{v_\ell\}_{\ell=0}^n$ with the property that (possibly only approximately) $Law(v_\ell) = \mu_\ell$. This idea is central to the algorithmic developments in the paper.

2.3.4. *Important frequently used notation*

The approximations we develop will be based on matching first- and second-order moments. In the service of designing these approximations, it is useful to define various first- and second-order statistics computed under the law of $(\widehat{v}_{n+1}, \widehat{y}_{n+1})$. First define the mean and covariance of $\widehat{v}_{n+1}$:

$$\boxed{\begin{aligned} \widehat{m}_{n+1} &= \mathbb{E}\widehat{v}_{n+1}, && (2.32a)\\ \widehat{C}_{n+1} &= \mathbb{E}((\widehat{v}_{n+1} - \widehat{m}_{n+1}) \otimes (\widehat{v}_{n+1} - \widehat{m}_{n+1})). && (2.32b) \end{aligned}}$$

Then define the mean of the predicted data, cross-covariance from predicted data to state and covariance of the data:

$$\boxed{\begin{aligned} \widehat{o}_{n+1} &= \mathbb{E}\widehat{y}_{n+1}, && (2.33a)\\ \widehat{C}^{vy}_{n+1} &= \mathbb{E}((\widehat{v}_{n+1} - \widehat{m}_{n+1}) \otimes (\widehat{y}_{n+1} - \widehat{o}_{n+1})), && (2.33b)\\ \widehat{C}^{yy}_{n+1} &= \mathbb{E}((\widehat{y}_{n+1} - \widehat{o}_{n+1}) \otimes (\widehat{y}_{n+1} - \widehat{o}_{n+1})). && (2.33c) \end{aligned}}$$

From these covariances we define the matrix

$$K_n = \widehat{C}^{vy}_{n+1} \left(\widehat{C}^{yy}_{n+1}\right)^{-1}. \tag{2.34}$$

This particular choice of K_n, known as the *Kalman gain*, plays a central role in the mean-field maps which underpin ensemble Kalman methods through their particle approximations. Note that if $\widehat{C}^{yy}_{n+1}$ is not invertible, then its action may still be defined through a pseudo-inverse.

It is sometimes useful to express the Kalman gain K_n in terms of the variable $\widehat{h}_{n+1} = h(\widehat{v}_{n+1})$ and without reference to predicted data $\widehat{y}_{n+1}$. For this purpose we define the following correlation matrices, computed under the law of $\widehat{v}_{n+1}$:

$$\boxed{\begin{aligned} \widehat{o}_{n+1} &= \mathbb{E}h(\widehat{v}_{n+1}), && (2.35a)\\ \widehat{C}^{vh}_{n+1} &= \mathbb{E}((\widehat{v}_{n+1} - \widehat{m}_{n+1}) \otimes (\widehat{h}_{n+1} - \widehat{o}_{n+1})), && (2.35b)\\ \widehat{C}^{hh}_{n+1} &= \mathbb{E}((\widehat{h}_{n+1} - \widehat{o}_{n+1}) \otimes (\widehat{h}_{n+1} - \widehat{o}_{n+1})), && (2.35c) \end{aligned}}$$

Then, in place of (2.33) and (2.34), we have

$$\widehat{C}^{vy}_{n+1} = \widehat{C}^{vh}_{n+1}, \quad \widehat{C}^{yy}_{n+1} = \widehat{C}^{hh}_{n+1} + \Gamma, \tag{2.36a}$$

$$K_n = \widehat{C}^{vh}_{n+1}\left(\widehat{C}^{hh}_{n+1} + \Gamma\right)^{-1}. \tag{2.36b}$$

Note that if $\Gamma \succ 0$, K_n is well-defined without recourse to the use of the pseudo-inverse.

Example 2.7. In the setting of the linear and Gaussian Example 2.6, we have

$$\widehat{C}^{vy}_{n+1} = \widehat{C}_{n+1}H^\top,$$
$$\widehat{C}^{yy}_{n+1} = H\widehat{C}_{n+1}H^\top + \Gamma,$$

and the mean update (2.28a) may be written as

$$m_{n+1} = \widehat{m}_{n+1} + K_n\left(y^\dagger_{n+1} - H\widehat{m}_{n+1}\right).$$

In particular, only the single covariance $\widehat{C}_{n+1}$ needs to be computed. Note the similarity of the resulting algorithm to the 3DVAR algorithm (2.8), in the linear Gaussian setting, since in this linear case $\widehat{m}_{n+1} = Mm_n$. It differs only through having a time-varying gain matrix K_n. □

This linear Gaussian setting provides some motivation for the Kalman gain. The origin of this key concept in the more general nonlinear and non-Gaussian setting will be described in the next subsection.

2.4. Gaussian projected filtering distribution

The *Gaussian projected filter* gives a Gaussian approximation of the true filtering distribution. It is defined by the following three steps:

(i) taking input Gaussian at time n as $Law(v_n)$ and pushing this measure forward under (2.29) to find (typically non-Gaussian) measure $Law(\widehat{v}_{n+1}, \widehat{y}_{n+1})$;

(ii) projecting this joint measure onto the nearest Gaussian (in a sense that we will make precise);

(iii) conditioning this Gaussian on the data $y^\dagger_{n+1}$ to find output Gaussian at time $n+1$.

We note that conditioning a Gaussian random variable on linear functionals of the random variable returns another Gaussian. Thus the algorithm maps Gaussians to Gaussians. The resulting approximation of the filtering distribution plays an important role in motivating the mean-field maps, introduced in Section 2.5, that underlie ensemble Kalman methods. It is also of interest as a method in its own right.

In what follows in this subsection we introduce the Gaussian projected approximation to the evolution (2.23). In the case where $\Psi(\cdot)$ and $h(\cdot)$ are linear, the

resulting formulae deliver exact solutions of the filtering cycle (2.23), leading to the Kalman filter as given in Example 2.6. The Gaussian projected filter also leads to a derivation of the Kalman gain (2.34), beyond the linear Gaussian setting; an alternative derivation, using the minimum variance approach, may be found in Appendix C.3.

To describe the Gaussian projected approximation to the filtering distribution, we define the map $\mathcal{G}$, definition of which uses $\mathfrak{P} = \mathfrak{P}(\mathbb{R}^d)$ and $\mathfrak{G} = \mathfrak{G}(\mathbb{R}^d)$ defined in Section 1.4.

Definition 2.8. Define $\mathcal{G}\colon \mathfrak{P} \mapsto \mathfrak{G}$ by

$$\mathcal{G}\mu = \mathsf{N}(m^\mu, C^\mu), \quad m^\mu = \mathbb{E}^\mu u, \quad C^\mu = \mathbb{E}^\mu((u - \mathbb{E}^\mu u) \otimes (u - \mathbb{E}^\mu u)),$$

where $u \sim \mu$. □

Thus the map $\mathcal{G}$ applied to measure μ simply computes the Gaussian with mean and covariance calculated with respect to the typically non-Gaussian measure μ. Notice that $\mathcal{G}$ is the identity on Gaussians. Furthermore, $\mathcal{G} \circ \mathcal{G} = \mathcal{G}$. We refer to this as a *projection* onto Gaussians because it corresponds to finding the closest point to given measure μ, with respect to a Kullback–Leibler divergence:[4]

$$\mathcal{G}\mu = \operatorname{argmin}_{\pi \in \mathfrak{G}} d_{\mathrm{KL}}(\mu||\pi). \tag{2.37}$$

We now use mapping $\mathcal{G}$ to find an approximation to the evolution (2.22) which generates measures remaining Gaussian; intuitively this will be a good approximation whilst the measures $\{\mu_n\}$ evolving under (2.23) remain close to Gaussian. To this end we consider random variable $v_n \sim \mu_n^G$, where the probability measure μ_n^G evolves according to

$$\mu_{n+1}^G = \mathcal{B}_n(\mathcal{G}\mathcal{Q}\mathcal{P}\mu_n^G), \quad \mu_0^G = \mathsf{N}(m_0, C_0). \tag{2.38}$$

This may be decomposed as follows:

$$\boxed{\begin{aligned} \widehat{\mu}_{n+1}^G &= \mathcal{P}\mu_n^G, && \text{(2.39a)} \\ \pi_{n+1}^G &= \mathcal{Q}\widehat{\mu}_{n+1}^G, && \text{(2.39b)} \\ \mu_{n+1}^G &= \mathcal{B}_n\left(\mathcal{G}\pi_{n+1}^G\right). && \text{(2.39c)} \end{aligned}}$$

Map (2.38) defines a nonlinear Markov process, similarly to (2.23). It also maps Gaussians into Gaussians. This fact follows from the fact that the nonlinear map $\mathcal{B}_n(\cdot)$, which represents conditioning, maps Gaussians into Gaussians. The map $\mu_n^G \mapsto \mu_{n+1}^G$ hence defines a deterministic mapping from the mean m_n and covariance C_n of μ_n^G into the mean m_{n+1} and covariance C_{n+1} of μ_{n+1}^G. We now identify this map explicitly.

For $v_n \sim \mu_n^G$ we introduce the random variables $\widehat{v}_{n+1}, \widehat{y}_{n+1}$ defined by (2.29). It then follows that $\widehat{v}_{n+1} \sim \widehat{\mu}_{n+1}^G = \mathcal{P}\mu_n^G$ and that $(\widehat{v}_{n+1}, \widehat{y}_{n+1}) \sim \pi_{n+1}^G = \mathcal{Q}\mathcal{P}\mu_n^G$. Note

[4] For details see Section 2.7.

also that $\widehat{\mu}^G_{n+1}$ and π^G_{n+1} are not Gaussian, but are defined by the Gaussian μ^G_n and hence completely determined by m_n and C_n. Thus the mean and covariance under $\widehat{\mu}^G_{n+1}$ are $(\widehat{m}_{n+1}, \widehat{C}_{n+1})$ given by (2.32). Furthermore, $\mathcal{G}\pi^G_{n+1}$ is defined by

$$\mathcal{G}\pi^G_{n+1} = \mathsf{N}\left(\begin{bmatrix}\widehat{m}_{n+1}\\ \widehat{o}_{n+1}\end{bmatrix}, \begin{bmatrix}\widehat{C}_{n+1} & \widehat{C}^{vy}_{n+1}\\ \left(\widehat{C}^{vy}_{n+1}\right)^\top & \widehat{C}^{yy}_{n+1}\end{bmatrix}\right), \tag{2.40}$$

where all relevant quantities are defined in Section 2.3.4.

We now condition the Gaussian $\mathcal{G}\pi^G_{n+1}$, on the second component of the vector taking value $y^\dagger_{n+1}$. From this we find the Gaussian measure $\mu^G_{n+1} = \mathcal{B}_n\left(\mathcal{G}\pi^G_{n+1}\right)$ characterized by mean m_{n+1} and covariance C_{n+1} given by the following lemma.

Lemma 2.9. Assume that $\Gamma \succ 0$. Let m_n and C_n denote the mean and covariance under the Gaussian projected filter. Consider equations (2.29) initialized at $v_n \sim \mathsf{N}(m_n, C_n)$, and then $(\widehat{m}_{n+1}, \widehat{C}_{n+1})$ defined by (2.32); furthermore, define the mean of the observed data and covariances given by (2.33). Then $\widehat{C}^{yy}_{n+1} \succ 0$ for all $n \in \mathbb{Z}^+$ and

$$\boxed{\begin{aligned} m_{n+1} &= \widehat{m}_{n+1} + \widehat{C}^{vy}_{n+1}\left(\widehat{C}^{yy}_{n+1}\right)^{-1}\left(y^\dagger_{n+1} - \widehat{o}_{n+1}\right), &(2.41a)\\ C_{n+1} &= \widehat{C}_{n+1} - \widehat{C}^{vy}_{n+1}\left(\widehat{C}^{yy}_{n+1}\right)^{-1}\left(\widehat{C}^{vy}_{n+1}\right)^\top, &(2.41b)\end{aligned}}$$

where $\{y^\dagger_n\}$ arises from a fixed realization of (2.1). ◇

Proof. We first note that $\widehat{C}^{yy}_{n+1} \succ 0$. Indeed, since by assumption $\Gamma \succ 0$ and by definition $\widehat{C}^{hh}_{n+1} \succeq 0$, then $\widehat{C}^{hh}_{n+1} + \Gamma \succ 0$. Hence by (2.36) we have that $\widehat{C}^{yy}_{n+1} \succ 0$. Now consider the distribution of the Gaussian $\mathcal{G}\pi^G_{n+1}$ given by (2.40). Conditioning the resulting joint random variable on $(v, y) \in \mathbb{R}^{d_v} \times \mathbb{R}^{d_y}$ on $y = y^\dagger_{n+1}$, it is possible to conclude from standard formulae for conditioned Gaussians that m_{n+1} and C_{n+1} are given by the expressions in (2.41). □

Equations (2.29), (2.32), (2.33) and (2.41) define a mapping from μ^G_n, characterized by (m_n, C_n), into μ^G_{n+1}, characterized by (m_{n+1}, C_{n+1}). They comprise an explicit set of formulae for the mapping (2.38): since Gaussians are determined by mean and covariance, the map on measures is completely determined by the map from (m_n, C_n) to (m_{n+1}, C_{n+1}).

We note that (2.41) can also be rewritten as

$$\boxed{\begin{aligned} m_{n+1} &= \widehat{m}_{n+1} + \widehat{C}^{vh}_{n+1}\left(\widehat{C}^{hh}_{n+1} + \Gamma\right)^{-1}\left(y^\dagger_{n+1} - \widehat{o}_{n+1}\right), &(2.42a)\\ C_{n+1} &= \widehat{C}_{n+1} - \widehat{C}^{vh}_{n+1}\left(\widehat{C}^{hh}_{n+1} + \Gamma\right)^{-1}\left(\widehat{C}^{vh}_{n+1}\right)^\top. &(2.42b)\end{aligned}}$$

Equations (2.29), (2.32), (2.35) and (2.42) then also define the mapping from μ^G_n into μ^G_{n+1} and also comprise an explicit set of formulae for the updates of the mean and covariance which characterize mapping (2.38).

Finally we note that, using the definition (2.34) of Kalman gain, we can rewrite (2.41) as

$$m_{n+1} = \widehat{m}_{n+1} + K_n\big(y^\dagger_{n+1} - \widehat{o}_{n+1}\big),$$
$$C_{n+1} = \widehat{C}_{n+1} - K_n\big(\widehat{C}^{vy}_{n+1}\big)^\top.$$

Thus we see how the Kalman gain arises naturally within the context of this Gaussian projected filter.

Remark 2.10. The preceding explicit formulae for (2.38) involve the computation of expectations under the non-Gaussian measure $Law(\widehat{v}_{n+1}, \widehat{y}_{n+1})$. For this reason they do not constitute an algorithm. A possible approach to algorithmic implementations involves quadrature to approximate the expectations, leading, for example, to the *unscented Kalman filter* approach; details may be found in Section 2.7. However, the formulation of explicit maps on Gaussians plays another, important, role in this paper: we use it as a way of explaining the sense in which the distribution of our mean-field models approximate evolution of measures under the true filtering distribution; see Section 2.5.6. The explicit map on means and covariances can also be used to derive the Kalman filter, which applies in the linear Gaussian setting and is presented above in Example 2.6. □

2.5. *Mean-field maps*

In the previous section we did not adopt the sample-path perspective, but rather chose to represent the evolution of the filtering distribution, approximately, as the evolution of Gaussians. In this subsection we introduce our first explicit instance of the sample-path perspective, finding an evolution in state space which approximately captures the evolution of the filtering distribution. Like the Gaussian projected filter it uses a Gaussian ansatz, but in a different way, leading to a state space evolution that is not Gaussian.

Note that elements of $\mathfrak{P}(\mathbb{R}^d)$ are *infinite-dimensional* objects, for any d. Thus the filtering distribution defines a nonlinear evolution in an infinite-dimensional space. This fact goes to the heart of the computational challenges faced when solving the filtering problem. These computational challenges are further exacerbated when $d \gg 1$. The manifold of Gaussians $\mathfrak{G}(\mathbb{R}^d)$ is finite-dimensional, because it is parametrized by the mean and covariance and hence has dimension $\frac{1}{2}d(d+3)$. The preceding subsection provides explicit *finite-dimensional maps* for the mean and covariance which characterize the Gaussian projected filter $\mu^G_n \mapsto \mu^G_{n+1}$, an approximation which is (intuitively) accurate when the true filter is close to Gaussian. Nonetheless, if $d \gg 1$ this method can still be prohibitive because the algorithm acts on a space of dimension that grows quadratically in d.

To address the issue that the Gaussian projected filter may not be efficient if $d \gg 1$, in this section we introduce a more ambitious aim: to find maps on finite-dimensional spaces of dimension d with the property that (possibly only approximately) their output is equal in law to the map on measures $\mu_n \mapsto \mu_{n+1}$

given by the filtering cycle. We achieve this by studying transport maps that achieve (2.30) or (2.31); we then weaken this requirement and ask only for transport maps that approximately achieve (2.30) or (2.31) in a manner that we will make precise. When combined with (2.29), either (2.30) or (2.31) gives a sample-path evolution that can be used as the basis of algorithms to solve the filtering problem. This transport map viewpoint leads us to the subject of mean-field maps, namely random maps that depend on the law of the state being mapped. When approximated by particle methods, these maps lead to methods that scale linearly with d, in contrast to the Gaussian projected filter, which scales quadratically.

This section is organized as follows. We start in Section 2.5.1 with an introduction to transport maps. Section 2.5.2 describes two distinct transport approaches which effect exact filtering: one based on the conditioning component of the overall Bayesian inference step, a transport between probability measures on different spaces; and the other based on the prior-to-posterior map that constitutes the Bayesian inference step of filtering, a transport between probability measures on the same space. We refer to these maps which effect exact filtering as *perfect transport*.[5] Sections 2.5.4 and 2.5.5 are concerned with approximations of these two perfect transports, and are motivated in Section 2.5.3 with an explicit example. In these approximations the pushforward under the transport map is designed to match only the first- and second-order moments of the target measure. Hence these approximations are closely related to, but different from, the previously defined Gaussian projected filter; we elaborate on this connection in Section 2.5.6. That subsection also includes Example 2.15, in which we identify mean-field formulations of the Kalman filter; recall that this filter applies only to linear Gaussian systems, and is defined in Example 2.6.

2.5.1. *Transport maps*

We start by setting up notation used throughout. Consider probability measures π and π' on $\mathbb{R}^d$ and $\mathbb{R}^{d'}$ respectively, and recall from Section 1.4 the definition of *pushforward* of a measure under under $T\colon \mathbb{R}^d \to \mathbb{R}^{d'}$: the statement $\pi' = T^\sharp \pi$ is a succinct way of stating that if $Law(v) = \pi$ and $v' = T(v)$, then $Law(v') = \pi'$. A *transport* $T\colon \mathbb{R}^d \to \mathbb{R}^{d'}$ from π to π' is a map with the property that the pushforward of probability measure π under T, $T^\sharp \pi$, is equal to probability measure π'. In the following we refer to π as the *source measure* and π' as the *target measure* defining the transport. In our setting, π' will be uniquely determined by π and an observed piece of finite-dimensional data. Thus T depends on π, and so we may view T as a mapping $\mathbb{R}^d \times \mathfrak{P} \to \mathbb{R}^{d'}$, suppressing, for the moment, explicit dependence on the

[5] Perfect transport should not be confused with *optimal transport*, which identifies among all (perfect) transport maps the one minimizing a certain cost functional such as that leading to the Wasserstein distance. See Section 2.7, and Theorem C.6, for more details. We use *perfect* here to distinguish from the *approximate* transport maps, based only on matching first and second moment; these approximate transport maps underpin ensemble Kalman methods.

observed data. We can compute the pushforward under T on any measure in $\mathfrak{P}$; but when we compute the pushforward on π we obtain π'. We emphasize that transport maps are not uniquely defined by their source and target measures, and require certain conditions for their existence, which we assume here to be satisfied. The underlying mathematical concept is that of *coupling of measures*. See Section 2.7 for discussion of transport, optimal transport and coupling.

We will also consider classes of *approximate transport maps* which do not achieve transport from π to π', but instead match first- and second-order moment information (we will be precise below). Such maps will also depend on π.

We now clarify an important notational issue. The dependence of (possibly approximate) transport T on a measure in $\mathfrak{P}$ does not affect the definition of pushforward; we employ the following general definition of pushforward for measure-dependent maps, taken to hold for all π_1, π_2 regardless of any assumed relationship between them:

$$T(\cdot;\pi_1)^\sharp \pi_2 = T(\cdot;\widetilde{\pi})^\sharp \pi_2 \Big|_{\widetilde{\pi}=\pi_1}; \tag{2.43}$$

in particular,

$$T(\cdot;\pi)^\sharp \pi = T(\cdot;\widetilde{\pi})^\sharp \pi \Big|_{\widetilde{\pi}=\pi}, \tag{2.44a}$$

$$T(\cdot;\pi)^\sharp (\mathcal{G}\pi) = T(\cdot;\widetilde{\pi})^\sharp (\mathcal{G}\pi) \Big|_{\widetilde{\pi}=\pi}. \tag{2.44b}$$

In the preceding, pushforward under $\widetilde{\pi}$-dependent map $T(\cdot,\widetilde{\pi})$ denotes regular pushforward with no relationship assumed between $\widetilde{\pi}$ and the measure being pushed forward. Note that we may define $\mathcal{T}\colon \mathfrak{P}(\mathbb{R}^d) \to \mathfrak{P}(\mathbb{R}^d)$ by $\mathcal{T}(\pi) = T(\cdot;\pi)^\sharp \pi$. The (approximate) transport maps just identified can be recast as mean-field maps when used in the context where source π is the distribution of the input to T.

2.5.2. *Perfect transport*

Consider the idea of finding a transport map that acts on the joint space of state and data, to effect conditioning with respect to the observed data. To this end we consider the dynamical system, assumed to hold for all $n \in \mathbb{Z}^+$:

$$\boxed{\begin{aligned} \widehat{v}_{n+1} &= \Psi(v_n) + \xi_n, && (2.45a)\\ \widehat{y}_{n+1} &= h(\widehat{v}_{n+1}) + \eta_{n+1}, && (2.45b)\\ v_{n+1} &= T^S\left(\widehat{v}_{n+1}, \widehat{y}_{n+1}; \pi_{n+1}, y^\dagger_{n+1}\right), && (2.45c) \end{aligned}}$$

where $\{y^\dagger_n\}$ arises from a fixed realization of (2.1). Recall that $\mu_n = Law(v_n)$, $\widehat{\mu}_{n+1} = Law(\widehat{v}_{n+1})$ and $\pi_{n+1} = Law(\widehat{v}_{n+1}, \widehat{y}_{n+1})$.

This is an example of the sample-path perspective, and (2.29), (2.30) in particular. The first two equations, which coincide with (2.29), effect the mappings from μ_n

to $\widehat{\mu}_{n+1}$ and from $\widehat{\mu}_{n+1}$ to π_{n+1}. Map[6] $T_n^S(\cdot,\cdot) := T^S(\cdot,\cdot;\pi_{n+1},y_{n+1}^\dagger)$ is then an explicit example of (2.30) defined to effect the desired conditioning of π_{n+1} on $y_{n+1}^\dagger$ in order to obtain μ_{n+1}. The letter S in T^S denotes the dependence of the map on the *stochastic* data $\widehat{y}_{n+1}$. Thus equations (2.45) define a mean-field stochastic dynamical system mapping v_n to v_{n+1}: stochastic because of the noise in (2.45a) and (2.45b), mean-field because the map T^S in (2.45c) depends on the law π_{n+1} of $(\widehat{v}_{n+1},\widehat{y}_{n+1})$, and hence on μ_n. The three update steps in this mean-field stochastic dynamical system lead to the following maps on measures:

$$\boxed{\begin{aligned}\widehat{\mu}_{n+1} &= \mathcal{P}\mu_n, &(2.46a)\\ \pi_{n+1} &= \mathcal{Q}\widehat{\mu}_{n+1}, &(2.46b)\\ \mu_{n+1} &= (T_n^S)^\sharp \pi_{n+1}. &(2.46c)\end{aligned}}$$

This is simply a restatement of (2.22), noting that $(T_n^S)^\sharp$ has been chosen so that pushforward corresponds to conditioning π_{n+1} on data $y_{n+1}^\dagger$ to obtain μ_{n+1}. In particular, $\mathcal{T}_n^S(\pi_{n+1}) := (T_n^S)^\sharp\pi_{n+1}$ has property $\mathcal{T}_n^S(\pi_{n+1}) = \mathcal{B}_n(\pi_{n+1})$. Note that the implied map from μ_n to μ_{n+1} is a nonlinear Markov process, because of the dependence of T_n^S on π_{n+1} and hence on μ_n. Furthermore, we have that $Law(v_\ell) = \mu_\ell$ for all $\ell \in \mathbb{Z}^+$. The important takeaway is that (2.45) defines a sample-path picture of the evolution of the filtering distribution: it provides a map in state space with law governed by the filter. We emphasize again that such a sample-path representation is not uniquely defined.

Now consider a different approach to transport for filtering: we seek a transport map that acts on the state space only to effect Bayes' theorem, i.e. mapping prior $\widehat{\mu}_{n+1}$ to posterior μ_{n+1}. To this end we consider the dynamical system

$$\boxed{\begin{aligned}\widehat{v}_{n+1} &= \Psi(v_n) + \xi_n, &(2.47a)\\ v_{n+1} &= T^D\left(\widehat{v}_{n+1};\widehat{\mu}_{n+1},y_{n+1}^\dagger\right), &(2.47b)\end{aligned}}$$

again assumed to hold for all $n \in \mathbb{Z}^+$. The first equation maps $v_n \sim \mu_n$ to $\widehat{v}_{n+1} \sim \widehat{\mu}_{n+1}$, thus giving a sample-path realization of (2.24a). In the second equation, the map $T_n^D(\cdot) := T^D(\cdot;\widehat{\mu}_{n+1},y_{n+1}^\dagger)$ is chosen so that if $\widehat{v}_{n+1} \sim \widehat{\mu}_{n+1}$ then $v_{n+1} \sim \mu_{n+1}$, thus giving a sample-path realization of (2.24b). Thus we have another instance of the sample-path perspective, and (2.29a), (2.31) in particular.

Equation (2.47) constitutes another mean-field stochastic dynamical system: stochastic because of the noise in (2.47a); mean-field because the map T^D in (2.47b) depends on the law of $\widehat{v}_{n+1}$ itself, and hence on μ_n. The symbol D distinguishes map T^D from map T^S: map T^D is *deterministic* in the sense that it does not require

[6] It is convenient to use both the notation $T^S(\cdot,\cdot;\pi_{n+1},y_{n+1}^\dagger)$, to be explicit about important dependencies in T^S, and to use the notation T_n^S, for succinct statement of certain formulae when dropping explicit dependence of T^S on π_{n+1} and $y_{n+1}^\dagger$. We will use analogous notation for other mean-field maps in what follows.

stochastic data $\widehat{y}_{n+1}$, in contrast to T^S. Therefore we again have that $Law(v_\ell) = \mu_\ell$ for all $\ell \in \mathbb{Z}^+$, where the probability measure μ_n evolves according to

$$\boxed{\begin{aligned} \widehat{\mu}_{n+1} &= \mathcal{P}\mu_n, && (2.48a)\\ \mu_{n+1} &= (T_n^D)^\sharp \widehat{\mu}_{n+1}. && (2.48b) \end{aligned}}$$

This is simply a restatement of (2.24), noting that $(T_n^D)^\sharp$ has been chosen so that the pushforward corresponds to the application of Bayes' theorem to incorporate data $y_{n+1}^\dagger$. In particular, $\mathcal{T}_n^D(\widehat{\mu}_{n+1}) := (T_n^D)^\sharp \widehat{\mu}_{n+1}$ has property $\mathcal{T}_n^D(\widehat{\mu}_{n+1}) = \mathcal{L}_n(\widehat{\mu}_{n+1})$. The evolution (2.48) is another nonlinear Markov process, now because of the dependence of T_n^D on $\widehat{\mu}_{n+1}$. Again, the underlying sample-path representation (2.47) is not uniquely defined.

The two transport maps T^S and T^D introduce an important conceptual approach to algorithms for filtering, but determining the maps can be as hard as, or harder than, solving the filtering problem itself. Thus, in the next two subsections, we turn to relaxations of the perfect transport effected by T^S and T^D. We instead seek mean-field maps which match only first- and second-order moment information; this relaxation allows for approximate transport maps with simple affine forms (in senses to be made precise). The perspective of matching first- and second-order moments naturally suggests working with Gaussians, and hence we also relate the approximate transport to the Gaussian projected filter.

2.5.3. Second-order transport: motivation

To motivate the more general ideas behind second-order transport, we first study an explicit example. It is well known how to transform samples from a unit centred Gaussian random variable on $\mathbb{R}$ into samples from a Gaussian random variable with mean $m \neq 0$ and variance $\sigma \neq 1$ by a simple scaling and shifting operation. An appropriate generalization of such a procedure suggests consideration of the map

$$\boxed{\begin{aligned} \widehat{v}_{n+1} &= \Psi(v_n) + \xi_n, && (2.49a)\\ \widehat{y}_{n+1} &= h(\widehat{v}_{n+1}) + \eta_{n+1}, && (2.49b)\\ v_{n+1} &= m_{n+1} + C_{n+1}^{1/2}\widehat{C}_{n+1}^{-1/2}(\widehat{v}_{n+1} - \mathbb{E}\widehat{v}_{n+1}), && (2.49c) \end{aligned}}$$

where m_{n+1}, $\widehat{C}_{n+1}$ and C_{n+1} are determined by (2.32), (2.33) and (2.41), using (2.49a) and (2.49b). Here $\{y_n^\dagger\}$ arises again from a fixed realization of (2.1). This is a specific instance of the sample-path perspective, and (2.29), (2.31) in particular. In this case the sample-path evolution of v_n has law which only approximates the true filtering law.

The map $v_n \mapsto v_{n+1}$ defined by (2.49) is a mean-field map because of the dependence of m_{n+1}, C_{n+1} and $\widehat{C}_{n+1}$ on $\mathcal{Q}\,Law(v_n)$. It may be viewed as an approximation to (2.45) which is exact when $\mathcal{Q}\,Law(v_n)$ is Gaussian. To demonstrate exactness on Gaussians it suffices to show that we obtain the desired mean and covariance after

application of the map. It is clear from (2.49c) that $\mathbb{E}v_{n+1} = m_{n+1}$ and also that

$$\begin{aligned}
&\mathbb{E}((v_{n+1} - m_{n+1}) \otimes (v_{n+1} - m_{n+1})) \\
&\quad = C_{n+1}^{1/2} \widehat{C}_{n+1}^{-1/2} \mathbb{E}((v_{n+1} - m_{n+1}) \otimes (v_{n+1} - m_{n+1})) \widehat{C}_{n+1}^{-1/2} C_{n+1}^{1/2} \\
&\quad = C_{n+1}^{1/2} \widehat{C}_{n+1}^{-1/2} \widehat{C}_{n+1} \widehat{C}_{n+1}^{-1/2} C_{n+1}^{1/2} \\
&\quad = C_{n+1}.
\end{aligned}$$

It is important to recognize that, in general, v_{n+1} defined by (2.49c) will not be Gaussian-distributed since $\widehat{v}_{n+1}$, defined by (2.49a), will not be Gaussian either. However, although (2.49) does not provide a closed iteration on Gaussians, the map from $\widehat{v}_{n+1}$ to v_{n+1} agrees with the same step in the Gaussian projected filter, at the level of first- and second-order moments. But, because it is not a closed iteration on $\mathfrak{G}(\mathbb{R}^{d_v})$, it is clearly not the same as the Gaussian projected filter.

Whilst the mean-field map from (2.49) is a relatively transparent way to achieve the goal of matching first- and second-order moments in the transport step, there is an uncountable set of ways of achieving this objective; the next two subsections demonstrate this, identifying all mean-field maps effecting approximate transport from within two specific classes of affine transformations. We will then highlight a small subset that have been used in practice, each of which is useful in certain specific contexts.

2.5.4. *Second-order transport: stochastic case*

The first class of approximate filters determined by mean-field maps have the sample-path form

$$\boxed{\begin{aligned}
\widehat{v}_{n+1} &= \Psi(v_n) + \xi_n, && \text{(2.50a)} \\
\widehat{y}_{n+1} &= h(\widehat{v}_{n+1}) + \eta_{n+1}, && \text{(2.50b)} \\
v_{n+1} &= \widetilde{T}^S\left(\widehat{v}_{n+1}, \widehat{y}_{n+1}; \pi_{n+1}, y_{n+1}^\dagger\right), && \text{(2.50c)}
\end{aligned}}$$

where $\{y_n^\dagger\}$ arises from a fixed realization of (2.1). We identify $\widetilde{T}_n^S : \mathbb{R}^{d_v} \times \mathbb{R}^{d_y} \to \mathbb{R}^{d_v}$, where $\widetilde{T}_n^S(\cdot) := \widetilde{T}^S(\cdot\,; \pi_{n+1}, y_{n+1}^\dagger)$ approximates an exact transport map $T_n^S(\cdot) = T^S(\cdot\,; \pi_{n+1}, y_{n+1}^\dagger)$, defined previously, by matching the first- and second-order moments.

We next introduce[7] $\pi = Law(\widehat{v}_{n+1}, \widehat{y}_{n+1})$ and assume that the exact and approximate transport maps satisfy, respectively,

$$(T_n^S)^\sharp \pi = \mathcal{B}_n(\pi), \tag{2.51a}$$

$$\mathcal{G}\left((\widetilde{T}_n^S)^\sharp \pi\right) = \mathcal{B}_n(\mathcal{G}\pi), \tag{2.51b}$$

[7] We temporarily drop explicit notational dependence on $n+1$ in π and in $y^\dagger$. This should not cause confusion as the approximate map we derive is concerned simply with finding a pushforward that approximates conditioning of $Law(\widehat{v}_{n+1}, \widehat{y}_{n+1})$ on $\widehat{y}_{n+1} = y^\dagger$.

for all measures π on the product space $\mathbb{R}^{d_v} \times \mathbb{R}^{d_y}$. Of course, T_n^S and $\widetilde{T}_n^S$ will depend on π and then pushforward is to be interpreted as in (2.43) and (2.44). It is intuitive that (2.51b) enforces map $\widetilde{T}_n^S$ to satisfy (2.51a) when π is Gaussian; we prove this in Lemma 2.11 below.

Perfect transport corresponds to asking that for all measures π on the space $\mathbb{R}^{d_v} \times \mathbb{R}^{d_y}$, (2.51a) holds; second-order transport relaxes this and asks only that (2.51b) holds. Whilst achieving (2.51a) may be harder than solving the filtering problem directly, we will show that achieving (2.51b) is straightforward and leads to computationally tractable methods. This gain in tractability comes at the price of only achieving (2.51b) in place of (2.51a). However, it is intuitive that this price will not be high for settings in which the filtering distribution, and the predictive distribution on state and data, is not too far from Gaussian; we flesh out this idea in Section 2.5.6 below.

Working to satisfy (2.51b) allows us to find tractable approximate second-order transport maps by seeking $\widetilde{T}^S$ in the form

$$\widetilde{T}^S(\widehat{v}_{n+1}, \widehat{y}_{n+1}; \pi, y^\dagger) := A\widehat{v}_{n+1} + B\widehat{y}_{n+1} + a. \tag{2.52}$$

We allow the matrices/vectors A, B, a to depend on $(\pi, y^\dagger)$; however, they are assumed to be independent of $(\widehat{v}_{n+1}, \widehat{y}_{n+1})$. Making this assumption ensures that the transport map is affine with respect to the realization of $(\widehat{v}_{n+1}, \widehat{y}_{n+1})$ (but not their law). This in turn leads to tractable computations to determine A, B, a on the basis of matching second-order moments of perfect transport. In addition to computational tractability, the affine form of the transport map $\widetilde{T}^S$ is motivated by the following, which shows that the approximate transport is perfect when applied to a Gaussian source.

Lemma 2.11. Consider approximate transport map $\widetilde{T}_n^S = \widetilde{T}^S(\widehat{v}_{n+1}, \widehat{y}_{n+1}; \pi, y^\dagger)$ with the form (2.52), assumed to satisfy (2.51b). Then $\widetilde{T}^S$ depends on π only through $\mathcal{G}\pi$. Furthermore,

$$(\widetilde{T}_n^S)^\sharp(\mathcal{G}\pi) = \mathcal{B}_n(\mathcal{G}\pi);$$

thus, if π is Gaussian, equation (2.51b) implies (2.51a). ◇

Proof. We first note that (2.51b) is equivalent to insisting that

$$\mathcal{G}\left((\widetilde{T}_n^S)^\sharp \mathcal{G}\pi\right) = \mathcal{B}_n(\mathcal{G}\pi) \tag{2.53}$$

for all measures π; this follows because, noting the definition (2.43) and consequence (2.44), the first and second moments of $(\widetilde{T}_n^S)^\sharp \mathcal{G}\pi$ and $(\widetilde{T}_n^S)^\sharp \pi$ agree, because of the affine form (2.52) assumed for $\widetilde{T}_n^S$. Recall that $(\widetilde{T}_n^S)$ depends on $(\pi, y^\dagger) = \left(Law(\widehat{v}_{n+1}, \widehat{y}_{n+1}), y_{n+1}^\dagger\right)$. From the identity (2.53), it is clear that $\widetilde{T}_n^S$ only depends on π through $\mathcal{G}\pi$ because changing $\pi \to \mathcal{G}\pi$ leaves the identity invariant, as $\mathcal{G} \circ \mathcal{G} = \mathcal{G}$. Now note that because Gaussians are preserved under affine transformations,

$$\left(\widetilde{T}^S(\widehat{v}_{n+1}, \widehat{y}_{n+1}; \pi, y^\dagger)\right)^\sharp (\mathcal{G}\pi) = \mathcal{G}\left(\left(\widetilde{T}^S(\widehat{v}_{n+1}, \widehat{y}_{n+1}; \pi, y^\dagger)\right)^\sharp \pi\right),$$

or, in compact notational form,

$$(\widetilde{T}_n^S)^\sharp(\mathcal{G}\pi) = \mathcal{G}\big((\widetilde{T}_n^S)^\sharp\pi\big). \tag{2.54}$$

The desired display in the lemma is then immediate from (2.51b). □

An affine transport map of the form (2.52), when combined with particle approximations, leads to practical implementable algorithms and achieves (2.51b) by ensuring that $(\widetilde{T}_n^S)^\sharp\pi$ has first and second moments which agree with those of the Gaussian projected filter; these are given by equations (2.32), (2.33) and (2.41) when π is the law of $(\widehat{v}_{n+1}, \widehat{y}_{n+1})$.

In Appendix C.1, we identify the (uncountable) set of all possible A, B, a which achieve the desired matching of first- and second-order moments. Here we focus on the two specific choices given in Example C.5 from that appendix. The first that we highlight corresponds to the choice

$$\widetilde{T}^S\big(\widehat{v}_{n+1}, \widehat{y}_{n+1}; \widetilde{\pi}_{n+1}, y_{n+1}^\dagger\big) := \widehat{v}_{n+1} + K_n\big(y_{n+1}^\dagger - \widehat{y}_{n+1}\big),$$

with $K_n = K(\pi_{n+1})$ given by (2.34). Thus we obtain the following mean-field dynamical system, which corresponds to (2.15) in the setting where the Kalman gain K_n is defined by (2.34):

$$\boxed{\begin{aligned} \widehat{v}_{n+1} &= \Psi(v_n) + \xi_n, && \text{(2.55a)}\\ \widehat{y}_{n+1} &= h(\widehat{v}_{n+1}) + \eta_{n+1}, && \text{(2.55b)}\\ v_{n+1} &= \widehat{v}_{n+1} + \widehat{C}_{n+1}^{vy}\big(\widehat{C}_{n+1}^{yy}\big)^{-1}\big(y_{n+1}^\dagger - \widehat{y}_{n+1}\big), && \text{(2.55c)} \end{aligned}}$$

where $\{y_n^\dagger\}$ arises from a fixed realization of (2.1), and equations (2.32) and (2.33) define the Kalman gain $K_n = \widehat{C}_{n+1}^{vy}(\widehat{C}_{n+1}^{yy})^{-1}$. We refer to this as *Kalman transport*, noting that it serves as a derivation of the Kalman gain, beyond the linear Gaussian setting. This is a specific instance of the sample-path perspective, and (2.29), (2.30) in particular. Again, this is a case in which the sample-path evolution for v_n has law which only approximates the true filtering law.

The second transport map from Example C.5 corresponds to the choice

$$\widetilde{T}^S\big(\widehat{v}_{n+1}, \widehat{y}_{n+1}; \pi_{n+1}, y_{n+1}^\dagger\big) := m_{n+1} + C_{n+1}^{1/2}\widehat{C}_{n+1}^{-1/2}(\widehat{v}_{n+1} - \mathbb{E}\widehat{v}_{n+1}),$$

leading to the mean-field map (2.49), recalling that m_{n+1}, $\widehat{C}_{n+1}$ and C_{n+1} are determined by (2.32), (2.33) and (2.41), using (2.49a) and (2.49b).

Remark 2.12. One important difference between the mean-field models (2.55) and (2.49) is that the former involves inversion of matrices in data space and the latter in state space. The relative dimensions of the two spaces plays a role in determining which mean-field model is more appropriate as the basis of algorithms. A second notable difference is that the mean-field model (2.49) does not require generation of the stochastic data $\widehat{y}_{n+1}$. This is because we may employ the identity

$\mathbb{E}\widehat{y}_{n+1} = \mathbb{E}h(\widehat{v}_{n+1})$ and use (2.35) and (2.36) to compute m_{n+1}, C_{n+1} and $\widehat{C}_{n+1}$. Motivated by this observation, the next subsection studies a wide class of approximate transport maps with the property that they do not require generation of stochastic data. □

2.5.5. *Second-order transport: deterministic case*

We now turn our attention to approximate filters defined by deterministic mean-field maps. We seek to approximate the exact transport (2.47) by mean-field maps with the sample-path form

$$\boxed{\begin{aligned} \widehat{v}_{n+1} &= \Psi(v_n) + \xi_n, && \text{(2.56a)} \\ v_{n+1} &= \widetilde{T}^D\left(\widehat{v}_{n+1}; \widehat{\mu}_{n+1}, y^\dagger_{n+1}\right), && \text{(2.56b)} \end{aligned}}$$

where $\{y^\dagger_n\}$ arises from a fixed realization of (2.1). As in the previous subsection we drop explicit n-dependence on the measure $\widehat{\mu}_{n+1}$ and on the data $y^\dagger_{n+1}$ when no confusion arises from doing so. To this end we define, for $\widehat{v}_{n+1}$ given by (2.29b), $\widehat{\mu} = Law(\widehat{v}_{n+1})$ and $y^\dagger = y^\dagger_{n+1}$. In the following, $T^D_n : \mathbb{R}^{d_v} \to \mathbb{R}^{d_v}$ is defined by $T^D_n(\cdot) = T^D(\cdot; \widehat{\mu}, y^\dagger)$ and $\widetilde{T}^D_n : \mathbb{R}^{d_v} \to \mathbb{R}^{d_v}$ is defined by $\widetilde{T}^D_n(\cdot) = \widetilde{T}^D(\cdot; \widehat{\mu}, y^\dagger)$; this is a useful notational convention for the reasons explained in the stochastic transport setting.

Analogously to the identities (2.51) in the previous subsection, we seek an approximation $\widetilde{T}^D$ which, in comparison with the true transport map T^D, satisfies

$$(T^D_n)^\sharp \widehat{\mu} = \mathcal{B}_n(\mathcal{Q}\widehat{\mu}), \tag{2.57a}$$

$$\mathcal{G}\left((\widetilde{T}^D_n)^\sharp \widehat{\mu}\right) = \mathcal{B}_n(\mathcal{G}\mathcal{Q}\widehat{\mu}), \tag{2.57b}$$

for all measures $\widehat{\mu}$ on the state space $\mathbb{R}^{d_v}$. As in the previous subsection, where we studied approximate stochastic transport, we again seek maps with a specific affine form. Concretely, the maps are assumed to be affine in the pair $(\widehat{v}_{n+1}, \widehat{h}_{n+1})$, with $\widehat{h}_{n+1} = h(\widehat{v}_{n+1})$, leading to the assumed form $\widetilde{T}^D_n(\cdot) = \widetilde{T}^D(\cdot; \mu, y^\dagger)$ with

$$\widetilde{T}^D(\widehat{v}_{n+1}, \widehat{h}_{n+1}; \mu, y^\dagger) := R\widehat{v}_{n+1} + S\widehat{h}_{n+1} + r, \tag{2.58}$$

for $(\widehat{\mu}, y^\dagger)$-dependent matrices/vectors R, S, r of appropriate dimensions. Note, however, that R, S, r are assumed to be independent of the realization $(\widehat{v}_{n+1}, \widehat{h}_{n+1})$, depending only on its law, so that the transport map is affine in $(\widehat{v}_{n+1}, \widehat{h}_{n+1})$. With this restriction, which will lead to practical implementable algorithms, we simply ask that (2.57b) holds: the first and second moments of the output map agree with those of the Gaussian projected filter, given by equations (2.32), (2.35) and (2.42).

As in the previous subsection, there are uncountably many choices of R, S, r which we identify in Appendix C.2; Example C.12 highlights two important cases. The first coincides with (2.49) since $S = 0$, but the second leads to a new mean-field

map. To formulate this new map we first define $\widetilde{K}_n = \widetilde{K}(\widehat{\mu})$ by

$$\widetilde{K}_n = \widehat{C}^{vh}_{n+1}\left(\left(\widehat{C}^{hh}_{n+1} + \Gamma\right) + \Gamma^{1/2}\left(\widehat{C}^{hh}_{n+1} + \Gamma\right)^{1/2}\right)^{-1}. \tag{2.59}$$

We then make the choice

$$\widetilde{T}^D(\widehat{v}_{n+1}, \widehat{h}_{n+1}; \widehat{\mu}, y^\dagger) := \widehat{v}_{n+1} - \widetilde{K}_n(\widehat{h}_{n+1} - \mathbb{E}\widehat{h}_{n+1}) + K_n(y^\dagger - \mathbb{E}\widehat{h}_{n+1}),$$

with K_n given by (2.36) and repeated here for convenience:

$$K_n = \widehat{C}^{vh}_{n+1}\left(\widehat{C}^{hh}_{n+1} + \Gamma\right)^{-1}.$$

The second mean-field map identified in Example C.12 is

$$\widehat{v}_{n+1} = \Psi(v_n) + \xi_n, \tag{2.60a}$$
$$\widehat{h}_{n+1} = h(\widehat{v}_{n+1}), \tag{2.60b}$$
$$v_{n+1} = \widehat{v}_{n+1} - \widetilde{K}_n(\widehat{h}_{n+1} - \mathbb{E}\widehat{h}_{n+1}) + K_n\left(y^\dagger_{n+1} - \mathbb{E}\widehat{h}_{n+1}\right), \tag{2.60c}$$

where K_n and $\widetilde{K}_n$ are computed under $Law(\widehat{v}_{n+1})$. This is another instance of the sample-path perspective, and (2.29), (2.31) in particular. Again this sample-path evolution for v_n has law which only approximates the true filtering law.

Remark 2.13. If the ensemble spread is such that the size of $\widehat{C}^{hh}_{n+1}$ is much smaller than the size of the observational covariance Γ, then we may invoke the approximation $\widehat{C}^{hh}_{n+1} + \Gamma \approx \Gamma$. With this approximation it follows that $\widetilde{K}_n \approx \frac{1}{2}K_n$ in (2.59). Some deterministic ensemble Kalman filters are derived from mean-field dynamics which exploit this approximation by setting $\widetilde{K}_n = \frac{1}{2}K_n$ in (2.60). We then replace (2.60c) with the compact update step

$$v_{n+1} = \widehat{v}_{n+1} + K_n\left(y^\dagger_{n+1} - \frac{1}{2}(\mathbb{E}\widehat{h}_{n+1} + \widehat{h}_{n+1})\right). \tag{2.61}$$

Such a formulation corresponds to the control-theoretic perspective of (2.15), with K_n given by (2.36) and the innovation (2.16) replaced by

$$\mathfrak{I}_n = y^\dagger_{n+1} - \frac{1}{2}(\mathbb{E}\widehat{h}_{n+1} + \widehat{h}_{n+1}). \tag{2.62}$$

Filters based on this mean-field dynamics thus invoke an additional approximation of perfect transport, over and above that stemming from matching only first and second moments: they assume further that the observational noise dominates ensemble variation. However, we will see that in the continuous-time limit described in Section 3, this form of the innovation arises naturally and does not constitute an additional approximation. □

2.5.6. *Second-order transport: summary*

It is helpful at this point to take stock of two approximations to filtering that we have introduced, Gaussian projected filtering and approximate transport, and

discuss their inter-relations. For simplicity we do this in the context of mean-field stochastic maps, but similar considerations extend to mean-field deterministic maps. In this subsection we also include Example 2.15, demonstrating the existence of mean-field maps for the Kalman filter which, recall, applies only in the linear Gaussian setting.

Recall that

$$\mu_{n+1} = \mathcal{B}_n(\mathcal{Q}\mathcal{P}\mu_n), \quad \mu_0 = \mathsf{N}(m_0, C_0), \tag{2.63a}$$

$$\mu^G_{n+1} = \mathcal{B}_n(\mathcal{G}\mathcal{Q}\mathcal{P}\mu^G_n), \quad \mu^G_0 = \mathsf{N}(m_0, C_0), \tag{2.63b}$$

With the goal of discussing the inter-relations between Gaussian projected filtering and approximate transport methods, we let μ^{MF} denote the measure associated with using the mean-field map $\widetilde{T}^S_n$ to approximate the conditioning step in (2.22). Using (2.51b) to rewrite the Gaussian projected filter, and using the construction of the stochastic mean-field model (2.50), we obtain

$$\mu^G_{n+1} = \mathcal{G}\big((\widetilde{T}^S_n)^\sharp(\mathcal{Q}\mathcal{P}\mu^G_n)\big), \quad \mu^G_0 = \mathsf{N}(m_0, C_0), \tag{2.64a}$$

$$\mu^{MF}_{n+1} = (\widetilde{T}^S_n)^\sharp(\mathcal{Q}\mathcal{P}\mu^{MF}_n), \quad \mu^{MF}_0 = \mathsf{N}(m_0, C_0). \tag{2.64b}$$

Remark 2.14. Equations (2.64) show that $\{\mu^{MF}_n\}$ is close to $\{\mu^G_n\}$ if the Gaussian projection in (2.64a) is close to the identity where it acts on the output of one step. Equations (2.63) show that $\{\mu^G_n\}$ is close to $\{\mu_n\}$ if the Gaussian projection in (2.63b) is close to the identity where it acts on the joint space of state and observation. Together these two facts suggest that $\{\mu^{MF}_n\}$, $\{\mu^G_n\}$ and $\{\mu_n\}$ are all close to one another if the two Gaussian projections can be viewed as being close to the identity map, where they appear in (2.63) and in (2.64). This provides a potential path for analysis of the mean-field model, away from the linear setting where it is exact. Note also that (2.64a) shows that the Gaussian projected filter evolves within the manifold of Gaussian probability measures; the mean-field model (2.64b) does not. □

Example 2.15. Assume that $v_0 \sim \mathsf{N}(m_0, C_0)$, that $\Gamma \succ 0$, and consider the Kalman filter setting of Example 2.6; in particular, (2.6) prevails, rendering Ψ and h linear. The mean-field stochastic dynamical system (2.55) then takes the form

$$\widehat{v}_{n+1} = Mv_n + \xi_n, \tag{2.65a}$$

$$\widehat{y}_{n+1} = H\widehat{v}_{n+1} + \eta_{n+1}, \tag{2.65b}$$

$$v_{n+1} = \widehat{v}_{n+1} + \widehat{C}_{n+1}H^\top(H\widehat{C}_{n+1}H^\top + \Gamma)^{-1}\big(y^\dagger_{n+1} - \widehat{y}_{n+1}\big), \tag{2.65c}$$

where $\{y^\dagger_n\}$ arises from a fixed realization of (2.25) and where C_n is the covariance of v_n and $\widehat{C}_{n+1} = MC_nM^\top + \Sigma$ is the covariance of $\widehat{v}_{n+1}$. The resulting dynamics give a sample-path representation of the Kalman filter in that $v_n \sim \mathsf{N}(m_n, C_n)$, where m_n, C_n are as given in Example 2.6. This follows because the map defined by (2.65) is well-defined, since $\Gamma \succ 0$. Lemma 2.11 shows that the approximate transport is exact in this Gaussian setting.

Similar ideas can be applied to (2.49) and (2.60) to determine other mean-field models with law equal to that of the Kalman filter. Furthermore, we observe that the formulation based on (2.49) can be symmetrized to obtain, in the linear Gaussian setting of (2.6),

$$\widehat{v}_{n+1} = Mv_n + \xi_n,\ n \in \mathbb{Z}^+, \tag{2.66a}$$

$$v_{n+1} = m_{n+1} + A_n(\widehat{v}_{n+1} - \widehat{m}_{n+1}), \tag{2.66b}$$

$$A_n = (C_{n+1})^{1/2}\left[(C_{n+1})^{1/2}\widehat{C}_{n+1}(C_{n+1})^{1/2}\right]^{-1/2}(C_{n+1})^{1/2}, \tag{2.66c}$$

with $\left(\widehat{m}_{n+1}, \widehat{C}_{n+1}, \widehat{C}^{vh}_{n+1}, \widehat{C}^{hh}_{n+1}\right)$ defined by (2.32) and (2.35), and (m_{n+1}, C_{n+1}) given by (2.41). We note that the second component of the map may be written in gradient form, and corresponds to an optimal transport from $\widehat{\mu}_{n+1}$ into μ_{n+1} in the sense of the Euclidean Wasserstein distance of optimal transportation (see Section 2.7 for details). Indeed, this is true for an entire family of weighted Wasserstein distances: see Example C.7. To recognize the gradient structure, define

$$\Phi_n(v) := \langle m_{n+1}, v\rangle + \frac{1}{2}\langle A_n(v - \widehat{m}_{n+1}), v - \widehat{m}_{n+1}\rangle$$

and note that then

$$\widehat{v}_{n+1} = Mv_n + \xi_n,\ n \in \mathbb{Z}^+, \tag{2.67a}$$

$$v_{n+1} = \nabla\Phi_n(\widehat{v}_{n+1}). \tag{2.67b}$$

□

2.6. *Ensemble Kalman methods*

The mean-field formulations from Section 2.5 provide clear insights into many of the design choices and mechanisms that underlie ensemble Kalman methods. In this subsection we take the mean-field models and use particle approximations to derive implementable numerical algorithms. When approximated by interacting particle systems, the mean-field formulations of ensemble Kalman methods lead to actionable algorithms.

We start, in Section 2.6.1, with the setting in which transport is perfect. These algorithms are not, in general, implementable since determining perfect transport is itself a difficult computational task and the subject of ongoing research; see Section 2.7. Thus we turn to particle approximations of the transports designed to match first- and second-order statistics. This leads to the (stochastic) *ensemble Kalman filter* in Section 2.6.2, and to (deterministic) *ensemble square root filters* in Section 2.6.3. The methods derived in this subsection involve approximating the Kalman gain by computing covariances under the empirical measure defined by the ensemble of particles. To avoid overloading notation, in the rest of this subsection C_{n+1}, $\widehat{C}_{n+1}$, $\widehat{C}^{vy}_{n+1}$ and $\widehat{C}^{yy}_{n+1}$ will denote covariances computed with expectation under the empirical measure. With this notation in place, in the rest of this specific subsection, K_n will directly refer to the particle approximation of the Kalman gain,

computed using the covariances with respect to the empirical measure, without further specification needed. Throughout this section $\{y_n^\dagger\}$ arises from a fixed realization of (2.1).

2.6.1. *Perfect particle filters*

The mean-field equations (2.45) could, in principle, be approximated through a particle approximation of the mean field leading to the following conceptual (because map T_n^S is not known explicitly) algorithm: let $\mathsf{J} = \{1, \ldots, J\}$ and consider, for $(n, j) \in \mathbb{Z}^+ \times \mathsf{J}$, the interacting particle dynamical system

$$\widehat{v}_{n+1}^{(j)} = \Psi\big(v_n^{(j)}\big) + \xi_n^{(j)}, \tag{2.68a}$$

$$\widehat{y}_{n+1}^{(j)} = h\big(\widehat{v}_{n+1}^{(j)}\big) + \eta_{n+1}^{(j)}, \tag{2.68b}$$

$$v_{n+1}^{(j)} = T^S\big(\widehat{v}_{n+1}^{(j)}, \widehat{y}_{n+1}^{(j)}; \pi_{n+1}^{\mathsf{J}}, y_{n+1}^\dagger\big), \tag{2.68c}$$

$$\pi_{n+1}^{\mathsf{J}} = \frac{1}{J}\sum_{j=1}^{J} \delta_{\left(\widehat{v}_{n+1}^{(j)}, \widehat{y}_{n+1}^{(j)}\right)}. \tag{2.68d}$$

This evolves the particles $\{v_n^{(j)}\}_{j\in\mathsf{J}}$ into $\{v_{n+1}^{(j)}\}_{j\in\mathsf{J}}$. Here the $\{\xi_n^{(j)}\}$ are, for each j, random variables given by the known distribution of ξ_n specified in (2.2) and, furthermore, are drawn independently with respect to each $(n, j) \in \mathbb{Z}^+ \times \mathsf{J}$. Similar considerations apply to the $\{\eta_n^{(j)}\}$ which, additionally, are independent of the $\{\xi_n^{(j)}\}$. It is intuitive that the large J limit of this system recovers the mean-field dynamics (2.45) and, in particular,

$$\mu_n^{\mathsf{J}} = \frac{1}{J}\sum_{j=1}^{J} \delta_{v_n^{(j)}} \approx \mu_n. \tag{2.69}$$

Applying a similar idea to (2.47) leads to the following conceptual (because map T_n^D is not known explicitly) algorithm. Consider, for $(n, j) \in \mathbb{Z}^+ \times \mathsf{J}$, the interacting particle dynamical system

$$\widehat{v}_{n+1}^{(j)} = \Psi\big(v_n^{(j)}\big) + \xi_n^{(j)}, \tag{2.70a}$$

$$v_{n+1}^{(j)} = T^D\big(\widehat{v}_{n+1}^{(j)}; \widehat{\mu}_{n+1}^{\mathsf{J}}, y_{n+1}^\dagger\big), \tag{2.70b}$$

$$\widehat{\mu}_{n+1}^{\mathsf{J}} = \frac{1}{J}\sum_{j=1}^{J} \delta_{\widehat{v}_{n+1}^{(j)}}. \tag{2.70c}$$

This evolves the particles $\{v_n^{(j)}\}_{j\in\mathsf{J}}$ into $\{v_{n+1}^{(j)}\}_{j\in\mathsf{J}}$. The same assumptions are made about the $\{\xi_n^{(j)}\}$ as for the preceding interacting particle dynamical system. It is again intuitive that the large J limit of this system recovers the mean-field dynamics (2.47), and the evolution (2.48). In particular, it is intuitive that (2.69) holds for this particle approximation too. We reiterate that in practice these algorithms are, in

general, not easy to use. More specifically, finding particle-based approximations $T^{S,J}$ and $T^{D,J}$ to the desired transport maps such that

$$\lim_{J\to\infty} T^{S,J} = T^S, \quad \lim_{J\to\infty} T^{D,J} = T^D$$

in an appropriate sense is a computationally challenging task and the subject of ongoing research. This leads to the next two subsections in which we replace T^S and T^D, in the interacting particles systems (2.68) and (2.70), by the previously introduced affine approximate transports $\widetilde{T}^S$ and $\widetilde{T}^D$, respectively.

2.6.2. *Stochastic ensemble Kalman filters*

Particle approximation of the mean-field dynamical system (2.55), effecting Kalman transport, bring us to the stochastic EnKF (*ensemble Kalman filter*). This method may be derived by writing down a particle approximation of the mean-field stochastic dynamics defined by (2.55). We evolve the particles $\{v_n^{(j)}\}_{j\in\mathsf{J}}$ into $\{v_{n+1}^{(j)}\}_{j\in\mathsf{J}}$ according to the following stochastic interacting particle system, holding for $(n, j) \in \mathbb{Z}^+ \times \mathsf{J}$:

$$\widehat{v}_{n+1}^{(j)} = \Psi\big(v_n^{(j)}\big) + \xi_n^{(j)},\ n \in \mathbb{Z}^+, \tag{2.71a}$$

$$\widehat{y}_{n+1}^{(j)} = h\big(\widehat{v}_{n+1}^{(j)}\big) + \eta_{n+1}^{(j)},\ n \in \mathbb{Z}^+, \tag{2.71b}$$

$$v_{n+1}^{(j)} = \widehat{v}_{n+1}^{(j)} + K_n\big(y_{n+1}^\dagger - \widehat{y}_{n+1}^{(j)}\big), \tag{2.71c}$$

$$\pi_{n+1}^{\mathsf{J}} = \frac{1}{J}\sum_{j=1}^{J} \delta_{\left(\widehat{v}_{n+1}^{(j)}, \widehat{y}_{n+1}^{(j)}\right)}. \tag{2.71d}$$

Here the Kalman gain from (2.34) is approximated using the empirical measure π_{n+1}^{J}, but it is still denoted by K_n to avoid proliferation of notation; details follow below. The same assumptions regarding $\{\xi_n^{(j)}\}$ and $\{\eta_{n+1}^{(j)}\}$ are made as for (2.68). We let $\mathbb{E}_n^{\mathsf{J}}$ denote expectation under π_n^{J}. For the basic implementation of EnKF (2.71) the desired covariance matrices, and Kalman gain (2.34), are then approximated by expectation under π_{n+1}^{J}, so that[8]

$$\begin{aligned}
\widehat{C}_{n+1}^{vy} &= \mathbb{E}_{n+1}^{\mathsf{J}}\big(\big(\widehat{v}_{n+1} - \mathbb{E}_{n+1}^{\mathsf{J}}\widehat{v}_{n+1}\big) \otimes \big(\widehat{y}_{n+1} - \mathbb{E}_{n+1}^{\mathsf{J}}\widehat{y}_{n+1}\big)\big),\\
\widehat{C}_{n+1}^{yy} &= \mathbb{E}_{n+1}^{\mathsf{J}}\big(\big(\widehat{y}_{n+1} - \mathbb{E}_{n+1}^{\mathsf{J}}\widehat{y}_{n+1}\big) \otimes \big(\widehat{y}_{n+1} - \mathbb{E}_{n+1}^{\mathsf{J}}\widehat{y}_{n+1}\big)\big),\\
K_n &= \widehat{C}_{n+1}^{vy}\big(\widehat{C}_{n+1}^{yy}\big)^{-1}.
\end{aligned}$$

Note that a pseudo-inverse may be required to define K_n. An alternative, avoiding the pseudo-inverse, is to use a particle approximation in formula (2.36b), leading

[8] The empirical covariance computations are often modified to accommodate the widely adopted convention of scaling by $1/(J-1)$, instead of $1/J$, in view of the matrix being computed from $J-1$ independent increments about the mean.

to the use of

$$\widehat{C}^{vh}_{n+1} = \mathbb{E}^{\mathsf{J}}_{n+1}\big(\big(\widehat{v}_{n+1} - \mathbb{E}^{\mathsf{J}}_{n+1}\widehat{v}_{n+1}\big) \otimes \big(h(\widehat{v}_{n+1}) - \mathbb{E}^{\mathsf{J}}_{n+1}h(\widehat{v}_{n+1})\big)\big), \tag{2.72a}$$

$$\widehat{C}^{hh}_{n+1} = \mathbb{E}^{\mathsf{J}}_{n+1}\big(\big(h(\widehat{v}_{n+1}) - \mathbb{E}^{\mathsf{J}}_{n+1}h(\widehat{v}_{n+1})\big) \otimes \big(h(\widehat{v}_{n+1}) - \mathbb{E}^{\mathsf{J}}_{n+1}h(\widehat{v}_{n+1})\big)\big), \tag{2.72b}$$

$$K_n = \widehat{C}^{vh}_{n+1}\big(\widehat{C}^{hh}_{n+1} + \Gamma\big)^{-1}, \tag{2.72c}$$

to compute the gain K_n. The advantage of this latter formulation is that it ensures positivity, and hence invertibility, of the covariance in data space, if Γ is positive definite. It is hence typically preferred.

Pseudo-code for the stochastic EnKF may be found as Algorithm 2 in Appendix A.

Example 2.16. We return to the set-up of Example 2.3, and now demonstrate performance of the stochastic EnKF on the same Lorenz '96 model. Indeed, we again study the Lorenz '96 (single-scale) model for unknown $v \in C(\mathbb{R}^+, \mathbb{R}^L)$ satisfying equations (2.9) with $L = 9, h_v = -0.8$ and $F = 10$ and function m as shown in Figure 2.1. We consider observations $\{y^\dagger_n\}_{n\in\mathbb{Z}^+}$ arising from the model

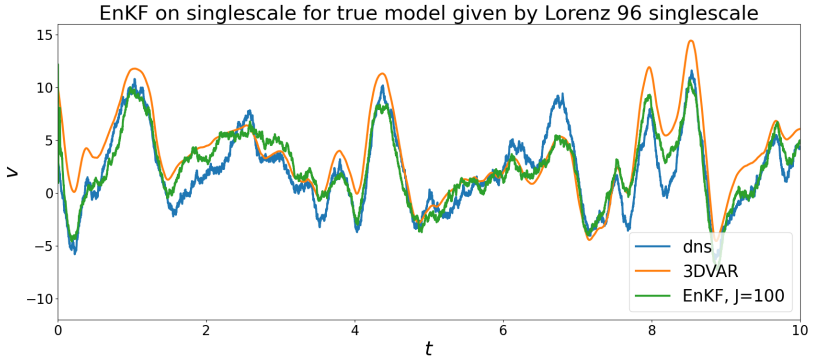
$$\begin{aligned} v^\dagger_{n+1} &= \Psi(v^\dagger_n) + \xi^\dagger_n, \\ y^\dagger_{n+1} &= h\big(v^\dagger_{n+1}\big) + \eta^\dagger_{n+1}, \end{aligned}$$

where Ψ is the solution operator for (2.9) over the observation time interval τ, and $\{\xi^\dagger_n\}_{n\in\mathbb{Z}^+}$, $\{\eta^\dagger_n\}_{n\in\mathbb{N}}$ are mutually independent Gaussian sequences defined by

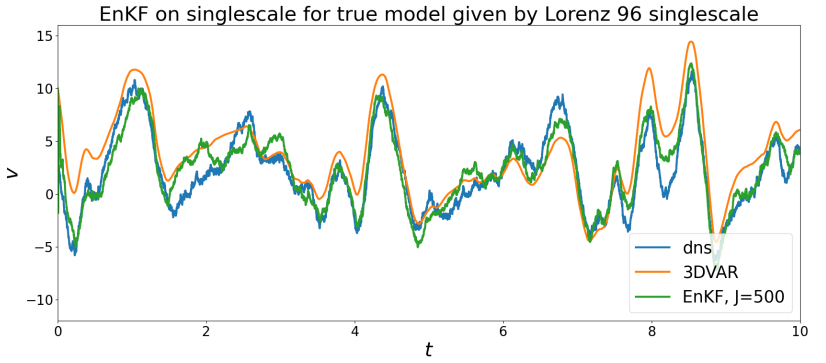
$$\xi^\dagger_n \sim \mathsf{N}(0, \sigma^2 I) \text{ i.i.d.}, \quad \eta^\dagger_n \sim \mathsf{N}(0, \gamma^2 I) \text{ i.i.d.},$$

with $\sigma = 0.1$ and $\gamma = 0.1$. We again assume that the observation function is linear: $h(v) = Hv$ for matrix $H\colon \mathbb{R}^9 \to \mathbb{R}^6$ defined by (2.10).

Figures 2.5(a) and 2.5(b) demonstrate the performance of stochastic EnKF in this experimental setting with $\tau = 10^{-3}$ and using $J = 10^2$ and $J = 5 \times 10^2$, respectively, against the performance of 3DVAR with no noise; note that the EnKF uses a time-varying estimate of the gain K_n, whilst 3DVAR uses the fixed K given in Example 2.3. These experiments illustrate that using sufficiently large ensembles, the ensemble Kalman filter outperforms 3DVAR on such a nonlinear filtering problem where the true state and observational noise levels are high. Here *outperforms* refers to mean squared error in recovery of the state. To demonstrate this improvement quantitatively, we compute (a) the mean squared error between the estimates yielded by 3DVAR and the true states, and (b) the mean squared error between the ensemble mean of stochastic EnKF and the true states. In particular we report time-averaged mean squared errors obtained from both 3DVAR and stochastic EnKF given by use of formula (2.17) from Example 2.5 using $t^* = 3$ and $T = 10$. An ensemble size of $J = 10^2$ yields $e_{\text{EnKF}} = 1.05 \times 10^0$, while for $J = 5 \times 10^2$ we obtain $e_{\text{EnKF}} = 5.24 \times 10^{-1}$. For comparison, the error obtained using 3DVAR is $e_{\text{3DVAR}} = 1.85 \times 10^0$. □



(a) EnKF with ensemble size $J = 10^2$



(b) EnKF with ensemble size $J = 5 \times 10^2$

Figure 2.5. In this experiment we set the noise levels $\sigma = 10^{-1}, \gamma = 10^{-1}$. We display the estimates of v_3 in time produced by EnKF (using ensemble average) and 3DVAR against the true dynamics using observation time interval $\tau = 10^{-3}$. Again 'dns' refers to direct numerical simulation. The results show that the EnKF provides a more accurate estimate of the trajectory, albeit at higher cost in terms of number of model evaluations.

2.6.3. Ensemble square root filters

The variants on the EnKF described in this subsection are known as ensemble square root filters; they are based on mean-field maps (2.49) and (2.60), approximated by interacting particle systems.

Remark 2.17. Square root filters are sometimes referred to as deterministic ensemble Kalman filters, to distinguish them from the ensemble Kalman filters described in the preceding subsection (see the discussion of this, and bibliographical information, in Section 1.1). However, we have already used the terminology 'stochastic' and 'deterministic' to distinguish between different variants on the mean-field models that we describe in Sections 2.5.4 and 2.5.5 respectively. With

the exception of the discussion in Section 1.1, we simply refer to square root filters for the methods introduced in this subsection. We note, however, that they are deterministic in the sense that they do not require generation of random variables.

We introduce two families of square root filters: of adjustment and transform type. We emphasize again that the choice of which method to use in practice is determined by implementation details such as the number of particles J, the dimension of state space d_v, and the dimension of the observation space d_y. These implementation details, although important, are not the focus of this paper. □

Ensemble adjustment Kalman filters. We introduce two different particle-based approximations of mean-field models. Both methods are examples of a general class of algorithms known as *ensemble adjustment Kalman filters*: EAKF. The starting point for the first of these EAKF methods is the mean-field map (2.49a), (2.49c), repeated here for convenience:

$$\widehat{v}_{n+1} = \Psi(v_n) + \xi_n, \tag{2.73a}$$

$$v_{n+1} = m_{n+1} + C_{n+1}^{1/2}\widehat{C}_{n+1}^{-1/2}(\widehat{v}_{n+1} - \mathbb{E}\widehat{v}_{n+1}), \tag{2.73b}$$

where m_{n+1}, $\widehat{C}_{n+1}$, and C_{n+1} are determined by (2.32), (2.35), (2.36) and (2.41).[9]

As in the previous subsection the methods evolve particle ensemble $\{v_n^{(j)}\}_{j\in\mathsf{J}}$ into $\{v_{n+1}^{(j)}\}_{j\in\mathsf{J}}$, via the predictive ensemble $\{\widehat{v}_{n+1}^{(j)}\}_{j\in\mathsf{J}}$. However, they do not employ simulated data $\{\widehat{y}_{n+1}^{(j)}\}_{j\in\mathsf{J}}$; rather, they make use of $\{\widehat{h}_{n+1}^{(j)}\}_{j\in\mathsf{J}}$, where $\widehat{h}_n^{(j)} = h(\widehat{v}_n^{(j)})$. To define the methods it helps to introduce new notation. Slightly modifying the notation in the preceding subsection, we now let $\mathbb{E}_n^{\mathsf{J}}$ denote expectation with respect to the empirical measure

$$\widehat{\mu}_n^{\mathsf{J}} = \frac{1}{J}\sum_{j=1}^{J}\delta_{\widehat{v}_n^{(j)}}.$$

We let $\widehat{v}_{n+1}$ denote the random variable with this distribution and, as before, $\widehat{h}_{n+1} = h(\widehat{v}_{n+1})$.

Next we define matrix $\widehat{V}_n$ comprising scaled ensemble deviations in state space:

$$\widehat{V}_n = \frac{1}{\sqrt{J}}\left(\widehat{v}_n^{(1)} - \mathbb{E}_n^{\mathsf{J}}\widehat{v}_n, \widehat{v}_n^{(2)} - \mathbb{E}_n^{\mathsf{J}}\widehat{v}_n, \ldots, \widehat{v}_n^{(J)} - \mathbb{E}_n^{\mathsf{J}}\widehat{v}_n\right) \in \mathbb{R}^{d_v\times J}.$$

We then define the analogous matrix $\widehat{H}_n$ in observation space:

$$\widehat{H}_n = \frac{1}{\sqrt{J}}\left(\widehat{h}_n^{(1)} - \mathbb{E}_n^{\mathsf{J}}\widehat{h}_n, \widehat{h}_n^{(2)} - \mathbb{E}_n^{\mathsf{J}}\widehat{h}_n, \ldots, \widehat{h}_n^{(J)} - \mathbb{E}_n^{\mathsf{J}}\widehat{h}_n\right) \in \mathbb{R}^{d_y\times J}.$$

[9] Although the identity (2.41) is derived in a subsection concerning the Gaussian projected filter, it is contained in Lemma 2.9, the proof of which simply concerns conditioning of Gaussians.

With this notation in hand, we have, with expectations computed under $\mathbb{E}_n^{\mathsf{J}}$,

$$\widehat{C}_n^{vh} = \widehat{V}_n \widehat{H}_n^\top, \tag{2.74a}$$

$$\widehat{C}_n^{hh} = \widehat{H}_n \widehat{H}_n^\top, \tag{2.74b}$$

and the ensemble-based approximation of the Kalman gain matrix is[10]

$$K_n = \widehat{V}_{n+1} \widehat{H}_{n+1}^\top \big(\widehat{H}_{n+1} \widehat{H}_{n+1}^\top + \Gamma \big)^{-1}. \tag{2.75}$$

This is a linear algebraic reformulation of the Kalman gain approximation resulting from (2.72). By making a particle approximation of (2.73), using (2.32), (2.35), (2.36) and (2.41), we obtain

$$\boxed{\begin{aligned}
\widehat{v}_{n+1}^{(j)} &= \Psi(v_n^{(j)}) + \xi_n^{(j)}, \ n \in \mathbb{Z}^+, && \text{(2.76a)} \\
\widehat{m}_{n+1} &= \mathbb{E}_{n+1}^{\mathsf{J}} \widehat{v}_{n+1}, && \text{(2.76b)} \\
m_{n+1} &= \widehat{m}_{n+1} + K_n \big(y_{n+1}^\dagger - \mathbb{E}_{n+1}^{\mathsf{J}} \widehat{h}_{n+1} \big), && \text{(2.76c)} \\
v_{n+1}^{(j)} &= m_{n+1} + C_{n+1}^{1/2} \widehat{C}_{n+1}^{-1/2} \big(\widehat{v}_{n+1}^{(j)} - \widehat{m}_{n+1} \big), && \text{(2.76d)}
\end{aligned}}$$

where K_n is given by (2.75); furthermore, $\widehat{C}_{n+1}$ and C_{n+1} are computed empirically using

$$\widehat{C}_{n+1} = \widehat{V}_{n+1} \widehat{V}_{n+1}^\top, \quad C_{n+1} = \widehat{V}_{n+1} \big(I + \widehat{H}_{n+1}^\top \Gamma^{-1} \widehat{H}_{n+1} \big)^{-1} \widehat{V}_{n+1}^\top. \tag{2.77}$$

The first of these two formulae follows similarly to (2.74); the second of these two formulae is derived as follows:[11]

$$\begin{aligned}
C_{n+1} &= \widehat{C}_{n+1} - \widehat{C}_{n+1}^{vh} \big(\widehat{C}_{n+1}^{hh} + \Gamma \big)^{-1} \big(\widehat{C}_{n+1}^{vh} \big)^\top && \text{(2.78a)} \\
&= \widehat{V}_{n+1} \widehat{V}_{n+1}^\top - \widehat{V}_{n+1} \widehat{H}_{n+1}^\top \big(\widehat{H}_{n+1} \widehat{H}_{n+1}^\top + \Gamma \big)^{-1} \widehat{H}_{n+1} \widehat{V}_{n+1}^\top && \text{(2.78b)} \\
&= \widehat{V}_{n+1} \big(I - \widehat{H}_{n+1}^\top \big(\widehat{H}_{n+1} \widehat{H}_{n+1}^\top + \Gamma \big)^{-1} \widehat{H}_{n+1} \big) \widehat{V}_{n+1}^\top && \text{(2.78c)} \\
&= \widehat{V}_{n+1} \big(I + \widehat{H}_{n+1}^\top \Gamma^{-1} \widehat{H}_{n+1} \big)^{-1} \widehat{V}_{n+1}^\top. && \text{(2.78d)}
\end{aligned}$$

The EAKF (2.76) takes as starting point (2.49). If instead we apply a particle approximation to the mean-field dynamical system (2.60), we obtain a second

[10] Here too, the empirical covariance computations are often modified to accommodate the widely adopted convention of scaling by $1/(J-1)$ instead of $1/J$.

[11] Using, in the last line, the identity $I - W^\top (WW^\top + I)^{-1} W = (I + W^\top W)^{-1}$, which holds for all (not necessarily square) matrices W.

version of the EAKF:

$$\widehat{v}_{n+1}^{(j)} = \Psi\big(v_n^{(j)}\big) + \xi_n^{(j)},\ n \in \mathbb{Z}^+, \tag{2.79a}$$
$$\widehat{h}_{n+1}^{(j)} = h\big(\widehat{v}_{n+1}^{(j)}\big), \tag{2.79b}$$
$$\widehat{m}_{n+1} = \mathbb{E}_{n+1}^{\mathsf{J}} \widehat{v}_{n+1}, \tag{2.79c}$$
$$m_{n+1} = \widehat{m}_{n+1} + K_n\big(y_{n+1}^{\dagger} - \mathbb{E}_{n+1}^{\mathsf{J}} \widehat{h}_{n+1}\big), \tag{2.79d}$$
$$v_{n+1}^{(j)} = m_{n+1} + \big(\widehat{v}_{n+1}^{(j)} - \widehat{m}_{n+1}\big) - \widetilde{K}_n\big(\widehat{h}_{n+1}^{(j)} - \mathbb{E}_{n+1}^{\mathsf{J}} \widehat{h}_{n+1}\big). \tag{2.79e}$$

By making an empirical approximation of the formula (2.59), the matrix $\widetilde{K}_n$ is defined using the identification

$$\widetilde{K}_n = \widehat{V}_{n+1} \widehat{H}_{n+1}^{\top} \big[\big(\widehat{H}_{n+1} \widehat{H}_{n+1}^{\top} + \Gamma\big) + \big(\widehat{H}_{n+1} \widehat{H}_{n+1}^{\top} + \Gamma\big)^{1/2} \Gamma^{1/2} \big]^{-1}.$$

Remark 2.18. The key difference between (2.76) and (2.79) is that the former involves inversion in state space, and the latter in data space. The relative size of the two dimensions dictates which is preferable. □

Ensemble transform Kalman filters. The two EAKFs just defined both involve application, and inversion, of matrices which are applied on the left and act on state space. A different class of algorithms, known as *ensemble transform Kalman filters* (ETKF), involve matrix multiplication from the right, and consequently inversions take place in the ensemble space of dimension J. In many applications this is far smaller than the dimension of the state or data spaces, and then use of this version of the methodology is preferred. The aim is to determine matrix $Z_n \in \mathbb{R}^{J \times J}$, and to derive Kalman gain K_n from Z_n, so that the following interacting particle system produces an ensemble of particles $\{v_{n+1}^{(j)}\}_{j \in \mathsf{J}}$ with empirical covariance C_{n+1} defined by the second item in display (2.77):

$$\widehat{v}_{n+1}^{(j)} = \Psi\big(v_n^{(j)}\big) + \xi_n^{(j)},\ n \in \mathbb{Z}^+, \tag{2.80a}$$
$$\widehat{m}_{n+1} = \mathbb{E}_{n+1}^{\mathsf{J}} \widehat{v}_{n+1}, \tag{2.80b}$$
$$m_{n+1} = \widehat{m}_{n+1} + K_n\big(y_{n+1}^{\dagger} - \mathbb{E}_{n+1}^{\mathsf{J}} \widehat{h}_{n+1}\big), \tag{2.80c}$$
$$v_{n+1}^{(j)} = m_{n+1} + \sum_{i=1}^{J} \big(\widehat{v}_{n+1}^{(i)} - \widehat{m}_{n+1}\big)(Z_n)_{ij}. \tag{2.80d}$$

To this end we define the matrix of ensemble deviations

$$V_n = \frac{1}{\sqrt{J}} \big(v_n^{(1)} - \mathbb{E}_{*,n}^{\mathsf{J}} v_n, v_n^{(2)} - \mathbb{E}_{*,n}^{\mathsf{J}} v_n, \ldots, v_n^{(J)} - \mathbb{E}_{*,n}^{\mathsf{J}} v_n\big) \in \mathbb{R}^{d_v \times J},$$

where the expectation $\mathbb{E}_{*,n}^{\mathsf{J}}$ is with respect to the empirical measure (2.69) and v_n is a random variable with this distribution. It then follows from (2.80d) that

$$V_{n+1} = \widehat{V}_{n+1} Z_n. \tag{2.81}$$

If we define

$$Z_n = \left(I + \widehat{H}_{n+1}^{\top}\Gamma^{-1}\widehat{H}_{n+1}\right)^{-1/2} \in \mathbb{R}^{J\times J}, \tag{2.82}$$

then, as desired, $V_{n+1}V_{n+1}^{\top} = C_{n+1}$ as defined by (2.77), by virtue of (2.78). The calculations in (2.78) can also be utilized to verify that the empirical Kalman gain matrix defined by (2.75) satisfies

$$K_n = \widehat{V}_{n+1}Z_n^2\widehat{H}_{n+1}^{\top}\Gamma^{-1}.$$

Using this formula for K_n in (2.76) leads to an algorithm which matches first- and second-order statistics of the Gaussian projected filter, at $n+1$, requiring only matrix inversions in space of dimension defined by the number of particles J.

We finally note that (2.80c) and (2.80d) can be combined into a single transformation step of the form

$$v_{n+1}^{(j)} = \sum_{i=1}^{J} \widehat{v}_{n+1}^{(i)}(S_n)_{ij}, \tag{2.83}$$

where $S_n \in \mathbb{R}^{J\times J}$ replaces the matrix Z_n in (2.80d) such that

$$m_{n+1} = \sum_{j=1}^{J} v_{n+1}^{(j)} = \sum_{i,j=1}^{J} \widehat{v}_{n+1}^{(i)}(S_n)_{ij}$$

holds in addition to (2.81) with Z_n replaced by S_n.

Remark 2.19. Formulation (2.83) has a number of attractive features. First, it clearly reveals that the analysis $\{v_{n+1}^{(j)}\}_{j\in\mathsf{J}}$ lies in the span of the space spanned by the predictions $\{\widehat{v}_{n+1}^{(j)}\}_{j\in\mathsf{J}}$, which is relevant whenever $J < d_v$. Second, all particle implementations of the ensemble Kalman filter and many of its extensions can be put into the framework (2.83) with the (possibly random) matrix S_n chosen appropriately. Third, it encodes a coupling between the prediction $\{\widehat{v}_{n+1}^{(j)}\}_{j\in\mathsf{J}}$ and the analysis $\{v_{n+1}^{(j)}\}_{j\in\mathsf{J}}$ at the level of their associated empirical measures $\widehat{\mu}_n^{\mathsf{J}}$ and μ_n^{J}, respectively. See the following bibliographical section for more details. □

2.7. *Bibliographical notes*

Ideas from feedback control underlie the material in Section 2.2, addressing Objective 1. Control theory is an enormous subject in its own right and we cannot do justice to it in this paper. For the study of linear control theory, as illustrated in Example 2.2, see Åström and Murray (2021) and Sontag (2013) for engineering and mathematical treatments respectively. For the control-theoretic approach to the state estimation problem see Luenberger (1964, 1971).

Our study of control-theoretic methods has focused on 3DVAR. Recent analysis of the 3DVAR method rests heavily on ideas arising from determining modes for dissipative evolution equations, an idea with roots in the paper by Foias and Prodi

(1967) and unified in the book by Temam (2012). The use of these ideas in data assimilation was introduced in Olson and Titi (2003) and developed further in Hayden, Olson and Titi (2011) and Foias, Mondaini and Titi (2016). The papers of Law and Stuart (2012), Law, Shukla and Stuart (2014), Law, Sanz-Alonso, Shukla and Stuart (2016*a*) and Sanz-Alonso and Stuart (2015) essentially establish the stability of these deterministic results to small noise perturbations; see also Moodey, Lawless, Potthast and van Leeuwen (2013) for related analysis. In the context of using observations to control the instability of chaotic systems, all of this work may be seen as building on the study of synchronization by Pecora and Carroll (1990), reviewed in Ashwin (2003).

In Section 2.3 we introduce the probabilistic approach to filtering, addressing Objective 2. In low-dimensional systems, particle filters provide a flexible and efficient tool for attacking probabilistic filtering; see Doucet *et al.* (2001). However, in this paper our focus is on high-dimensional problems and Kalman-based methods specifically. The books by Reich and Cotter (2015), Asch, Bocquet and Nodet (2016), Law, Stuart and Zygalakis (2015), Harlim and Majda (2010), Abarbanel (2013) and Evensen, Vossepoel and van Leeuwen (2022) provide overviews of a variety of filtering methods, and ensemble Kalman methods from Section 2.6 in particular.

The Kalman filter (Kalman 1960) from Example 2.6 led to arguably the first systematic analysis of an algorithm for incorporation of discrete-time data into estimation of a discrete-time stochastic dynamical system; it applies only to linear Gaussian systems. The monograph by Jazwinski (2007) provides an introduction to nonlinear filtering in both discrete and continuous time; in particular, it discusses the extended Kalman filter (ExKF), found by applying the Kalman filter to a linearization of the state and data dynamics. However, the ExKF does not work well in high dimensions (Ghil *et al.* 1981), motivating the use of mean-field maps, as introduced in Section 2.5, and the ensemble-based methods from Section 2.6 which approximate them. The approximation of mean-field maps by interacting particle systems is reviewed in Sznitman (1991).

We refer the reader to the excellent monographs by Asch *et al.* (2016) and Evensen *et al.* (2022), which emphasize important implementation details not covered in this paper, relating to these ensemble Kalman methods; these include techniques such as inflation and localization that are central to the success of these methods in high dimensions. We also refer to Vetra-Carvalho *et al.* (2018) for a comprehensive review of the algorithmic details of ensemble Kalman methods. Here we point to two implementation details that are of particular practical importance. The first concerns the use of ensemble square root filters from Section 2.6.3. The matrix Z_n in (2.81) is not uniquely defined by the requirement $V_{n+1}V_{n+1}^\top = C_{n+1}$. Formula (2.82) constitutes one possible choice, which leads to a symmetric Z_n. See Livings, Dance and Nichols (2008) for more details. Second, finite particle implementations of the stochastic EnKF from Section 2.6.2 entail that the random realizations $\widehat{y}_{n+1}^{(j)}$ appear in the Kalman gain K_n as well as in the innovation term in (2.71c).

As first observed by Houtekamer and Mitchell (2005), this leads to a systematic underestimation of the ensemble spread, which vanishes in the $J \to \infty$ limit, but can affect the performance of the EnKF for small particle sizes. In Section 5.6 we will highlight the same effect when discussing finite particle implementations (Nüsken and Reich 2019, Garbuno-Inigo, Nüsken and Reich 2020*b*) of the ensemble Kalman sampler for Bayesian inversion, based on the mean-field model proposed in Garbuno-Inigo, Hoffmann, Li and Stuart (2020*a*).

Particle-based extensions of the classical Kalman filter to nonlinear filtering problems include the unscented Kalman filter and the ensemble Kalman filter. While this paper focuses primarily on ensemble Kalman filter techniques, the unscented Kalman filter is an approach based on application of quadrature to the Gaussian projected filter from Section 2.4; see Julier, Uhlmann and Durrant-Whyte (2000) as well as Särkkä and Svensson (2023). A discussion and evaluation in the context of ensemble square root filters may be found in Wang, Bishop and Julier (2004).

Much of the development of ensemble Kalman methods reflects the historical roots of the subject in the geophysical sciences, the atmosphere-ocean sciences in particular, including Lagrangian data assimilation, and in the modelling of subsurface flow (Burgers *et al.* 1998, Houtekamer and Mitchell 1998, Anderson 2001, Bishop *et al.* 2001, Whitaker and Hamill 2002, Tippett *et al.* 2003, Hunt *et al.* 2007, Li and Reynolds 2009, Sakov *et al.* 2012, Bocquet and Sakov 2014, Evensen 2019, Bocquet and Sakov 2012, Bocquet *et al.* 2017, Gurumoorthy *et al.* 2017, Sampson, Carrassi, Aydoğdu and Jones 2021, Kuznetsov, Ide and Jones 2003, Salman, Kuznetsov, Jones and Ide 2006). We also mention the randomized maximum likelihood (RML) approach to Bayesian inference, which is closely related to the analysis step of a stochastic EnKF and was also developed primarily through application in the geophysical sciences (Kitanidis 1995, Oliver, Cunha and Reynolds 1997, Oliver, Reynolds and Liu 2008).

There is also a body of literature concerning the analysis and development of ensemble methods with an emphasis on applications in complex and turbulent flows: Grooms, Lee and Majda (2014, 2015), Robinson, Grooms and Kleiber (2018), Lee, Majda and Qi (2017), Gottwald and Majda (2013), Kelly, Majda and Tong (2015), Tong, Majda and Kelly (2016*a*,*b*), Kelly *et al.* (2015), Harlim, Mahdi and Majda (2014), Majda and Tong (2018), Fertig, Harlim and Hunt (2007) and Harlim and Hunt (2007*a*,*b*). The conceptual fluid dynamics models of Lorenz (1996) (often referred to, collectively, as Lorenz '96 models) have been particularly influential in germinating this body of work, and we use them exclusively in our illustrative Examples 2.3, 2.16, 2.5, 4.23 and B.1. Furthermore, we will make use of the relationship between the multiscale and single-scale version of the model as developed in Fatkullin and Vanden-Eijnden (2004).

Recently ensemble Kalman methods have been developed for potential use in machine learning (Haber, Lucka and Ruthotto 2018, Kovachki and Stuart 2019, Guth, Schillings and Weissmann 2022, Grooms 2021, Gottwald and Reich 2021,

Yang and Grooms 2021, Pidstrigach and Reich 2023); see also Bocquet *et al.* (2017) and Chen, Sanz-Alonso and Willett (2022*b*) for research at the intersection of machine learning with ensemble Kalman methodology.

In this survey we started with mean-field equations, in Section 2.5, and then discretized the mean-field limit using J particles in subsequent subsections. It is of interest to demonstrate that the discrete formulations which arise actually converge to the mean-field equations in the $J \to \infty$ limit. This has indeed been established for the ensemble Kalman filter when applied in the linear Gaussian setting in which the mean-field limit exactly recovers the filtering distribution (Le Gland, Monbet and Tran 2011, Mandel, Cobb and Beezley 2011, Kwiatkowski and Mandel 2015) and indeed some of these results also apply in the nonlinear setting. In the continuous-time setting, long-time error estimates, exploiting ergodicity of the Kalman–Bucy filter itself and propagation of chaos ideas to extend to the ensemble Kalman approximations, are developed in Del Moral and Tugaut (2018); see Section 3.7 for further details concerning continuous time. Law, Tembine and Tempone (2016*b*) have studied related work concerning the mean-field limit of ensemble Kalman methods in the context of non-Gaussian problems. Ding, Li and Lu (2021) and Ding and Li (2021*a*,*b*) studied particle approximation of mean-field limits beyond the Gaussian setting, primarily in the context of the solution of inverse problems; see the discussion in Section 5.6. Hoel, Law and Tempone (2016) and Chernov *et al.* (2021) studied the use of multilevel approximation of the mean-field limit, coupling ensemble approximations at different levels of space or time discretization.

In Section 2.5.3 we introduce the idea of second-order transport: approximations of the perfect transport maps that effect filtering. The non-uniqueness of second-order transport maps is studied in continuous time, for linear Gaussian stochastic differential equations, in Taghvaei and Mehta (2020); this work is closely related to our analysis in the first two subsections of Appendix C. Non-Gaussian extensions are discussed in Taghvaei and Hosseini (2022). We also highlight that it is possible to construct second-order transport maps $\widetilde{T}$, which satisfy conditions different from (2.51b) and (2.57b), respectively. For example, one could request that

$$\mathcal{G}\big((\widetilde{T}_n^S)^\sharp \pi\big) = \mathcal{G}\mathcal{B}_n(\pi).$$

This approximation has been utilized by Lei and Bickel (2011). The analogous deterministic approach has been put forward by Tödter and Ahrens (2015) and has been explored further, for example, in Acevedo, de Wiljes and Reich (2017). Alternatively, one can also replace the definition (2.37) of the Gaussian projection operator G. To this end, recall that the Kullback–Leibler divergence between probability measures π_1 and π_2 on $\mathbb{R}^d$ is defined as

$$d_{\mathrm{KL}}(\pi_1||\pi_2) = \int \pi_1(du) \log \frac{\mathrm{d}\pi_1}{\mathrm{d}\pi_2}(u);$$

in particular, it is not symmetric in its two arguments. Gaussian variational inference (Bishop 2011) is, for example, based on the definition

$$\mathcal{G}\mu = \operatorname{argmin}_{\pi \in \mathfrak{G}} d_{\mathrm{KL}}(\pi||\mu) \tag{2.84}$$

in place of (2.37); note, however, that the minimizer of (2.84) may not be unique, whilst the minimizer of (2.37) is always unique.

While we follow the moment matching perspective on the derivation of ensemble Kalman filter methods in this survey, we mention in passing that there is an alternative perspective based on linear minimum variance estimators. See van Leeuwen (2020) in the context of the stochastic ensemble Kalman filter, and Appendix C.3 as well as Lei and Bickel (2011) for nonlinear extensions. The Bayes linear methodology is an alternative approach of interest (Goldstein and Wooff 2007).

Even in the mean-field limit $J \to \infty$, the ensemble Kalman filter provides approximations only to approximate transport-based filters. They are only exact in the linear Gaussian setting (Le Gland *et al.* 2011); recent works develop new tools of analysis to extend this to nonlinear filtering problems that are close to Gaussian (Carrillo, Hoffmann, Stuart and Vaes 2024, Calvello, Monmarché, Stuart and Vaes 2024). As mentioned in Section 1.1, sequential Monte Carlo methods can be designed to be consistent with the underlying nonlinear filtering problem, as defined, for example, by perfect transport-based filters. Foundational analysis of these particle methods is undertaken in Crisan, Del Moral and Lyons (1999) and Del Moral (2004); but we reiterate that, in contrast to ensemble Kalman-based methods, they do not scale well to high dimensions. We also point to Del Moral, Doucet and Jasra (2006) for an application of the sequential Monte Carlo method to Bayesian inference problems; the approach therein is closely related to iterative implementations of the EnKF that were subsequently developed in Li and Reynolds (2009), Gu and Oliver (2007) and Sakov *et al.* (2012).

Despite only providing approximations to the exact filtering distribution, i.e. from the perspective of Objective 2, accuracy and stability results for the ensemble Kalman filter, viewed as a state estimator and hence from the perspective of Objective 1, have been derived. See, for example, González-Tokman and Hunt (2013), Kelly, Law and Stuart (2014), Tong *et al.* (2016*a*,*b*) and Del Moral and Horton (2023). Mechanisms for finite-time filter divergence have also been identified (Gottwald and Majda 2013, Kelly *et al.* 2015).

Extending the ensemble Kalman filter to strongly nonlinear and high-dimensional state estimation problems constitute an area of active ongoing research. The current state of the art has been summarized in van Leeuwen *et al.* (2019) in the context of high-dimensional geophysical applications. Extensions of the transport framework (2.47), which build on approximating the perfect transport maps T^D in (2.47b) in an asymptotically consistent manner, include the work by Reich (2013), Cheng and Reich (2015), Spantini *et al.* (2022) and Zech and Marzouk (2022). In an alternative line of research there have been several proposals to construct

hybrid methods, which aim to adaptively bridge between the ensemble Kalman and particle filters, including the work by Stordal *et al.* (2011), Frei and Künsch (2013), Chustagulprom, Reich and Reinhardt (2016) and Nerger (2022).

In this context, the transformation formula (2.83) proves to be rather useful since most existing particle-based methods can be covered by appropriate choices of S_n, where S_n is typically the realization of a random matrix. See Reich and Cotter (2015) for more details. For example, a resampling step in a sequential Monte Carlo method gives rise to a matrix S_n with a single non-zero entry equal to one in each of its columns. More generally it holds that, for all $j \in \mathsf{J}$,

$$\sum_{i\in\mathsf{J}}(S_n)_{ij} = 1.$$

In particular, the matrix S_n can be chosen to correspond to an optimal coupling between two discrete random variables (Reich 2013), which builds a link between filtering and optimal transport also explored in Corenflos, Thornton, Deligiannidis and Doucet (2021). The subject of optimal transport is given a comprehensive treatment in Villani (2008); see also Villani (2021). Computational aspects of the subject including entropy-regularized optimal transport are discussed in Cuturi (2013) and Peyré and Cuturi (2019). Entropy-regularization is linked to the Schrödinger bridge problem, and connections with data assimilation are developed in Reich (2019). See also Section 1.1 for discussion of transport-based methodologies within the context of ensemble Kalman methods. See Example C.7 and Reich and Cotter (2015) for the connection between the map (2.67) and optimal transport. For a derivation of the standard formulae for mean and covariance of conditioned Gaussians used in the proof of Lemma 2.9, see Eaton (2007).

Finally we observe that we do not discuss, in this paper, the *smoothing* approach to state estimation from data in model (2.1). This approach aims at finding the entire sequence $\{v_\ell\}_{\ell=0}^n$ from the data $Y_n^\dagger$. Thus state estimates depend on data in their future. To read about smoothing, see Evensen *et al.* (2022) and Sanz-Alonso, Stuart and Taeb (2023). We note here that there is a smoothing counterpart of the 3DVAR algorithm (see Remark 2.1) known as 4DVAR because, for physical systems, it uses data distributed in the three space dimensions and one time dimension.

3. State estimation: continuous time

This section is devoted to deriving, and studying properties of, continuous-time analogues of concepts introduced in the previous Section 2. We start in Section 3.1 by defining the set-up. Thereafter the subsections mirror those from Section 2, describing the relevant continuous-time analogues; in particular, we conclude in Section 3.7 with bibliographical notes.

All problems arising in practice are implemented as algorithms in discrete time, so it is important to establish motivation for the continuous-time formulations.

There are two primary reasons for introducing them. The first is that continuous-time limits of the discrete algorithms provide a way to understand and interpret the behaviour of the discrete algorithms; results about accuracy, stability and uncertainty quantification, which shed light on the relative merits of different algorithmic approaches, are often cleanest in the continuous-time setting. The second is that many problems arising in practice involve physical processes which evolve in continuous time; the data informing these models is typically discrete in time, but when the observations take place at very high frequency, it is insightful to consider the idealization of continuous-time data. Both of these motivations underlie the developments in this section.

3.1. Set-up

We start by recalling the discrete-time set-up (2.1) for state-observation coevolution: for all $n \in \mathbb{Z}^+$ we have

$$\begin{aligned} v_{n+1} &= \Psi(v_n) + \xi_n, \\ y_{n+1} &= h(v_{n+1}) + \eta_{n+1}; \end{aligned}$$

we assume that $v_0, \{\xi_n\}_{n\in\mathbb{Z}^+}$ and $\{\eta_n\}_{n\in\mathbb{N}}$ are mutually independent Gaussians defined by

$$v_0 \sim \mathsf{N}(m_0, C_0), \quad \xi_n \sim \mathsf{N}(0, \Sigma) \text{ i.i.d.}, \quad \eta_n \sim \mathsf{N}(0, \Gamma) \text{ i.i.d.}$$

We introduce a small increment in time, denoted by Δt. From the map $\Psi(\cdot)$ defining the systematic component of the state dynamics, we now define an infinitesimal analogue $f(\cdot)$. We also introduce the rescaled observation operators $\mathsf{h}(\cdot)$ from the original nonlinear observation operator $h(\cdot)$, and we introduce state/observational covariances $(\mathsf{\Gamma}, \mathsf{\Sigma})$ by rescaling (Γ, Σ):

$$\Psi(v) = v + \Delta t f(v), \quad h(v) = \Delta t \mathsf{h}(v), \tag{3.1a}$$

$$\Sigma = \Delta t \mathsf{\Sigma}, \quad \Gamma = \Delta t \mathsf{\Gamma}. \tag{3.1b}$$

By virtue of our assumptions on Ψ and h, functions f and h are assumed to be known measurable functions (with respect to the Borel algebra), bounded on compact sets. In the linear setting $\Psi(\cdot) = M\cdot$, $h(\cdot) = H\cdot$ we will also introduce an infinitesimal vector field $f(\cdot) = F\cdot$ for matrix F, and rescaled linear observation operator H:[12]

$$M = I + \Delta t F, \quad H = \Delta t \mathsf{H}. \tag{3.2}$$

The observation $\{y_n\}$ is best thought of, in the scalings we introduce, as capturing increments of a process $\{z_n\}$. To capture this, and extend it to the specific realization of the data appearing in the algorithms, and the artificial data used in some

[12] Note the difference, conceptual and notational, between the discrete-time objects $(h(\cdot), H, \Gamma, \Sigma)$ and the related continuous-time objects $(\mathsf{h}(\cdot), \mathsf{H}, \mathsf{\Gamma}, \mathsf{\Sigma})$.

algorithms, we introduce the variables $z_n, z_n^\dagger, \widehat{z}_n$ by assuming that

$$y_{n+1} := z_{n+1} - z_n = \Delta z_{n+1}, \tag{3.3a}$$
$$y_{n+1}^\dagger := z_{n+1}^\dagger - z_n^\dagger = \Delta z_{n+1}^\dagger, \tag{3.3b}$$
$$\widehat{y}_{n+1} := \widehat{z}_{n+1} - \widehat{z}_n = \Delta \widehat{z}_{n+1}. \tag{3.3c}$$

Note that $z_n, z_n^\dagger, \widehat{z}_n$ have dimension $d_z = d_y$. We assume that $z_0 = z_0^\dagger = \widehat{z}_0 = 0$. Then $z_n, z_n^\dagger, \widehat{z}_n$ are uniquely defined from $y_n, y_n^\dagger, \widehat{y}_n$, respectively.

In the following we define $t_n = n\Delta t$. With the scalings above in hand, we may view the state v_n and observation y_n as relating to approximations of continuous-time processes $v(\cdot)$ and $z(\cdot)$: $v_n \approx v(t_n)$, $z_n \approx z(t_n)$. We also introduce continuous-time process $\widehat{v}(\cdot)$, which will be used in the prediction steps of algorithms, and $\big(z^\dagger(\cdot), \widehat{z}(\cdot)\big)$, which denotes the continuous-time observed data that we are conditioning on, and predicted data, respectively. We assume that $z(0) = z^\dagger(0) = \widehat{z}(0) = 0$. Under the rescalings above, and in the limit $\Delta t \to 0$, the data assimilation problem may be reformulated in terms of SDEs. Furthermore, the related mappings on measures, and discrete-time algorithms that stem from them, may be reformulated in terms of SPDEs and SDEs respectively; we now go on to identify these continuous-time stochastic processes.

Applying the rescalings in (3.1) and the reparametrization of y_{n+1} in (3.3a), we obtain the system

$$v_{n+1} = v_n + \Delta t\, f(v_n) + \xi_n, \tag{3.4a}$$
$$z_{n+1} = z_n + \Delta t \mathsf{h}(v_{n+1}) + \eta_{n+1}, \tag{3.4b}$$

for all $n \in \mathbb{Z}^+$, where we assume that $v_0, \{\xi_n\}_{n\in\mathbb{Z}^+}$ and $\{\eta_n\}_{n\in\mathbb{N}}$ are mutually independent Gaussians defined by

$$v_0 \sim \mathsf{N}(m_0, C_0), \quad \xi_n \sim \mathsf{N}(0, \Delta t\Sigma) \text{ i.i.d.}, \quad \eta_n \sim \mathsf{N}(0, \Delta t\Gamma) \text{ i.i.d.}$$

Note that (3.4) is a variant on the Euler–Maruyama discretization of a vector-valued SDE. Indeed, by taking the $\Delta t \to 0$ limit it is clear that the natural continuous-time analogue of equations (2.1) is the SDE

$$\boxed{\begin{aligned} \mathrm{d}v &= f(v)\,\mathrm{d}t + \sqrt{\Sigma}\,\mathrm{d}W, \quad v_0 \sim \mathsf{N}(m_0, C_0), \\ \mathrm{d}z &= \mathsf{h}(v)\,\mathrm{d}t + \sqrt{\Gamma}\,\mathrm{d}B, \quad z(0) = 0, \end{aligned}} \tag{3.5a, 3.5b}$$

taken to hold for all $t \in \mathbb{R}^+$. The vector fields $f(\cdot)$ and $\mathsf{h}(\cdot)$ describe the systematic, deterministic components of the dynamics and observation processes and are assumed to be known. The systematic components of the model are subjected to white noise defined through the independent standard Brownian motions W and B, in $\mathbb{R}^{d_v}$ and $\mathbb{R}^{d_z}$ respectively, and correlated across the state and data spaces via

the covariances Σ, Γ. The initial condition for v is Gaussian and independent of W and B. Analogous to the discrete-time setting, we assume that

$$C_0 \geq 0, \quad \Sigma \geq 0, \quad \Gamma > 0. \tag{3.6}$$

Note that, for each fixed $t \in \mathbb{R}^+$, the state $v(t) \in \mathbb{R}^{d_v}$ and the observations $z(t) \in \mathbb{R}^{d_z}$.

Throughout we use † again to denote a specific realization of a process, as in the discrete-time setting. We assume that we have available to us a sample path $\{z^\dagger(t)\}_{t\in\mathbb{R}^+}$ of the observation coordinates of a realization of the SDE (3.5). From this sample path we wish to recover the true realization of the state $\{v^\dagger(t)\}_{t\in\mathbb{R}^+}$ which gave rise to it. These observation and state sample paths are generated by $v_0^\dagger$, $\{W^\dagger\}_{t\in\mathbb{R}^+}$ and $\{B^\dagger\}_{t\in\mathbb{R}^+}$, specific realizations of the initial condition and the Brownian motions driving the state and observation components of (3.5). We also introduce $Z^\dagger(t) = \{z^\dagger(s)\}_{0\leq s\leq t}$.

Analogously to the discrete-time setting, it is natural to establish two distinct objectives, both related to recovery of the state from the observation.

Objective 1. Design an algorithm that produces output $v(t)$ from $Z^\dagger(t)$ so that $\{v(t)\}_{t\in\mathbb{R}^+}$ estimates $\{v^\dagger(t)\}_{t\in\mathbb{R}^+}$, the true signal generated by (3.5a).

Objective 2. Design an algorithm which estimates the distribution of random variable $v(t)|Z^\dagger(t)$, the conditional distribution defined by (3.5).

As in discrete time, we are interested in Markovian formulations that update the estimate $v(t)$, or the distribution $v(t)|Z^\dagger(t)$, sequentially as the data is acquired. All of the algorithms we describe depend only on the increments of the process $z^\dagger(t)$, hence the fixing of $z^\dagger(0) = 0$ is immaterial. In the next two subsections we describe control-theoretic and probabilistic approaches to this problem which, respectively, provide the basis for algorithms addressing Objectives 1 and 2. Following the road map from the previous section in the discrete-time setting, we then proceed to study exact transport leading to mean-field equations related to Objective 2; we then study second-order approximations of exact transport, and finally reach ensemble Kalman methods through particle approximations.

3.2. Control theory perspective

As for the time-discrete problem, we again start with the control-theoretic approach based on the small uncertainty assumption. Specifically we assume that the three covariances appearing in (3.6) are small so that the states and observations can be well approximated as deterministic. In this setting we derive a continuous-time analogue of the 3DVAR methodology.

Applying the rescalings (3.1) to (2.3) we obtain, in the deterministic setting,

$$\widehat{v}_{n+1} = v_n + \Delta t\, f(v_n), \tag{3.7a}$$

$$\widehat{z}_{n+1} = \widehat{z}_n + \Delta t \mathsf{h}(\widehat{v}_{n+1}), \tag{3.7b}$$

$$v_{n+1} = \widehat{v}_{n+1} + K\left(\Delta z_{n+1}^\dagger - \Delta \widehat{z}_{n+1}\right). \tag{3.7c}$$

The observed increments $\Delta z^\dagger_{n+1}$ are defined in (3.3b), and are derived from a specific fixed realization of (3.5). For fixed observed increments $\Delta z^\dagger_{n+1}$, equations (3.7) define a deterministic map $v_n \mapsto v_{n+1}$. Taking the continuous-time limit, and eliminating $\widehat{z}$, we obtain the following estimator v for $v^\dagger$ given $Z^\dagger$,

$$\boxed{\mathrm{d}v = f(v)\,\mathrm{d}t + K(\mathrm{d}z^\dagger - \mathsf{h}(v)\,\mathrm{d}t),} \tag{3.8}$$

where $z^\dagger$ (the data) is obtained from a specific fixed realization of (3.5):

$$\mathrm{d}v^\dagger = f(v^\dagger)\,\mathrm{d}t + \sqrt{\Sigma}\,\mathrm{d}W^\dagger, \quad v^\dagger(0) \sim \mathsf{N}(m_0, C_0), \tag{3.9a}$$

$$\mathrm{d}z^\dagger = \mathsf{h}(v^\dagger)\,\mathrm{d}t + \sqrt{\Gamma}\,\mathrm{d}B^\dagger, \quad z^\dagger(0) = 0. \tag{3.9b}$$

Equation (3.8) defines a continuous-time analogue of the 3DVAR algorithm (2.4) and *gain matrix* K should be viewed as a parameter to be chosen. The equation has the form of a controlled ordinary differential equation (ODE); typically it is initialized with $v(0) \sim \mathsf{N}(m_0, C_0)$.

We now include the effect of uncertainty, allowing for non-zero covariances in (3.6). Accounting for noise in the expressions for $\widehat{v}_{n+1}$ and $\widehat{z}_{n+1}$ in (3.7), we obtain the following rescaling of (2.15):

$$\begin{aligned}
\widehat{v}_{n+1} &= v_n + \Delta t f(v_n) + \xi_n, \\
\widehat{z}_{n+1} &= \widehat{z}_n + \Delta t \mathsf{h}(\widehat{v}_{n+1}) + \eta_{n+1}, \\
v_{n+1} &= \widehat{v}_{n+1} + K_n\left(\Delta z^\dagger_{n+1} - \Delta \widehat{z}_{n+1}\right),
\end{aligned}$$

for all $n \in \mathbb{Z}^+$, where we assume v_0, $\{\xi_n\}_{n\in\mathbb{Z}^+}$ and $\{\eta_n\}_{n\in\mathbb{N}}$ are mutually independent Gaussians defined by (3.1). We may now formally take the $\Delta t \to 0$ limit and obtain the following continuous-time analogue of (2.15), namely the controlled SDE formulation

$$\boxed{\mathrm{d}v = f(v)\,\mathrm{d}t + \sqrt{\Sigma}\,\mathrm{d}W + K(\mathrm{d}z^\dagger - \mathrm{d}\widehat{z}),} \tag{3.10a}$$

$$\boxed{\mathrm{d}\widehat{z} = \mathsf{h}(v)\,\mathrm{d}t + \sqrt{\Gamma}\,\mathrm{d}B.} \tag{3.10b}$$

where $z^\dagger$ is given by (3.9). The standard Brownian motions $W, W^\dagger, B$ and $B^\dagger$, in $\mathbb{R}^{d_v}$, $\mathbb{R}^{d_v}$, $\mathbb{R}^{d_z}$ and $\mathbb{R}^{d_z}$ respectively, are all independent of one another. As in the discrete-time analogue, encapsulated in (2.15), the choice of a (now *time-dependent*) gain matrix K remains to be determined and is crucial for the success of such a methodology; and as in discrete time, a time-evolving gain matrix is often desirable. To determine K we adopt a mean-field perspective, as we did in discrete time. To this end we now discuss the evolution of probability measures describing the conditional distribution of $v(t)|Z^\dagger(t)$.

3.3. *Probabilistic perspective*

We start, in Section 3.3.1, by discussing the unconditioned dynamics and introducing the Fokker–Planck equation associated with the state space evolution. In Section 3.3.2 we take the formulation of the filtering iteration in discrete time, from Section 2.3, and take a continuous-time limit to derive the Kushner–Stratonovich equation; we study the linear Gaussian setting, and the Kalman–Bucy filter, as a special case. In Section 3.3.3 we introduce the *sample-path perspective*, central to the algorithmic approach developed in this paper. Section 3.3.4 defines notation that will be useful throughout the remainder of this section on continuous-time data assimilation.

The derivation of continuous-time limits in the previous two subsections is relatively straightforward. However, there is an important practical and theoretical issue which we need to address. As mentioned at the start of Section 3, continuous-time observations $z^\dagger(t)$ are typically an idealization of discrete-time data collected at instances $\tau_k = k\delta$, $\delta > 0$, $k \in \mathbb{N}$ only.[13] In order to make use of continuous-time algorithms and theory it is then useful to construct a continuous-time approximation $z^{\dagger,\delta}$; to be concrete we will use piecewise linear interpolation. With this set-up we need to deal with two small parameters: the time-step Δt used in (3.1) to obtain a continuous-time limit, and the data sampling interval δ. The following remark addresses choices that we make in these notes about the manner in which we take the limit $(\Delta t, \delta) \to 0$.

Remark 3.1. There are results for continuous-time filtering which imply that the desired limiting equation can be found in either Itô or Stratonovich forms by considering different orders of the limits $\Delta t \to 0$ and $\delta \to 0$; see the bibliographical notes in Section 3.7. Many theoretical results are derived by first taking $\delta \to 0$ and then $\Delta t \to 0$. From a practical and theoretical perspective, however, it is sometimes more convenient to first consider the limit $\Delta t \to 0$ followed by the limit $\delta \to 0$. We will utilize the latter sequence of limits in the following subsection in order to derive a set of evolution equations for the conditional probability measure $\mu(v, t)$ solving Objective 2. These equations will in turn guide our choice of the gain matrix K in (3.10a). We emphasize that the choice about the order in which to take the limits $(\Delta t, \delta) \to 0$ is problem-dependent and should be considered carefully whenever continuous-time modelling is employed.

When the data $z^\dagger$ arises itself from numerical simulations of a continuous problem, then it is most convenient to set $\delta = \Delta t$, and hence $\tau_k = t_k$; this is implicitly used in the derivation of the continuous-time sample-path equations in the two preceding subsections. □

[13] Non-equally spaced data is also of relevance in this context, but we do not consider it here.

3.3.1. Unconditioned dynamics

First consider the continuous-time limit of the evolution associated with $\mathcal{P}$ from (2.18b), (2.18c) which, with the scaling adopted in this section, is defined by

$$(\mathcal{P}\mu)(\mathrm{d}v) = \left(\int_{u\in\mathbb{R}^{d_v}} p(u,v)\mu(\mathrm{d}u)\right)\mathrm{d}v, \tag{3.11a}$$

$$p(u,v) = \frac{1}{(2\pi\Delta t)^{d_v/2}\sqrt{\det(\Sigma)}}\exp\left(-\frac{1}{2\Delta t}|v-u-\Delta t f(u)|_\Sigma^2\right). \tag{3.11b}$$

Now view r_n given by (2.18a) as approximating $r(t_n)$. Recall that the underlying continuous-time limit of the sample-path evolution is given by the SDE (3.5a). Thus the time evolution of the probability density $r(\cdot,t)$ is given by the Fokker–Planck equation[14]

$$\partial_t r = -\nabla\cdot(rf) + \frac{1}{2}\nabla\cdot(\nabla\cdot(r\Sigma)). \tag{3.12}$$

Note that, as in discrete time, this evolution is linear and decoupled from the state space evolution (3.5a). We refer to the latter as a *continuous-time Markov process.*

3.3.2. The filtering distribution

As in discrete time, we now consider the evolution equation for the state conditioned on observations. Our starting point here is the iteration on measures, the filtering cycle, defined by (2.24), under the scalings (3.1). We assume that δ is an integer multiple of Δt so that the $\{\tau_k\}$ are a subset of the $\{t_n\}$. In what follows we will first fix δ and let $\Delta t \to 0$; in order to obtain the integer multiple property we thus consider $\Delta t \to 0$ along a subsequence. We replace the true observation path $z^\dagger(t)$ with its piecewise linear approximation $z^{\dagger,\delta}(t)$ based on linear interpolation of values $\{z^\dagger(\tau_k)\}$. To be precise we assume that the derivative is *càdlàg*.[15] We then have that the implied observation increments $\Delta z_{n+1}^\dagger$ are constant over the time intervals $[t_n, t_{n+1})$ and are given by

$$\Delta z_{n+1}^\dagger = \frac{\mathrm{d}z^{\dagger,\delta}}{\mathrm{d}t}(t_n)\Delta t. \tag{3.13}$$

In this setting, the operators $\mathcal{P}$ and $\mathcal{L}_n$, defined by (2.18) and (2.21), respectively, become

$$(\mathcal{P}\mu)(\mathrm{d}v) = \left(\int_{u\in\mathbb{R}^{d_v}} p(u,v)\mu(\mathrm{d}u)\right)\mathrm{d}v, \tag{3.14a}$$

$$\mathcal{L}_n(\mu)(\mathrm{d}v) = q(v,\Delta z_{n+1}^\dagger)\mu(\mathrm{d}v)\Big/\left(\int_{\mathbb{R}^{d_v}} q\left(v,\Delta z_{n+1}^\dagger\right)\mu(\mathrm{d}v)\right), \tag{3.14b}$$

[14] Here, and in what follows, we use the standard notation from continuum mechanics for the divergence of vector and second-order tensor fields, and for the gradient of scalar and vector fields; see Section 3.7 for references.

[15] Continuous from the right, limits exist from the left.

where, from (2.18) and (2.19) with the scalings (3.1),(3.3b),

$$p(u,v) = \frac{1}{(2\pi\Delta t)^{d_v/2}\sqrt{\det(\Sigma)}} \exp\left(-\frac{1}{2\Delta t}|v-u-\Delta t f(u)|_\Sigma^2\right), \tag{3.15a}$$

$$q(v,\Delta z) = \frac{1}{(2\pi\Delta t)^{d_y/2}\sqrt{\det(\Gamma)}} \exp\left(-\frac{1}{2\Delta t}|\Delta z-\Delta t \mathsf{h}(v)|_\Gamma^2\right). \tag{3.15b}$$

With these formulae in hand we may now derive the continuous-time analogue of (2.24), for $\mu(v,t)$.

For ease of exposition we assume that μ has density ρ and derive the equation satisfied by ρ. To do this we employ the *split-step principle*: we find the continuous-time evolution equation associated with each of $\mathcal{P}$ and $\mathcal{L}_n$ (equations (3.14a) and (3.14b) respectively) separately, and then add the right-hand sides of the resulting evolution equations to obtain the desired continuous-time limit resulting from the composition of $\mathcal{L}_n$ and $\mathcal{P}$. We use r as a dummy variable to denote the density being evolved, for both of the split-steps, and in both discrete (r_n) and continuous ($r(t)$) time, in what follows.

First recall that the continuous-time limit of the evolution associated with $\mathcal{P}$, as defined by (3.14a) and (3.15a), is given by the Fokker–Planck equation (3.12). Secondly, consider the second component of the split-step argument: we determine a continuous-time limit of the evolution associated with $\mathcal{L}_n$ as described by (3.14b) and (3.15b). The following lemma presents an evolution equation for r associated with $\mathcal{L}_n$, describing how observation of the piecewise continuous interpolated data $z^{\dagger,\delta}(t)$ changes the density $r(t,v)$.

Lemma 3.2. Assume that $\Gamma \succ 0$. The continuous-time limit of the evolution associated with $\mathcal{L}_n$, as described by (3.14b) and (3.15b), is given by

$$\partial_t r = \left\langle \mathsf{h} - \mathbb{E}\mathsf{h}, \frac{\mathrm{d}z^{\dagger,\delta}}{\mathrm{d}t}\right\rangle_\Gamma r - \frac{1}{2}\left\{|\mathsf{h}|_\Gamma^2 - \mathbb{E}|\mathsf{h}|_\Gamma^2\right\} r. \tag{3.16}$$

◇

Proof. Consider the discrete-time evolution

$$r_{n+1} = \mathcal{L}_n r_n, \tag{3.17}$$

where $\mathcal{L}_n$ is defined by (3.14b) and (3.15b). By Taylor expansion we have

$$\exp\left(-\frac{1}{2\Delta t}\left|\Delta z_{n+1}^\dagger - \Delta t \mathsf{h}(v)\right|_\Gamma^2\right) = 1 - \frac{\Delta t}{2}\left|\frac{\Delta z_{n+1}^\dagger}{\Delta t} - \mathsf{h}(v)\right|_\Gamma^2 + O(\Delta t^2).$$

Then, using expressions (3.14b) and (3.15b), we obtain

$$(\mathcal{L}_n r_n)(v) = \frac{1}{C(\Delta t)}\left(1 - \frac{\Delta t}{2}\left|\frac{\Delta z_{n+1}^\dagger}{\Delta t} - \mathsf{h}(v)\right|_\Gamma^2 + O(\Delta t^2)\right) r_n(v), \tag{3.18}$$

where

$$C(\Delta t) = \int_{\mathbb{R}^{d_v}} \left(1 - \frac{\Delta t}{2}\left|\frac{\Delta z_{n+1}^{\dagger}}{\Delta t} - \mathsf{h}(v)\right|_{\Gamma}^2 + O(\Delta t^2)\right) r_n(v)\,\mathrm{d}v.$$

By integrating and noting that r_n is a density, it follows that

$$C(\Delta t) = 1 - \frac{\Delta t}{2}\mathbb{E}\left|\frac{\Delta z_{n+1}^{\dagger}}{\Delta t} - \mathsf{h}(v)\right|_{\Gamma}^2 + O(\Delta t^2), \tag{3.19}$$

where expectation is with respect to v distributed as random variable with density r_n. Hence, combining (3.18) and (3.17), we obtain, to leading order in Δt,

$$r_{n+1} = r_n\left\{1 + \Delta t\left\langle \mathsf{h} - \mathbb{E}\mathsf{h}, \frac{\mathrm{d}z^{\dagger,\delta}}{\mathrm{d}t}\right\rangle_{\Gamma} - \frac{\Delta t}{2}\left\{|\mathsf{h}|_{\Gamma}^2 - \mathbb{E}|\mathsf{h}|_{\Gamma}^2\right\} + O(\Delta t^2)\right\};$$

here we have used that the data increments $\Delta z_{n+1}^{\dagger}$ are given by (3.13) for fixed δ. Taking the continuous-time limit $\Delta t \to 0$ with fixed observation interval $\delta > 0$ leads to the evolution equation (3.16). □

Now taking the $\delta \to 0$ limit in (3.16), we obtain the following non-local nonlinear stochastic evolution equation for density $r(v, t)$:

$$\mathrm{d}r = \langle \mathsf{h} - \mathbb{E}\mathsf{h}, \circ \mathrm{d}z^{\dagger}\rangle_{\Gamma} r - \frac{1}{2}\left\{|\mathsf{h}|_{\Gamma}^2 - \mathbb{E}|\mathsf{h}|_{\Gamma}^2\right\} r\,\mathrm{d}t. \tag{3.20}$$

Here $\circ$ denotes Stratonovich integration; this form of integration arises in the limit $\delta \to 0$ because the equation is derived by making a smooth approximation $z^{\dagger,\delta}$ of $z^{\dagger}$ and passing to the limit. Recall that $z^{\dagger}$ is given by (3.9). The equation is non-local and nonlinear because $\mathbb{E}$ denotes expectation at time t with respect to density $r(\cdot, t)$.

We now wish to invoke the split-step principle and combine the evolutions (3.12) and (3.20). However, before doing this we proceed to convert equation (3.20) into its more common Itô representation. For this, the following lemma is crucial. In proving it we will use the concepts of quadratic variation and covariation. For an introduction to these concepts consult the lecture notes by Eberle, which are referenced in Section 3.7. We note that quadratic variation and covariation are first defined between scalar-valued process and can then be lifted to define (i) the covariation of an inner product $\langle x, y\rangle$ between vector processes x and y, which is scalar-valued, (ii) the quadratic variation of vector process x, which is matrix-valued, and (iii) the covariation of vector process x with scalar process z, which is vector-valued.

Use of quadratic variation and covariation leads to a succinct, streamlined proof. Furthermore, in Appendix D we provide explicit calculations for the reader who is interested in understanding the details of the conversion by means of the definitions of Itô and Stratonovich integrals as limits. In particular, the concepts underlying the quadratic variation and covariation calculations in the following are derived from first principles in Lemma D.1.

Lemma 3.3. Assume that $\Gamma \succ 0$ and that $z^\dagger$ is given by (3.9). The Itô and Stratonovich interpretations of the stochastic forcing term in (3.20) are related through

$$\begin{aligned} \mathrm{d}r &= \langle \mathsf{h} - \mathbb{E}\mathsf{h}, \circ \mathrm{d}z^\dagger \rangle_\Gamma r - \frac{1}{2}\big\{|\mathsf{h}|_\Gamma^2 - \mathbb{E}|\mathsf{h}|_\Gamma^2\big\} r \,\mathrm{d}t \\ &= \langle \mathsf{h} - \mathbb{E}\mathsf{h}, \mathrm{d}z^\dagger - \mathbb{E}\mathsf{h}\,\mathrm{d}t \rangle_\Gamma r. \end{aligned}$$ ◇

Proof. Using the formula for the Itô–Stratonovich conversion between two semi-martingales, we have

$$\langle \mathsf{h} - \mathbb{E}\mathsf{h}, \circ \mathrm{d}z^\dagger \rangle_\Gamma r = \langle \mathsf{h} - \mathbb{E}\mathsf{h}, \mathrm{d}z^\dagger \rangle_\Gamma r + \frac{1}{2}\,\mathrm{d}\big[\langle (\mathsf{h} - \mathbb{E}\mathsf{h})r, z^\dagger \rangle_\Gamma\big], \tag{3.21}$$

where $[\langle \cdot, \cdot \rangle]$ denotes covariation of the inner product $\langle \cdot, \cdot \rangle$. Note that (3.9b) implies that the quadratic variation of $z^\dagger$ is given by the matrix identity $[z^\dagger, z^\dagger] = t\Gamma$. Furthermore, because of (3.20), the conversion formula (3.21) and the fact that $\mathrm{d}[z^\dagger, z^\dagger] = \Gamma\,\mathrm{d}t$, it follows that the covariation between scalar r and vector $z^\dagger$ satisfies the vector identity

$$\mathrm{d}[r, z^\dagger] = (\mathsf{h} - \mathbb{E}\mathsf{h}) r \,\mathrm{d}t. \tag{3.22}$$

Consider $\psi : \mathbb{R} \to \mathbb{R}^{d_y}$ so that derivative $\psi'(r)$ may be identified with an element in $\mathbb{R}^{d_y}$. Then, for any such differentiable ψ,

$$\mathrm{d}[\langle \psi(r), z^\dagger \rangle_\Gamma] = \langle \psi'(r), \mathrm{d}[r, z^\dagger] \rangle_\Gamma. \tag{3.23}$$

We now apply this identity in the setting where $\psi(r) = (\mathsf{h} - \mathbb{E}\mathsf{h})r$, noting that r defines the expectation in this definition. Thus

$$\psi'(r)\,\delta r = (\mathsf{h} - \mathbb{E}\mathsf{h})\,\delta r - \left(\int \mathsf{h}\,\delta r \,\mathrm{d}v\right) r.$$

Hence the covariation in (3.21) satisfies, using (3.23) for the first equality and (3.22) for the second equality,

$$\begin{aligned} \mathrm{d}[\langle (\mathsf{h} - \mathbb{E}\mathsf{h})r, z^\dagger \rangle_\Gamma] &= \langle \mathsf{h} - \mathbb{E}\mathsf{h}, \mathrm{d}[r, z^\dagger] \rangle_\Gamma - \left(\int \langle \mathsf{h}, \mathrm{d}[r, z^\dagger] \rangle_\Gamma \,\mathrm{d}v\right) r \\ &= |\mathsf{h} - \mathbb{E}\mathsf{h}|_\Gamma^2 r \,\mathrm{d}t - \mathbb{E}\big(\langle \mathsf{h}, \mathsf{h} - \mathbb{E}\mathsf{h} \rangle_\Gamma\big) r \,\mathrm{d}t \\ &= \big\{|\mathsf{h} - \mathbb{E}\mathsf{h}|_\Gamma^2 - \mathbb{E}|\mathsf{h} - \mathbb{E}\mathsf{h}|_\Gamma^2\big\} r \,\mathrm{d}t. \end{aligned}$$

Using this identity in (3.21) and rearranging, we find that

$$\langle \mathsf{h} - \mathbb{E}\mathsf{h}, \circ \mathrm{d}z^\dagger \rangle_\Gamma r - \frac{1}{2}\big\{|\mathsf{h}|_\Gamma^2 - \mathbb{E}|\mathsf{h}|_\Gamma^2\big\} r \,\mathrm{d}t = \langle \mathsf{h} - \mathbb{E}\mathsf{h}, \mathrm{d}z^\dagger - \mathbb{E}\mathsf{h}\,\mathrm{d}t \rangle_\Gamma r,$$

which in turn leads to the following desired Itô representation of (3.20):

$$\mathrm{d}r = \langle \mathsf{h} - \mathbb{E}\mathsf{h}, \mathrm{d}z^\dagger - \mathbb{E}\mathsf{h}\,\mathrm{d}t \rangle_\Gamma r. \tag{3.24}$$

□

Using the Itô form of the equation from the preceding lemma, the Fokker–Planck equation (3.12) and the split-step principle deliver the following.

Theorem 3.4. Assume that $\Gamma > 0$ and that $z^\dagger$ is given by (3.9). The time evolution of the density $\rho(\cdot, t)$ for the random variable $v(t)|Z^\dagger(t)$ is characterized by the Itô SPDE

$$\boxed{\mathrm{d}\rho = -\nabla \cdot (\rho f)\,\mathrm{d}t + \frac{1}{2}\nabla \cdot (\nabla \cdot (\rho \Sigma))\,\mathrm{d}t + \langle \mathsf{h} - \mathbb{E}\mathsf{h}, \mathrm{d}z^\dagger - \mathbb{E}\mathsf{h}\,\mathrm{d}t\rangle_\Gamma \rho.} \tag{3.25}$$

◇

Equation (3.25) is known as the Kushner–Stratonovich equation. Equation (3.25) is to be interpreted in the Itô sense with respect to the driving noise $z^\dagger(t)$. The equation is nonlinear, unlike the unconditioned dynamics governed by the Fokker–Planck equation (3.12). The corresponding Stratonovich formulation follows immediately from (3.20) in combination with (3.12):

$$\mathrm{d}\rho = \left(-\nabla \cdot (\rho f) + \frac{1}{2}\nabla \cdot (\nabla \cdot (\rho \Sigma)) - \frac{1}{2}\left\{|\mathsf{h}|_\Gamma^2 - \mathbb{E}|\mathsf{h}|_\Gamma^2\right\}\rho\right)\mathrm{d}t \tag{3.26a}$$

$$+ \langle \mathsf{h} - \mathbb{E}\mathsf{h}, \circ \mathrm{d}z^\dagger\rangle_\Gamma \rho. \tag{3.26b}$$

Example 3.5. We consider the setting where

$$f(\cdot) = F\cdot, \quad \mathsf{h}(\cdot) = \mathsf{H}\cdot \tag{3.27}$$

in equations (3.5) so that

$$\boxed{\begin{aligned} \mathrm{d}v &= Fv\,\mathrm{d}t + \sqrt{\Sigma}\,\mathrm{d}W, \quad v(0) \sim \mathsf{N}(m_0, C_0), && \text{(3.28a)}\\ \mathrm{d}z &= \mathsf{H}v\,\mathrm{d}t + \sqrt{\Gamma}\,\mathrm{d}B, \quad z(0) = 0. && \text{(3.28b)} \end{aligned}}$$

Now consider data $z^\dagger(t)$ generated by

$$\mathrm{d}v^\dagger = Fv^\dagger\,\mathrm{d}t + \sqrt{\Sigma}\,\mathrm{d}W^\dagger, \quad v^\dagger(0) \sim \mathsf{N}(m_0, C_0), \tag{3.29a}$$

$$\mathrm{d}z^\dagger = \mathsf{H}v^\dagger\,\mathrm{d}t + \sqrt{\Gamma}\,\mathrm{d}B^\dagger, \quad z^\dagger(0) = 0. \tag{3.29b}$$

We are interested in the filtering distribution for $v(t)|Z^\dagger(t)$.

Because of the linearity and additive Gaussian noise, this filtering distribution is Gaussian and is given by the Kalman–Bucy filter, the continuous-time analogue of the Kalman filter described in Example 2.6. Indeed, if $\Gamma > 0$ in (3.28), then the density $\rho(\cdot, t)$ associated with the random variable $v(t)|Z^\dagger(t)$, with $z^\dagger$ given by (3.29), is Gaussian with mean $m(\cdot)$ and covariance $C(\cdot)$ given by

$$\boxed{\begin{aligned} \mathrm{d}m &= Fm\,\mathrm{d}t + C\mathsf{H}^\top\Gamma^{-1}(\mathrm{d}z^\dagger - \mathsf{H}m\,\mathrm{d}t), \quad m(0) = m_0, && \text{(3.30a)}\\ \mathrm{d}C &= FC\,\mathrm{d}t + CF^\top\,\mathrm{d}t + \Sigma\,\mathrm{d}t - C\mathsf{H}^\top\Gamma^{-1}\mathsf{H}C\,\mathrm{d}t, \; C(0) = C_0. && \text{(3.30b)} \end{aligned}}$$

These equations for the mean and covariance of the Kalman–Bucy filter may be obtained by taking the continuous-time limit of the Kalman filter from Example 2.6 under the scalings (3.1b), (3.2) and (3.3b).

The Kalman–Bucy filter also yields an exact solution of the Kushner–Stratonovich equation (3.25) under (3.27). Indeed, if $\rho(\cdot, 0)$ is initialized at the Gaussian with mean m_0 and covariance C_0 then $\rho(\cdot, t)$ solving the Kushner–Stratonovich equation (3.25) has solution given by the Gaussian $\mathsf{N}\big(m(t), C(t)\big)$. □

3.3.3. *The sample-path perspective*

Similarly to the discrete-time setting, a key idea in developing algorithms for filtering in continuous time is the *sample-path perspective.* We will seek to identify (for implementable algorithms only approximately) mean-field SDEs with the property that their solution $v(t)$, has $Law(v(t))$ equal to that given by the density r governed by the Kushner–Stratonovich equation (3.25). Analogously to discrete time, the mean-field SDEs will couple to the solution of the equation for evolution of the density; the resulting processes are termed *nonlinear Markov processes.* We will first introduce Gaussian projected filters and then discuss mean-field models derived through the sample-path perspective. To these ends we now introduce some important notational conventions.

3.3.4. *Important frequently used notation*

The notation we define here is used primarily to discuss exact (and approximate) mean-field models, here evolving in continuous time, to (approximately) solve the filtering problem. In discrete time, expectations may be taken under the law of the (possibly approximate) discretely evolving mean-field model, or under the predictive distribution found from pushing this law forward under the model. This distinction disappears in continuous time, and we simply need to compute expectations under the (possibly approximate) continuously evolving mean-field models that we will introduce later. To this end we define, with $m := \mathbb{E}v$,

$$C := \mathbb{E}((v - m) \otimes (v - m)), \tag{3.31a}$$

$$C^{vf} := \mathbb{E}((v - m) \otimes (f(v) - \mathbb{E}f(v))), \tag{3.31b}$$

$$C^{v\mathsf{h}} := \mathbb{E}((v - m) \otimes (\mathsf{h}(v) - \mathbb{E}\mathsf{h}(v))). \tag{3.31c}$$

All expectations are under the mean-field model for v. The covariances should be viewed as functions of time t. In deriving continuous-time models from discrete-time models, we will also use discrete-time analogues, computed under the law of random variable v_n and denoted by C_n, C_n^{vf} and $C_n^{v\mathsf{h}}$.

3.4. *Gaussian projected filtering distribution*

As in discrete time, the Gaussian projected filtering distribution plays an important conceptual role in understanding later filtering algorithms. The evolution equations

for the mean m and the covariance matrix C follow naturally from a continuous-time limit of the associated discrete-time filtering formulations. We summarize the resulting equations in the following.

Theorem 3.6. Assume that $\Gamma \succ 0$, and that $z^\dagger$ is given by (3.9). Consider the discrete-time Gaussian projected filter, namely the map from (m_n, C_n) to (m_{n+1}, C_{n+1}) defined by choosing $v_n \sim \mathsf{N}(m_n, C_n)$ and then using (2.29), (2.32), (2.35) and (2.42). Under the rescalings (3.1), and in the limit $\Delta t \to 0$, we obtain the following continuous-time limit of this map:

$$\boxed{\begin{aligned} \mathrm{d}m &= \mathbb{E} f(v)\,\mathrm{d}t + C^{v\mathsf{h}}\Gamma^{-1}(\mathrm{d}z^\dagger - \mathbb{E}\mathsf{h}(v)\,\mathrm{d}t), && \text{(3.32a)}\\ \mathrm{d}C &= C^{vf}\,\mathrm{d}t + (C^{vf})^\top\,\mathrm{d}t + \Sigma\,\mathrm{d}t - C^{v\mathsf{h}}\Gamma^{-1}(C^{v\mathsf{h}})^\top\,\mathrm{d}t, && \text{(3.32b)}\end{aligned}}$$

where the expectations and covariances are computed under $\mathsf{N}\big(m(t), C(t)\big)$, using (3.31). ◇

Remark 3.7. Note that the preceding equation implicitly defines gain matrix

$$K = C^{v\mathsf{h}}\Gamma^{-1}. \tag{3.33}$$

This specific choice of gain will play a central role in what follows. □

Proof of Theorem 3.6. The Gaussian projected filter is defined by evolution of mean and covariance given by equations (2.32), (2.35) and (2.42), repeated and reordered here for convenience:

$$\begin{aligned}
\widehat{m}_{n+1} &= \mathbb{E}\Psi(v_n),\\
\widehat{C}_{n+1} &= \mathbb{E}((\Psi(v_n) - \widehat{m}_{n+1}) \otimes (\Psi(v_n) - \widehat{m}_{n+1})) + \Sigma,\\
m_{n+1} &= \widehat{m}_{n+1} + \widehat{C}^{vh}_{n+1}\big(\widehat{C}^{hh}_{n+1} + \Gamma\big)^{-1}\big(y^\dagger_{n+1} - \mathbb{E}\widehat{h}_{n+1}\big),\\
C_{n+1} &= \widehat{C}_{n+1} - \widehat{C}^{vh}_{n+1}\big(\widehat{C}^{hh}_{n+1} + \Gamma\big)^{-1}\big(\widehat{C}^{vh}_{n+1}\big)^\top,\\
\widehat{C}^{vh}_{n+1} &= \mathbb{E}((\widehat{v}_{n+1} - \mathbb{E}\widehat{v}_{n+1}) \otimes (\widehat{h}_{n+1} - \mathbb{E}\widehat{h}_{n+1})),\\
\widehat{C}^{hh}_{n+1} &= \mathbb{E}((\widehat{h}_{n+1} - \mathbb{E}\widehat{h}_{n+1}) \otimes (\widehat{h}_{n+1} - \mathbb{E}\widehat{h}_{n+1})).
\end{aligned}$$

Recall that $\widehat{h}_{n+1} := h(\widehat{v}_{n+1})$. Expectations in the prediction step are with respect to the law of $v_n \sim \mathsf{N}(m_n, C_n)$, and expectations in the analysis step are with respect to the law of $\widehat{v}_{n+1}$ given by (2.29a), assuming that $v_n \sim \mathsf{N}(m_n, C_n)$; thus $Law(\widehat{v}_{n+1}) = \mathcal{P}\, Law(v_n)$.

We now impose the rescalings (3.1) on these equations. The reader's attention is drawn to the fact that $h = \Delta t\mathsf{h}$ when reading the next formula in order to understand the scalings with Δt; a similar notational shift from regular font to mathsf occurs in other formulae that follow it, and is crucial to understanding orders of magnitude with respect to Δt. Under the rescalings and using that $\xi_{n+1} = O(\Delta t^{1/2})$ and has

mean zero,

$$\begin{aligned}\widehat{C}^{vh}_{n+1} &= \Delta t\mathbb{E}((v_n - \mathbb{E}v_n + \xi_n + O(\Delta t)) \otimes (\mathsf{h}(v_n) - \mathbb{E}\mathsf{h}(v_n) + D\mathsf{h}(v_n)\xi_n + O(\Delta t)))\\ &= \Delta t\mathbb{E}((v_n - \mathbb{E}v_n) \otimes (\mathsf{h}(v_n) - \mathbb{E}\mathsf{h}(v_n))) + O(\Delta t^2)\\ &= \Delta t C^{v\mathsf{h}}_n + O(\Delta t^2).\end{aligned}$$

Similarly

$$\begin{aligned}\widehat{C}^{hh}_{n+1} &= \Delta t^2\mathbb{E}\big((\mathsf{h}(v_n) - \mathbb{E}\mathsf{h}(v_n) + O(\Delta t^{1/2})) \otimes (\mathsf{h}(v_n) - \mathbb{E}\mathsf{h}(v_n) + O(\Delta t^{1/2}))\big)\\ &= O(\Delta t^2).\end{aligned}$$

Thus, since $\Gamma = \Delta t\Gamma$,

$$\widehat{C}^{vh}_{n+1}\big(\widehat{C}^{hh}_{n+1} + \Gamma\big)^{-1} = C^{v\mathsf{h}}_n\Gamma^{-1} + O(\Delta t), \tag{3.34a}$$

$$\widehat{C}^{vh}_{n+1}\big(\widehat{C}^{hh}_{n+1} + \Gamma\big)^{-1}\widehat{C}^{vh}_{n+1} = \Delta t C^{v\mathsf{h}}_n\Gamma^{-1}\big(C^{v\mathsf{h}}_n\big)^\top + O(\Delta t^2). \tag{3.34b}$$

Furthermore, since $\Psi(v_n) = v_n + \Delta t f(v_n)$ and $\widehat{m}_{n+1} = \mathbb{E}v_n + \Delta t\mathbb{E}f(v_n)$,

$$\mathbb{E}((\Psi(v_n) - \widehat{m}_{n+1}) \otimes (\Psi(v_n) - \widehat{m}_{n+1})) = C_n + \Delta t C^{vf}_n + \Delta t\big(C^{vf}_n\big)^\top + O(\Delta t^2).$$

Using these approximations, and the fact that $\Delta z^\dagger_{n+1} = O(\Delta t^{1/2})$, we obtain

$$\begin{aligned}\widehat{m}_{n+1} &= m_n + \Delta t\mathbb{E}f(v_n),\\ \widehat{C}_{n+1} &= C_n + \Delta t C^{vf}_n + \Delta t\big(C^{vf}_n\big)^\top + \Delta t\Sigma + O(\Delta t^2),\\ m_{n+1} &= \widehat{m}_{n+1} + C^{v\mathsf{h}}_n\Gamma^{-1}\big(\Delta z^\dagger_{n+1} - \Delta t\mathbb{E}\mathsf{h}(\widehat{v}_{n+1})\big) + O(\Delta t^{3/2}),\\ C_{n+1} &= \widehat{C}_{n+1} - \Delta t C^{v\mathsf{h}}_n\Gamma^{-1}\big(C^{v\mathsf{h}}_n\big)^\top + O(\Delta t^2),\\ C^{vf}_n &= \mathbb{E}((v_n - m_n) \otimes (f(v_n) - \mathbb{E}f(v_n))),\\ C^{v\mathsf{h}}_n &= \mathbb{E}((v_n - m_n) \otimes (\mathsf{h}(v_n) - \mathbb{E}\mathsf{h}(v_n))).\end{aligned}$$

In the following, $C^{v\mathsf{h}}, C^{vf}$ are functions of time, defined by (3.31), and $C^{v\mathsf{h}}_n, C^{vf}_n$ are discrete-time analogues computed under the law of v_n.

Combining the prediction and analysis steps, we find that

$$\begin{aligned}m_{n+1} &= m_n + \Delta t\mathbb{E}f(v_n) + C^{v\mathsf{h}}_n\Gamma^{-1}\big(\Delta z^\dagger_{n+1} - \Delta t\mathbb{E}\mathsf{h}(\widehat{v}_{n+1})\big) + O(\Delta t^{3/2}),\\ C_{n+1} &= C_n + \Delta t C^{vf}_n + \Delta t\big(C^{vf}_n\big)^\top + \Delta t\Sigma - \Delta t C^{v\mathsf{h}}_n\Gamma^{-1}\big(C^{v\mathsf{h}}_n\big)^\top + O(\Delta t^2).\end{aligned}$$

Taking the $\Delta t \to 0$ limit, we deduce the continuous-time analogue of equations (2.32) and (2.35), namely (3.32) as desired. □

Example 3.8. In the linear setting (3.27), the Gaussian projected filter (3.32) reduces to the Kalman–Bucy filter from Example 3.5. □

3.5. Mean-field evolution equations

In the preceding subsection we approximated the evolution of the filtering distribution by the evolution of a Gaussian. Here we take a different approach,

seeking a sample-path perspective, identifying mean-field SDEs which (possibly only approximately) have solutions with law given by the filtering distribution. We develop a continuous-time analogue of the discrete time mean-field approach from Section 2.5.

In Section 3.5.1 we describe work concerning the derivation of explicit mean-field SDEs equal in law to the filtering distribution; this is a departure from our discussion of this topic in discrete time where no explicit maps were identified in the general setting. Sections 3.5.2 and 3.5.3 describe a variety of explicit approximate mean-field SDEs, based on matching first- and second-order moment information, and arising from rescaling of the discrete-time setting.

3.5.1. Perfect transport

Here we seek a mean-field dynamical system with law given by that of the Kushner–Stratonovich equation. The analogue of the discrete-time transport map T_n^S, is to find transport evolution equations for v and $\widehat{z}$ in the form of a mean-field SDE. We will seek to achieve this in the specific sample-path form

$$\mathrm{d}v = f(v)\,\mathrm{d}t + \sqrt{\Sigma}\,\mathrm{d}W + a(v;\rho)\,\mathrm{d}t + K(v;\rho)(\mathrm{d}z^\dagger - \mathrm{d}\widehat{z}), \tag{3.35a}$$

$$\mathrm{d}\widehat{z} = \mathsf{h}(v)\,\mathrm{d}t + \sqrt{\Gamma}\,\mathrm{d}B, \tag{3.35b}$$

where ρ is the time-dependent density of v. Note that (3.35b) simply generates simulated data from the state v. On the other hand, (3.35a) combines the underlying known model evolution with a nudged innovation term, based on the difference between the observed and simulated data, and a correction to the drift f. The nudging and correction terms are defined by gain K and drift correction a. The goal is to choose drift correction a and gain K such that the induced time evolution of the density ρ of v agrees with the density ρ given by the Kushner–Stratonovich equation (3.25). We thus find a controlled SDE, similar in form to that proposed in (3.10) but with an additional drift term and with mean-field dependence.

Remark 3.9. We emphasize that, as in the discrete-time transport formulation, there is a considerable degree of non-uniqueness in the choice of transport in general, and here in the choice of a and K specifically. We make specific, simple, choices in the theorem that follows. Working in continuous time enables very explicit identification of exact mean-field models; this is not possible in discrete time. □

The following theorem identifies a sample-path perspective that exactly captures evolution of the filtering distribution.

Theorem 3.10. Assume that $\Gamma \succ 0$ and that there exists $K = K(v;\rho)$ satisfying the identity

$$-\nabla\cdot(\rho K^\top) = \Gamma^{-1}(\mathsf{h} - \mathbb{E}\mathsf{h})\rho. \tag{3.36}$$

Consider the stochastic mean-field dynamics given by

$$\boxed{\begin{aligned} \mathrm{d}v &= f(v)\,\mathrm{d}t + \sqrt{\Sigma}\,\mathrm{d}W + \nabla\cdot(K\Gamma K^\top)\,\mathrm{d}t - K\Gamma\nabla\cdot K^\top\,\mathrm{d}t + K(\mathrm{d}z^\dagger - \mathrm{d}\widehat{z}), \quad (3.37a)\\ \mathrm{d}\widehat{z} &= \mathsf{h}(v)\,\mathrm{d}t + \sqrt{\Gamma}\,\mathrm{d}B, \quad (3.37b)\end{aligned}}$$

where $z^\dagger$ is given by (3.9). Assume that solution v has law with smooth density and that the Kushner–Stratonovich equation (3.25) has smooth density as solution. Then the law of v has density given by the Kushner–Stratonovich equation (3.25). ◇

Proof. To simplify calculations we will first look at choosing a and K to get agreement, at the level of densities of v, with (3.24). A straightforward modification, using the split-step principle again, then provides the generalization to (3.25). This split-step approach enables us to consider only the case where $f(\cdot) \equiv 0$ and $\Sigma = 0$.

Note that, in (3.35), $z^\dagger$ is a fixed, given, trajectory and we are interested in the evolution of the probability density induced for v by the randomness over the distribution on trajectories $\widehat{z}$. Using (3.35b) in (3.35a), and recalling that it suffices to set f and Σ to zero, we obtain

$$\mathrm{d}v = a(v;\rho)\,\mathrm{d}t + K(v;\rho)(\mathrm{d}z^\dagger - \mathsf{h}(v)\,\mathrm{d}t - \sqrt{\Gamma}\,\mathrm{d}B). \tag{3.38}$$

Applying Fokker–Planck analysis, modified to the mean-field setting, shows that the time evolution of the density ρ of v under (3.35) is provided by the nonlinear (because of dependence of a, K on ρ) SPDE[16] to be interpreted in the Itô sense:

$$\mathrm{d}\rho = -\nabla\cdot(\rho(a - K\mathsf{h}))\,\mathrm{d}t - \langle\nabla\cdot(\rho K^\top), \mathrm{d}z^\dagger\rangle + \nabla\cdot(\nabla\cdot(\rho K\Gamma K^\top))\,\mathrm{d}t. \tag{3.39}$$

Note that, although $z^\dagger$ is a fixed trajectory, it contributes to the diffusion term in this Fokker–Planck equation, explaining the factor 1 rather than $1/2$. This arises as a contribution from the quadratic variation of the path of $z^\dagger$ to the evolution of ρ. For further details on the derivation of (3.39), see Lemma D.2 in Appendix D. To get agreement with (3.24), a and K have to be chosen such that

$$\begin{aligned} &\rho\langle(\mathsf{h} - \mathbb{E}\mathsf{h}), \Gamma^{-1}(\mathrm{d}z^\dagger - \mathbb{E}\mathsf{h}\,\mathrm{d}t)\rangle \\ &\quad = -\nabla\cdot(\rho(a - K\mathsf{h}))\,\mathrm{d}t - \langle\nabla\cdot(\rho K^\top), \mathrm{d}z^\dagger\rangle + \nabla\cdot(\nabla\cdot(\rho K\Gamma K^\top))\,\mathrm{d}t. \end{aligned}$$

Equating the two terms involving the data $z^\dagger$ shows immediately that $K(v;\rho)$ has to satisfy the (vector-valued) PDE (3.36). From this, equating the terms that do not

[16] Here we use the standard convention from continuum mechanics that the divergence of a matrix is to be interpreted via computation of derivatives with respect to the second index; see Section 3.7 for references to relevant continuum mechanics textbooks. Strictly speaking, equation (3.39) is an SPDE only if we now view $z^\dagger$ as a random variable rather than a fixed realization.

involve the data $z^\dagger$, it follows that $a(v;\rho)$ has to satisfy

$$\begin{aligned}\nabla\cdot(\rho a) &= \nabla\cdot(\nabla\cdot(\rho K\Gamma K^\top)) + \nabla\cdot(\rho K(\mathsf{h}-\mathbb{E}\mathsf{h}))\\ &= \nabla\cdot(\nabla\cdot(\rho K\Gamma K^\top)) - \nabla\cdot(K\Gamma\nabla\cdot(\rho K^\top))\\ &= \nabla\cdot(\rho(\nabla\cdot(K\Gamma K^\top) - K\Gamma\nabla\cdot K^\top)).\end{aligned}$$

A natural choice for a is provided by asking that the term on which the divergence acts is zero. This yields

$$a = \nabla\cdot(K\Gamma K^\top) - K\Gamma\nabla\cdot K^\top, \tag{3.40}$$

a solution for a with no explicit dependence on ρ; note, however, that a does depend on ρ implicitly through the dependence of K on ρ.

With these choices of a, K, and applying the split-step principle so that the mean-field dynamics are consistent with (3.25) rather than (3.24), we obtain a version of the *feedback particle filter*; in particular, we find the equation in its stochastic mean-field formulation given by (3.37), where $K = K(v;\rho)$ solves (3.36) and ρ evolves according to the Kushner–Stratonovich equation (3.25), an equation which also defines the law of v. □

Remark 3.11. There is an interesting interpretation of the contribution a to the drift term in (3.37): it is simply the Itô-to-Stratonovich-like correction with regard to the v-dependence in $K(v;\rho)$. However, there is a subtlety that the equation for ρ itself depends on the data $z^\dagger$ and the correction does not account for the ρ-dependence of $K(v;\rho)$: the Stratonovich correction is only with respect to v-dependence of the drift, while an Itô interpretation is retained with respect to the ρ dependence. We refer to Appendix D.3 for discussion of this unusual form of stochastic integration. There we also derive the full Stratonovich correction (with respect to both v and ρ dependence) of the exact mean-field model. □

Remark 3.12. The mean-field equations (3.37) require a gain $K(v;\rho)$ which satisfies (3.36). Let $\mathbb{E}$ denote expectation with respect to random variable v distributed according to probability measure with density ρ. Using appropriately regular test functions $\psi:\mathbb{R}^{d_v}\to\mathbb{R}^{d_v}$, chosen to have mean-zero under $\mathbb{E}$, equation (3.36) can be rephrased in the weak form[17]

$$\mathbb{E}(K^\top\nabla\psi) = \Gamma^{-1}C^{\mathsf{h}\psi}; \tag{3.41}$$

here $C^{\mathsf{h}\psi}$ denotes the covariance between h under $\mathbb{E}$ and, in what follows, we also let $C^{v\mathsf{h}}$ denote covariance between v and h. The particular choice $\psi(v) = v - \mathbb{E}v$ leads to

$$\mathbb{E}K = C^{v\mathsf{h}}\Gamma^{-1}. \tag{3.42}$$

[17] Recall that the conventions from continuum mechanics that we use to define the divergence and gradient of vector fields are discussed in Section 3.7.

Note, in particular, that this identity can be satisfied by making the *constant gain* ansatz that K is independent of v and is then given by

$$K = C^{v\mathsf{h}}\Gamma^{-1}, \tag{3.43a}$$

$$C^{v\mathsf{h}} = \mathbb{E}((v - \mathbb{E}v) \otimes (\mathsf{h}(v) - \mathbb{E}\mathsf{h}(v))). \tag{3.43b}$$

More generally speaking, we note that (3.41) is amenable to numerical approximations. This will be discussed further in Section 3.6.1 below. □

We close this subsection with the mean-field formulation using deterministic innovations:

$$\mathrm{d}v = f(v)\,\mathrm{d}t + \sqrt{\Sigma}\,\mathrm{d}W + \frac{1}{2}(\nabla\cdot(K\Gamma K^\top) - K\Gamma\nabla\cdot K^\top)\,\mathrm{d}t + K\left(\mathrm{d}z^\dagger - \frac{1}{2}(\mathsf{h}+\mathbb{E}\mathsf{h})\,\mathrm{d}t\right), \tag{3.44}$$

where $z^\dagger$ is given by (3.9). The relationship between this equation and (3.37) is analogous to the relationship between (2.61) and (2.15) already encountered in the discrete-time setting. See Section 3.7 for discussion of the historical development of the various exact mean-field models presented here.

3.5.2. *Second-order transport: stochastic case*

Recall that in the discrete-time setting there is an uncountable set of maps that effect approximate transport, in the sense of matching the first- and second-order statistics of the analysis map, using either stochastic or deterministic models. These are elucidated in Appendix C. In the main text, however, we have concentrated on a handful of examples. In this and the next subsection we study continuous-time analogues of some of these examples, starting with the Kalman transport map (2.55) recalled here, and reformulated, for convenience:[18]

$$\widehat{v}_{n+1} = \Psi(v_n) + \mathsf{N}(0,\Sigma), \tag{3.45a}$$

$$\widehat{y}_{n+1} = h(\widehat{v}_{n+1}) + \mathsf{N}(0,\Gamma), \tag{3.45b}$$

$$v_{n+1} = \widehat{v}_{n+1} + \widehat{C}^{vh}_{n+1}\left(\widehat{C}^{hh}_{n+1} + \Gamma\right)^{-1}\left(y^\dagger_{n+1} - \widehat{y}_{n+1}\right). \tag{3.45c}$$

We now apply the rescaling (3.1) to obtain, using (3.34a),

$$\begin{aligned}
\widehat{v}_{n+1} &= v_n + \Delta t f(v_n) + \sqrt{\Delta t}\mathsf{N}(0,\Sigma),\\
\widehat{z}_{n+1} &= \widehat{z}_n + \Delta t\mathsf{h}(\widehat{v}_{n+1}) + \sqrt{\Delta t}\mathsf{N}(0,\Gamma),\\
v_{n+1} &= \widehat{v}_{n+1} + C^{v\mathsf{h}}_n\Gamma^{-1}\left(\Delta z^\dagger_{n+1} - \Delta\widehat{z}_{n+1}\right) + O(\Delta t^{3/2}).
\end{aligned}$$

[18] The Gaussian notation used in the first two equations is shorthand for the first two equations appearing in (2.15), together with the assumptions detailed following those equations. Thus we use $\mathsf{N}(0,\Sigma)$ to denote an i.i.d. realization from the stated Gaussian distribution, and similarly for other variables. We use variants of this notation in what follows.

The preceding calculation, leading to $O(\Delta t^{3/2})$ error, uses the fact that the noises entering the equations for $\widehat{v}_{n+1}$, $\widehat{z}_{n+1}$ and $z^\dagger_{n+1}$ are independent. Taking the continuous-time limit, we obtain

$$\boxed{\begin{aligned} \mathrm{d}v &= f(v)\,\mathrm{d}t + \sqrt{\Sigma}\,\mathrm{d}W + C^{v\mathsf{h}}\Gamma^{-1}(\mathrm{d}z^\dagger - \mathrm{d}\widehat{z}), && (3.46a)\\ \mathrm{d}\widehat{z} &= \mathsf{h}(v)\,\mathrm{d}t + \sqrt{\Gamma}\,\mathrm{d}B. && (3.46b) \end{aligned}}$$

Again W, B are independent standard Brownian motions of appropriate dimensions and $z^\dagger$ is given by (3.9). This is an instance of a sample-path perspective that leads to approximation of the true filtering evolution.

Remark 3.13. Note the recurrence of the continuous time gain K first identified in Remark 3.7. In this subsection we have derived the mean-field equations (3.46) from the associated stochastic discrete-time formulation (3.45). However, there is another way to derive the gain, as outlined in Remark 3.12. □

3.5.3. Second-order transport: deterministic case

We may apply a similar analysis to that in Section 3.5.2 but instead working in the setting of deterministic transport, starting from the discrete-time transport (2.60). We may rewrite and reformulate this equation here to obtain

$$\begin{aligned} \widehat{v}_{n+1} &= \Psi(v_n) + \mathsf{N}(0,\Sigma), \quad \widehat{h}_{n+1} = h(\widehat{v}_{n+1}),\\ v_{n+1} &= \widehat{v}_{n+1} - \widetilde{K}_n(\widehat{h}_{n+1} - \mathbb{E}\widehat{h}_{n+1}) + K_n\left(y^\dagger_{n+1} - \mathbb{E}\widehat{h}_{n+1}\right), \end{aligned}$$

where

$$K_n = \widehat{C}^{vh}_{n+1}\left(\widehat{C}^{hh}_{n+1} + \Gamma\right)^{-1}, \quad \widetilde{K}_n = \widehat{C}^{vh}_{n+1}\left(\left(\widehat{C}^{hh}_{n+1} + \Gamma\right) + \Gamma^{1/2}\left(\widehat{C}^{hh}_{n+1} + \Gamma\right)^{1/2}\right)^{-1}.$$

Applying the rescalings (3.1) gives

$$\begin{aligned} \widehat{v}_{n+1} &= v_n + \Delta t\, f(v_n) + \sqrt{\Delta t}\mathsf{N}(0,\Sigma),\\ v_{n+1} &= \widehat{v}_{n+1} + C^{v\mathsf{h}}_n\Gamma^{-1}\left(\Delta z^\dagger_{n+1} - \frac{1}{2}(\mathsf{h}(\widehat{v}_{n+1}) + \mathbb{E}\mathsf{h}(\widehat{v}_{n+1})\Delta t\right) + O(\Delta t^2), \end{aligned}$$

where we have used

$$K_n = C^{v\mathsf{h}}_n\Gamma^{-1} + O(\Delta t), \quad \widetilde{K}_n = \frac{1}{2}C^{v\mathsf{h}}_n\Gamma^{-1} + O(\Delta t).$$

Taking the continuous-time limit, we obtain

$$\boxed{\mathrm{d}v = f(v)\,\mathrm{d}t + \sqrt{\Sigma}\,\mathrm{d}W + C^{v\mathsf{h}}\Gamma^{-1}\left(\mathrm{d}z^\dagger - \frac{1}{2}(\mathsf{h}(v) + \mathbb{E}\mathsf{h}(v))\,\mathrm{d}t\right),} \tag{3.47}$$

where, once again, $z^\dagger$ is given by (3.9).

These mean-field equations can also be derived directly from the general mean-field equation (3.44), assuming a deterministic innovation, invoking the constant

gain approximation $K = K(\rho)$ and using (3.42). The reasoning is similar to that in Remark 3.13.

Example 3.14. We now consider mean-field formulations of the Kalman–Bucy filter from Example 3.5. First consider the evolution of random variable v initialized at a Gaussian and satisfying

$$\mathrm{d}v = Fv\,\mathrm{d}t + \sqrt{\Sigma}\,\mathrm{d}W + CH^\top\Gamma^{-1}(\mathrm{d}z^\dagger - \mathrm{d}\widehat{z}), \tag{3.48a}$$

$$\mathrm{d}\widehat{z} = Hv\,\mathrm{d}t + \sqrt{\Gamma}\,\mathrm{d}B, \tag{3.48b}$$

where C is the covariance of v and where $z^\dagger$ is given by (3.29). Then direct calculation shows that the mean and covariance of v satisfy equations (3.30). Thus we have identified a sample-path perspective leading to a mean-field SDE for v with law equal to that of the Kalman–Bucy filter.

Similar considerations, starting at (3.47), give the following variant of the preceding mean-field Kalman–Bucy filter:

$$\mathrm{d}v = Fv\,\mathrm{d}t + \sqrt{\Sigma}\,\mathrm{d}W + CH^\top\Gamma^{-1}\left(\mathrm{d}z^\dagger - \frac{1}{2}H(v+m)\,\mathrm{d}t\right), \tag{3.49}$$

where (m, C) are the mean and covariance of v and where $z^\dagger$ is given by (3.29). Again direct computations verify this. An important observation highlighted by this example is that mean-field models consistent with a given measure evolution are not typically unique. □

3.6. Ensemble Kalman methods

We now discuss several particle approximations of the mean-field equations derived in the preceding subsections. We start with the mean-field equations (3.37), based on perfect transport, before considering particle approximations for continuous-time approximate sample path, and hence transport, formulations. Throughout this subsection, $z^\dagger$ is given by (3.29).

3.6.1. Perfect particle filters

As in Section 2.6.1, we define $\mathsf{J} := \{1, \ldots, J\}$. The desired approximation to the mean-field model (3.37) evolves particle ensemble $\{v^{(j)}\}_{j\in\mathsf{J}}$ and its associated empirical measure

$$\mu^{\mathsf{J}}(t) = \frac{1}{J}\sum_{j=1}^{J}\delta_{v^{(j)}} \tag{3.50}$$

according to the interacting particle system

$$\mathrm{d}v^{(j)} = f(v^{(j)})\,\mathrm{d}t + a^{(j)}\,\mathrm{d}t + \sqrt{\Sigma}\,\mathrm{d}W^{(j)} + K^{(j)}(\mathrm{d}z^\dagger - \mathrm{d}\widehat{z}^{(j)}), \tag{3.51a}$$

$$\mathrm{d}\widehat{z}^{(j)} = \mathsf{h}(v^{(j)})\,\mathrm{d}t + \sqrt{\Gamma}\,\mathrm{d}B^{(j)}, \tag{3.51b}$$

where $z^\dagger$ is given by (3.9) and the $\{W^{(j)}\}_{j\in\mathsf{J}}$ and $\{B^{(j)}\}_{j\in\mathsf{J}}$ are mutually independent collections of i.i.d. Brownian motions in $\mathbb{R}^{d_v}$ and $\mathbb{R}^{d_y}$ respectively. The interaction between the particles arises from the gain matrices

$$K^{(j)} := K^{\mathsf{J}}(v^{(j)}; \mu^{\mathsf{J}}), \tag{3.52}$$

$j \in \mathsf{J}$, where the matrix-valued function K^{J} approximates the solution K of (3.36); the drift term $a^{(j)}$ is defined by (3.40) with $K(v)$ replaced by $K^{\mathsf{J}}(v;\mu^{\mathsf{J}})$ and then evaluated at $v = v^{(j)}$. The key analysis question underlying the particle methodology is to show that the empirical measure (3.50) approximates the law of v satisfying (3.37). This type of question is widely studied for numerous problems in the physical, biological and social sciences; see Section 3.7.

Remark 3.15. In practice, this particle approximation of the perfect particle filter is impractical except in low-dimensional systems. This is because of the challenge of numerically approximating equation (3.36) for K, or its weak formulation (3.41). In this remark we discuss various approximation approaches that make this problem tractable.

We start with the numerical approximation of the weak formulation (3.41). Let $\mathcal{F}$ denote a space of vector-valued functions mapping $\mathbb{R}^{d_v}$ into $\mathbb{R}^d$, and that are mean zero with respect to expectation $\mathbb{E}$ under density ρ.[19] We make the ansatz that $K^\top = \nabla\Psi$, for some $\Psi \in \mathcal{F}$. Then (3.41) can be rewritten as

$$\mathbb{E}(\nabla\Psi\nabla\psi) = \Gamma^{-1}C^{\mathsf{h}\psi},$$

assumed to hold for all $\psi \in \mathcal{F}$.

In order to obtain the desired numerical approximation K^{J}, let $\mathbb{E}^{\mathsf{J}}$ denote expectation with respect to the empirical measure (3.50) and consider $(K^{\mathsf{J}})^\top = \nabla\Psi^{\mathsf{J}}$ and

$$\mathbb{E}^{\mathsf{J}}(\nabla\Psi^{\mathsf{J}}\nabla\psi) = \Gamma^{-1}C^{\mathsf{h}\psi} \tag{3.53}$$

with the correlation $C^{\mathsf{h}\psi}$ approximated by

$$C^{\mathsf{h}\psi} = \mathbb{E}^{\mathsf{J}}((\mathsf{h}(v) - \mathbb{E}^{\mathsf{J}}\mathsf{h}(v)) \otimes (\psi(v) - \mathbb{E}^{\mathsf{J}}\psi(v))).$$

The final step in the numerical approximation is to choose an appropriate finite-dimensional subspace $\mathcal{F}^L \subset \mathcal{F}$ of dimension $L \ll J$ and to determine $\Psi^{\mathsf{J}} \in \mathcal{F}^L$ such that (3.53) holds for all $\psi \in \mathcal{F}^L$.

Alternatively we may return to the strong formulation (3.36). We again make the ansatz $K^\top = \nabla\Psi$, giving rise to the differential equation

$$\mathcal{L}_\rho\Psi = -\Gamma^{-1}(\mathsf{h} - \mathbb{E}\mathsf{h})$$

[19] We will consider different choices for d but use the same notation for the space.

with differential operator $\mathcal{L}_\rho$[20] defined by

$$\mathcal{L}_\rho \Psi = \rho^{-1} \nabla \cdot (\rho \nabla \Psi) = \Delta \Psi + \nabla \Psi \nabla \log \rho.$$

Note that in the case of scalar observations, $\mathcal{L}_\rho$ is the infinitesimal generator of a diffusion process with invariant density ρ; in the vector case the same statement holds component by component. This fact may be used to approximate the action of $\mathcal{L}_\rho$ via its heat semigroup, and this may be used as the basis for numerical approximations. See Section 3.7 for references to the relevant literature. □

The two general approaches to approximating the gain matrix in (3.52), outlined in the preceding remark, are interesting theoretically and perhaps hold promise in the future as computer power grows, but they lead to very expensive computations. To address this we now return to the constant gain approximation from Remark 3.12, now in the particle setting. In the current context the identification of K^{J} arises from (3.53), by making the choice $\Psi(v) = (K^{\mathsf{J}})^\top (v - \mathbb{E}^{\mathsf{J}} v)$ and $\psi(v) = v - \mathbb{E}^{\mathsf{J}} v$. These choices result in the following approximation of (3.43):

$$K^{\mathsf{J}} = C^{v\mathsf{h}} \Gamma^{-1}, \tag{3.54a}$$

$$C^{v\mathsf{h}} = \mathbb{E}^{\mathsf{J}}((v - \mathbb{E}^{\mathsf{J}} v) \otimes (\mathsf{h}(v) - \mathbb{E}^{\mathsf{J}} \mathsf{h}(v))). \tag{3.54b}$$

Note that the constant gain approximation (3.54) also implies that the drift term $a^{(j)}$ in (3.51) vanishes. We summarize numerical implementation details in the following two subsections.

3.6.2. Stochastic ensemble Kalman filters

We now consider the mean-field model (3.46) and its numerical approximation. Since the drift correction a does not appear here, the only significant difference, in comparison with the interacting particle approximation (3.51), arises from the choice of the interaction term, $K^{(j)}$. This gain is now independent of j and determined by (3.54). In summary, we obtain the following SDEs: for $j \in \mathsf{J} := \{1, \dots, J\}$ we have

$$\boxed{\begin{aligned} \mathrm{d}v^{(j)} &= f(v^{(j)})\,\mathrm{d}t + \sqrt{\Sigma}\,\mathrm{d}W^{(j)} + C^{v\mathsf{h}} \Gamma^{-1} (\mathrm{d}z^\dagger - \mathrm{d}\widehat{z}^{(j)}), && (3.55a) \\ \mathrm{d}\widehat{z}^{(j)} &= \mathsf{h}(v^{(j)})\,\mathrm{d}t + \sqrt{\Gamma}\,\mathrm{d}B^{(j)}. && (3.55b) \end{aligned}}$$

Here, again, $z^\dagger$ is given by (3.9) and the $\{W^{(j)}\}_{j \in \mathsf{J}}$ and $\{B^{(j)}\}_{j \in \mathsf{J}}$ are mutually independent collections of i.i.d. Brownian motions. The $\{v^{(j)}\}_{j=1}^J$ provide a time-evolving ensemble which approximates the filtering distribution via (3.50); the equations are derived based on use of second-order transport approximation of perfect transport. Each equation for $v^{(j)}$ evolves according to the underlying dynamics model, together with a nudging term based on the difference between simulated data $\widehat{z}^{(j)}$ and observed data $z^\dagger$. The gain couples the particles together.

[20] Not to be confused with operator $\mathcal{L}_n$ defined by viewing Bayes' theorem, within filtering, as a prior-to-posterior map at time n.

3.6.3. Deterministic ensemble Kalman filters

Similarly we may make an empirical approximation of the mean-field model (3.47) as follows: for $j \in \mathsf{J} := \{1, \dots, J\}$ we consider

$$\boxed{\mathrm{d}v^{(j)} = f(v^{(j)})\,\mathrm{d}t + \sqrt{\Sigma}\,\mathrm{d}W^{(j)} + C^{v\mathsf{h}}\Gamma^{-1}\left(\mathrm{d}z^{\dagger} - \frac{1}{2}\left(\mathsf{h}(v^{(j)}) + \mathbb{E}^{\mathsf{J}}\mathsf{h}(v)\right)\mathrm{d}t\right).} \tag{3.56}$$

The notation is as in the preceding subsection and, in particular, the formula (3.50) again gives the particle approximation of the approximate filter; now the relevant approximate filter is defined by the distribution of (3.56).

3.7. Bibliographical notes

The entirety of Section 3 is framed in the language of SDEs; see Evans (2012) and Øksendal (2013) for background in this area. In passing from discrete to continuous time we often invoke ideas from the numerical solution of SDEs; see Kloeden and Platen (1991) and Higham (2001) for introductions to this area. The conventions from continuum mechanics, that we use to define the divergence and gradient of vector fields, are the same as those adopted, and described in detail, in Gonzalez and Stuart (2008) and Gurtin (1982).

Next we review literature in the control-theoretic approach introduced in Section 3.2. For a general overview of control theory in continuous time see Sontag (2013). Our presentation of control-theoretic methods for data assimilation focuses on 3DVAR. Theoretical analysis of the continuous-time 3DVAR algorithm (3.8) may be found in Law *et al.* (2014), Blömker, Law, Stuart and Zygalakis (2013), Azouani, Olson and Titi (2014), Gesho, Olson and Titi (2016), Olson and Titi (2003), Mondaini and Titi (2018) and Larios and Pei (2024).

We now review the probabilistic approach introduced in Section 3.3. The Kalman–Bucy filter (Kalman and Bucy 1961) contains what is perhaps the first systematic derivation and analysis of an algorithm for the incorporation of continuous-time data into estimation of a sample path of an SDE. Its extension to nonlinear and non-Gaussian distributions is provided by the Kushner–Stratonovich equation (3.25). A heuristic derivation of both the Kalman–Bucy filter as well as the Kushner–Stratonovich equation can be found in Jazwinski (2007), while Bain and Crisan (2008) cover the field of continuous-time filtering in full detail.

The idea of Strang splitting, which we use to derive the Kushner–Stratonovich equation, originates in Strang (1968); for an overview of splitting methods see McLachlan and Quispel (2002). Furthermore, we rely on robustness results for continuous-time filtering (Clark and Crisan 2005), which imply that smooth approximations $z^{\dagger,\delta}$ to stochastic observations $z^{\dagger}$ are justified and that the order of taking limits $\Delta t \to 0$ and $\delta \to 0$ can be accounted for by appropriate Stratonovich-to-Itô correction terms. An introduction to the required covariation formulae used

in the proof of Lemma 3.3 and identity (3.23), which follows from covariation of stochastic integrals, can be found in Eberle (2019). We note that robustness results do not carry over to associated mean-field equations and filtering problems with correlated noise (Coghi, Nilssen, Nüsken and Reich 2023).

The numerical approximation of the Kushner–Stratonovich equation (3.25) has a long history. Crisan and Lyons (1999) have demonstrated how (3.25) can be approximated by a particle method; this is a generalization of the bootstrap particle filter to continuous time (Crisan *et al.* 1999). See also Bain and Crisan (2008) for a detailed discussion of alternative approximation techniques. Hu, Kallianpur and Xiong (2002) discussed the solution of the unnormalized and linear version of the Kushner–Stratonovich equation known as the Zakai equation.

A mean-field approach to the Kushner–Stratonovich equation (3.25) appeared first in the work of Crisan and Xiong (2010), which utilizes robustness results and smoothed data $z^{\dagger,\delta}$ in the $\delta \to 0$ limit. The mean-field equations (3.44) were proposed in Yang *et al.* (2013) (in the one-dimensional setting), while the stochastic counterpart appeared first in Reich (2019). The mathematical relationship between the various mean-field formulations has been analysed in Pathiraja, Reich and Stannat (2021). Numerical implementations of the mean-field equations (3.37) are discussed in Taghvaei, de Wiljes, Mehta and Reich (2017), while Taghvaei, Mehta and Meyn (2020) provide a detailed analysis of the diffusion map approximation discussed after (3.15). The constant gain approximation $K = C^{vh}\Gamma$, which corresponds to the ensemble Kalman filter, arises as a particular scaling limit from the diffusion map approach, as discussed, for example, in Taghvaei *et al.* (2017). See also the recent survey by Taghvaei and Mehta (2023).

The continuous-time ensemble Kalman filter formulations (3.55) and (3.56) appeared first in Bergemann and Reich (2012). These formulations are based on earlier work on homotopy formulations of the Bayesian inference step by Bergemann and Reich (2010*b*) and Reich (2011). See also the subsequent derivations in Law *et al.* (2015), which contains a unified derivation of the Kalman–Bucy filter, continuous-time 3DVAR and continuous-time ensemble Kalman methods, starting from their discrete-time counterparts.

Rigorous derivation of continuous-time ensemble Kalman filter formulations from their discrete-time counterparts can be found in Lange and Stannat (2021*a*,*b*), Blömker, Schillings and Wacker (2018) and Blömker, Schillings, Wacker and Weissmann (2022). Derivation and analysis of the properties of continuous-time limits in the context of solving inverse problems may be found in Schillings and Stuart (2017); see Section 5.6. Numerical time-stepping methods for continuous-time ensemble Kalman filter formulations are analysed in Amezcua, Kalnay, Ide and Reich (2014).

Ding *et al.* (2021) and Ding and Li (2021*a*,*b*) have undertaken a systematic analysis of the link between interacting particle systems and mean-field systems in continuous time, mostly focused on the solution of inverse problems; however, the methods developed are more widely applicable. Similar to the stochastic ensemble

Kalman filter, particle implementation (3.55) leads to undesirable correlations via the synthetic data $\widehat{z}^{(j)}$, which appear both in the gain $K^{\mathsf{J}} = C^{\nu\mathsf{h}}\Gamma^{-1}$ and the innovation term $\mathrm{d}\mathfrak{I} = \mathrm{d}z^{\dagger} - \mathrm{d}\widehat{z}^{(j)}$, effectively giving rise to coloured noise. These numerically induced correlations vanish in the limit $J \to \infty$.

Well-posedness, stability and accuracy results for the ensemble Kalman filter first appeared in Kelly *et al.* (2014), where the incompressible Navier–Stokes equations were studied, using the continuous-time formulations of stochastic ensemble Kalman methods first identified in Reich (2011) and reviewed in Law *et al.* (2015). Subsequent analyses of related issues for variants on the continuous-time ensemble Kalman filter may be found in de Wiljes, Reich and Stannat (2018) and de Wiljes and Tong (2020). Del Moral and Tugaut (2018) have initiated a line of research related to the use of particle approximations of mean-field models to understand long-time error estimates for particle systems approximating the Kalman–Bucy filter, the setting in which mean-field ensemble Kalman filters exactly reproduce the true filtering distribution; see Bishop, Del Moral and Pathiraja (2018), Bishop, Del Moral, Kamatani and Remillard (2019), Bishop and Del Moral (2019), Bishop, Del Moral and Niclas (2020) and, for an overview, Bishop and Del Moral (2023). The control perspective deployed in Law *et al.* (2014) and Azouani *et al.* (2014) to study filter stability and accuracy has been unified and extended to ensemble Kalman filter formulations in Biswas and Branicki (2024), utilizing a particular form of covariance localization and additive inflation.

The classical Kalman filter can be viewed from the perspective of minimum variance estimation (discussed in discrete time in Section C.3) and optimal control (Kalman and Bucy 1961). This perspective has recently been extended to nonlinear filtering in Kim and Mehta (2024*a*,*b*), which opens up new perspectives for developing and analysing numerical algorithms.

4. Inverse problems: discrete time

In this section we adapt the ideas of Section 2 to solve inverse problems. We start in Section 4.1 with statement of the inverse problem, followed in Sections 4.2 and 4.3 by discussion of optimization and Bayesian approaches respectively; these are analogous to the presentation of control and probabilistic approaches to the data assimilation problem in Sections 2.2 and 2.3.

The basic methodology we highlight is to formulate filtering problems which (possibly only approximately) solve the inverse problem. Section 4.4 is devoted to Bayesian probabilistic filtering methods that solve the inverse problem by morphing the prior into the posterior in a finite time; Section 4.5 discusses filtering methods which work on infinite-time horizons, exhibiting exponential convergence to approximate solutions of the optimization or Bayesian formulations of the problem from arbitrary starting points. In both Sections 4.4 and 4.5 we demonstrate the use of Gaussian projected filtering and ensemble Kalman methods to solve the filtering

problems that arise. In Section 4.6 we present numerical examples illustrating ensemble Kalman methods for inverse problems. We conclude in Section 4.7 with bibliographical notes.

Remark 4.1. In Sections 4.4 and 4.5 we present only mean-field statements of the Gaussian projected filter and ensemble Kalman-based methods. The reader can generalize the ideas in Section 2, based on interacting particle system approximations, to derive implementable algorithms from the ensemble-based mean-field algorithms introduced. Similarly, the unscented Kalman filter can be used to derive implementable algorithms from the Gaussian projected filters introduced here, as discussed in Section 2.7. Pseudo-code for the schemes implemented in the numerical examples of Section 4.6 may be found in Appendix A. □

4.1. Set-up

This subsection and Section 5 are entirely devoted to solution of the inverse problem of finding unknown parameter $u \in \mathbb{R}^{d_u}$ from data $w \in \mathbb{R}^{d_w}$, when w is related to u via the equation

$$\boxed{w = G(u) + \gamma.} \tag{4.1}$$

Here $G\colon \mathbb{R}^{d_u} \to \mathbb{R}^{d_w}$ is the *forward model* and γ represents noise polluting the data. We assume that G is measurable with respect to the Borel algebra on input and output spaces, and is bounded on compact subsets of $\mathbb{R}^{d_u}$. The basic methodology we highlight is to formulate filtering problems which (possibly only approximately) solve the inverse problem.

Remark 4.2. The filtering problem from Section 2.1 requires solution of an inverse problem, defined by (2.1b), at each step n; this inverse problem is a specific instance of (4.1). Indeed, the map $\mathcal{L}_n$ in (2.24b) denotes Bayesian solution of this inverse problem, a concept we will define, in the more general setting of this section, in Section 4.3. Furthermore, the smoothing problem, referred to at the very end of Section 2.7, can also be formulated as an instance of the general inverse problem (4.1). □

We will work in a setting where we assume a probabilistic model for the joint random variable (u, w). We then assume that we have available to us $w^\dagger$, the observation coordinate of a specific realization $(u^\dagger, w^\dagger)$ under this probabilistic model. This realization is itself generated by $\gamma^\dagger$, a specific realization of the observational noise. We will consider two approaches to the inverse problem.

Objective 1. Design an algorithm producing output u from $w^\dagger$ so that u estimates $u^\dagger$, the true state underlying the data.

Objective 2. Design an algorithm which estimates the distribution of random variable $u|w^\dagger$.

In Section 4.2 we define an optimization approach to determine an approximation of $u^\dagger$ from $w^\dagger$, addressing Objective 1. In Section 4.3 we define the Bayesian probabilistic formulation, which also underpins the algorithms derived in the remainder of the section, addressing Objective 2.

4.2. *Optimization formulation*

Given matrices $C_0 \succ 0$, $\Gamma \succ 0$, vector m_0 and the specific data realization, namely $w^\dagger$, we may define the nonlinear least-squares loss function Φ and its Tikhonov-regularized counterpart Φ_R as follows:

$$\Phi(u) = \frac{1}{2}|w^\dagger - G(u)|^2_\Gamma, \tag{4.2a}$$

$$\Phi_R(u) = \Phi(u) + \frac{1}{2}|u - m_0|^2_{C_0}. \tag{4.2b}$$

Minimization of Φ_R constitutes a solution to the inverse problem. The specific weighted norms used in the least-squares loss Φ, and the regularization leading to Φ_R, are best understood from the probabilistic formulation in the following subsection; however, the remainder of this subsection can be understood without recourse to this probabilistic formulation.

The Tikhonov regularized least-squares problem associated with the inverse problem (4.1) may be viewed as an unregularized least-squares problem arising from the modified inverse problem

$$w_R = G_R(u) + \gamma_R,$$

where we write

$$w_R := \begin{pmatrix} w \\ m_0 \end{pmatrix}, \quad G_R(u) := \begin{pmatrix} G(u) \\ u \end{pmatrix}, \quad \Gamma_R := \begin{pmatrix} \Gamma & 0 \\ 0 & C_0 \end{pmatrix}, \tag{4.3}$$

for γ_R being the generalized observation error. In particular, the cost functional (4.2b) can be rewritten as

$$\Phi_R(u) = \frac{1}{2}|w^\dagger_R - G_R(u)|^2_{\Gamma_R}, \tag{4.4}$$

for

$$w^\dagger_R := \begin{pmatrix} w^\dagger \\ m_0 \end{pmatrix}. \tag{4.5}$$

Remark 4.3. A building block in many algorithms for minimization of $\Psi\colon \mathbb{R}^{d_u} \to \mathbb{R}^+$ is *gradient descent*. In basic form, this is an iteration for sequence $\{u_n\}_{n\in\mathbb{Z}^+}$ defined by

$$u_{n+1} = u_n - \alpha\nabla\Psi(u_n). \tag{4.6}$$

To solve the inverse problem, we can use iteration (4.6) with $\Psi = \Phi$ or $\Psi = \Phi_R$.

In Section 4.5.2 we develop derivative-free *affine-invariant* algorithms[21] based on Gaussian projected and ensemble Kalman filters. These algorithms offer an alternative to (4.6); they only approximately minimize the least-squares objective, in general. However, in the quadratic case they reproduce an exact mean-field gradient descent algorithm, as we will show as this section unfolds. □

4.3. Bayesian formulation

We now consider the *Bayesian* approach to this inverse problem. We again assume $C_0 \succ 0$ and $\Gamma \succ 0$, as in the previous subsection. To be concrete we assume *prior* $u \sim \mathsf{N}(m_0, C_0)$, that $\gamma \sim \mathsf{N}(0, \Gamma)$ and that u and γ are independent. It then follows that the likelihood $w|u \sim \mathsf{N}(G(u), \Gamma)$. Application of Bayes' theorem shows that the *posterior distribution* on $u|w^\dagger$ is distributed according to measure μ given by

$$\mu(\mathrm{d}u) = \frac{1}{\mathcal{Z}} \exp(-\Phi_R(u))\,\mathrm{d}u, \tag{4.7a}$$

$$\mathcal{Z} = \int_{\mathbb{R}^{d_u}} \exp(-\Phi_R(u))\,\mathrm{d}u. \tag{4.7b}$$

We note that the least-squares-based optimization approaches to the inverse problem introduced in the preceding subsection can now be explicitly linked to the probabilistic formulation of the inverse problem. In particular, minimizing Φ is referred to as the *maximum likelihood approach*, whilst minimizing Φ_R is called the *maximum a posteriori approach*.

Example 4.4. Consider the case of linear forward model

$$G(\cdot) = L \cdot . \tag{4.8}$$

Thus Φ_R is quadratic and it is straightforward to show that the Hessian of Φ_R is greater than or equal to, in the sense of quadratic forms, C_0^{-1}. Since $C_0^{-1} \succ 0$, we deduce that Φ_R has a unique critical point and this critical point is a global minimizer. It is then natural to define

$$L_R := \begin{pmatrix} L \\ I \end{pmatrix} \tag{4.9}$$

and note that, with this definition, $G_R(\cdot) = L_R \cdot$. We then have

$$\Phi_R(u) = \frac{1}{2} |w_R^\dagger - L_R u|^2_{\Gamma_R}.$$

Since Φ_R is quadratic we deduce that the posterior μ is Gaussian. We denote the

[21] See Remark 4.18 for a discussion of the implications of affine invariance, in the discrete-time setting. We will study affine invariance in detail in continuous time in Section 5; see Definition 5.11 and Remark 5.12.

mean by m_{post} and the covariance by C_{post}. Matrix C_{post} is readily defined via its precision, the Hessian of Φ_R:

$$C_{\text{post}}^{-1} = \mathsf{S} = L_R^\top \Gamma_R^{-1} L_R. \tag{4.10}$$

As discussed above, $C_{\text{post}}^{-1} \succ 0$ so that $C_{\text{post}} \succ 0$; in particular, C_{post} is hence indeed invertible. The minimizer of Φ_R is at the mean m_{post} of the posterior, which solves the *normal equations*

$$C_{\text{post}}^{-1} m_{\text{post}} = L_R^\top \Gamma_R^{-1} w_R^\dagger. \tag{4.11}$$

This representation of the posterior, which is Gaussian, is in terms of the precision matrix and the mean. There is an alternative and useful representation formula for the posterior covariance and mean, derived as follows. Consider the Gaussian random variable (u, w) defined by choosing $u \sim \mathsf{N}(m_0, C_0)$ and $w|u \sim \mathsf{N}(Lu, \Gamma)$. Then the solution of the Bayesian inverse problem (4.7) is given in this linear setting by the distribution of $u|w^\dagger$. By using standard conditioning formulae for Gaussian random variables, we obtain

$$m_{\text{post}} = m_0 + C_0 L^\top (L C_0 L^\top + \Gamma)^{-1} (w^\dagger - L m_0), \tag{4.12a}$$

$$C_{\text{post}} = C_0 - C_0 L^\top (L C_0 L^\top + \Gamma)^{-1} L C_0. \tag{4.12b}$$

These formulae are equivalent to (m_1, C_1) found from the Kalman filter Bayesian update step (2.28) of Example 2.6, with the choices $M = I$, $\Sigma = 0$, $H = L$ and $\Gamma = \Gamma$. □

Remark 4.5. A commonly used methodology for sampling from target distribution μ on $\mathbb{R}^{d_u}$ is MCMC. At abstract level this defines a Markov chain for density ρ_n given by transition kernel $\mathcal{K}(\alpha)$, where α describes hyper-parameters that define the specific method used. Thus

$$\rho_{n+1} = \mathcal{K}(\alpha)\rho_n. \tag{4.13}$$

This is a linear iteration for the density ρ_n, designed to converge to the target density, defined by the posterior, as $n \to \infty$. In Section 4.5.3 we show how mean-field ensemble Kalman methods, which induce a nonlinear iteration on density ρ_n, may be used as an alternative to (4.13), defining an approximate Bayesian posterior by iterating to infinity. Furthermore, this iteration will be shown to be exact for Gaussian posteriors, and to benefit from affine invariance.[22] □

4.4. Finite-time algorithms

The basic idea used in this subsection, to address the solution of inverse problems, is rooted in a sequential formulation of Bayesian inference. From this sequential

[22] Recall that discussion of the implications of affine invariance may be found in Remark 4.18, in the discrete-time setting, and that in Section 5 we will study affine invariance in detail in continuous time.

formulation we derive a filtering problem whose solution, at a particular time, gives the desired posterior. Section 4.4.1 describes the formulation, Section 4.4.2 the use of Gaussian projected filters and Section 4.4.3 the use of ensemble Kalman methods.

4.4.1. Formulation

To understand this sequential approach we define, for integer $N > 1$,

$$\Phi_{R,n}(u) = \frac{n}{N}\Phi(u) + \frac{1}{2}|u - m_0|^2_{C_0}, \tag{4.14}$$

noting that $\Phi_{R,N}(u) = \Phi_R(u)$. Now consider the sequence μ_n of probability measures with negative log density given (up to an additive constant with respect to variation of u) by $\Phi_{R,n}(u)$. Then the Bayesian inference problem (4.7) can be reformulated as a sequence of N Bayesian inference steps where the prior μ_n is morphed into posterior μ_{n+1} using the data likelihood $\exp(-\Phi(u)/N)$ in each step:

$$\mu_{n+1}(\mathrm{d}u) \propto \exp\left(-\frac{1}{N}\Phi(u)\right)\mu_n(\mathrm{d}u). \tag{4.15}$$

Hence

$$\mu_n(\mathrm{d}u) \propto \exp\left(-\frac{n}{N}\Phi(u)\right)\mu_0(\mathrm{d}u). \tag{4.16}$$

Thus

$$\mu_N(\mathrm{d}u) \propto \exp(-\Phi(u))\mu_0(\mathrm{d}u). \tag{4.17}$$

The initial prior is set to $\mu_0 = \mathsf{N}(m_0, C_0)$ and the Nth posterior μ_N delivers the desired Bayesian solution to the inverse problem, given in (4.7).

Remark 4.6. Here we have introduced an iteration index n to morph from prior to posterior. In what follows we will identify n with an *artificial* time and then import ideas from filtering to solve the inverse problem. Notice that, in this approach, the single inverse problem (4.17) of interest, is replaced by N inverse problems of the form (4.15). This can be beneficial because each of the N inverse problems (4.15) is easier to solve than the single inverse problem (4.17), because the defining change of measure is closer to the identity.

A variant on this idea is to morph from prior to posterior by (possibly artificially) considering the data w as sequentially acquired and incrementally including components of the data at each step n, again leading to a sequence of measures $\{\mu_n\}_{n=0}^N$, with μ_N equal to the posterior. □

Given this sequence of measures μ_n, it is possible to identify a stochastic dynamical system with filtering distribution μ_n. Application of any filtering method to this filtering problem, and ensemble Kalman filters in particular, then provides a method to approximate the posterior distribution. These sequential formulations of Bayesian inversion are well known and have, for example, been exploited in

the use of sequential Monte Carlo methods for Bayesian inference; see Section 4.7 for details.

Remark 4.7. To employ sequential Monte Carlo methods based on (4.15) it is necessary to invoke some form of approximation. The resulting outcome of such approximations depends on the choice of positive integer N. Empirically it is found that choosing $N \gg 1$ gives better approximations of the desired posterior (4.17); however, this must be traded against the additional cost of taking N steps.

Standard particle filter-based sequential Monte Carlo methods do not always scale well to high dimensions, in the same way that the particle filter for state estimation scales poorly. Consequently, ensemble Kalman variants of sequential Monte Carlo methods have an important place in the field, and we will deploy these, and variants of them, after introducing the stochastic dynamical system underlying sequential Monte Carlo. □

Our first step is to show how to realize (4.15) via a filtering problem which we refer to as a *transport problem*:[23] it transports the prior initial condition into the desired posterior, through a discrete-time evolution. See Theorem 4.8 below and note that in Section 5.4.1 analogous developments are made in continuous time. The transport problem is exact: it introduces no approximations. To derive algorithms we proceed to discuss approximations to the transport. We follow discussion of exact transport with study of the Gaussian projected filter, applied in this inverse problem context, in Section 4.4.2, and the continuous-time analogue in Section 5.4.2. Study of Gaussian projection is then followed by discussion of the application of mean-field Kalman transport algorithms to the transport problem in Section 4.4.3; continuous-time analogues are covered in Section 5.4.3.

In the following development we define Δt so that

$$\boxed{N\Delta t = 1.} \tag{4.18}$$

Now consider the combined state-observation system in the form

$$u_{n+1} = u_n, \tag{4.19a}$$

$$w_{n+1} = G(u_{n+1}) + \frac{1}{\sqrt{\Delta t}}\gamma_{n+1}, \tag{4.19b}$$

for $n \in \{0, \ldots, N-1\}$, where $\{\gamma_{n+1}\}_{n=0}^{N-1}$ is an i.i.d. sequence with variance $\mathsf{N}(0, \Gamma)$. It is intuitive that, since $N\Delta t = 1$, $u_0 \sim \mathsf{N}(m_0, C_0)$ and the observed data $w_{n+1}^\dagger = w^\dagger$ for all $n \in \{0, \ldots, N-1\}$, then u_N conditioned on $W_N^\dagger := \{w_{n+1}^\dagger\}_{n=0}^{N-1}$ is distributed as μ, defined as in (4.7). Indeed, using (4.19b) for $n \in \{0, \ldots, N-1\}$ corresponds to making N independent noisy observations of $G(u_0)$, all with noise variance $\Delta t^{-1}\Gamma$; this is statistically equivalent to a single noisy observation of $G(u_0)$ with

[23] Note that we have introduced transport ideas in Section 2 to underpin algorithms which approximate the Bayesian inference that defines the analysis step in filtering.

noise variance Γ. Since u_0 is initialized as $\mathsf{N}(m_0, C_0)$ (the prior), the problem reduces to the Bayesian inverse problem for $u|w^\dagger$.

This intuition may be substantiated by using the discussion around equations (4.14) and (4.15). In order to avail ourselves of the results from Section 2, we first rescale the observation equation in (4.19) to obtain

$$\boxed{\begin{aligned} u_{n+1} &= u_n, && \text{(4.20a)}\\ y_{n+1} &= \Delta t G(u_{n+1}) + \eta_{n+1}, && \text{(4.20b)} \end{aligned}}$$

with $\eta_{n+1} \sim \mathsf{N}(0, \Delta t\Gamma)$ and $y_{n+1} = \Delta t w_{n+1}$. Denote the observed data by $Y_n^\dagger = \{y_\ell^\dagger\}_{\ell=1}^n$, where $y_\ell^\dagger = \Delta t w_\ell^\dagger$. We may then show the following.

Theorem 4.8. Consider the dynamical system (4.20) and assume that $C_0 \succ 0$, $\Gamma \succ 0$ and $N\Delta t = 1$. Assume also that $u_0 \sim \mathsf{N}(m_0, C_0)$ and $\eta_{n+1} \sim \mathsf{N}(0, \Delta t\Gamma)$; furthermore, assume that $\{\eta_n\}_{n=1}^N$ forms an i.i.d. sequence, independent of u_0. Then μ_n, the law of $u_n|Y_n^\dagger$ defined by (4.20), satisfies (4.15), and in particular μ_N is equal to the posterior distribution μ, if the data is chosen as $y_n^\dagger = \Delta t\, w^\dagger$, for $n \in \{1, \dots, N\}$. ◇

Proof. Let μ_n be the law of $u_n|Y_n^\dagger$. Since (in the notation of Section 2) $\widehat{\mu}_{n+1} = \mu_n$ we see that the mapping μ_n to μ_{n+1} is simply given by Bayes' theorem: $\mu_{n+1} = \mathcal{L}_n(\mu_n)$. This observation yields the following identity, expressed in terms of ρ_n the Lebesgue density of measure μ_n:

$$\begin{aligned} \log \rho_{n+1} - \log \rho_n &= -\frac{1}{2\Delta t}|y_n^\dagger - \Delta t G(u)|_\Gamma^2 + \text{const.} \\ &= -\frac{\Delta t}{2}|w^\dagger - G(u)|_\Gamma^2 + \text{const.} \end{aligned}$$

Summing over $n \in \{0, \dots, N-1\}$, using the fact that

$$\log \rho_0 = -\frac{1}{2}|u - m_0|_{C_0}^2 + \text{const.},$$

we deduce that

$$\log \rho_n = -\Phi_{R,n} + \text{const.},$$

with $\Phi_{R,n}$ given by (4.14). Choosing $n = N$ gives the desired result concerning the posterior. □

Thus we may approach the problem of (approximately) sampling from μ by (approximately) solving the filtering problem defined by (4.20), for μ_n, until discrete time $n = N$. In particular, we may use the Gaussian projected filter or ensemble Kalman methods to approximate this filtering problem. In the next two subsections we consider, respectively, these two approximation methods.

4.4.2. Algorithms: Gaussian projected filter

We now apply the ideas from Section 2.4, which concerns the Gaussian projected filter in the general setting, to the specific setting of the stochastic dynamical system (4.20). Let $\mathbb{E}$ denote expectation under $u \sim \mathsf{N}(m_n, C_n)$ and define

$$C_n^{uG} = \mathbb{E}((u - \mathbb{E}u) \otimes (G(u) - \mathbb{E}G(u))), \tag{4.21a}$$

$$C_n^{GG} = \mathbb{E}((G(u) - \mathbb{E}G(u)) \otimes (G(u) - \mathbb{E}G(u))). \tag{4.21b}$$

Noting that prediction under (4.20a) is trivial, it follows that the predicted mean and covariance satisfy $\widehat{m}_{n+1} = m_n$ and $\widehat{C}_{n+1} = C_n$. Hence, using (2.42) in the specific setting of (4.20) yields

$$m_{n+1} = m_n + \Delta t C_n^{uG} \left(\Gamma + \Delta t C_n^{GG}\right)^{-1} (w^\dagger - \mathbb{E}G(u)), \tag{4.22a}$$

$$C_{n+1} = C_n - \Delta t C_n^{uG} \left(\Gamma + \Delta t C_n^{GG}\right)^{-1} \left(C_n^{uG}\right)^\top. \tag{4.22b}$$

Note that the difference between the data $w^\dagger$ and the mean of $G(u)$ under the Gaussian at step n acts as a forcing term in the evolution of the mean from n to $n+1$, promoting a Gaussian which agrees with the data. This forcing term is weighted by covariance information. The covariance of the Gaussian projected filter is non-increasing from step to step since $\langle u, C_{n+1}u\rangle \le \langle u, C_n u\rangle$ for all $u \in \mathbb{R}^{d_u}$; this reflects the fact that more information is received at each step $n \mapsto n+1$ as the unknown u is repeatedly observed.

Example 4.9. In the setting of the linear inverse problem (4.4), where $G(u) = Lu$, the Gaussian projected filter equations (4.22) reduce to

$$m_{n+1} = m_n + \Delta t C_n L^\top (\Gamma + \Delta t L C_n L^\top)^{-1} (w^\dagger - L m_n), \tag{4.23a}$$

$$C_{n+1} = C_n - \Delta t C_n L^\top (\Gamma + \Delta t L C_n L^\top)^{-1} L C_n. \tag{4.23b}$$

These equations may be iterated to map from (m_0, C_0) directly to (m_n, C_n), obtaining

$$m_n = m_0 + n\Delta t C_0 L^\top (\Gamma + n\Delta t L C_0 L^\top)^{-1} (w^\dagger - L m_0), \tag{4.24a}$$

$$C_n = C_0 - n\Delta t C_0 L^\top (\Gamma + n\Delta t L C_0 L^\top)^{-1} L C_0. \tag{4.24b}$$

These formulae may also be obtained by applying Bayes' formula, in the linear setting, to (4.16) and using (4.18), namely $N\Delta t = 1$. The Gaussian posterior measure $\mu = \mathsf{N}(m_{\text{post}}, C_{\text{post}})$ given by (4.10) and (4.11) may now be found by choosing mean and covariance $(m_{\text{post}}, C_{\text{post}}) = (m_N, C_N)$. This follows from Section 2 because the Gaussian projected filter recovers the Kalman filter, which is exact for linear Gaussian problems. ∎

4.4.3. Algorithms: ensemble Kalman filter

Recall that mean-field models lead to ensemble Kalman methods through particle approximation. In this section we simply highlight use of one of the mean-field

models, in the context of inverse problems, namely the stochastic Kalman transport approach from Section 2.5.4 and its deterministic variant from Section 2.5.5. We leave details of particle approximations of these mean-field models to the reader, and to Appendix A for related pseudo-code.

Employing the state-observation model (4.20) within the stochastic Kalman transport model (2.55), we obtain the mean-field dynamical system, for i.i.d. unit Gaussian sequence $\{\xi_n\}$ in $\mathbb{R}^{d_w}$,

$$u_{n+1} = u_n + \Delta t C_n^{uG}\left(\Delta t C_n^{GG} + \Gamma\right)^{-1}\left(w^\dagger - G(u_n) - \sqrt{\frac{\Gamma}{\Delta t}}\xi_n\right), \tag{4.25a}$$

$$C_n^{uG} = \mathbb{E}((u_n - \mathbb{E}u_n) \otimes (G(u_n) - \mathbb{E}G(u_n))), \tag{4.25b}$$

$$C_n^{GG} = \mathbb{E}((G(u_n) - \mathbb{E}G(u_n)) \otimes (G(u_n) - \mathbb{E}G(u_n))). \tag{4.25c}$$

Note that, here, expectation $\mathbb{E}$ is computed under the law of u_n itself. As for the Gaussian projected filter (4.22), the evolution promotes a distribution which is compatible with the data, here with a forcing term, weighted by covariance information, applied to the evolution of the state u_n; but unlike the mean-field evolution equation for the mean, there is additional noise for the state evolution.

Recall Theorem 4.8. Since the ensemble Kalman transport algorithm used here provides an approximation of the filtering distribution for the dynamical system (4.20), it follows that the random variable u_N provides an approximation to the posterior distribution μ, provided that $u_0 \sim \mathsf{N}(m_0, C_0)$, the prior distribution. This statement can be made exact in the linear case, as the following example shows.

Example 4.10. Assume $G(u) = Lu$ for $L \in \mathbb{R}^{d_w \times d_u}$ so that the posterior distribution of the Bayesian inverse problem μ is given in Example 4.4. Then the solution of the mean-field model (4.25) satisfies $u_N \sim \mu$. This is a specific instance of what we observed in Example 2.15, namely that the mean-field model reproduces the Kalman filter on linear Gaussian problems. We note also that the Gaussian projected filter is identical to the Kalman filter in this case; see Example 4.9.

Although it is implicit in Example 4.4, in this specific inverse problem context we demonstrate the equivalence with the Kalman filter explicitly. To do this we note that, in the linear setting, equation (4.25) defines a closed evolution in the set of Gaussian probability measures. The updates for the mean m_n and covariance C_n of u_n then coincide with the Kalman filter, and hence the Gaussian projected filter, in this linear case given by equations (4.23). Indeed, by taking the expectation under the law of u_n of (4.25a), it is readily checked that in the linear setting we obtain (4.23a) for the mean update. To obtain the evolution equation of the covariance, recall that

$$C_{n+1} = \mathbb{E}((u_{n+1} - m_{n+1}) \otimes (u_{n+1} - m_{n+1})). \tag{4.26}$$

Substituting into (4.26) the expression for u_{n+1}, given by (4.25a) in the linear setting $G(u) = Lu$, and the expression for m_{n+1}, given by (4.23a), and then computing the

expectation yields (4.23b). It follows from the calculations of Example 4.9 that $u_N \sim \mu$. □

We conclude this subsection by stating the corresponding deterministic transport formulation. We employ the approximation (2.61). This holds in our case provided Δt is small enough. Choosing K_n as implicitly defined in (4.25), we obtain, with expectation $\mathbb{E}$ computed under the law of u_n itself, the following mean-field model:

$$\boxed{\begin{aligned} u_{n+1} &= u_n + \Delta t C_n^{uG}\left(\Delta t C_n^{GG} + \Gamma\right)^{-1}\left(w^\dagger - \frac{1}{2}(G(u_n) + \mathbb{E}G(u_n))\right), && \text{(4.27a)}\\ C_n^{uG} &= \mathbb{E}((u_n - \mathbb{E}u_n) \otimes (G(u_n) - \mathbb{E}G(u_n))), && \text{(4.27b)}\\ C_n^{GG} &= \mathbb{E}((G(u_n) - \mathbb{E}G(u_n)) \otimes (G(u_n) - \mathbb{E}G(u_n))). && \text{(4.27c)} \end{aligned}}$$

4.5. Infinite-time algorithms

Algorithms which (approximately) transport prior to posterior in finite time, as described in the preceding subsection, are attractive. However, they can be quite rigid as they do not benefit from strong stability to perturbations. An alternative, pursued in this section, is to seek algorithms which converge to the desired solution on an infinite time horizon, from arbitrary starting points, and which exhibit exponential stability. This is hard to achieve in general, but can be achieved exactly for Gaussian problems. When applied beyond the Gaussian setting this hence leads to a methodology consistent with the application of ensemble Kalman filter approximations, which themselves invoke a Gaussian ansatz and yet are used beyond the Gaussian setting. Section 4.5.1 describes the infinite time horizon formulation. In Section 4.5.2 we consider this infinite time horizon perspective for the solution of optimization problems associated with the inverse problem (4.1). Section 4.5.3 considers the same perspective for Bayesian inversion.

4.5.1. Formulation

To motivate what follows, we consider algorithms that solve the optimization problem by extending ideas from the previous section to iterate a filtering problem over an infinite time horizon. To explain this idea, recall the identity (4.16), restated here for convenience:

$$\mu_n(\mathrm{d}u) \propto \exp\left(-\frac{n}{N}\Phi(u)\right)\mu_0(\mathrm{d}u).$$

We note that if we evaluate this identity at $n = N$ then we obtain the Bayesian posterior distribution; the resulting algorithms are based on solving the associated filtering problem on interval $n = 0, 1, \ldots, N$. Now we observe that if, instead, we iterate $n \to \infty$ for fixed N (and hence fixed Δt), then μ_n will converge to a sum of Dirac measures supported at global minimizers of Φ that are contained in the support of μ_0. Thus we can iterate algorithms such as the Gaussian projected filter

from Section 4.4.2, or the ensemble Kalman filter from Section 4.4.3, to $n = \infty$, in order to obtain an approximate solution of the optimization problem for Φ, within the support of μ_0.

In the following example we study both the Gaussian projected filter and the ensemble Kalman filter in the linear Gaussian setting. We consider their application when we iterate $n \to \infty$ for fixed N. The two algorithms coincide in this linear Gaussian setting. Studying their properties gives insight into the proposed iterative approach to optimization. In particular it motivates the use of *regularization* and *variance inflation*, as introduced following the example.

Example 4.11. We consider the setting of Example 4.9, where the linear inverse problem with $G(u) = Lu$ is considered, making the additional assumption that $LC_0L^\top$ has full rank. The Gaussian projected filter equations (4.22), iterated over n steps, give the single-step update (4.24). Since $LC_0L^\top$ has full rank, this delivers the following closed-form update in the image of L:

$$Lm_n = Lm_0 + \Delta t LC_0L^\top \left(\frac{1}{n}\Gamma + \Delta t LC_0L^\top\right)^{-1} (w^\dagger - Lm_0), \tag{4.28a}$$

$$LC_nL^\top = LC_0L^\top - \Delta t LC_0L^\top \left(\frac{1}{n}\Gamma + \Delta t LC_0L^\top\right)^{-1} LC_0L^\top. \tag{4.28b}$$

If we fix Δt and let $n \to \infty$, then we see that

$$\begin{aligned} Lm_n &= w^\dagger + O(1/n), \\ LC_nL^\top &= O(1/n). \end{aligned}$$

We notice from the previous example that there are two issues when letting $n \to \infty$ for fixed Δt. First, the mean converges to a point m_∞ solving $Lm_\infty = w^\dagger$, rather than a minimizer of the regularized functional $\Phi_R(\cdot)$. Secondly, the convergence rate is only of order $1/n$. □

We now seek to address the two problems identified in this example, to develop improved methodology.

Regularization. The first problem identified in Example 4.11, namely that regularization disappears when taking $n \to \infty$, can be addressed by considering the iteration defined by

$$\mu_n(\mathrm{d}u) \propto \exp(-n\Delta t \Phi_R(u))\, \mu_0(\mathrm{d}u) \tag{4.29}$$

instead of the iteration defined by (4.16). Recalling G_R, Γ_R defined by (4.3) and assuming that $\Gamma_R \succ 0$, then a sequence of measures μ_n given by (4.29) may be generated by the filtering distribution associated with the following modification of (4.20):

$$u_{n+1} = u_n, \tag{4.30a}$$

$$y_{n+1} = \Delta t G_R(u_{n+1}) + \eta_{n+1}. \tag{4.30b}$$

Here $\eta_{n+1} \sim \mathsf{N}(0, \Delta t \Gamma_R)$, and we consider the setting where the observed data is $y^\dagger_{n+1} = \Delta t w^\dagger_R$. Note that $w^\dagger_R$ is the fixed vector defined in (4.5). We may apply sequential filtering techniques, such as the Gaussian projected filter and the ensemble Kalman filter, to the filtering problem defined by (4.30). Since μ_n converges to a Dirac delta distribution centred about the minimizer of Φ_R, within the support of μ_0, this addresses the first problem. However, the rate of convergence remains of order $1/n$ so that the second problem is not addressed; this is verified explicitly for the linear case in the forthcoming Example 4.15.

Variance inflation. The second problem identified in Example 4.11 is algebraic convergence. The root cause of the algebraic convergence is the collapse of the covariance C_n to zero. Therefore, in order to accelerate the convergence rate, we need to modify the sequential update steps to ensure that the covariance of (approximate) filters does not collapse to zero; at the same time we must ensure that the mean m_n still converges, exactly in the linear setting and approximately in the general nonlinear case, to the minimizer of Φ_R as $n \to \infty$.

In order to achieve this non-collapsing covariance we modify (4.30) by adding a form of variance inflation to the evolution of the parameter u_n, and consider the stochastic dynamical system given by

$$\boxed{\begin{aligned} u_{n+1} &= u_n + \xi_n, && \text{(4.31a)} \\ y_{n+1} &= \Delta t G_R(u_{n+1}) + \eta_{n+1}. && \text{(4.31b)} \end{aligned}}$$

Here $\xi_n \sim \mathsf{N}(0, \beta \Delta t \Sigma_n)$, $\beta \geq 0$ and $\eta_{n+1} \sim \mathsf{N}(0, \Delta t \Gamma_R)$ for covariance inflation matrix Σ_n to be defined. Note that if $\beta = 0$ we simply recover (4.30).

We now discuss the choice of Σ_n. Because the true covariance that we wish to recover is that of filtering distribution, it is natural that Σ_n is defined in terms of the covariance of the filter. Defining $Y^\dagger_n = \{y^\dagger_\ell\}_{\ell=1}^n$ with $y^\dagger_\ell := \Delta t w^\dagger_R$, we may consider the filtering distribution defined by random variable $u_n | Y^\dagger_n$. We let C_n denote the covariance under this filtered random variable. We then set $\Sigma_n := C_n$.

Remark 4.12. With this choice of C_n, equation (4.31) defines a form of mean-field model for state-observation evolution. Previously in this paper the state-observation models we have considered have not been of mean-field type; we only introduced mean-field models as the basis of sample-path-based algorithms to (approximately) solve a filtering problem. Here, in contrast, the mean-field dependence of the proposed model (4.31), with $\Sigma_n := C_n$, is through the filtering distribution associated with $u_n | Y^\dagger_n$. Thus, even before we develop mean-field models to approximate the law of the filtering distribution via sample-path-based algorithms, the underlying state-observation model is linked to filtering.

Although the filtering distribution depends on the history $Y^\dagger_n$, equation (4.31) can be rendered Markovian by coupling it to the evolution of the filtering distribution $\mu_n \mapsto \mu_{n+1}$, and noting that C_n is computed under μ_n.

We note that, in practice, identification of the exact covariance of the filtering distribution is not possible. Thus, in the Gaussian projected filter and ensemble Kalman filter that follow, we will use approximations of C_n; however, we will also denote these approximations by C_n to avoid proliferation of notation. □

We now derive implementable algorithms to approximate the filtering distribution defined by (4.31).

Gaussian projected filter. We begin by applying the ideas from Section 2.4, which concerns the Gaussian projected filter in the general setting, to the specific setting of the stochastic dynamical system (4.31). Using (2.42) in the specific setting of (4.31) yields

$$\widehat{m}_{n+1} = m_n, \tag{4.32a}$$
$$\widehat{C}_{n+1} = (1+\beta\Delta t)C_n, \tag{4.32b}$$
$$m_{n+1} = \widehat{m}_{n+1} + \Delta t \widehat{C}^{uG}_{R,n+1}\left(\Delta t \widehat{C}^{GG}_{R,n+1} + \Gamma_R\right)^{-1}\left(w^\dagger_R - \widehat{\mathrm{o}}_{n+1}\right), \tag{4.32c}$$
$$C_{n+1} = \widehat{C}_{n+1} - \Delta t \widehat{C}^{uG}_{R,n+1}\left(\Delta t \widehat{C}^{GG}_{R,n+1} + \Gamma_R\right)^{-1}\left(\widehat{C}^{uG}_{R,n}\right)^\top. \tag{4.32d}$$

Here we define

$$\widehat{\mathrm{o}}_{n+1} = \mathbb{E}G_R(\widehat{u}_{n+1}), \tag{4.33a}$$
$$\widehat{C}^{uG}_{R,n+1} = \mathbb{E}((\widehat{u}_{n+1} - \mathbb{E}\widehat{u}_{n+1}) \otimes (G_R(\widehat{u}_{n+1}) - \mathbb{E}G_R(\widehat{u}_{n+1}))), \tag{4.33b}$$
$$\widehat{C}^{GG}_{R,n+1} = \mathbb{E}((G_R(\hat{u}_{n+1}) - \mathbb{E}G_R(\widehat{u}_{n+1})) \otimes (G_R(\widehat{u}_{n+1}) - \mathbb{E}G_R(\widehat{u}_{n+1}))), \tag{4.33c}$$

where, in (4.32), all expectations are with respect to $\widehat{u}_{n+1} \sim \mathsf{N}(\widehat{m}_{n+1}, \widehat{C}_{n+1})$. Note that we have used the covariance of the Gaussian projected filter to define the variance inflation required to determine (4.32b), since we do not have the covariance under the true filtering distribution.

Ensemble Kalman filter. Instead of the Gaussian projected filter, we may use the ensemble Kalman filter. We use the covariance C_n of the ensemble Kalman filter to define the variance inflation since, again, we do not have the covariance under the true filtering distribution. With these considerations in hand, application of the stochastic Kalman transport mean-field model (2.55) to (4.31) yields

$$\widehat{u}_{n+1} = u_n + \xi_n, \tag{4.34a}$$
$$\widehat{y}_{n+1} = \Delta t G_R(\widehat{u}_{n+1}) + \eta_{n+1}, \tag{4.34b}$$
$$u_{n+1} = \widehat{u}_{n+1} + \widehat{C}^{uG}_{R,n+1}\left(\Delta t \widehat{C}^{GG}_{R,n+1} + \Gamma_R\right)^{-1}\left(\Delta t w^\dagger_R - \widehat{y}_{n+1}\right). \tag{4.34c}$$

Here $\xi_n \sim \mathsf{N}(0, \beta\Delta t C_n)$, $\eta_{n+1} \sim \mathsf{N}(0, \Delta t \Gamma_R)$ and expectations appearing in (4.33), to define $\left(\widehat{C}^{uG}_{R,n+1}, \widehat{C}^{GG}_{R,n+1}\right)$, are computed under the law of $\widehat{u}_{n+1}$.

Remark 4.13. Recall the mean-field dynamical system (4.31) and consider its filtering distribution. In Section 4.5.2 we show that in the linear case the mean of the filtering distribution converges to the posterior mean of the underlying Bayesian inverse problem, and hence to a minimizer of the Tikhonov regularized least-squares function Φ_R. For $\beta = 0$ convergence is algebraic, whilst it is exponential for $\beta > 0$. Furthermore, in Section 4.5.3, we show that for a particular choice of β, in the linear case the filtering distribution converges to the Bayesian posterior distribution defined by the inverse problem.

Recall that the Gaussian projected filter (4.32) and the mean-field ensemble Kalman filter (4.34) exactly reproduce the evolution of the filtering distribution, for linear Gaussian problems. As a consequence everything stated in this remark for the filtering distribution applies also to the law defined by (4.32) and by (4.34). □

We also note that it is possible to use corresponding deterministic transport formulations in place of (4.34). We employ the approximation (2.61). This holds in our case provided Δt is small enough. Choosing K_n as implicitly defined in (4.34), we obtain the following mean-field model:

$$\boxed{\begin{aligned}
\widehat{u}_{n+1} &= u_n + \xi_n, && \text{(4.35a)}\\
u_{n+1} &= \widehat{u}_{n+1} + \Delta t K_n \left(w_R^\dagger - \frac{1}{2}\left(G_R(\widehat{u}_{n+1}) + \widehat{o}_{n+1}\right)\right), && \text{(4.35b)}\\
K_n &= \widehat{C}^{uG}_{R,n+1}\left(\Delta t \widehat{C}^{GG}_{R,n+1} + \Gamma_R\right)^{-1}. && \text{(4.35c)}
\end{aligned}}$$

Here $\xi_n \sim \mathsf{N}(0, \beta\Delta t C_n)$. All expectations used to define $\left(\widehat{C}^{uG}_{R,n+1}, \widehat{C}^{GG}_{R,n+1}, \widehat{o}_{n+1}\right)$ are given by (4.33), computed under the law of $\widehat{u}_{n+1}$. This also exactly solves the filtering problem defined by (4.31), in the linear Gaussian setting.

Remark 4.14. We note that it is possible to replace (4.34) with the mean-field model

$$\boxed{\begin{aligned}
\widehat{u}_{n+1} &= u_n + \frac{\gamma}{2}(u_n - \mathbb{E}u_n), && \text{(4.36a)}\\
\widehat{y}_{n+1} &= \Delta t G_R(\widehat{u}_{n+1}) + \eta_{n+1}, && \text{(4.36b)}\\
u_{n+1} &= \widehat{u}_{n+1} + \widehat{C}^{uG}_{R,n+1}\left(\Delta t \widehat{C}^{GG}_{R,n+1} + \Gamma_R\right)^{-1}\left(\Delta t w_R^\dagger - \widehat{y}_{n+1}\right). && \text{(4.36c)}
\end{aligned}}$$

Here the expectation on u_n is with respect to the approximate filtering distribution generated by this model. Now note that the predictive mean and covariance defined by (4.34) are governed by (4.32a), (4.32b). The same equations govern the evolution of predictive mean and covariance of (4.36) provided that we choose γ to be the unique positive solution of the equation $\gamma + \gamma^2/4 = \Delta t\beta$. Thus the resulting methodology will coincide with the Gaussian projected filter and with the ensemble Kalman filter on linear Gaussian problems. □

4.5.2. Algorithms for optimization formulation

Recall that we have introduced the non-standard mean-field dynamical system (4.31). We have also described how the filtering distribution of the dynamical system may be approximated by the Gaussian projected filter (4.32) and the mean-field ensemble Kalman filter (4.34). In this subsection we substantiate the statements made about these algorithms in Remark 4.13. We initially discuss algorithms with algebraic convergence, for $\beta = 0$; and then we introduce a generalization of the analysis to $\beta > 0$ which allows us to obtain exponential convergence.

Algebraic convergence. Here we consider the setting of (4.31) where $\beta = 0$. Recall that for this choice of β, the stochastic dynamical system (4.31) reduces to (4.30). The following example illustrates that in the linear case where $G_R(\cdot) = L_R\cdot$, we recover an algebraic rate of convergence of the filtering distribution to a Dirac distribution at the posterior mean, when fixing Δt and taking $n \to \infty$. This is analogous to Example 4.11, which considers algorithms based on L, not L_R.

Example 4.15. Assume that u_0 is initialized at a Gaussian $\mathsf{N}(m_0, C_0)$ and assume also that $C_0, \Gamma_R \succ 0$. Consider the setting where $G_R(\cdot) = L_R\cdot$ for matrix $L_R \in \mathbb{R}^{(d_w+d_u)\times d_u}$. Now consider the filtering distribution $u_n|Y_n^\dagger$ given in (4.30), with $Y_n^\dagger$ data defined $y_n^\dagger = \Delta t w_R^\dagger$, with $w_R^\dagger$ as in (4.5).

The desired filtering distribution is Gaussian $\mathsf{N}(m_n, C_n)$. We now show that the iteration $(m_n, C_n) \mapsto (m_{n+1}, C_{n+1})$ converges to the posterior distribution, and does so at an algebraic rate. The reader should compare this with Example 4.11 which, using filtering based on G rather than G_R, and again in the linear case, also results in algebraic convergence; furthermore, convergence is only in the image space under the forward map L.

To prove convergence to the Dirac distribution we first identify the update equations for (m_n, C_n). Note that the predictive mean $\widehat{m}_{n+1}$ and covariance $\widehat{C}_{n+1}$ defined by (4.30a) trivially satisfy

$$\begin{aligned}\widehat{m}_{n+1} &= m_n,\\ \widehat{C}_{n+1} &= C_n.\end{aligned}$$

To find (m_{n+1}, C_{n+1}) it is again convenient to derive the formulae using precision rather than covariance matrices. To this end we view the Gaussian $\mathsf{N}(\widehat{m}_{n+1}, \widehat{C}_{n+1})$ as prior distribution for the linear inverse problem defined by (4.30b) with data realization $y_{n+1}^\dagger = \Delta t w_R^\dagger$. Note that the likelihood, being linear and Gaussian, is conjugate to the prior, so that the resulting posterior, which is the filtering distribution on $u_{n+1}|Y_{n+1}^\dagger$, is Gaussian with mean and covariance (m_{n+1}, C_{n+1}), which can be found by completing the square:

$$\begin{aligned}C_{n+1}^{-1} &= \widehat{C}_{n+1}^{-1} + \Delta t L_R^\top \Gamma_R^{-1} L_R,\\ C_{n+1}^{-1} m_{n+1} &= \widehat{C}_{n+1}^{-1} \widehat{m}_{n+1} + \Delta t L_R^\top \Gamma_R^{-1} w_R^\dagger.\end{aligned}$$

We therefore find that

$$C_n^{-1} = C_0^{-1} + n\Delta t L_R^\top \Gamma_R^{-1} L_R. \tag{4.37}$$

Recall that the posterior covariance $C_{\text{post}} = (L_R^\top \Gamma_R^{-1} L_R)^{-1}$ from (4.10) is positive definite. Hence it follows that the covariance converges to zero algebraically fast: $C_n = O(1/n)$.

Now note that

$$C_n^{-1} m_n = C_0^{-1} m_0 + n\Delta t L_R^\top \Gamma_R^{-1} w_R^\dagger,$$

so that

$$m_n = \left(C_0^{-1} + n\Delta t L_R^\top \Gamma_R^{-1} L_R\right)^{-1} \left(C_0^{-1} m_0 + n\Delta t L_R^\top \Gamma_R^{-1} w_R^\dagger\right).$$

We deduce that since the posterior mean is given by (4.11),

$$m_n = C L_R^\top \Gamma_R^{-1} w_R^\dagger + O(1/n) = m_{\text{post}} + O(1/n), \tag{4.38}$$

again exhibiting algebraic convergence. Combining (4.37) and (4.38) yields the result. □

Exponential convergence. Example 4.15 shows that when $\beta = 0$, filtering based on (4.31) leads to convergence to the posterior distribution at an algebraic rate, in the linear setting, when fixing Δt and taking $n \to \infty$. We now exhibit, when $\beta > 0$, an exponential rate of convergence to a Gaussian with correct posterior mean and a β-dependent scaled posterior covariance.

Proposition 4.16. Assume that u_0 is initialized at a Gaussian $\mathsf{N}(m_0, C_0)$ and assume also that $C_0, \Gamma_R \succ 0$. Consider the setting where $G_R(\cdot) = L_R \cdot$ for matrix $L_R \in \mathbb{R}^{(d_w+d_u)\times d_u}$. Now consider the filtering distribution $u_n | Y_n^\dagger$ defined by (4.31) for $\beta > 0$, with data $Y_n^\dagger$ defined by $y_n^\dagger = \Delta t w_R^\dagger$, where $w_R^\dagger$ is defined in (4.5). Then the filtering distribution is Gaussian $\mathsf{N}(m_n, C_n)$ for all $n \geq 1$. For any fixed $\Delta t > 0$, the mean and covariance converge at an exponential rate $(1 + \Delta t\beta)^{-n}$, as $n \to \infty$, to the limits $m_\infty = m_{\text{post}}$ and

$$C_\infty = \frac{\beta}{1 + \beta\Delta t} C_{\text{post}},$$

where $(m_{\text{post}}, C_{\text{post}})$ are the posterior mean (4.11) and covariance (4.10). ◇

Remark 4.17. Motivated by this proposition we may use the Gaussian projected filter (4.32) or the stochastic or deterministic Kalman transport algorithms, (4.34) and (4.35) respectively, to approximate the filtering distribution implied by (4.31). In so doing we generate approximate solutions of the Tikhonov regularized optimization problem defined by (4.1). Furthermore, the exact solution is recovered in the linear Gaussian setting. □

Remark 4.18. It is a remarkable fact that the rate of convergence is independent of the properties of the limiting Gaussian posterior distribution, and in particular of the conditioning of the posterior covariance. This desirable property is a result of

the *affine invariance* of the Gaussian projected filter and ensemble Kalman methods that we deploy in this subsection. Affine invariance is a subject we will study in more detail in the context of continuous-time approaches to inversion, developed in Section 5. In this context we note that Definition 5.11 may be extended to discrete-time algorithms. □

Proof of Proposition 4.16. We first identify the update equations for (m_n, C_n). Note that the predictive mean $\widehat{m}_{n+1}$ and covariance $\widehat{C}_{n+1}$ defined by (4.31a) satisfy

$$\widehat{m}_{n+1} = m_n, \tag{4.39a}$$

$$\widehat{C}_{n+1} = (1 + \beta\Delta t)C_n. \tag{4.39b}$$

To find (m_{n+1}, C_{n+1}) it is convenient to derive the formulae using precision rather than covariance matrices. To this end we view the Gaussian $\mathsf{N}(\widehat{m}_{n+1}, \widehat{C}_{n+1})$ as prior distribution for the linear inverse problem defined by (4.31b) conditioned on specific realization of the data $w^\dagger_{n+1} = w^\dagger_R$. Note that the likelihood, since linear and Gaussian, is conjugate to the prior so that the filtering distribution on $u_{n+1}|W^\dagger_{n+1}$ is Gaussian with mean and covariance (m_{n+1}, C_{n+1}) which can be found by completing the square

$$C_{n+1}^{-1} = \widehat{C}_{n+1}^{-1} + \Delta t L_R^\top \Gamma_R^{-1} L_R, \tag{4.40a}$$

$$C_{n+1}^{-1} m_{n+1} = \widehat{C}_{n+1}^{-1} \widehat{m}_{n+1} + \Delta t L_R^\top \Gamma_R^{-1} w^\dagger_R. \tag{4.40b}$$

Combining (4.39) and (4.40) shows that (m_n, C_n) update according to the formulae

$$C_{n+1}^{-1} = \left(\frac{1}{1+\beta\Delta t}\right) C_n^{-1} + \Delta t L_R^\top \Gamma_R^{-1} L_R,$$

$$C_{n+1}^{-1} m_{n+1} = \left(\frac{1}{1+\beta\Delta t}\right) C_n^{-1} m_n + \Delta t L_R^\top \Gamma_R^{-1} w^\dagger_R.$$

We can therefore write

$$C_n^{-1} = \left(\frac{1}{1+\beta\Delta t}\right)^n C_0^{-1} + \left(\sum_{k=0}^{n-1} \left(\frac{1}{1+\beta\Delta t}\right)^k\right) \Delta t L_R^\top \Gamma_R^{-1} L_R,$$

and so

$$C_n^{-1} = \left(\frac{1}{1+\beta\Delta t}\right)^n C_0^{-1} + \frac{1+\beta\Delta t}{\beta}\left(1 - \left(\frac{1}{1+\beta\Delta t}\right)^n\right) L_R^\top \Gamma_R^{-1} L_R. \tag{4.41}$$

Recall the posterior covariance $C_{\text{post}} = (L_R^\top \Gamma_R^{-1} L_R)^{-1}$ given in (4.10). It is clear that the precision converges exponentially fast to C_{post}^{-1}, scaled by $(1 + \beta\Delta t)/\beta$, and hence that the covariance converges exponentially fast to the appropriately scaled C_{post}.

Similarly we may write the expression for the mean as

$$C_n^{-1} m_n = \left(\frac{1}{1+\beta\Delta t}\right)^n C_0^{-1} m_0 + \frac{1+\beta\Delta t}{\beta}\left(1 - \left(\frac{1}{1+\beta\Delta t}\right)^n\right) L_R^\top \Gamma_R^{-1} w_R^\dagger,$$

so that the exponential convergence of the mean to the steady state m_{post} given by (4.11) may be deduced, using the expression (4.41). □

Remark 4.19. Recall the Hessian S of Φ_R, defined in (4.10). Using the formulae for the predictive mean and covariance, it is also easy to deduce that the expression for m_{n+1} is given by

$$m_{n+1} = \left(\frac{1}{1+\beta\Delta t}\right) \cdot m_n + \left(\frac{\beta\Delta t}{1+\beta\Delta t}\right) \cdot \mathsf{S}^{-1} L_R^\top \Gamma_R^{-1} w_R^\dagger,$$

and hence

$$m_{n+1} = m_n + \left(\frac{\beta\Delta t}{1+\beta\Delta t}\right) \cdot \mathsf{S}^{-1}\left(\mathsf{S} m_n + L_R^\top \Gamma_R^{-1} w_R^\dagger\right). \tag{4.42}$$

Since $\nabla\Phi_R(u) = \mathsf{S}u + L_R^\top \Gamma_R^{-1} w_R^\dagger$ and $D^2\Phi_R(u) = \mathsf{S}$, the iteration (4.42) may be viewed as a Gauss–Newton scheme for minimizing Φ_R. □

4.5.3. Algorithms for Bayesian formulation

Again recall that we have introduced the non-standard mean-field dynamical system (4.31) and shown how the filtering distribution of the dynamical system may be approximated by the Gaussian projected filter (4.32) and the mean-field ensemble Kalman filter (4.34). In this subsection we substantiate the statements made about these algorithms in Remark 4.13 in relation to Bayesian sampling. In particular we show that they exactly recover the posterior in the linear Gaussian setting by choosing

$$\beta = \frac{1}{1-\Delta t}. \tag{4.43}$$

The following is a direct corollary of Proposition 4.16. As in Remark 4.18, we note that the rate of convergence, in this case to the posterior distribution, is universal across all Gaussian posteriors.

Corollary 4.20. Assume that u_0 is initialized at a Gaussian $\mathsf{N}(m_0, C_0)$ and assume also that $C_0, \Gamma_R > 0$. Consider the setting where $G_R(\cdot) = L_R \cdot$ for matrix $L_R \in \mathbb{R}^{(d_w+d_u)\times d_u}$. Now consider the filtering distribution $u_n | Y_n^\dagger$ defined by (4.31) for β given by (4.43), with data $Y_n^\dagger$ defined by $y_n^\dagger = \Delta t w_R^\dagger$, where $w_R^\dagger$ is defined in (4.5). Then the filtering distribution is Gaussian $\mathsf{N}(m_n, C_n)$ for all $n \geq 1$. For any fixed $\Delta t > 0$ the mean and covariance converge at an exponential rate $(1-\Delta t)^n$, as $n \to \infty$, to the limits $m_\infty = m_{\text{post}}$ and $C_\infty = C_{\text{post}}$, where $(m_{\text{post}}, C_{\text{post}})$ are the posterior mean (4.11) and covariance (4.10). ◇

Remark 4.21. Motivated by this corollary, in the context of (4.31) we may use the Gaussian projected filter (4.32) or the stochastic or deterministic Kalman transport algorithms, (4.34) and (4.35) respectively, to generate approximate solutions of the Bayesian inverse problem defined by (4.1), in the general nonlinear setting. Of course, because these algorithms employ Gaussian approximations, this does not produce the exact filtering distribution. Furthermore, to make resulting algorithms tractable we will predict in (4.31a) using the covariance C_n of the Gaussian projected filter or the Kalman transport algorithm, rather than the covariance under the true filtering distribution, which is not tractable: see the final paragraph in Remark 4.12. We note, however, that in the linear Gaussian case all the algorithms are exact and hence recover the true posterior. □

4.6. Ensemble Kalman methods for inversion: examples

In this section we have so far concentrated entirely on mean-field models, leaving the details of deriving particle approximations to the reader. In this subsection, however, we make a brief foray into finite particle ensemble approximations of the mean-field models introduced in our discussion of inverse problems. The algorithms we employ can be found in Appendix A as Algorithms 3, 4 and 5.

Algorithms 3 and 4 are based on finite particle approximations of the mean-field model in (4.25). Algorithm 3 is based on iterating until N satisfying $N\Delta t = 1$ and, at that time-step, aims to approximate the posterior. Algorithm 4 performs the same iteration but to N_∞ assumed to satisfy $N_\infty \Delta t \gg 1$ so that it approximately solves an optimization problem; see the discussion at the start of Section 4.5.1. Algorithm 5 is based on (4.34) with β given by (4.43), and aims to approximate the posterior by iterating to N_∞ : $N_\infty \Delta t \gg 1$.

In Example 4.22 we study a one-dimensional nonlinear inverse problem; working in one dimension enables comparison of the true posterior distribution with approximations arising from the various ensemble Kalman inversion schemes described in preceding subsections. We also demonstrate an optimization approach to inversion, in the context of Example 4.22. Subsequently, in Example 4.23, we return to the setting of the Lorenz '96 dynamical system, now to estimate unknown parameters rather than the state; our focus is on studying ensemble Kalman methods from the perspective of the optimization approach to the parameter estimation problem.

Example 4.22. Recall Example 4.9 concerning the linear Gaussian inverse problem; there we show that the exact posterior is obtained either by iterating (4.23) N times or by evaluating (4.24) at $n = N$. Although derived in the context of the Gaussian projected filter, the example also applies to mean-field ensemble Kalman methods since they, like the Gaussian projected filter, are exact for linear Gaussian problems. However, in the nonlinear case this equivalence does not hold exactly, because of approximations that are made by the Gaussian projected and ensemble Kalman filtering methods. In this example we study the effect of these

approximations by examining the behaviour of particle-based ensemble Kalman methods applied to a nonlinear inverse problem.

We consider the setting of (4.1) with a nonlinear forward map $G : \mathbb{R} \to \mathbb{R}$, given by

$$G(u) = \frac{7}{12}u^3 - \frac{7}{2}u^2 + 8u. \tag{4.44}$$

The observational noise is assumed to be of the form $\eta \sim \mathsf{N}(0, 1)$. Assuming observation $w^\dagger = 2$ results in likelihood

$$\frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}(G(u) - 2)^2\right). \tag{4.45}$$

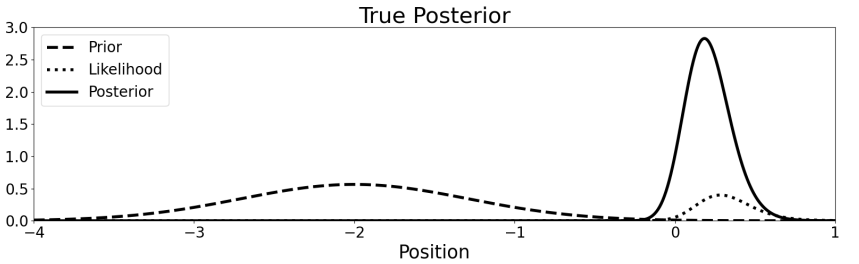
Furthermore, assuming a Gaussian prior of mean -2 and variance $1/2$, the posterior on $u|w^\dagger$ is proportional to

$$\frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}(G(u) - 2)^2 - (u + 2)^2\right). \tag{4.46}$$

We note that the forward map $G(u)$ is monotonic and the posterior (4.46) is unimodal: in Figure 4.1(a) we display the posterior distribution. We also show the Gaussian prior and the likelihood (4.45).

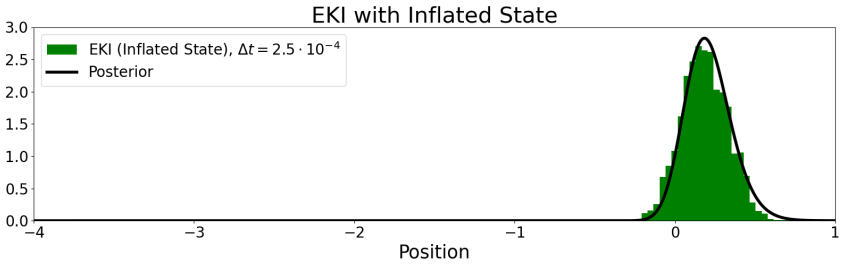
Figure 4.1(b) shows the ensemble Kalman inversion iteration (4.25) from Section 4.4, which is designed to transport prior to posterior in finite time, fixing N iterations and Δt so that $N\Delta t = 1$; see Algorithm 3. Recall that the iteration exactly recovers the posterior, in the mean-field limit, when applied to linear Gaussian inverse problems (Example 4.9), but that here the inverse problem is nonlinear and non-Gaussian. We employ this algorithm with $J = 2 \times 10^3$ ensemble members. We run Algorithm 3 with two choices of N: $N = 4 \times 10^3$ and hence $\Delta t = 2.5 \times 10^{-4}$, and with $N = 1$ and hence $\Delta t = 1$. Figure 4.1(b) shows that the one-step approach, with $N = 1$, leads to a very poor approximation of the posterior. In contrast, the scheme with $N = 4 \times 10^3$ yields reasonable approximation quality of the posterior; see Remark 4.6. In this panel we also run Algorithm 4 for the finite number of iterations $N_\infty = 10^6$, with $\Delta t = 2.5 \times 10^{-4}$. The result reflects the theoretical interpretation: iterating $n \to \infty$ leads to solution of the optimization formulation of ensemble Kalman inversion and, as discussed in Section 4.5, results in convergence to a Dirac centred at the minimizer of the least-squares functional Φ, the maximum likelihood estimate found by maximizing (4.45); this simply delivers the point $G^{-1}(2)$.

Figure 4.1(c) shows the ensemble Kalman inversion iteration (4.34) from Section 4.5, namely Algorithm 5. This is designed to approximate the true posterior, when $1/\beta = 1 - \Delta t$: indeed, Corollary 4.20 shows that in the linear setting the mean-field model (4.34) converges to the true posterior in limit $n \to \infty$, with this choice of β. Our numerical results, which are conducted with this choice of β, show that use of the Algorithm 5, when applied to the nonlinear inverse problem, produces an excellent posterior approximation; this demonstrates that the linear



(a)

(b)



(c)

Figure 4.1. The plots display the results obtained for the inverse problem described by the one-dimensional nonlinear forward map (4.44). (a) The prior, likelihood and posterior. (b) Comparison of the true posterior PDF with the approximation obtained using the finite ensemble Kalman inversion iteration (4.25), as detailed in Algorithm 3, with $\Delta t = 2.5 \times 10^{-4}$ and $\Delta t = 1$, iterated to time $n = N$ where $N\Delta t = 1$. Panel (b) also includes results found from applying the optimization Algorithm 4, iterating over $N_\infty = 10^6$ steps with $\Delta t = 2.5 \times 10^{-4}$; in this case, the ensemble approaches the Dirac measure supported on the minimizer of the unregularized least-squares loss Φ, given by the peak of the likelihood. (c) The posterior approximation using the ensemble Kalman inversion iteration with covariance inflation as described by the mean-field model (4.34), as detailed in Algorithm 5, with $\beta = 1/(1 - \Delta t)$ for parameter $\Delta t = 2.5 \times 10^{-4}$; the algorithm is iterated over $N_\infty = 10^5$ steps. It is clear that this scheme yields the highest-quality posterior approximation.

theory is indicative of the behaviour of the algorithm beyond the linear Gaussian setting. □

Example 4.23. As in the examples from Section 2, we again consider the Lorenz '96 (single-scale) model for $v \in C(\mathbb{R}^+, \mathbb{R}^L)$ satisfying the equations

$$\dot{v}_\ell = -v_{\ell-1}(v_{\ell-2} - v_{\ell+1}) - v_\ell + u + h_v m(v_\ell), \quad \ell = 1 \dots L, \tag{4.47a}$$

$$v_{\ell+L} = v_\ell, \quad \ell = 1 \dots L. \tag{4.47b}$$

As before, we set $L = 9$, $h_v = -0.8$ and $u = 10$. We recall that function m is shown in Figure 2.1. In Section 2 we focused on recovering the state v from partial and noisy observations. Here we concentrate on recovering the parameter u.[24]

Our objective is to recover parameter u from time-averaged data. We assume that the system is ergodic so that infinite time-averages produce averages over the invariant measure. Furthermore, we assume that convergence in time, of averages, is governed by a central limit theorem. We let $G_T \colon \mathbb{R} \to \mathbb{R}^2$ denote the mean and variance, defined via averaging over time T and over the L components of v, of the state of system (4.47). In principle G_T depends also on initialization, but this effect is negligible for T large, and zero for $T = \infty$, by ergodicity. In particular, with system state $v^\dagger$ evolving according to

$$v^\dagger_{n+1} = \Psi(v^\dagger_n), \tag{4.48}$$

where Ψ is the solution operator for (4.47) over the observation time interval τ, with true parameter $u = u^\dagger$, the action of forward operator G_T on u is defined as follows:

$$G_T(u) = \begin{pmatrix} w_1 \\ w_2 \end{pmatrix},$$

with, for $M\tau = T$,

$$w_1 = \frac{1}{L}\sum_{l=1}^{L} \bar{v}^\dagger_l, \quad w_2 = \frac{1}{L \cdot M}\sum_{n=1}^{M}\sum_{l=1}^{L} \left(v^\dagger_{n;l} - \bar{v}^\dagger_l\right)^2, \quad \bar{v}^\dagger_l = \frac{1}{M}\sum_{n=1}^{M} v^\dagger_{n;l},$$

where we have used $v^\dagger_{n;l}$ to denote the lth variable in vector $v^\dagger_n$.

We consider finding u from an observation $w \in \mathbb{R}^2$ arising from the model

$$w = G_\infty(u) + \gamma. \tag{4.49}$$

In practice the specific realization $w^\dagger$, from which we invert to find u, is found by integrating G_T to a finite time $T = 100$, not $T = \infty$. Variable $\gamma \sim \mathrm{N}(0, \Gamma)$ accounts for the resulting central limit theorem correction. To solve the parameter estimation problem for u, we use ensemble Kalman methods in Algorithms 3, 4 and 5. We do not have access to G_∞ and so, instead, the algorithms are implemented by using

[24] We have used the notation u instead of F, for the forcing parameter, to align with the notation for the unknown parameter used throughout the section concerning inverse problems.

G_T with $T = 10$, initialized after a burn-in time of duration $t^* = 10$. The burn-in phase itself results from an initial condition chosen at random from a Gaussian distribution with mean 0 and standard deviation 40. In the experiments shown we take $\Gamma = \sigma^2 I$, with $\sigma = 10^{-1}$. All the ensemble Kalman inversion schemes are initialized from a prior Gaussian of mean 0 and standard deviation 10, and use an ensemble of size $J = 30$.

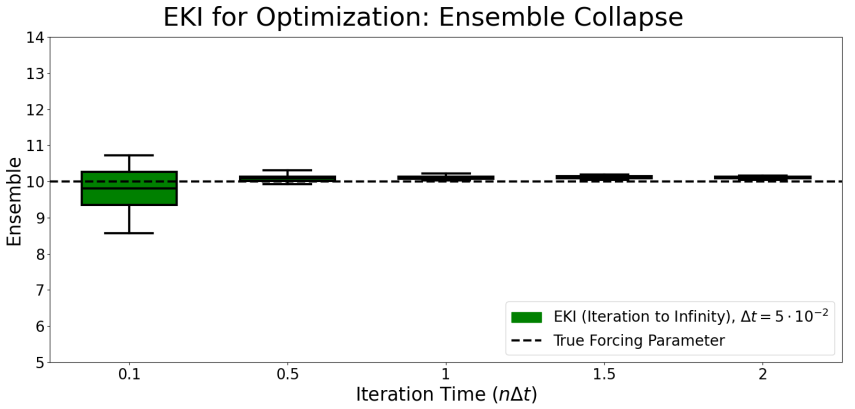
In Figure 4.2 we display the ensemble approximations obtained via application of the EKI methodology for optimization, namely Algorithm 4. In practice, the scheme is run for a finite number N_∞ of iterations. Indeed, in Figure 4.2(a) the scheme is applied with $\Delta t = 5 \times 10^{-2}$ and run up to $N_\infty = 40$ iterations. On the other hand, in Figure 4.2(b) the scheme is run for 20 iterations with $\Delta t = 1$. In both settings ensemble collapse occurs as the number of iterations grow. The ensemble mean, displayed as the central line in each box plot, converges to a point yielding a qualitatively good estimate of the true forcing parameter, with an error of $O(10^{-1})$.

In Figure 4.3 we show an application of Algorithm 3 with Δt set to 5×10^{-2}, running for $N = 20$ steps ($N\Delta t = 1$), and of Algorithm 5 with $1/\beta = 1 - \Delta t$ and $\Delta t = 5 \times 10^{-2}$ for $N_\infty = 40$ steps. For linear inverse problems, the output of both algorithms at these specific steps delivers the posterior distribution exactly in the mean-field limit, by Example 4.9 and Corollary 4.20. As noted in the previous paragraph, such a posterior approximation should be interpreted with caution for this nonlinear inverse problem. However, we show in Figure 4.3 that the ensemble means accurately predict the true forcing up to an error of $O(10^{-1})$ and that, furthermore, the two ensembles are similar. However, interpreting the posterior distributions in this case is harder as we do not have access to the true posterior. We note that in Example 4.22 we were able to demonstrate that Algorithm 5 delivered a better posterior approximation than Algorithm 3 and it would be interesting to determine whether such a conclusion holds more generally. □

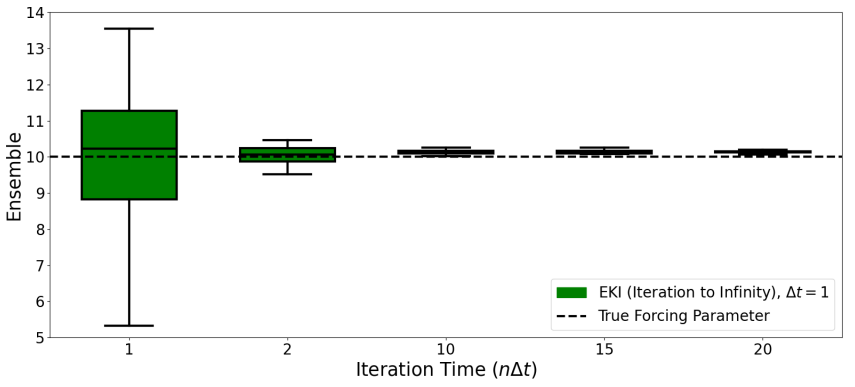
4.7. Bibliographical notes

Sequential Monte Carlo methods may be used to approximately morph one probability distribution (source) into another (target), using empirical approximation and a discrete-time homotopy (Del Moral *et al.* 2006, Chopin and Papaspiliopoulos 2020). In general the methodology does not scale well to high-dimensional problems (Beskos, Crisan and Jasra 2014). However, some success has been achieved in this direction (Kantas, Beskos and Jasra 2014), and a basic underlying theory is described in Beskos, Jasra, Muzaffer and Stuart (2015). Our presentation in this paper is confined to the setting of ensemble Kalman methods because of their empirical success and scalability to high dimensions.

The development of ensemble Kalman methods for inverse problems was pioneered in the study of reservoir simulation, in the context of learning subsurface properties from localized flow measurements (Chen and Oliver 2012, Gu and Oliver 2007, Li and Reynolds 2009, Emerick and Reynolds 2013*a*,*b*, Evensen



(a)



(b)

Figure 4.2. The figure displays box and whisker plots for the ensembles produced by Algorithm 4 with $\Delta t = 5 \times 10^{-2}$ (a) and $\Delta t = 1$ (b). The box and whisker plots represent the ensembles by depicting the ensemble mean, as a line within the shaded region. Furthermore, the edges of the boxes represent the first and third quartiles, i.e. the values of ensemble members corresponding to the median of the first half of the samples, and the median of the second half of the samples, respectively. Finally, the whiskers mark the furthest samples lying within a distance from the box of 1.5 times the distance between the first and third quartiles (the interquartile range). In both cases we note ensemble collapse onto a value close to the truth underlying the data.

2018). Subsequent work has studied parameter estimation in chaotic dynamical systems, such as those arising in weather forecasting using ensemble methods for joint state and parameter estimation (Pulido *et al.* 2018, Bocquet, Brajard, Carrassi and Bertino 2020, Gottwald and Reich 2021), and by matching to time-averaged statistics (Schneider, Lan, Stuart and Teixeira 2017, Cleary *et al.* 2021, Dunbar,

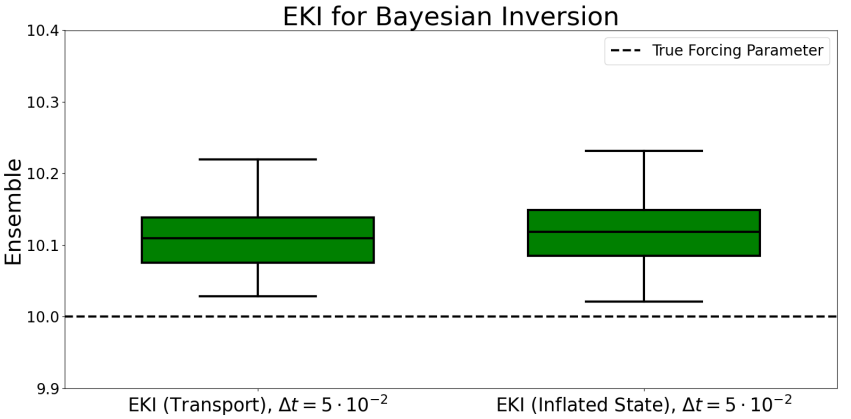


Figure 4.3. The figure displays box and whisker plots for the Bayesian posterior ensemble approximation produced by Algorithm 3 with $\Delta t = 5 \times 10^{-2}$ and Algorithm 5 with $\Delta t = 5 \times 10^{-2}$. The box and whisker plots represent the ensembles by depicting the ensemble mean, as a line within the shaded region. Furthermore, the edges of the boxes represent the first and third quartiles, i.e. the values of ensemble members corresponding to the median of the first half of the samples, and the median of the second half of the samples, respectively. Finally, the whiskers mark the furthest samples lying within a distance from the box of 1.5 times the distance between the first and third quartiles (the interquartile range). The posterior mean in both cases is close to the true value of the parameter underlying the data; the size of the posterior spread is also similar. Note, however, that in this problem we do not have a true posterior distribution against which to compare.

Garbuno-Inigo, Schneider and Stuart 2021), motivated by climate modelling. The chaotic dynamics that underlie weather and climate models lead to complicated energy landscapes for minimization and sampling (Lea, Allen and Haine 2000). Huang, Schneider and Stuart (2022*b*) and Dunbar, Duncan, Stuart and Wolfram (2022) demonstrate the benefits of using ensemble methods for such problems, rather than computing exact derivatives: the ensemble approach effectively works in a smoothed energy landscape.

The idea of transporting prior to posterior, as developed in Section 4.4.1, is widely used in the statistics literature; see Chopin and Papaspiliopoulos (2020) and Del Moral *et al.* (2006) for a unified perspective and for citations to earlier works which characterize the deformation of one measure to the other either through incrementally building up the available data, or through a temperature-like annealing parameter in the likelihood. In the context of data assimilation, and ensemble Kalman methods in particular, these ideas were developed by Li and Reynolds (2009), Gu and Oliver (2007), Daum *et al.* (2010), Reich (2011) and Sakov *et al.* (2012).

Iglesias, Law and Stuart (2013) highlight how the invariant subspace property of finite ensemble Kalman methods (Anderson 2001) may be viewed, in the context

of inverse problems, as a form of regularization: the iteration remains in the linear span of the initial ensemble. Thus it is possible to study maximum likelihood estimation, regularized by restriction to this subspace. More recent methodology is developed in the context of optimization in Huang *et al.* (2022*b*), although the methodology therein is not affine-invariant; it is also developed for sampling in Huang, Huang, Reich and Stuart (2022*a*), resulting in an affine-invariant methodology. See those papers for details concerning uniqueness of, and exponential convergence to, steady-state solutions in the linear Gaussian setting. Here we have adopted the approach of Huang *et al.* (2022*a*) to both optimization and sampling, which yields exponential convergence under both settings, as discussed in Proposition 4.16. A geometric picture of iterative ensemble Kalman methods for inverse problems has been developed by Qian and Beattie (2024), who consider the fundamental observed and unobserved subspaces defined by linear inverse problems and their interaction with the invariant subspace defined by ensemble Kalman iteration.

The stochastic perturbations utilized in (4.31) are closely related to multiplicative ensemble inflation methods, as widely used in ensemble Kalman filter implementations (Asch *et al.* 2016, Evensen *et al.* 2022). The effects of additive inflation and variable step-size implementations of ensemble Kalman inversion have been studied in Chada and Tong (2022) and in Weissmann, Chada, Schillings and Tong (2022).

The optimization and sampling approaches for inverse problems can in principle be combined with ideas from stochastic annealing (Kushner and Yin 2003) and stochastic gradient descent (Goodfellow, Bengio and Courville 2016). Haber *et al.* (2018), Kovachki and Stuart (2019) and Pidstrigach and Reich (2023) demonstrate the use of ensemble Kalman methods for inversion, when combined with stochastic gradient descent, and mini-batching in particular, as well as the application of ensemble Kalman methods beyond the setting of the L_2-loss functions Φ and Φ_R considered here.

However, despite the growing use of ensemble Kalman methods to solve inverse problems, it is important to appreciate that all ensemble Kalman-based methods invoke approximations which amount to matching only first- and second-order statistics, at some point in the algorithmic development. For this reason the methods are intuitively only useful as samplers for problems with posterior distribution close to a Gaussian. This idea is carefully explained in Ernst, Sprungk and Starkloff (2015) where the mean-field limit of ensemble Kalman methods for inverse problems is compared with the desired posterior distribution; as is the case for state estimation, analysis is required to justify use of ensemble Kalman methods beyond the linear and Gaussian regime. We also highlight that our analysis in this paper, which focuses on the mean-field limit, does not capture important aspects of the performance of ensemble Kalman methods at finite ensemble size, and the important practical issue of covariance localization; these issues are studied in Al-Ghattas and Sanz-Alonso (2023). Furthermore, Tong and Morzfeld (2023) study the relationship between localization, in the solution of inverse problems using

ensemble Kalman methods, and the subspace property explained in Iglesias *et al.* (2013, Theorem 2.1). As an alternative to localization, the concept of dropout combined with variable step-size has been shown to lead to optimal algorithmic performance in Liu, Reich and Tong (2025).

The idea of using ensemble methods for performing the optimization step within variational data assimilation was introduced in Zupanski (2005). The connection between iterative applications of the ensemble Kalman filter and optimization were first investigated in Iglesias *et al.* (2013), and developed to include constraints in Albers *et al.* (2019) and Chada, Schillings and Weissmann (2019), and Tikhonov regularization in Chada, Stuart and Tong (2020). Recall from Section 4.5.2 the algebraic rates of convergence arising in the basic optimization method arising from iterating to infinity. This undesirable feature of optimization methods based on statistical linearization of mean-field gradient descent can be ameliorated to some extent by the use of adaptive time-steps, connections to the Levenberg–Marquardt algorithm and the use of stopping criteria; see Iglesias (2015, 2016) and Iglesias and Yang (2021). Recent work of Parzer and Scherzer (2022) has developed a systematic theory for early stopping using ensemble Kalman inversion including incorporation of Nyström methodology. Other interacting particle system approaches to optimization have been proposed, including feedback particle (Zhang 2013, Zhang, Taghvaei and Mehta 2017), unscented Kalman approaches (Huang *et al.* 2022*a*,*b*) and consensus-based optimization (Tsianos, Lawlor and Rabbat 2012, Carrillo *et al.* 2018, Fornasier, Huang, Pareschi and Sünnen 2020, Ha, Jin and Kim 2021).

Finally we note that Kalman methods have been related to approximate Bayesian computation (ABC) methodologies, utilizing a linear regression ansatz (Sisson, Fan and Beaumont 2018, Nott, Marshall and Ngoc 2012). Such methods are in turn closely related to Bayes linear and best linear unbiased estimators (BLUE) as discussed in Goldstein and Rougier (2006), Goldstein and Wooff (2007), Lei and Bickel (2011), Nott *et al.* (2012), Snyder (2014), Goldstein (2014), Reich and Cotter (2015) and Latz (2016). BLUE is discussed in more detail in Appendix C.3.

5. Inverse problems: continuous time

In this section we derive continuous-time limits of the ideas developed in the preceding section for the solution of inverse problems. As a consequence the ideas may also be viewed as adaptations of Section 3 to the solution of inverse problems. We start in Section 5.1 by recalling the inverse problem, followed in Sections 5.2 and 5.3 by discussion of the optimization and Bayesian approaches respectively, focusing on gradient flows; this flow perspective provides a conceptual basis for thinking about the algorithms for inverse problems that we will subsequently develop. Section 5.4 is devoted to Bayesian probabilistic filtering methods which solve the inverse problem by morphing the prior into the posterior in finite time. Section 5.5 discusses filtering methods which work on infinite-time horizons,

exhibiting exponential convergence to approximate solutions of the optimization or Bayesian formulations of the problem, from arbitrary starting points. Analogously to Section 4, we demonstrate application of Gaussian projected filtering and ensemble Kalman methods to solve the filtering problems defined in Sections 5.4 and 5.5; the comments from Remark 4.1 apply here too. We conclude in Section 5.6 with bibliographical notes.

5.1. *Set-up*

Recall the inverse problem (4.1) of recovering u from w, where

$$\boxed{w = G(u) + \gamma,}$$

introduced in full detail in Section 4.1. Under the assumptions laid out in Section 4.3, and in particular Gaussianity and independence of unknown parameter u and noise η, and given a specific realization $w^\dagger$ of data w, we have a posterior distribution on the random variable $u|w^\dagger$ which is defined by[25]

$$\boxed{\begin{aligned} \mu(\mathrm{d}u) &= \frac{1}{\mathcal{Z}} \exp(-\Phi_R(u))\,\mathrm{d}u, && \text{(5.1a)} \\ \mathcal{Z} &= \int_{\mathbb{R}^{d_u}} \exp(-\Phi_R(u))\,\mathrm{d}u. && \text{(5.1b)} \end{aligned}}$$

Here

$$\boxed{\begin{aligned} \Phi(u) &= \frac{1}{2}|w^\dagger - G(u)|^2_\Gamma, && \text{(5.2a)} \\ \Phi_R(u) &= \Phi(u) + \frac{1}{2}|u - m_0|^2_{C_0}, && \text{(5.2b)} \end{aligned}}$$

for prior mean vector vector m_0, prior covariance matrix $C_0 \succ 0$ and noise covariance matrix $\Gamma \succ 0$.

Rather than solving the Bayesian inverse problem, which can be prohibitively expensive, optimization methods may be developed to find a point estimate of $u|w^\dagger$ as minimizer of Φ over a compact set, or as minimizer of Φ_R over the whole space $\mathbb{R}^{d_u}$. The next two sections show, respectively, how we may develop continuous-time gradient flows which minimize Φ_R, or Φ, and gradient flows which find the posterior distribution μ.

5.2. *Optimization formulation: gradient flows*

The goal of this subsection is to study gradient flows to minimize an objective function. For us particular focus is on the choice of Φ or Φ_R as objective, but since some of the considerations are quite general, we frame aspects of the discussion in a general setting.

[25] The normalization constant $\mathcal{Z}$ is the probability of the observed data under the model, sometimes called the *evidence*.

Deterministic viewpoint. Consider the standard gradient descent, applied to an energy function $\Psi\colon \mathbb{R}^{d_u} \to \mathbb{R}^+$, namely

$$\frac{\mathrm{d}u}{\mathrm{d}t} = -\nabla\Psi(u). \tag{5.3}$$

Note that this may be found as the continuous-time limit of the discrete-time gradient descent algorithm (4.6), choosing $\alpha = \Delta t$, letting $u(n\Delta t) = u_n$ and taking the limit $\Delta t \to 0$.

Along solutions of (5.3),

$$\frac{\mathrm{d}}{\mathrm{d}t}\Psi(u) = \left\langle \nabla\Psi(u), \frac{\mathrm{d}u}{\mathrm{d}t} \right\rangle \tag{5.4a}$$

$$= -\left|\frac{\mathrm{d}u}{\mathrm{d}t}\right|^2. \tag{5.4b}$$

Equation (5.3) is said to possess a gradient flow structure in parameter space $\mathbb{R}^{d_u}$ because the vector field driving the evolution of u is tangential to the gradient of the *energy* $\Psi(u)$ in the standard Euclidean *metric*; this is the geometric reason for the non-increasing property of $\Psi(u)$ along trajectories.

Geometric perspective on the space of probability densities. We now define gradient flow structure from a probabilistic viewpoint, studying evolution of probability densities. Again we need both an energy and a metric. To this end we introduce some notation that will be useful in the probabilistic formulation of (5.3). It will also be used more generally in subsequent discussion of other gradient flows on the space of probability density functions.

We denote the manifold of all smooth probability density functions on $\mathbb{R}^{d_u}$ by $\mathfrak{P}_+ = \mathfrak{P}_+(\mathbb{R}^{d_u})$. We may then define the tangent space $T_\rho\mathfrak{P}_+$ to $\mathfrak{P}_+$, at $\rho \in \mathfrak{P}_+$, by

$$T_\rho\mathfrak{P}_+ = \left\{\sigma \in C^\infty(\mathbb{R}^{d_u})\colon \int_{\mathbb{R}^{d_u}} \sigma(u)\,\mathrm{d}u = 0\right\}. \tag{5.5}$$

Given the tangent space $T_\rho\mathfrak{P}_+$, we define its dual[26]

$$T_\rho\mathfrak{P}_+^\star = \left\{\psi \in C^\infty(\mathbb{R}^{d_u})\colon \int_{\mathbb{R}^{d_u}} \psi(u)\,\rho(u)\,\mathrm{d}u = 0\right\}. \tag{5.6}$$

In this article, for simplicity of exposition, we will define underlying metric structure via the positive operator $\mathcal{M}(\rho)\colon T_\rho\mathfrak{P}_+ \to T_\rho\mathfrak{P}_+^\star$. A precise mathematical treatment requires further assumptions on the considered set $\mathfrak{P}_+$. See the bibliography for relevant literature on this topic. Operator $\mathcal{M}(\rho)$ may be linked to an underlying Riemannian metric tensor $g_\rho\colon T_\rho\mathfrak{P}_+ \times T_\rho\mathfrak{P}_+ \to \mathbb{R}$; however, since this

[26] This informal definition of tangent space, and its dual space, requires careful handling for probability measures on non-compact manifolds, such as $\mathbb{R}^{d_u}$; see citations to the literature in Section 5.6. The dual is also known as the cotangent space.

metric tensor plays no role in our presentation, we will work directly with $\mathcal{M}(\rho)$, and with its inverse $\mathcal{M}(\rho)^{-1}: T_\rho\mathfrak{P}_+^\star \to T_\rho\mathfrak{P}_+$. Note that $\mathcal{M}(\rho)^{-1}$ maps into the tangent space $T_\rho\mathfrak{P}_+$, implying that it maps into functions that integrate to zero over $\mathbb{R}^{d_u}$; see (5.5).

Probabilistic viewpoint. We now demonstrate gradient structure inherent in the probabilistic viewpoint on the ODE (5.3), arising from allowing the initial condition $u(0)$ to be random. We will show that the Liouville equation governing the evolution of the probability density function associated with the random variable $u(t)$ also has a gradient structure. In so doing we must exhibit an appropriate energy and metric.

We assume that $u(t)$ has smooth probability density $\rho(u,t)$ for all $t \geq 0$. Then ρ satisfies the Liouville equation

$$\partial_t \rho = \nabla \cdot (\rho \nabla \Psi). \tag{5.7}$$

The energy and metric defining the gradient structure for equation (5.7) are

$$\mathcal{E}(\rho) := \int_{\mathbb{R}^{d_u}} \Psi(u)\rho(u)\,\mathrm{d}u, \tag{5.8a}$$

$$\mathcal{M}(\rho)^{-1}\psi := -\nabla \cdot (\rho \nabla \psi) \in T_\rho\mathfrak{P}_+. \tag{5.8b}$$

Operator $\mathcal{M}(\rho)$ corresponds to an underlying *Wasserstein-2 metric structure.*

The standard variational derivative[27] of $\mathcal{E}$ is given by

$$\frac{\delta \mathcal{E}}{\delta \rho} = \Psi.$$

The restriction of this variational derivative to the dual space $T_\rho\mathfrak{P}_+^\star$ is provided by

$$\frac{\delta \mathcal{E}}{\delta \rho}_{|T_\rho\mathfrak{P}_+^*} = \Psi - \mathbb{E}[\Psi]. \tag{5.9}$$

In the context of the Wasserstein-2 metric structure as presented here it is not necessary to distinguish between these two formulations of the variational derivative; however, the second formulation allows for unique solvability of the elliptic equation required to define $\mathcal{M}(\rho)$. For the Fisher–Rao metric structure considered later in Section 5.4, the second definition will play a more direct role. Hence we can rewrite (5.7) as

$$\partial_t \rho = \nabla \cdot \left(\rho \nabla \frac{\delta \mathcal{E}}{\delta \rho} \right). \tag{5.10}$$

This may be written abstractly as

$$\partial_t \rho = -\mathcal{M}(\rho)^{-1} \frac{\delta \mathcal{E}}{\delta \rho}(\rho). \tag{5.11}$$

[27] The variational derivative is identified by writing $\mathcal{E}(\rho+\sigma) - \mathcal{E}(\rho)$ as a linear operator acting on σ (plus higher-order terms in σ for energies $\mathcal{E}(\rho)$ which are not linear in ρ).

From this it follows that

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{E}(\rho) = \left\langle \frac{\delta \mathcal{E}}{\delta \rho}(\rho), \frac{\partial \rho}{\partial t} \right\rangle = -\left\langle \mathcal{M}(\rho)\frac{\partial \rho}{\partial t}, \frac{\partial \rho}{\partial t} \right\rangle \leq 0. \tag{5.12}$$

Hence the energy is decreasing along trajectories and the gradient structure is apparent.

Remark 5.1. It is interesting to compare the gradient structure (5.12) on $\mathfrak{P}_+$ to the gradient flow structure on $\mathbb{R}^{d_u}$ defined by (5.4). The state space gradient flow on $\mathbb{R}^{d_u}$ ensures decrease of $\Psi(u(t))$ along trajectories whilst the probability space gradient flow on $\mathfrak{P}_+$ ensures decrease of the expected value of $\Psi(u(t))$ across a distribution of trajectories found from random initialization of the state space problem. □

5.3. *Bayesian formulation: gradient flows*

In this section we study the Langevin SDE, for standard Brownian motion W,

$$\boxed{\mathrm{d}u = -\nabla \Phi_R(u)\,\mathrm{d}t + \sqrt{2}\,\mathrm{d}W.} \tag{5.13}$$

This is a noisy version of (5.3) in the case where $\Psi = \Phi_R$. It may also be found as a sample-path instantiation of the continuous-time limit of the MCMC algorithm (4.13), typically arising when α is the standard deviation of the proposal, choosing $\alpha = \Delta t$, letting $\rho(\cdot, n\Delta t) = \rho_n(\cdot)$ and sending $\Delta t \to 0$; see Section 5.6 for details.

The probability density function for the SDE (5.13) is governed by the Fokker–Planck equation

$$\partial_t \rho = \nabla \cdot (\rho \nabla \Phi_R) + \nabla \cdot (\nabla \rho) \tag{5.14a}$$

$$= \nabla \cdot (\rho \nabla \Phi_R + \rho \nabla \ln \rho), \tag{5.14b}$$

This equation has the density of the Bayesian posterior distribution (5.1) as steady state. This can be seen by noting that the right-hand side is divergence of a quantity which is zero if

$$\nabla(\Phi_R + \ln \rho) = 0.$$

This quantity can in turn be made zero by choosing

$$\rho \propto \exp(-\Phi_R),$$

so that ρ is given by the posterior distribution (5.1). As a consequence of the fact that the posterior is a steady state of the Fokker–Planck equation (5.14), the Langevin SDE (5.13) plays an important role in understanding algorithms for Bayesian inversion.

The machinery we established in the previous subsection, concerning gradient flows in the space of probability measures, is very powerful and demonstrates that any evolution equation of type (5.10) with appropriate potential $\mathcal{E}$ induces a gradient flow on $\mathfrak{P}_+$ with respect to the Wasserstein-2 metric. The Fokker–Planck

equation (5.14) associated with the Langevin equation (5.13) may be cast in this framework by making the choice

$$\mathcal{E}(\rho) = \int (\Phi_R + \ln \rho)\rho \, \mathrm{d}u \tag{5.15}$$

with variational derivative given by

$$\frac{\delta \mathcal{E}}{\delta \rho} = \Phi_R + \ln \rho.$$

Remark 5.2. We observe that, for $\mathrm{KL}[\cdot\|\cdot]$ denoting the *Kullback–Leibler divergence*,

$$\mathrm{KL}[\rho\|\pi] = \int \rho \log\left(\frac{\rho}{\pi}\right) \mathrm{d}u = \mathcal{E}(\rho) + \log \mathcal{Z}, \tag{5.16}$$

where π is the posterior density associated with posterior measure μ given by (4.7)[28] and the normalization constant $\mathcal{Z}$ is defined in (5.1b). It is thus possible to choose the energy to be $\mathrm{KL}[\rho\|\pi]$, since shifts by a constant in the energy do not change the evolution equations (5.10) and (5.11).

By the property of a divergence, the global minimizer of $\mathcal{E}(\rho)$ is attained at $\rho = \pi$ and hence solves the Bayesian inverse problem. It is thus of considerable value to have identified a gradient flow to minimize $\mathcal{E}(\rho)$ since such minimizers solve the Bayesian inverse problem. Theory concerning equation (5.14) as a gradient flow is contained in Section 5.6. □

In summary, the Fokker–Planck equation may be written in the abstract gradient form (5.11). We choose

$$\mathcal{E}(\rho) := \mathcal{E}(\rho) = \int (\Phi_R + \ln \rho)\rho \, \mathrm{d}u, \tag{5.17a}$$

$$\mathcal{M}(\rho)^{-1}\psi := -\nabla \cdot (\rho \nabla \psi) \in T_\rho \mathfrak{P}_+. \tag{5.17b}$$

This should be compared with (5.8) with the choice $\Psi = \Phi_R$: the metric structure defined by $\mathcal{M}$ is the same, but the energy $\mathcal{E}(\rho)$ has an additional term accounting for the Brownian motion appearing in (5.13).

5.4. Finite-time algorithms

The idea used in this subsection, to address the solution of inverse problems, is a continuous-time analogue of Section 4.4. From a sequential formulation of Bayesian inference we derive a filtering problem whose solution, at a particular time, gives the desired posterior. Section 5.4.1 is devoted to the formulation, Section 5.4.2 to the use of Gaussian projected filters and Section 5.4.3 to the use of ensemble Kalman methods.

[28] This notation is used throughout Section 5 and is not to be confused with the notation used for the joint law of state and data in previous sections.

5.4.1. Formulation

Employing the reparametrization (3.3a) in (4.20) and taking the continuum limit yields the SDE

$$\boxed{\begin{aligned} \mathrm{d}u &= 0, && (5.18a)\\ \mathrm{d}z &= G(u)\,\mathrm{d}t + \sqrt{\Gamma}\,\mathrm{d}B, && (5.18b)\end{aligned}}$$

with B a standard Brownian motion in $\mathbb{R}^{d_z}$. Given a specific realization of the observation process $z^\dagger(\cdot)$, we define $Z^\dagger(t) = \{z^\dagger(s)\}_{0\le s\le t}$, and consider the filtering distribution for the random variable $u(t)|Z^\dagger(t)$. However, there is a twist on the standard filtering setting: we are interested in the case where the data has constant derivative $\mathrm{d}z^\dagger(t)/\mathrm{d}t = w^\dagger$. Since the path $z^\dagger$ has zero quadratic variation the probability distribution is found by setting $z^\dagger(t) = tw^\dagger$ within the Stratonovich formulation of the non-local evolution equation for the density. Referring to (3.20), we see that this yields the following evolution for density $\rho(u,t)$ of $u(t)|Z^\dagger(t)$:

$$\partial_t\rho = \langle G - \mathbb{E}G, w^\dagger\rangle_\Gamma\rho - \frac{1}{2}\big\{|G|_\Gamma^2 - \mathbb{E}|G|_\Gamma^2\big\}\rho. \tag{5.19}$$

Here $\mathbb{E}$ denotes integration with respect to density $\rho(\cdot,t)$ so that the equation is non-local with respect to variable u and nonlinear with respect to density ρ. This is the analogue of the Kushner–Stratonovich equation for the filtering problem defined by (5.18), since the unconditioned variable u has trivial dynamics and since we are studying the case where the data $z^\dagger(t) = tw^\dagger$ has zero quadratic variation and is in fact differentiable. We may then show the following.

Theorem 5.3. Consider the dynamical system (5.18), and assume that $C_0 \succ 0$, $\Gamma \succ 0$, $u_0 \sim \mathsf{N}(m_0, C_0)$, and that u_0 is independent of Brownian motion B. Let $\rho(\cdot,t)$ denote the probability density function associated with the random variable $u(t)|Z^\dagger(t)$ evolving according to (5.18), with data chosen as $z^\dagger(t) = tw^\dagger$, for $t \in (0,1)$. Then the density $\rho(\cdot,t)$ satisfies (5.19), or equivalently for Φ given by (4.2a),

$$\partial_t\rho = -(\Phi - \mathbb{E}\Phi)\rho. \tag{5.20}$$

Furthermore, this equation has solution given by the formulae

$$\rho(u,t) = \frac{1}{\mathcal{Z}(t)}\exp\big(-t\Phi(u)\big)\rho_0(u), \tag{5.21a}$$

$$\mathcal{Z}(t) = \int_{\mathbb{R}^{d_u}} \exp\big(-t\Phi(u)\big)\rho_0(u)\,\mathrm{d}u; \tag{5.21b}$$

in particular, $\rho(\cdot,1)$ is equal to π, the density of the posterior distribution μ. ◇

Proof. Let ρ_0 denote the probability density function of the prior $\mathsf{N}(m_0, C_0)$. Recall that $\mu(t)$ has density given by equation (5.19) with initial condition $\rho|_{t=0} = \rho_0$. Now note that, recalling definition (5.2a) of Φ,

$$\langle G - \mathbb{E}G, w^\dagger\rangle_\Gamma \rho - \frac{1}{2}\left\{|G|_\Gamma^2 - \mathbb{E}|G|_\Gamma^2\right\}\rho \tag{5.22a}$$

$$= \left(\langle w^\dagger, G\rangle_\Gamma - \frac{1}{2}|G|_\Gamma^2\right)\rho - \mathbb{E}\left(\langle w^\dagger, G\rangle_\Gamma - \frac{1}{2}|G|_\Gamma^2\right)\rho \tag{5.22b}$$

$$= -(\Phi - \mathbb{E}\Phi)\rho. \tag{5.22c}$$

This establishes the equivalence of (5.20) with (5.19).

Note now that (5.21b) gives

$$\frac{\mathrm{d}\mathcal{Z}}{\mathrm{d}t} = -\mathbb{E}\Phi\,\mathcal{Z},$$

where expectation is under ρ. Hence it follows that differentiating (5.21a) gives (5.20). Since this is equivalent to (5.19) and since (5.19) characterizes the law of $u(t)|Z^\dagger(t)$ when $\mathrm{d}z^\dagger/\mathrm{d}t = w^\dagger$, the result is proved. □

The theorem establishes that the evolution equation (5.19) can be rewritten in the form (5.20), from which a gradient structure is apparent. Indeed, equation (5.20) can be written in the abstract form (5.11) with

$$\mathcal{E}(\rho) := \int \Phi\rho\,\mathrm{d}u, \tag{5.23a}$$

$$\mathcal{M}(\rho)^{-1}\psi := \rho\,\psi \in T_\rho\mathfrak{P}_+ \tag{5.23b}$$

for $\psi \in T_\rho\mathfrak{P}_+^*$. Note that the metric structure differs from what we have seen in (5.17). In particular we no longer use the Wasserstein-2 metric: the metric defined by the choice (5.23b) of $\mathcal{M}(\cdot)$ is known as the *Fisher–Rao metric*. The Fisher–Rao metric requires a more careful consideration of the variational derivative of $\mathcal{E}(\rho)$ as provided by (5.9). In particular, using

$$\psi = \frac{\delta\mathcal{E}}{\delta\rho}_{|T_\rho\mathfrak{P}_+^*} = \Phi - \mathbb{E}[\Phi] \in T_\rho\mathfrak{P}_+^*$$

in (5.23b) implies that the integral of the right-hand side of (5.23b) over $\mathbb{R}^{d_u}$ is zero so that it is, indeed, an element of the tangent space $T_\rho\mathfrak{P}_+$ given by (5.5).

Remark 5.4. The gradient flows defined by (5.8) and (5.17) both arose from considering the evolution equation for the probability density function associated with an evolution equation for $u \in \mathbb{R}^{d_u}$, the state space, namely equations (5.3) and (5.13) respectively. In contrast, while (5.20) describes the evolution of the density $\rho(\cdot, t)$, we did not derive it directly as the evolution equation for a random variable $u(t)$ in state space. However, we may seek to find such an evolution. To this end,

we postulate a mean-field differential equation

$$\frac{\mathrm{d}u}{\mathrm{d}t} = g(u, \rho).$$

Here ρ is the probability density function associated with the random variable u; since u is governed by a deterministic ordinary differential equation, the randomness in u originates from the initial condition $u(0)$. We now ask how to choose g so that the evolution equation for the probability density of $u(t)$, started from a random initialization, evolves according to (5.20). The evolution of this density will satisfy the associated nonlinear Liouville equation

$$\partial_t u = -\nabla \cdot (\rho g(\cdot, \rho)). \tag{5.24}$$

Equating (5.20) and (5.24), we obtain the condition

$$\nabla \cdot (\rho g(\cdot, \rho)) = (\Phi - \mathbb{E}(\Phi))\rho \tag{5.25}$$

on g. Whether or not this equation can be solved for g depends on properties of ρ, of Φ, and hence G; furthermore, even if solvable, the solution may not be unique. We note that, writing $g(u, \rho)$ as the gradient of a ρ-dependent potential E, so that $g(u, \rho) = \nabla_u E(u, \rho)$ renders (5.25) as a linear divergence form of elliptic equation for E. This elliptic equation is parametrized by probability density function ρ and should be viewed as holding everywhere on $\mathbb{R}^{d_u}$.

We have sought a state space model for u, which is an ordinary differential equation, albeit of mean-field type. It is also possible to seek stochastic evolution equations, such as birth–death processes or mean-field stochastic differential equations. □

5.4.2. Algorithms: Gaussian projected filter

To further elucidate the structure of Gaussian projected filtering for the inverse problem, we study its continuous-time formulation from Section 3.4 when applied to the specific state-observation model (5.18). To this end, first recall the discrete-time model (4.20), which has continuous-time limit (5.18). Taking the limit $\Delta t \to 0$ in (4.22), the Gaussian projected filter for (4.20), we obtain the following evolution equations for mean and covariance:

$$\boxed{\begin{aligned} \frac{\mathrm{d}m}{\mathrm{d}t} &= C^{uG}\Gamma^{-1}(w^\dagger - \mathbb{E}G(u)), && \text{(5.26a)} \\ \frac{\mathrm{d}C}{\mathrm{d}t} &= -C^{uG}\Gamma^{-1}(C^{uG})^\top, && \text{(5.26b)} \\ C^{uG} &= \mathbb{E}((u - \mathbb{E}u) \otimes (G(u) - \mathbb{E}G(u))). && \text{(5.26c)} \end{aligned}}$$

Here all expectations are computed under $u(t) \sim \mathsf{N}(m(t), C(t))$. This is the continuous-time Gaussian projected filter for the inverse problem (4.1).

As in the discrete-time case, the evolution for the mean promotes a Gaussian which is compatible with the data, through an innovation term which is weighted

by covariance information. We now illustrate the filter by considering equations (5.26) in the setting of linear G, when they may be solved exactly.

Example 5.5. Consider the setting of Example 4.4 in which $G(\cdot) = L\cdot$. The Gaussian projected filter (5.26) becomes

$$\frac{\mathrm{d}m}{\mathrm{d}t} = CL^\top \Gamma^{-1}(w^\dagger - Lm), \tag{5.27a}$$

$$\frac{\mathrm{d}C}{\mathrm{d}t} = -CL^\top \Gamma^{-1} LC. \tag{5.27b}$$

Note that this coincides with the Kalman–Bucy filter (3.30) for the specific filtering problem defined by (5.18) and with observed data $\mathrm{d}z^\dagger/\mathrm{d}t = w^\dagger$.

Now note that (5.19), the Kushner–Stratonovich equation, is solved by the Kalman–Bucy filter in the linear setting where $G(u) = Lu$. It follows that the solution ρ is given by the Gaussian $\mathsf{N}(m(t), C(t))$, where $m(t), C(t)$ solve the Gaussian projected filter equations (5.27). In particular, the posterior measure μ is Gaussian and given by $\mathsf{N}(m(1), C(1))$. To see this explicitly, note that, from Theorem 5.3, and in particular equation (5.21), in the linear case $G(\cdot) = L\cdot$, the solution of (5.19) is given by

$$\rho(u,t) \propto \exp\left(-\frac{t}{2}|w^\dagger - Lu|^2_\Gamma - \frac{1}{2}|u - m_0|^2_{C_0}\right). \tag{5.28}$$

Completing the square shows that this density corresponds to Gaussian $\mathsf{N}(m(t), C(t))$ with mean and covariance satisfying

$$C(t)^{-1}m(t) = tL^\top \Gamma^{-1} w^\dagger + C_0^{-1} m_0, \tag{5.29a}$$

$$C(t)^{-1} = C_0^{-1} + tL^\top \Gamma^{-1} L. \tag{5.29b}$$

Note that since $C_0 \succ 0$ it follows that $C_0^{-1} \succ 0$, and hence (5.29b) shows that, for all $t \geq 0$, $C(t)^{-1} \succ 0$ and hence that $C(t) \succ 0$ for all $t \geq 0$; hence $C(t)$ is well-defined by (5.29b) and $m(t)$ is well-defined by (5.29a). It simply remains to show that $m(t)$ and $C(t)$ given by these formulae solve (5.27) when $m(0) = m_0$ and $C(0) = C_0$.

We thus turn our attention to equations (5.27). Note that $C(t)$ solving (5.27b) satisfies $C(0) = C_0 \succ 0$. Hence $C(0)^{-1} \succ 0$. Thus, by continuity, $C(t)$ remains invertible for some positive interval of time $t \in [0, \tau)$ and, on this interval, direct computation with (5.27b) shows that

$$\frac{\mathrm{d}C^{-1}}{\mathrm{d}t} = L^\top \Gamma^{-1} L. \tag{5.30}$$

From this it follows by integration that $C^{-1}(t) \geq C_0^{-1} \succ 0$ for all t and hence that we may take $\tau = \infty$. Furthermore, the integration also shows that the solution of (5.30), solving (5.27b), delivers (5.29b) as desired.

We then notice that, from (5.27a),

$$C^{-1}\frac{\mathrm{d}m}{\mathrm{d}t} = L^\top \Gamma^{-1} w^\dagger - L^\top \Gamma^{-1} L m$$
$$= L^\top \Gamma^{-1} w^\dagger - \frac{\mathrm{d}C^{-1}}{\mathrm{d}t} m.$$

It follows that

$$\frac{\mathrm{d}}{\mathrm{d}t}(C^{-1}m) = L^\top \Gamma^{-1} w^\dagger$$

and integration, together with use of the initial conditions, shows that (5.27a) delivers the desired identity (5.29a).

It is also useful to write (5.29) using an explicit formula for $C(t)$ rather than the precision $C(t)^{-1}$. To this end, fix any $t > 0$ and consider the Gaussian random variable (u, w) defined by choosing $u \sim \mathsf{N}(m_0, C_0)$ and $w|u = \mathsf{N}(Lu, t^{-1}\Gamma)$. Then the density ρ given in (5.28) is the solution of the Bayesian inverse problem defined by the distribution of $u|w$. The mean and covariance may be found from (4.12) by replacing Γ by $t^{-1}\Gamma$ to yield

$$m(t) = m_0 + C_0 L^\top (L C_0 L^\top + t^{-1}\Gamma)^{-1}(w^\dagger - L m_0), \tag{5.31a}$$
$$C(t) = C_0 - C_0 L^\top (L C_0 L^\top + t^{-1}\Gamma)^{-1} L C_0. \tag{5.31b}$$

We also observe that the expression (5.31b) for $C(t)$ may be derived from (5.29b) by use of the Woodbury matrix identity. □

Remark 5.6. We obtained (5.26) as the continuous-time limit of its discrete-time formulation (4.32). However, the same evolution equations can be derived from the gradient flow (5.20) through a sequence of approximations. This perspective is outlined in Appendix E. □

5.4.3. Algorithms: ensemble Kalman filter

We now study the inverse problem using Kalman transport from Section 4.4.3, taking the continuous-time limit. We consider the specific state-observation model (5.18), and recall the discrete-time model (4.20) which has continuous-time limit (5.18). Taking the limit $\Delta t \to 0$ in (4.25), the ensemble Kalman filter for (4.20), we obtain the following evolution equations:

$$\mathrm{d}u = C^{uG}\Gamma^{-1}(w^\dagger\,\mathrm{d}t - \mathrm{d}\widehat{z}), \tag{5.32a}$$
$$\mathrm{d}\widehat{z} = G(u)\,\mathrm{d}t + \sqrt{\Gamma}\,\mathrm{d}B, \tag{5.32b}$$
$$C^{uG} = \mathbb{E}((u - \mathbb{E}u) \otimes (G(u) - \mathbb{E}G(u))). \tag{5.32c}$$

Here $B \in \mathbb{R}^{d_w}$ is a standard Brownian motion and expectation is under the law of u itself. As in the discrete-time case, the evolution for the state u promotes a solution which is compatible with the data, through an innovation term which is weighted by covariance information.

Remark 5.7. To obtain the resulting continuous-time formulation we may also start from the continuous-time state estimation methodology from Section 3.4 and apply it to the specific state-observation model (5.18). The SDE (5.32) may then be seen as a consequence of (3.46) applied to this state-observation model. However, special care is required in deriving the equation this way since the observations $z^\dagger$ in Section 3 were assumed to have non-vanishing quadratic variation; in contrast, in this section we have $\mathrm{d}z^\dagger/\mathrm{d}t = w^\dagger$, with $w^\dagger$ constant, and hence zero quadratic variation. □

To obtain further insight into the mean-field dynamical system (5.32), we once again consider the linear setting.

Example 5.8. Consider the SDE (5.32) in the setting where $G(u) = Lu$ for matrix $L \in \mathbb{R}^{d_w \times d_u}$. Then $u(1) \sim \mu$ where μ is given in Example 4.4. To see this recall that the Gaussian projected filter is exact in the linear setting, by Example 5.5, and hence delivers the desired posterior at time $t = 1$, by Theorem 5.3. Thus it suffices to show that the mean and covariance of u from (5.32) satisfy the Gaussian projected filter in the linear setting, given by (5.27). We first note that

$$\mathrm{d}u = CL^\top \Gamma^{-1}(w^\dagger \,\mathrm{d}t - Lu\,\mathrm{d}t - \sqrt{\Gamma}\,\mathrm{d}B), \tag{5.33}$$

where C is the covariance of u. By the Itô formula, $m = \mathbb{E}u$ satisfies (5.27a). It follows that $e = u - m$ satisfies

$$\mathrm{d}e = -CL^\top \Gamma^{-1} Le\,\mathrm{d}t - CL^\top \Gamma^{-1/2}\,\mathrm{d}B.$$

A second use of the Itô formula shows that $C = \mathbb{E}(e \otimes e)$ satisfies (5.27b). The desired result is established. □

We can also derive continuous limits of deterministic transports. Taking the $\Delta t \to 0$ limit in (4.27) results in the mean-field ODE formulation

$$\boxed{\begin{aligned} \frac{\mathrm{d}u}{\mathrm{d}t} &= C^{uG}\Gamma^{-1}\left(w^\dagger - \frac{1}{2}(G(u) + \mathbb{E}G(u))\right), && \text{(5.34a)} \\ C^{uG} &= \mathbb{E}((u - \mathbb{E}u) \otimes (G(u) - \mathbb{E}G(u))). && \text{(5.34b)} \end{aligned}}$$

Again the state evolution has a covariance-weighted forcing term which promotes evolution towards the data. As before, we may study this formulation in the linear setting.

Example 5.9. Consider the mean-field model (5.34) in the setting where $G(u) = Lu$ for matrix $L \in \mathbb{R}^{d_w \times d_u}$. Then $u(1) \sim \mu$, where the posterior distribution is given in Example 4.4. To show this we note that the Gaussian projected filter is exact in the linear setting, by Example 5.5, and hence delivers the desired posterior at time $t = 1$, by Theorem 5.3. Thus it suffices to show that the mean and covariance of u from (5.34) satisfy the Gaussian projected filter in the linear setting; this is given

by (5.27). We first note that the mean under (5.34) satisfies

$$\frac{\mathrm{d}m}{\mathrm{d}t} = CL^\top \Gamma^{-1}(w^\dagger - Lm),$$

which is (5.27a). Using this, it also follows that $e = u - m$ satisfies

$$\frac{\mathrm{d}e}{\mathrm{d}t} = -\frac{1}{2}CL^\top \Gamma^{-1} Le,$$

from which it follows that the variance satisfies (5.27b). □

Remark 5.10. Note that the nonlinear Liouville equation associated with the mean-field model (5.34) has the form

$$\partial_t \rho = -\nabla \cdot (\rho g_{\mathrm{KF}}), \tag{5.35a}$$

$$g_{\mathrm{KF}}(u,\rho) = C^{uG}\Gamma^{-1}\left(w^\dagger - \frac{1}{2}(G(u) + \mathbb{E}G(u))\right). \tag{5.35b}$$

This evolution equation *approximates* the evolution of the filtering distribution, except in the linear Gaussian setting when it is exact. On the other hand, as in Remark 5.4, we may seek a mean-field differential equation of the form

$$\frac{\mathrm{d}u}{\mathrm{d}t} = g(u,\rho), \tag{5.36}$$

which *exactly* replicates the filtering distribution in general. To do this requires that we choose g to solve (5.25):

$$\nabla \cdot (\rho g(\cdot,\rho)) = (\Phi - \mathbb{E}(\Phi))\rho.$$

In the linear Gaussian setting we can identify a solution of this equation by asking that (5.36) replicates (5.34), since we know the latter is exact in the linear and Gaussian setting.

In order to derive this result, we note that (5.35b) takes the form

$$g_{\mathrm{KF}}(u,\rho) = CL^\top \Gamma^{-1}\left(w^\dagger - \frac{1}{2}L(u+m)\right) \tag{5.37}$$

in the linear setting and ρ is Gaussian with mean m and covariance C. Hence the right-hand side of (5.35a) can now be evaluated explicitly, giving rise to

$$\begin{aligned}\nabla \cdot (\rho g_{\mathrm{KF}}) &= -\rho(u-m)^\top C^{-1} CL^\top \Gamma^{-1}\left(w^\dagger - \frac{1}{2}L(u+m)\right) + c_1\rho \\ &= \frac{1}{2}|Lu - w^\dagger|^2_\Gamma \rho + c_2\rho\end{aligned}$$

with normalization constants

$$c_1 = -\frac{1}{2}\mathbb{E}((u-m)^\top L^\top \Gamma^{-1} L(u+m))$$

and

$$c_2 = -\frac{1}{2}\mathbb{E}\left(|Lu - w^\dagger|_\Gamma^2\right).$$

Hence we have shown that g_{KF} satisfies (5.25) for $\Phi(u) = \frac{1}{2}|Lu - w^\dagger|_\Gamma^2$ in the linear Gaussian setting. See Section 5.6 for more details. □

5.5. *Infinite-time algorithms*

We now develop the ideas in Section 4.5 in the continuous-time setting. In particular we study algorithms posed on the infinite time horizon to solve the optimization problem of minimizing Φ_R given by (5.2), or to find the Bayesian posterior distribution given by (5.1). In Section 5.5.1 we consider this infinite time horizon perspective for the solution of optimization problems associated with the inverse problem (4.1). Section 5.5.2 considers the same perspective for Bayesian inversion.

5.5.1. *Algorithms for optimization formulation*

This rather lengthy subsection is broken into paragraphs concerning *preconditioned gradient flow*, *statistical linearization*, *gradient descent and statistical linearization*, *algebraic convergence* and *exponential convergence*. The initial development on preconditioning enables us to introduce *affine invariance* and the discussion of statistical linearization enables us to connect preconditioned gradient descent with ensemble Kalman methods. Then, as in Section 4.5.2, where similar issues are discussed in discrete time, we initially discuss algorithms with algebraic convergence. We then introduce generalizations which allow us to obtain exponential convergence.

Preconditioned gradient flow. We start by generalizing the standard gradient descent introduced in Section 5.2. Given objective function $\Psi\colon \mathbb{R}^{d_u} \to \mathbb{R}^+$ and given symmetric positive definite preconditioner $B \in \mathbb{R}^{d_u \times d_u}$, we introduce the equation

$$\frac{\mathrm{d}u}{\mathrm{d}t} = -B\nabla\Psi(u). \tag{5.38}$$

Note that, along solutions of (5.38),

$$\frac{\mathrm{d}}{\mathrm{d}t}\Psi(u) = \left\langle \nabla\Psi(u), \frac{\mathrm{d}u}{\mathrm{d}t}\right\rangle \tag{5.39a}$$

$$= -\left|\frac{\mathrm{d}u}{\mathrm{d}t}\right|_B^2. \tag{5.39b}$$

Equation (5.38) possesses a gradient flow structure in parameter space $\mathbb{R}^{d_u}$ with respect to a Euclidean metric weighted by B; this weighted metric changes the underlying geometry of the gradient flow, in comparison to the standard setting of equation (5.3), but it once again leads to the non-increasing property of $\Psi(u)$ along trajectories.

Now consider the affine transformation $u \mapsto \tilde{u}$ given by

$$\tilde{u} = Au + b, \tag{5.40}$$

where A is an invertible matrix and b a vector. An important issue in all vector space optimization problems is the relative scaling of the components of the vector. A highly desirable feature of an algorithm is that it should be insensitive to such scaling issues. This can be addressed by looking at differences between (i) the algorithm for u rewritten in terms of $\tilde{u}$ given by (5.40), and (ii) the same algorithm applied directly to variable $\tilde{u}$ optimizing $\tilde{\Psi}(\tilde{u})$, with the latter defined by

$$\tilde{\Psi}(\tilde{u}) = \Psi(A^{-1}(\tilde{u} - b)). \tag{5.41}$$

Definition 5.11. When the two ways (i) and (ii) of using reparametrization (5.40) lead to the same algorithm, for all choices of A, b, we say that the algorithm is *affine-invariant*. □

Remark 5.12. Affine-invariant algorithms are highly desirable as they are not sensitive to the scaling of the variables. At the optimum, which is unknown *a priori*, this scaling is not known and hence cannot be used to improve algorithms. Hence algorithms which are blind to such scaling are highly desirable. □

Applying the transformation (5.40) to (5.38) leads to

$$\frac{\mathrm{d}\tilde{u}}{\mathrm{d}t} = -AB\nabla\Psi(A^{-1}(\tilde{u} - b)), \tag{5.42}$$

the descent approach underlying algorithm viewed as in (i). In contrast, applying the same gradient descent to $\tilde{\Psi}$ given by (5.41) leads to the descent approach underlying algorithms viewed as in (ii):

$$\frac{\mathrm{d}\tilde{u}}{\mathrm{d}t} = -BA^{-\top}\nabla\Psi(A^{-1}(\tilde{u} - b)). \tag{5.43}$$

The two equations (5.42) and (5.43) only agree if

$$AB = BA^{-\top}$$

and such an identity cannot hold for all A, for a fixed B. Thus the basic gradient descent (5.38) is not affine-invariant. However, it is a remarkable fact that, by generalizing (5.38) to allow for mean-field dependence, we can achieve affine invariance.

To this end, consider the mean-field generalization of (5.38),

$$\frac{\mathrm{d}u}{\mathrm{d}t} = -B(\rho)\nabla\Psi(u), \tag{5.44}$$

where $\rho(\cdot, t)$ is the probability density function associated with the law of u, assuming that $u_0 = u(0)$ is drawn at random from probability density function $\rho_0(\cdot)$. If we assume that $B(\cdot)$ is positive definite symmetric for all possible input densities, then arguments similar to those above show that $\Psi(u)$ is non-increasing along

trajectories of (5.44). Furthermore, similar arguments show that the algorithm is affine-invariant if, for all invertible $A \in \mathbb{R}^{d_u \times d_u}$,

$$AB(\rho) = B(\tilde{\rho})A^{-\top},$$

where $\tilde{\rho}$ is the density of $\tilde{u}$ related to u, with density ρ, by (5.40). This identity holds for all invertible $A \in \mathbb{R}^{d_u \times d_u}$ if $B(\cdot)$ is chosen to be the covariance associated with its argument. We have thus discovered the affine-invariant mean-field gradient descent

$$\boxed{\frac{\mathrm{d}u}{\mathrm{d}t} = -C\nabla\Psi(u),} \tag{5.45}$$

where C is the covariance operator under the law of $u(t)$, and $u(0)$ is chosen at random from probability measure with density ρ_0.

Algorithms based on solving (5.45) are not themselves ensemble Kalman methods, although we have drawn inspiration from the power of mean-field methods to motivate the approach. In the next paragraph we introduce the idea of statistical linearization, leading to a variety of ensemble Kalman methods for optimization; and in the paragraph following it we use this idea to approximate (5.45) by an ensemble Kalman version of gradient descent which obviates the need for computing adjoints of the forward model $G(\cdot)$ in the setting where $\Psi(\cdot) = \Phi(\cdot)$ given by (5.2).

Statistical linearization. A basic building block in the Gaussian projected filter and mean-field ensemble Kalman models for inverse problems that we have presented in Sections 4.4.2 and 4.4.3, and their continuous-time analogues in Sections 5.4.2 and 5.4.3, is the object

$$C^{uG} = \mathbb{E}((u - \mathbb{E}u) \otimes (G(u) - \mathbb{E}G(u))), \tag{5.46}$$

here viewed as evolving in continuous time. Also of interest is the regularized analogue of C^{uG} arising when G_R, as defined in (4.3), is used in place of G. In a methodology based on exact properties only for first- and second-order statistics, it is natural that C^{uG} should appear when solving the inverse problem: the correlation between the parameter u, which we wish to estimate, and $G(u)$, which we observe, albeit polluted by additive noise. One way of understanding the role of C^{uG} in algorithms for inversion is through the idea of statistical linearization, providing a link between ensemble methods and derivatives of the objective function. The underlying principle is that the differences used in ensemble methods, and covariances in particular, act as a surrogate for derivatives.

The expectation defining (5.46) is computed, for the algorithms we consider, under the distribution of a Gaussian (for the Gaussian projected filter) or a more general distribution (for the mean-field ensemble Kalman model). To get some insight into the connection between ensemble methods and derivatives, we first consider the setting where $u \sim \mathsf{N}(m, C)$. Note that such u can be written as $u = m + \sqrt{C}\xi$ where $\xi \sim \mathsf{N}(0, I)$. With this assumption on u, (5.46) may be

reformulated as

$$C^{uG} = \mathbb{E}\big((\sqrt{C}\xi) \otimes \big(G(m + \sqrt{C}\xi) - \mathbb{E}G(m + \sqrt{C}\xi)\big)\big), \quad \xi \sim \mathsf{N}(0, I). \tag{5.47}$$

Using this we obtain the following connection between C^{uG} and derivatives of G.

Lemma 5.13. Assume that the second derivative of G is small: there is $\epsilon \ll 1$ such that

$$\sup_{u \in \mathbb{R}^{d_u}} |D^2 G(u)[\zeta, \zeta]| \le \epsilon |\zeta|^2.$$

Then C^{uG} given by (5.47) satisfies

$$C^{uG} = CDG(m)^\top + O(\epsilon).$$

Thus, when $C^{-1} \succ \lambda I$, for some $\lambda > 0$ independent of ϵ,

$$DG(m) = (C^{uG})^\top C^{-1} + O(\epsilon). \tag{5.48}$$

◇

Proof. Note that

$$G(m + \sqrt{C}\xi) = G(m) + DG(m)\sqrt{C}\xi + O(\epsilon),$$
$$\mathbb{E}G(m + \sqrt{C}\xi) = G(m) + O(\epsilon).$$

From (5.47),

$$C^{uG} = \mathbb{E}\big((\sqrt{C}\xi) \otimes (DG(m)\sqrt{C}\xi + O(\epsilon))\big), \quad \xi \sim \mathsf{N}(0, I),$$

and the desired result follows. □

Remark 5.14. Another perspective on the preceding lemma is via Stein's identity. This states that

$$C^{uG} = C(\mathbb{E}DG)^\top$$

for (5.47) when expectation is computed under a Gaussian measure $\mathsf{N}(m, C)$. The identity can be verified via integration by parts. Given the assumption stated in Lemma 5.13, $\mathbb{E}DG(u) = DG(m) + O(\epsilon)$ and the approximation result (5.48) also follows. □

When D^2G is indeed small, it is reasonable to use (5.48) as the basis for an approximation to $DG(\cdot)$ at points in an $O(1)$ ball around the mean. We now take this idea further and consider random variable $u \in \mathbb{R}^{d_u}$ (not necessarily Gaussian) and compute C and C^{uG} as covariance of u and cross-covariance of u with $G(u)$ respectively. We refer to use of the approximation

$$DG(u) \approx (C^{uG})^\top C^{-1}, \quad \text{for all } u \in \mathbb{R}^{d_u}, \tag{5.49}$$

as *statistical linearization.* The approximation can be invoked to replace the derivative within any standard optimization or sampling algorithm to solve the

inverse problem (4.1). Doing so results in a mean-field algorithm; that algorithm in turn can be approximated by particle methods. Statistical linearization allows the conversion of standard single particle optimization and sampling algorithms for inverse problems, with dependence on the derivative of the forward model, into derivative-free interacting particle system optimizers and samplers. An important application of this methodological approach is in the development of ensemble Kalman approximations of Gauss–Newton and Levenberg–Marquardt algorithms; pointers to the literature will be given in Section 5.6. We now turn to the use of statistical linearization in gradient descent, perhaps the most basic setting in which it can be used for optimization.

Gradient descent and statistical linearization. We provide further insight into the statistical linearization approach from the previous subsection by applying it in the context of gradient descent. Recall Φ and Φ_R defined in (5.2). We consider the regularized least-squares function so that $\Psi(\cdot) = \Phi_R(\cdot)$ in (5.3); but similar ideas may be developed in the unregularized setting where $\Psi(\cdot) = \Phi(\cdot)$. Recall that in the regularized setting

$$\Phi_R(u) = \Phi(u) + \frac{1}{2}|u - m_0|^2_{C_0}.$$

Note that

$$\nabla\Phi(u) = DG(u)^\top \Gamma^{-1}(G(u) - w^\dagger), \tag{5.50a}$$

$$\nabla\Phi_R(u) = DG(u)^\top \Gamma^{-1}(G(u) - w^\dagger) - C_0^{-1}(m_0 - u). \tag{5.50b}$$

Thus we obtain, from (5.3),

$$\frac{\mathrm{d}u}{\mathrm{d}t} = -\nabla\Phi_R(u) = DG(u)^\top \Gamma^{-1}(w^\dagger - G(u)) + C_0^{-1}(m_0 - u).$$

Similarly, the covariance preconditioned gradient flow (5.45) becomes

$$\frac{\mathrm{d}u}{\mathrm{d}t} = -C\nabla\Phi_R(u) = CDG(u)^\top \Gamma^{-1}(w^\dagger - G(u)) + CC_0^{-1}(m_0 - u). \tag{5.51}$$

We now approximate this equation using statistical linearization.

From (5.49) we deduce the equivalent (assuming C is invertible) approximation

$$CDG(u)^\top \approx C^{uG} \quad \text{for all } u \in \mathbb{R}^{d_u}.$$

Combining this with (5.50), we obtain

$$C\nabla\Phi(u) \approx C^{uG}\Gamma^{-1}(G(u) - w^\dagger). \tag{5.52}$$

Making this approximation in (5.51) gives the following ensemble Kalman approximation of mean-field gradient descent for Φ_R:

$$\boxed{\frac{\mathrm{d}u}{\mathrm{d}t} = C^{uG}\Gamma^{-1}(w^\dagger - G(u)) + CC_0^{-1}(m_0 - u).} \tag{5.53}$$

Remark 5.15. It may be verified that the affine invariance of (5.45) is preserved under the statistical linearization ansatz. To see this, note that if $G(\cdot) = L\cdot$, then the covariance matrix $C^{uG} = CL^\top$ so that

$$C^{uG}\Gamma^{-1}(w^\dagger - G(u)) = CL^\top\Gamma^{-1}(w^\dagger - Lu).$$

From this it follows that, in this linear setting,

$$C^{uG}\Gamma^{-1}(w^\dagger - G(u)) = -C\nabla\Phi(u),$$
$$\Phi(u) = \frac{1}{2}|Lu - w^\dagger|^2_\Gamma.$$

Hence (5.32) is also affine-invariant. Similar arguments also show that (5.34) and (5.53) are affine-invariant. In fact, this property holds for all the ensemble Kalman approaches to inverse problems developed in the previous section (discrete time) and the current section (continuous time). □

Algebraic convergence. We now observe that statistical linearization is exact for linear problems, and provide explicit calculations in this linear case. Recall G_R, Γ_R defined by (4.3) and assume that $\Gamma_R \succ 0$.

Example 5.16. Statistical linearization is exact in the linear setting: if $G(u) = Lu$ then $DG(u) = (C^{uG})^\top C^{-1}$. Thus, in the setting of Example 4.4, the preconditioned gradient flow (5.45) with $\Psi = \Phi_R$ reduces to (5.53) with $G(\cdot) = L\cdot$:

$$\frac{\mathrm{d}u}{\mathrm{d}t} = CL^\top\Gamma^{-1}(w^\dagger - Lu) + CC_0^{-1}(m_0 - u). \tag{5.54}$$

This equation leads to the following closed equations for evolution of the mean and covariance:

$$\frac{\mathrm{d}m}{\mathrm{d}t} = CL^\top\Gamma^{-1}(w^\dagger - Lm) + CC_0^{-1}(m_0 - m), \tag{5.55a}$$
$$\frac{\mathrm{d}C}{\mathrm{d}t} = -2CL^\top\Gamma^{-1}LC - CC_0^{-1}C = -2CL_R^\top\Gamma_R^{-1}L_RC. \tag{5.55b}$$

It is readily verified that the precision satisfies equation

$$\frac{\mathrm{d}C^{-1}}{\mathrm{d}t} = 2L_R^\top\Gamma_R^{-1}L_R \tag{5.56}$$

and hence the precision grows linearly in time to infinity. As a consequence, the covariance decays to zero at algebraic rate $O(1/t)$. □

Exponential convergence. To obtain exponential convergence we must overcome covariance collapse; to this end we study the idea of covariance inflation, introduced in discrete time in (4.31), in the continuous-time setting. Again recall G_R, Γ_R defined by (4.3) and assume that $\Gamma_R \succ 0$. We then consider the continuous-time

limit of (4.31) to obtain

$$\boxed{\begin{aligned} \mathrm{d}u &= \sqrt{\beta\Sigma}\,\mathrm{d}W, && \text{(5.57a)}\\ \mathrm{d}z &= G_R(u)\,\mathrm{d}t + \sqrt{\Gamma_R}\,\mathrm{d}B, && \text{(5.57b)} \end{aligned}}$$

where W and B are independent standard Brownian motions on $\mathbb{R}^{d_u}$ and R^{d_w} respectively. To determine Σ set $Z^\dagger(t) = \{z^\dagger(s)\}_{s\in[0,T]}$ and define $\Sigma = C$, where C is the covariance of random variable $u(t)|Z^\dagger(t)$. When we apply the SDE (5.57) to solve the inverse problem, we take $z^\dagger(s) := sw_R^\dagger$ for all $s \geq 0$.

Remark 5.17. As for the discrete-time model (4.31), equation (5.57) defines an unusual form of mean-field model through dependence on the filtering distribution. As a consequence, the filtering distribution is determined by a non-standard variant of the Kushner–Stratonovich equation which takes the form

$$\partial_t\rho = \frac{1}{2}\nabla\cdot(\nabla\cdot(\rho\mathcal{C}(\rho))) - \frac{1}{2}\left\{|G_R|^2_{\Gamma_R} - \mathbb{E}|G_R|^2_{\Gamma_R}\right\}\rho + \left\langle G_R - \mathbb{E}G_R, w_R^\dagger\right\rangle_{\Gamma_R}\rho.$$

This may be derived from (3.26) with $f \equiv 0$, $\Sigma = \mathcal{C}(\rho)$ (the covariance under ρ), $h = G_R$ and $\mathrm{d}z^\dagger(t) := w_R^\dagger\,\mathrm{d}t$. Note that appearance of the covariance matrix $\mathcal{C}(\rho)$ adds a further non-local nonlinearity which is not present in the density evolution (5.19) that arises without covariance inflation. □

We now derive continuous-time limits of the Gaussian projected filter (4.32) and the ensemble Kalman filters (4.34) and (4.35), derived in Section 4.5 for solution of the inverse problem (4.1). Starting with the Gaussian projected filter and taking the limit $\Delta t \to 0$, we obtain

$$\boxed{\begin{aligned} \frac{\mathrm{d}m}{\mathrm{d}t} &= C_R^{uG}\Gamma_R^{-1}\left(w_R^\dagger - \mathbb{E}G_R(u)\right), && \text{(5.58a)}\\ \frac{\mathrm{d}C}{\mathrm{d}t} &= \beta C - C_R^{uG}\Gamma_R^{-1}\left(C_R^{uG}\right)^\top, && \text{(5.58b)}\\ C_R^{uG} &= \mathbb{E}((u-\mathbb{E}u)\otimes(G_R(u)-\mathbb{E}G_R(u))), && \text{(5.58c)} \end{aligned}}$$

where all expectations are with respect to $u(t) \sim \mathsf{N}(m(t), C(t))$. Using the explicit form of Φ_R, this set of equations may be shown to be equivalent to

$$\begin{aligned} \frac{\mathrm{d}m}{\mathrm{d}t} &= C^{uG}\Gamma^{-1}(w^\dagger - \mathbb{E}G(u)) + CC_0^{-1}(m_0 - m), && \text{(5.59a)}\\ \frac{\mathrm{d}C}{\mathrm{d}t} &= \beta C - C^{uG}\Gamma^{-1}(C^{uG})^\top - CC_0^{-1}C. && \text{(5.59b)} \end{aligned}$$

As in discrete time, we are using the approximation of the covariance of the filtering distribution implied by the Gaussian projected filter; this is since we do not have access to the exact covariance.

Example 5.18. To highlight links with preconditioned gradient descent, we now investigate these equations in the linear Gaussian setting. With $G(\cdot) = L\cdot$, we

obtain from (5.59) the following equations for the mean and covariance matrix evolution:

$$\frac{\mathrm{d}m}{\mathrm{d}t} = CL^\top \Gamma^{-1}(w^\dagger - Lm) + CC_0^{-1}(m_0 - m), \tag{5.60a}$$

$$\frac{\mathrm{d}C}{\mathrm{d}t} = \beta C - CL^\top \Gamma^{-1} LC - CC_0^{-1} C = \beta C - CL_R^\top \Gamma_R^{-1} L_R C. \tag{5.60b}$$

These evolution equations for mean and covariance are now compared to the corresponding equations derived in Example 5.16, concerning statistical linearization. Equations (5.55) in that example arise from statistical linearization of preconditioned gradient descent, and exactly recover preconditioned gradient descent in the linear setting. We note that the evolution equations (5.55a) and (5.60a) for the mean agree, while the covariance matrices C defined by (5.55b) and (5.60b) undergo different evolutions. □

In the proposition which follows, we now show that the different evolution equation for the covariance (5.60b) leads to exponential convergence; this should be contrasted with the algebraic convergence resulting from (5.55b).

Proposition 5.19. Assume that u_0 is initialized at a Gaussian $\mathsf{N}(m_0, C_0)$ and assume also that $C_0, \Gamma_R > 0$. Consider the setting where $G_R(\cdot) = L_R \cdot$ for matrix $L_R \in \mathbb{R}^{(d_w+d_u)\times d_u}$. Now consider the filtering distribution $u(t)|Z^\dagger(t)$ defined by (5.57) for $\beta > 0$, with data $Z^\dagger(t)$ defined by $z^\dagger(s) = sw_R^\dagger$, where $w_R^\dagger$ is defined in (4.5). Then the filtering distribution is Gaussian $\mathsf{N}(m(t), C_(t))$ for all $t \geq 0$. The mean and covariance converge at an exponential rate $\exp(-\beta t)$, as $t \to \infty$, to the limits $m_\infty = m_{\mathrm{post}}$ and $C_\infty = \beta C_{\mathrm{post}}$, where $(m_{\mathrm{post}}, C_{\mathrm{post}})$ are the posterior mean (4.11) and covariance (4.10). ◇

Remark 5.20. It is a remarkable fact that the rate of convergence is independent of the properties of the limiting Gaussian posterior distribution, and in particular of the conditioning of the posterior covariance. This property is the continuous-time analogue of what we observed in Remark 4.18. The desirable universal convergence rate property is a result of the affine invariance of the Gaussian projected filter and ensemble Kalman methods studied in this subsection. □

Proof of Proposition 5.19. We first note that $\beta > 0$ implies that equation (5.60b) for the covariance has two equilibria: an unstable one at $C = 0$ and a stable one satisfying

$$C_\infty = \beta \left(L_R^\top \Gamma_R^{-1} L_R \right)^{-1} = \beta C_{\mathrm{post}}. \tag{5.61}$$

The exponential convergence to C_∞ for $\beta > 0$ can be best seen by considering the evolution for the precision matrix C^{-1}:

$$\frac{\mathrm{d}C^{-1}}{\mathrm{d}t} = -\beta C^{-1} + L_R^\top \Gamma_R^{-1} L_R. \tag{5.62}$$

Direct calculation of the time-derivative of $C^{-1}(t)m(t)$ shows that

$$\frac{\mathrm{d}}{\mathrm{d}t}(C^{-1}m) = -\beta C^{-1}m + L_R^\top \Gamma_R^{-1} w_R^\dagger.$$

Thus $C^{-1}(t)m(t)$ converges exponentially fast to limit $\frac{1}{\beta}L_R^\top \Gamma_R^{-1} w_R^\dagger$. Since $C(t)$ itself converges exponentially fast to βC_{post}, it follows that $m(t)$ converges exponentially fast to posterior mean m_{post} given by (4.11). □

Remark 5.21. For $\beta = 0$, we find algebraic convergence which we also found in equation (5.56), from Example 5.16, for the preconditioned gradient descent formulation. It should be noted that (5.56) contains a pre-factor of two which is not present in (5.62) so, even when $\beta = 0$, the evolution for mean and covariance does not coincide with that arising from preconditioned gradient descent. □

We now turn to ensemble Kalman methods, starting with the stochastic version. We take continuous-time limits in (4.34) to obtain

$$\boxed{\begin{aligned} \mathrm{d}u &= \sqrt{\beta C}\,\mathrm{d}W + C_R^{uG}\Gamma_R^{-1}\left(w_R^\dagger\,\mathrm{d}t - \mathrm{d}\widehat{z}\right), && \text{(5.63a)}\\ \mathrm{d}\widehat{z} &= G_R(u)\,\mathrm{d}t + \sqrt{\Gamma_R}\,\mathrm{d}B, && \text{(5.63b)} \end{aligned}}$$

where C_R^{uG} is computed using (5.58c), and C is the regular covariance, both under the law of u. As in discrete time, we are using the approximation of the covariance of the filtering distribution implied by the Gaussian projected filter; this is since we do not have access to the exact covariance. Here W and B are independent standard Brownian motions on $\mathbb{R}^{d_u}$ and R^{d_w} respectively. The corresponding deterministic transport formulation is found by taking the continuous-time limit in (4.35) to obtain

$$\boxed{\mathrm{d}u = \sqrt{\beta C}\,\mathrm{d}W + C_R^{uG}\Gamma_R^{-1}\left(w_R^\dagger - \frac{1}{2}(G_R(u) - \mathbb{E}G_R(u))\right)\mathrm{d}t,} \qquad \text{(5.64)}$$

where W is again a standard Brownian motion on $\mathbb{R}^{d_u}$ and where C_R^{uG} is computed using (5.58c), and C is the regular covariance, both under the law of u.

Remark 5.22. We note that it is possible to replace (5.57) by the filtering problem associated with the model

$$\boxed{\begin{aligned} \mathrm{d}u &= \frac{1}{2}\beta(u - \mathbb{E}u)\,\mathrm{d}t, && \text{(5.65a)}\\ \mathrm{d}z &= G_R(u)\,\mathrm{d}t + \sqrt{\Gamma_R}\,\mathrm{d}B. && \text{(5.65b)} \end{aligned}}$$

This results in the same Gaussian projected filter and ensemble Kalman methods as before, in the linear Gaussian setting. We note that using the filtering problem associated with (5.65) leads to analogous ensemble Kalman methods to (5.64) or (5.63); these can be obtained by directly replacing the term $\sqrt{\beta C}\,\mathrm{d}W$ with $\frac{1}{2}\beta(u - \mathbb{E}u)\,\mathrm{d}t$ in (5.64) and (5.63). □

5.5.2. Algorithms for Bayesian formulation

In Section 4.5.3 we derived Gaussian projected filter and ensemble Kalman methods that converge exponentially fast to the exact posterior distribution of the inverse problem (4.1) in the linear, Gaussian setting. This is achieved by choosing $\beta = (1 - \Delta t)^{-1}$ as stated in (4.43). A similar argument in the continuous-time setting thus leads to the choice $\beta = 1$. We have the following corollary of Proposition 5.19.

Corollary 5.23. Assume that u_0 is initialized at a Gaussian $\mathsf{N}(m_0, C_0)$ and assume also that $C_0, \Gamma_R \succ 0$. Consider the setting where $G_R(\cdot) = L_R \cdot$ for matrix $L_R \in \mathbb{R}^{(d_w+d_u)\times d_u}$. Now consider the filtering distribution $u(t)|Z^\dagger(t)$ defined by (5.57) for $\beta = 1$, with data $Z^\dagger(t)$ defined by $z^\dagger(s) = sw_R^\dagger$, where $w_R^\dagger$ is defined in (4.5). Then the filtering distribution is Gaussian $\mathsf{N}(m(t), C(t))$ for all $t \geq 0$. The mean and covariance converge at an exponential rate, as $t \to \infty$, to the limits $m_\infty = m_{\text{post}}$ and $C_\infty = C_{\text{post}}$, where $(m_{\text{post}}, C_{\text{post}})$ are the posterior mean (4.11) and covariance (4.10). ◇

Preconditioned gradient flows. While the above extension to Bayesian inference problems is straightforward and leads to an exact recovery of the posterior distribution in the linear Gaussian setting, the resulting methods deliver only approximations in the general nonlinear setting. This, of course, is a theme throughout this article. In the remainder of this subsection we take a different approach to deriving algorithms for sampling which are exact in the linear Gaussian setting. We start by considering preconditioned gradient descent, and its Langevin analogue; we then note that application of statistical linearization gives approximations of these evolutions which are exact in the linear Gaussian setting, and hence also exhibit desirable convergence properties, in the linear Gaussian setting.

First recall the preconditioned gradient descent (5.45) for $\Psi = \Phi_R$ resulting in

$$\boxed{\frac{\mathrm{d}u}{\mathrm{d}t} = -C\nabla\Phi_R(u).} \tag{5.66}$$

This methodology for optimization may be extended to a sampling methodology by considering the preconditioned mean-field Langevin SDE defined by

$$\boxed{\mathrm{d}u = -C\nabla\Phi_R(u)\,\mathrm{d}t + \sqrt{2C}\,\mathrm{d}W.} \tag{5.67}$$

Remark 5.24. We note here that it may be verified that (5.67) is invariant under affine transformations of type (5.40). Thus (5.67) provides an attractive generalization of standard Langevin dynamics (5.13) for sampling from the posterior distribution, because of the properties of affine-invariant algorithms highlighted in Remark 5.12. Although the desired posterior distribution is approached only in the limit $t \to \infty$, the fact that the convergence is exponential, with universal rate across all linear Gaussian inverse problems, makes the approach potentially competitive. Theory concerning this equation is discussed in Section 5.6.

Use of the statistical linearization approximation (5.49), which we discuss next, converts both (5.66) and (5.67) into mean-field ensemble Kalman methods which, through particle approximations, lead to implementable derivative-free methods. □

Inexact gradients. It is possible to apply statistical linearization (5.52) to the preconditioned Langevin equation (5.67), resulting in the evolution equation

$$\boxed{\mathrm{d}u = C_R^{uG}\Gamma_R^{-1}\left(w_R^{\dagger} - G_R(u)\right)\mathrm{d}t + \sqrt{2C}\,\mathrm{d}W.} \tag{5.68}$$

Here C_R^{uG} is computed using (5.58c), and C is the regular covariance, both under the law of u. Like (5.63) and (5.64), this equation converges exponentially fast to the posterior distribution in the linear Gaussian case. However, the form of the mean-field stochastic differential equation is fundamentally different: here the Brownian noise arises in state space $\mathbb{R}^{d_u}$, whereas in the ensemble Kalman methods it appears in data space $\mathbb{R}^{d_y}$.

Exact gradients. This paragraph concerns analysis of (5.66) and (5.67) when the exact gradients of Φ, and hence Φ_R, are available. We start by considering the geometric properties of (5.66). We assume that $u(t)$ has smooth probability density $\rho(u,t)$ for all $t \geq 0$ and recall that we denote the manifold of all smooth probability density functions on $\mathbb{R}^{d_u}$ by $\mathfrak{P}_+$. Then ρ satisfies the Liouville equation

$$\partial_t \rho = \nabla \cdot (\rho C \nabla \Phi_R). \tag{5.69}$$

Again the appearance of the covariance matrix $C = \mathcal{C}(\rho)$ renders (5.69) a nonlinear and non-local partial differential equation on $\mathfrak{P}_+$. We will show that the evolution of ρ on $\mathfrak{P}_+$ has gradient flow structure of the form given in (5.11).

In order to see this gradient structure we need to identify the energy functional which is being minimized and then introduce a metric structure in which (5.69) is a gradient flow. In this case we may choose

$$\mathcal{E}(\rho) := \mathcal{E}(\rho) = \int \Phi_R \rho \,\mathrm{d}u, \tag{5.70a}$$

$$\mathcal{M}(\rho)^{-1}\psi := -\nabla \cdot (\rho \mathcal{C}(\rho)\nabla\psi) \in T_\rho \mathfrak{P}_+. \tag{5.70b}$$

We refer to the *Kalman–Wasserstein metric* as the metric induced by this metric tensor, generalizing the Wasserstein-2 metric introduced in Section 5.3. Note that the variational derivative of $\mathcal{E}$ is given by

$$\frac{\delta \mathcal{E}}{\delta \rho} = \Phi_R.$$

Hence we can rewrite (5.69) as

$$\partial_t \rho = \nabla \cdot \left(\rho \mathcal{C}(\rho) \nabla \frac{\delta \mathcal{E}}{\delta \rho}\right). \tag{5.71}$$

Note that

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{E}(\rho) &= \int_{\mathbb{R}^{d_u}} \frac{\delta \mathcal{E}}{\delta \rho}\partial_t \rho \,\mathrm{d}u \\ &= -\int_{\mathbb{R}^{d_u}} \rho \left| \mathcal{C}(\rho)^{1/2}\nabla \frac{\delta \mathcal{E}}{\delta \rho}\right|^2 \mathrm{d}u \\ &= -\int_{\mathbb{R}^{d_u}} \left\langle \nabla \frac{\delta \mathcal{E}}{\delta \rho}\,,\, \mathcal{C}(\rho)\nabla \frac{\delta \mathcal{E}}{\delta \rho}\right\rangle \rho \,\mathrm{d}u \\ &\leq 0.\end{aligned}$$

It is interesting to compare the gradient structure on $\mathfrak{P}_+$ to the gradient flow structure on $\mathbb{R}^{d_u}$ defined by (5.38) with $B = \mathcal{C}(\rho)$. The state space gradient flow on $\mathbb{R}^{d_u}$ ensures decrease of $\Phi_R(u(t))$ along trajectories, whilst the probability space gradient flow on $\mathfrak{P}_+$ ensures decrease of the expected value of $\Phi_R(u(t))$ across a distribution of trajectories found from random initialization of the state space problem.

The preceding calculations demonstrate that *any* evolution equation of type (5.71) with appropriate energy functional $\mathcal{E}$ induces a gradient flow on $\mathfrak{P}_+$. In particular, using the energy functional

$$\mathcal{E}(\rho) = \int (\Phi_R + \ln \rho)\rho \,\mathrm{d}u$$

defined in (5.15), we observe that the associated evolution equation (5.71) becomes

$$\partial_t \rho = \nabla \cdot (\rho \mathcal{C}(\rho)\nabla \Phi_R) + \nabla \cdot (\mathcal{C}(\rho)\nabla \rho). \tag{5.72}$$

This nonlinear and non-local Fokker–Planck equation governs evolution of the probability density funtion for the mean-field SDE (5.67). Now recall Remark 5.2, in which we note that $\mathrm{KL}[\rho\|\pi]$ and $\mathcal{E}(\rho)$ differ by a constant, and where π is the posterior density associated with posterior measure μ given by (4.7). Thus the global minimizer of the gradient flow associated with (5.15) is attained at $\rho = \pi$ and hence solves the Bayesian inverse problem.

Finally, it is also useful to see the nonlinear and non-local Fokker–Planck equation (5.72) written in the abstract gradient form (5.11). In this case we may choose

$$\mathcal{E}(\rho) \coloneqq \mathcal{E}(\rho) = \int (\Phi_R + \ln \rho)\rho \,\mathrm{d}u, \tag{5.73a}$$

$$\mathcal{M}(\rho)^{-1}\psi \coloneqq -\nabla \cdot (\rho \mathcal{C}(\rho)\nabla \psi) \in T_\rho \mathfrak{P}_+. \tag{5.73b}$$

Remark 5.25. The Kalman–Wasserstein gradient flow structure for the Fokker–Planck equation associated with (5.67) is not maintained under the statistical linearization leading to (5.68). The resulting Fokker–Planck equation implied by (5.68) is not of gradient descent type in $\mathfrak{P}_+$, except in the case where $G(u)$ is linear. □

5.6. *Bibliographical notes*

The notion of gradient flow plays a central role in this paper. The subject is enormous and we cannot do justice to it here. We point the reader to the book by Hirsch, Smale and Devaney (2013) for the study of gradient flows in Euclidean space and to Ambrosio, Gigli and Savaré (2008) for gradient flows in metric spaces, including spaces of probability measures. The paper by Chen *et al.* (2023) contains an overview of gradient flows for probability measures, focused on applications to Bayesian inversion.

Continuous-time limits of ensemble Kalman filters were first derived in Bergemann and Reich (2010*a*) and Bergemann and Reich (2010*b*) and further explored in the context of continuous-time transport in Reich (2011). A connection between the non-stochastic Kalman transport equations (5.34) and preconditioned gradient descent was first identified in Bergemann and Reich (2010*b*) for finite ensemble sizes, and in Reich and Cotter (2015) for the mean-field limit. Pidstrigach and Reich (2023) have adopted the continuous-time setting. See also Yang, Blom and Mehta (2014) for related formulations based on the feedback particle filter approach to continuous-time filtering.

Schillings and Stuart (2017, 2018) studied the use of ensemble Kalman methods for optimization problems, taking a continuous-time limit, making a connection to preconditioned gradient descent and exploiting an invariant subspace property (see e.g. Iglesias *et al.* 2013, Theorem 2.1) inherent in the basic form of the ensemble Kalman methodology. This led to work on approximate sampling from the preconditioned Langevin equation in Garbuno-Inigo *et al.* (2020*a*,*b*), Nüsken and Reich (2019) and Liu, Stuart and Wang (2022); in particular, Nüsken and Reich (2019) and Garbuno-Inigo *et al.* (2020*b*) demonstrated a finite ensemble size correction to the mean-field limit introduced in Garbuno-Inigo *et al.* (2020*a*). Garbuno-Inigo *et al.* (2020*a*) used the non-standard Kalman–Wasserstein metric, first introduced in Reich and Cotter (2015), to provide a framework to analyse the preconditioned Langevin equation. It remains open to fully develop the mathematical foundations of gradient flows using this metric. These papers demonstrate the role played by the ensemble in preconditioning the dynamics. This makes a link to the important paper by Goodman and Weare (2010), which introduced the concept of affine-invariant ensemble samplers, an idea developed further in Leimkuhler, Matthews and Weare (2018).

Chada, Chen and Sanz-Alonso (2021) review the optimization perspective and provide a unifying framework, going beyond gradient descent-based methods. In particular, framing ensemble Kalman-based optimization methods in terms of statistical linearization, as we do in Section 5.5.1, originates in that paper. Reich and Weissmann (2021) and Pavliotis, Stuart and Vaes (2022) show how to construct a derivative-free Langevin sampler using localized ensembles; the use of localized ensembles to train neural networks may be found in Haber *et al.* (2018). An alternative interacting particle system approach to solving inverse problems and

optimization tasks is the use of consensus-based methods (Tsianos *et al.* 2012, Ha *et al.* 2021, Fornasier, Klock and Riedl 2024, Carrillo *et al.* 2018, Carrillo, Hoffmann, Stuart and Vaes 2022, Chen, Jin and Lyu 2022*a*, Pinnau, Totzeck, Tse and Martin 2017, Fornasier *et al.* 2020). Ding *et al.* (2021) and Ding and Li (2021*a*,*b*) have undertaken a systematic analysis of the link between interacting particle systems and mean-field systems, mostly focused on the solution of inverse problems; however, the methods developed are more widely applicable.

Recall that (5.26) is a derivative-free approach to approximately solving the Bayesian inverse problem by application of the Gaussian projected filter. In Appendix E we show that these equations may also be derived from the Fisher–Rao gradient flow (5.20), deriving equations for the mean and covariance evolution from it, and then by invoking a number of approximations. In Appendix E, as a step in this derivation, we obtain the equations

$$\frac{\mathrm{d}m}{\mathrm{d}t} = -\mathbb{E}(\Phi(u)(u-m)), \tag{5.74a}$$

$$\frac{\mathrm{d}C}{\mathrm{d}t} = -\mathbb{E}(\Phi(u)(u-m)(u-m)^{\top}) + C\mathbb{E}(\Phi(u)), \tag{5.74b}$$

which may be viewed as a closed evolution when expectation is computed under the Gaussian defined by (m, C). These closed equations also define a derivative-free approach to approximate the Bayesian inverse problem, different from (5.26). It is natural to ask which of (5.74) and (5.26) is preferable; computational experiments underpinning the work in Chen *et al.* (2023) indicate that (5.26) is more robust in various settings. Discussion of the Fisher–Rao gradient flow projected into the manifold of Gaussians is contained in Chen *et al.* (2023).

Theory showing that the Fokker–Planck equation (5.14) arises as a continuous-time limit of the MCMC method (4.13) was initiated in Gelman, Gilks and Roberts (1997), for the random walk Metropolis algorithm, and followed up for the Metropolis-adjusted Langevin equation in Roberts and Rosenthal (1998). See Roberts and Rosenthal (2001) for an overview of this field. Most of the analysis is done at the level of weak convergence of sample paths and hence works directly with (5.13) rather than with its density, which is governed by (5.14).

6. Conclusions and open problems

This paper presents a unifying perspective on the derivation, interpretation and analysis of ensemble Kalman methods through use of the ideas of mean-field models, second-order approximate transport and particle approximation. Both state estimation and parameter estimation (inverse problems) are studied; similar ideas may be developed for joint parameter-state estimation problems but are not discussed in this paper. The ideas have been presented in discrete time and, through specific parametric scalings, continuous-time limits have been identified. Our unifying approach constitutes a novel presentation of the subject, and creates a framework for the mathematical development of the subject area.

Ensemble Kalman methods have been enormously impactful in the geosciences, where they originated, and are starting to be used in numerous other application domains. However, if they are to realize their potential for widespread adoption and application, many research challenges remain. These challenges are both in mathematical analysis and in the development of methodology. One of the biggest challenges is the following: some theory, and abundant numerical evidence, show that ensemble Kalman methods perform well at state estimation and at parameter estimation. However, there is very little theory, or empirical evidence, which identifies situations in which the statistical information in the ensemble constitutes valid approximate Bayesian inference. The recent papers by Carrillo *et al.* (2024) and Calvello *et al.* (2024) make first steps in this direction. In mathematical analysis a number of other substantial challenges are presented by ensemble Kalman methods, which we list here.

- For state estimation, determine conditions under which the filtering distribution is well-approximated by mean-field models based on second-order transport. Find sharp error estimates and appropriate metrics for the analysis. Furthermore, identify which problems satisfy these conditions.
- For inverse problems, determine conditions under which the optimizer of a (to-be-identified) loss function is well-approximated by the mean or sample path of mean-field models based on second-order transport, in both the transport and iterative approaches to inversion. Find sharp error estimates.
- For inverse problems, determine conditions under which the Bayesian posterior is well-approximated by mean-field models based on second-order transport, in both the transport and iterative approaches to inversion. Find sharp error estimates and appropriate metrics for the analysis. Furthermore, identify which inverse problems satisfy these conditions.
- For all of the preceding three scenarios derive error bounds for particle approximations of the mean-field models. When low-rank structure is present in covariances prove that ensemble Kalman methods can correctly identify it, and exploit the low-rank structure in the analysis of particle approximations.
- In all of the particle methods arising above, compare the cost/error trade-off with that arising for other methods, to determine when ensemble Kalman methods are competitive.
- All of the algorithms in this paper are studied in idealized scenarios, in the absence of widely employed techniques such as covariance inflation and localization. Developing analyses which account for covariance inflation and localization will be highly desirable.

On the methodology side there are also a number of significant challenges, which we also list here.

- Given the ability to compute an ensemble of evaluations of the combined state-observation dynamical system, determine the optimal (in terms of cost/error trade-off) way to combine the ensemble to either estimate the state given an observation sequence, or the filtering distribution.
- Given the ability to compute an ensemble of evaluations of the forward model, what is the optimal way to combine them to either estimate the parameter given an observation, or the posterior distribution, for the corresponding inverse problem.
- What role might be played by machine learning in addressing the design of algorithms, and in particular in addressing the preceding questions.
- Develop an overarching interacting particle and mean-field framework that subsumes ensemble Kalman and alternative particle filters as well as derivative-free and consensus-based optimization methods, and use the framework to create new methods.
- Develop principles for the deployment of covariance inflation, and generalizations of localization, so that the resulting methodology is widely applicable and does not need application-specific principles to be applied. Different ideas may be needed in the inverse problem setting.
- Expand the preceding scenarios beyond the additive error models discussed in the survey, to include classification and other machine learning tasks.

Acknowledgements

EC is grateful to the Kortschak Scholars Program within the CMS Department at Caltech for financial support. The work of SR is supported by Deutsche Forschungsgemeinschaft (DFG) – Project-ID 318763901 – SFB1294. The work of AMS is supported by NSF award AGS1835860, and by a Department of Defense Vannevar Bush Faculty Fellowship, which also supports EC. In addition, AMS and EC are supported by the SciAI Center, funded by the Office of Naval Research (ONR), under grant no. N00014-23-1-2729. The work of EC was also supported by the Resnick Sustainability Institute.

The authors are grateful to Dmitry Burov for helpful advice regarding the numerical experiments; they are also grateful to Arnaud Vadeboncoeur, Eviatar Bach, Ricardo Baptista and Daniel Sanz-Alonso for helpful discussions which improved this paper. The authors thank Mark Asch for careful reading of the paper, and useful feedback. Finally AMS is grateful for hospitality offered at the University of Chicago, by Guillaume Bal, Peter McCullagh, Daniel Sanz-Alonso and Rebecca Willett, where he delivered lectures based on a preliminary version of this paper in May 2022.

A. Pseudo-code

In this appendix we provide pseudo-code describing several of the algorithms that we present and deploy in this paper. Algorithms 1 and 2, 3DVAR and the ensemble Kalman filter (EnKF) respectively, are applied in the context of the problem of state estimation for discrete-time dynamical systems presented in Sections 2.2 and 2.6. The scheme 3DVAR is employed in Examples 2.3, 2.5, 2.16 and B.1. The ensemble Kalman filter is applied in Example 2.16. Ensemble Kalman methods for inversion, as shown in Algorithms 3, 4 and 5, are presented in Sections 4.4 and 4.5, and applied within Examples 4.22 and 4.23. We refer to Algorithm 3 as Ensemble Kalman Inversion (Transport), as it arises from the approach to inversion described in Section 4.4.3; we refer to Algorithm 4 as Ensemble Kalman Inversion (Iteration to Infinity), as it arises from the approach to inversion described in Section 4.5; Algorithm 5 corresponds to an ensemble Kalman inversion method employing inflation by the covariance computed under the filtering distribution, as outlined in Section 4.5.3. We refer to Algorithm 5 as Ensemble Kalman Inversion (Inflated State).

Algorithm 1 3DVAR

Input: Initial $v_0 \in \mathbb{R}^{d_v}$ and fixed gain matrix $K \in \mathbb{R}^{d_v \times d_y}$.
for $n = 0$ to $N - 1$ **do**
Prediction:

$$\widehat{v}_{n+1} = \Psi(v_n).$$

Analysis:

$$v_{n+1} = \widehat{v}_{n+1} + K\left(y^{\dagger}_{n+1} - h(\widehat{v}_{n+1})\right).$$

end for
Output: Estimates $\{v_n\}_{n=0}^{N}$.

Algorithm 2 EnKF

Input: Ensemble size J, initial ensemble $\{v_0^{(j)}\}_{j=1}^J$.

for $n = 0$ to $N - 1$ **do**

Prediction: for $j = 1, \dots, J$ do

$$\widehat{v}_{n+1}^{(j)} = \Psi(v_n^{(j)}) + \xi_n^{(j)}, \quad \xi_n^{(j)} \sim \mathrm{N}(0, \Sigma).$$

Compute

$$\widehat{m}_{n+1} = \frac{1}{J}\sum_{j=1}^{J} \widehat{v}_{n+1}^{(j)}, \quad \widehat{o}_{n+1} = \frac{1}{J}\sum_{j=1}^{J} h\big(\widehat{v}_{n+1}^{(j)}\big),$$

$$\widehat{C}_{n+1}^{vh} = \frac{1}{J}\sum_{j=1}^{J} \big(\widehat{v}_{n+1}^{(j)} - \widehat{m}_{n+1}\big) \otimes \big(\widehat{v}_{n+1}^{(j)} - \widehat{o}_{n+1}\big),$$

$$\widehat{C}_{n+1}^{hh} = \frac{1}{J}\sum_{j=1}^{J} \big(\widehat{v}_{n+1}^{(j)} - \widehat{o}_{n+1}\big) \otimes \big(\widehat{v}_{n+1}^{(j)} - \widehat{o}_{n+1}\big),$$

$$K_{n+1} = \widehat{C}_{n+1}^{vh}\big(\widehat{C}_{n+1}^{hh} + \Gamma\big)^{-1}.$$

Analysis: for $j = 1, \dots, J$ do

$$\widehat{y}_{n+1}^{(j)} = h\big(\widehat{v}_{n+1}^{(j)}\big) + \eta_n^{(j)}, \quad \eta_n^{(j)} \sim \mathrm{N}(0, \Gamma),$$
$$v_{n+1}^{(j)} = \widehat{v}_{n+1}^{(j)} + K_{n+1}\big(y_{n+1}^{(j)} - h\big(\widehat{v}_{n+1}^{(j)}\big)\big).$$

end for

Output: Ensembles $\{v_n^{(j)}\}_{j=1}^J$ for $n = 0, \dots, N$.

Algorithm 3 Ensemble Kalman Inversion (Transport to Finite Time)

Input: Data $w^\dagger$, N and Δt such that $N\Delta t = 1$, ensemble size J, initial ensemble $\{u_0^{(j)}\}_{j=1}^J$.

for $n = 0$ to $N-1$ **do**

Prediction: for $j = 1, \dots, J$ do

$$\widehat{u}_{n+1}^{(j)} = u_n^{(j)}.$$

Compute

$$\widehat{m}_{n+1} = \frac{1}{J}\sum_{j=1}^{J} \widehat{u}_{n+1}^{(j)}, \quad \widehat{o}_{n+1} = \frac{1}{J}\sum_{j=1}^{J} G\big(\widehat{u}_{n+1}^{(j)}\big),$$

$$\widehat{C}_{n+1}^{uG} = \frac{1}{J}\sum_{j=1}^{J} \big(\widehat{u}_{n+1}^{(j)} - \widehat{m}_{n+1}\big) \otimes \big(G\big(\widehat{u}_{n+1}^{(j)}\big) - \widehat{o}_{n+1}\big),$$

$$\widehat{C}_{n+1}^{GG} = \frac{1}{J}\sum_{j=1}^{J} \big(G\big(\widehat{u}_{n+1}^{(j)}\big) - \widehat{o}_{n+1}\big) \otimes \big(G\big(\widehat{u}_{n+1}^{(j)}\big) - \widehat{o}_{n+1}\big).$$

Analysis: for $j = 1, \dots, J$ do

$$\widehat{w}_{n+1}^{(j)} = G\big(\widehat{u}_{n+1}^{(j)}\big) + \eta_n^{(j)}, \quad \eta_n^{(j)} \sim \mathrm{N}\left(0, \frac{\Gamma}{\Delta t}\right),$$

$$u_{n+1}^{(j)} = \widehat{u}_{n+1}^{(j)} + \Delta t \widehat{C}_{n+1}^{uG}\big(\Delta t \widehat{C}_{n+1}^{GG} + \Gamma\big)^{-1}\big(w^\dagger - \widehat{w}_{n+1}^{(j)}\big).$$

end for

Output: Ensemble $\{u_N^{(j)}\}_{j=1}^J$.

Algorithm 4 Ensemble Kalman Inversion (Iteration to Infinity)

Input: Data $w^\dagger$, N_∞, Δt, ensemble size J, initial ensemble $\{u_0^{(j)}\}_{j=1}^J$.
while $n < N_\infty$ **do**
Prediction: for $j = 1, \dots, J$ do

$$\widehat{u}_{n+1}^{(j)} = u_n^{(j)}.$$

Compute

$$\widehat{m}_{n+1} = \frac{1}{J}\sum_{j=1}^{J} \widehat{u}_{n+1}^{(j)}, \quad \widehat{o}_{n+1} = \frac{1}{J}\sum_{j=1}^{J} G\left(\widehat{u}_{n+1}^{(j)}\right),$$

$$\widehat{C}_{n+1}^{uG} = \frac{1}{J}\sum_{j=1}^{J} \left(\widehat{u}_{n+1}^{(j)} - \widehat{m}_{n+1}\right) \otimes \left(G\left(\widehat{u}_{n+1}^{(j)}\right) - \widehat{o}_{n+1}\right),$$

$$\widehat{C}_{n+1}^{GG} = \frac{1}{J}\sum_{j=1}^{J} \left(G\left(\widehat{u}_{n+1}^{(j)}\right) - \widehat{o}_{n+1}\right) \otimes \left(G\left(\widehat{u}_{n+1}^{(j)}\right) - \widehat{o}_{n+1}\right).$$

Analysis: for $j = 1, \dots, J$ do

$$\widehat{w}_{n+1}^{(j)} = G\left(\widehat{u}_{n+1}^{(j)}\right) + \eta_n^{(j)}, \quad \eta_n^{(j)} \sim \mathrm{N}\left(0, \frac{\Gamma}{\Delta t}\right),$$

$$u_{n+1}^{(j)} = \widehat{u}_{n+1}^{(j)} + \Delta t \widehat{C}_{n+1}^{uG}\left(\Delta t \widehat{C}_{n+1}^{GG} + \Gamma\right)^{-1}\left(w^\dagger - \widehat{w}_{n+1}^{(j)}\right).$$

end while
Output: Ensemble $\{u_{N_\infty}^{(j)}\}_{j=1}^J$.

Algorithm 5 Ensemble Kalman Inversion (Inflated State)

Input: Data $w_R^\dagger$, N_∞, Δt, ensemble size J, initial ensemble $\{u_0^{(j)}\}_{j=1}^J$.
while $n < N_\infty$ **do**
Prediction: Compute

$$m_n = \frac{1}{J}\sum_{j=1}^{J} u_n^{(j)}, \quad C_n = \frac{1}{J}\sum_{j=1}^{J}(u_n^{(j)} - m_n) \otimes (u_n^{(j)} - m_n)$$

for $j = 1, \ldots, J$ do

$$\widehat{u}_{n+1}^{(j)} = u_n^{(j)} + \xi_n^{(j)}, \quad \xi_n^{(j)} \sim \mathrm{N}\left(0, \frac{\Delta t}{1-\Delta t} C_n\right).$$

Compute

$$\widehat{m}_{n+1} = \frac{1}{J}\sum_{j=1}^{J} \widehat{u}_{n+1}^{(j)}, \quad \widehat{o}_{n+1} = \frac{1}{J}\sum_{j=1}^{J} G_R\big(\widehat{u}_{n+1}^{(j)}\big),$$

$$\widehat{C}_{R,n+1}^{uG} = \frac{1}{J}\sum_{j=1}^{J}\big(\widehat{u}_{n+1}^{(j)} - \widehat{m}_{n+1}\big) \otimes \big(G_R\big(\widehat{u}_{n+1}^{(j)}\big) - \widehat{o}_{n+1}\big),$$

$$\widehat{C}_{R,n+1}^{GG} = \frac{1}{J}\sum_{j=1}^{J}\big(G_R\big(\widehat{u}_{n+1}^{(j)}\big) - \widehat{o}_{n+1}\big) \otimes \big(G_R\big(\widehat{u}_{n+1}^{(j)}\big) - \widehat{o}_{n+1}\big).$$

Analysis: for $j = 1, \ldots, J$ do

$$\widehat{y}_{n+1}^{(j)} = \Delta t G_R\big(\widehat{u}_{n+1}^{(j)}\big) + \eta_n^{(j)}, \quad \eta_n^{(j)} \sim \mathrm{N}(0, \Delta t \Gamma_R),$$
$$u_{n+1}^{(j)} = \widehat{u}_{n+1}^{(j)} + \widehat{C}_{R,n+1}^{uG}\big(\Delta t \widehat{C}_{R,n+1}^{GG} + \Gamma_R\big)^{-1}\big(\Delta t w_R^\dagger - \widehat{y}_{n+1}^{(j)}\big).$$

end while
Output: Ensemble $\{u_{N_\infty}^{(j)}\}_{j=1}^J$.

B. Lorenz '96 models

To illustrate the problems of both state estimation and parameter estimation we use, throughout this paper, variants on the Lorenz '96 model. In particular, we use both the Lorenz '96 multiscale system, introduced in Section B.1, and a single-scale closure derived from it in the scale-separated case, described in Sections B.2 and 2.7. If we (i) generate data with the single-scale model, and assimilate using the same model, we are able to test algorithms in their basic (perfect model) form; on the other hand, if we (ii) generate data with the multiscale model, and assimilate using the single-scale model, this allows us to study the effect of model misspecification on data assimilation. Section B.3 contains an example of type (ii), whilst Examples 2.3, 2.5 and 2.16 are of type (i).

B.1. Lorenz '96 multiscale model

Let $v \in C(\mathbb{R}^+, \mathbb{R}^L)$ and $w \in C(\mathbb{R}^+, \mathbb{R}^{L\times J})$. We write down an ODE for (v, w) in which each variable $v_\ell \in \mathbb{R}$ is coupled to a subgroup of fast variables $w_\ell = \{w_{\ell,j}\}_{j=1}^J \in \mathbb{R}^J$; further (discrete) boundary condition couplings impose periodicity in L and link $\{w_{\ell,j}\}_{j=1}^J$ to $\{w_{\ell+1,j}\}_{j=1}^J$. The particular form of the system of ODEs is as given in Fatkullin and Vanden-Eijnden (2004). For $\ell = 1 \dots L$ and $j = 1 \dots J$, the ODEs are

$$\dot{v}_\ell = g_\ell(v) + h_v \overline{w}_\ell, \quad \overline{w}_\ell = \frac{1}{J}\sum_{j=1}^J w_{\ell,j}, \tag{B.1a}$$

$$\dot{w}_{\ell,j} = \frac{1}{\epsilon} r_j(v_\ell, w_\ell), \tag{B.1b}$$

where

$$g_\ell(v) \coloneqq -v_{\ell-1}(v_{\ell-2} - v_{\ell+1}) - v_\ell + F, \tag{B.2a}$$

$$r_j(v_\ell, w_\ell) \coloneqq -w_{\ell,j+1}(w_{\ell,j+2} - w_{\ell,j-1}) - w_{\ell,j} + h_w v_\ell, \tag{B.2b}$$

and we impose the boundary conditions

$$v_{\ell+L} = v_\ell, \quad w_{\ell+L,j} = w_{\ell,j}, \quad w_{\ell,j+J} = w_{\ell+1,j}. \tag{B.3}$$

Here $\epsilon > 0$ is a scale-separation parameter, $h_v, h_w \in \mathbb{R}$ govern the couplings between the fast and slow systems, and $F > 0$ provides a constant forcing.

B.2. Lorenz '96 single-scale model

Let $v = \{v_\ell\}_{\ell=1}^L$, $w = \{w_\ell\}_{\ell=1}^L$ and $\overline{w} = \{\overline{w}_\ell\}_{\ell=1}^L$. Then we may write (B.1) in the form

$$\dot{v} = G(v) + h_v \overline{w}, \tag{B.4a}$$

$$\dot{w} = \frac{1}{\epsilon} R(v, w), \tag{B.4b}$$

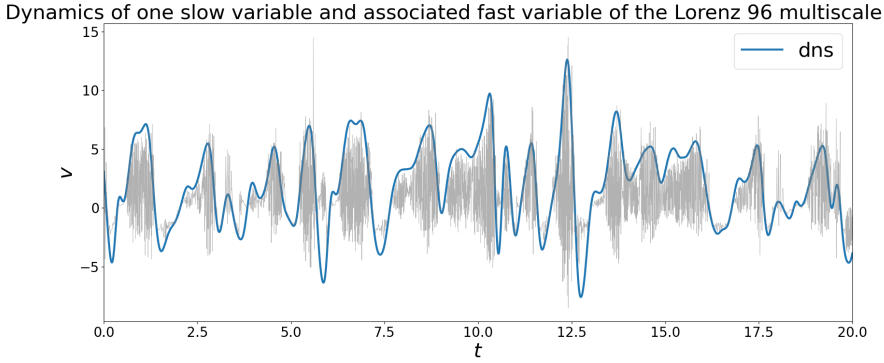


Figure B.1. Dynamics of a slow variable and an associated fast variable. Here 'dns', in blue, labels the slow variable computed by direct numerical simulation; the related fast variable is shown in grey.

for suitable definitions of G, R. If $\epsilon \ll 1$ then the dynamics for the w governed (B.4b) evolve on a much faster time-scale than the dynamics for the v governed by (B.4a). Thus it is a reasonable approximation to think of v as frozen in (B.1b). If we assume that the dynamics of w with v frozen are ergodic with invariant measure $\mu^v(\mathrm{d}w)$ (a measure in w, parametrized by v), then the averaging principle (Vanden-Eijnden 2003, Abdulle, Weinan, Engquist and Vanden-Eijnden 2012, Pavliotis and Stuart 2008) suggests that we may make the following approximation of $\overline{w}$:

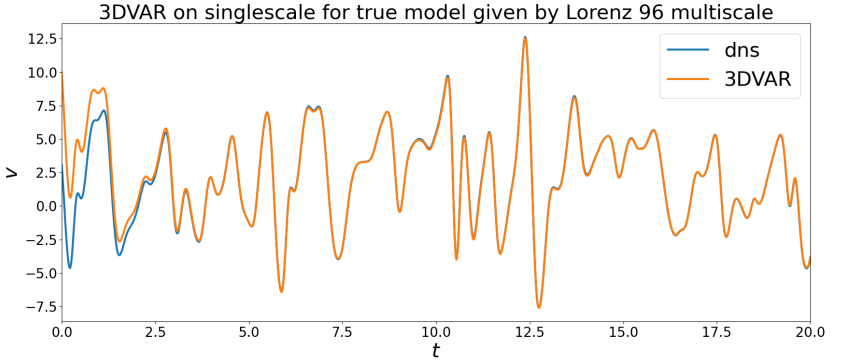
$$M(v) := \int \overline{w}\, \mu^v(\mathrm{d}w).$$

If we also invoke the approximation $M_\ell(v) \approx m(v_\ell)$, which is empirically shown to be valid for large J in Fatkullin and Vanden-Eijnden (2004), then we arrive at the single-scale Lorenz '96 model (2.9). This program of analysis for the Lorenz '96 model, and studies of the validity of the resulting single-scale approximation, was established in the paper by Fatkullin and Vanden-Eijnden (2004). The function $m(\cdot)$ is not given explicitly, but may be fitted to data in various different ways, as explained in Fatkullin and Vanden-Eijnden (2004). Figure 2.1 shows such an m, fitted using Gaussian process regression methodology as detailed in Section 4.3 of Burov, Giannakis, Manohar and Stuart (2021).

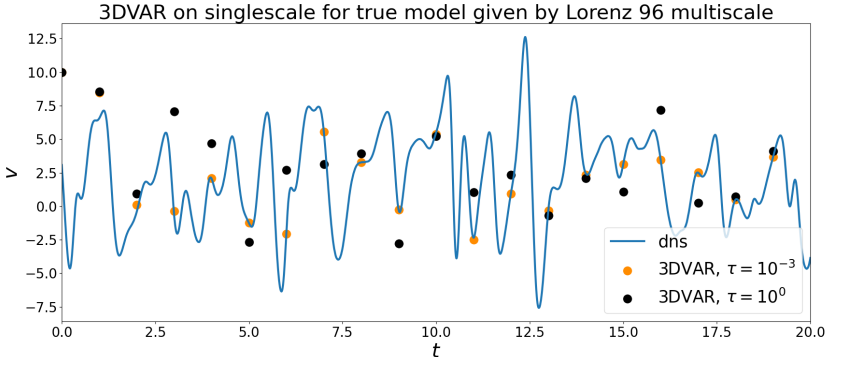
B.3. Example

The following example relates to the discussion in Remark 2.4.

Example B.1. Throughout this example we set $L = 9$, $J = 8$, $h_v = -0.8$, $h_w = 1$, $F = 10$ and $\epsilon = 2^{-7}$. Note that equation (B.1a) has the form of the single-scale Lorenz model (2.9) for v with the function $m(\cdot)$ replaced by a coupling to the fast variables w governed by (B.1b). In Figure B.1 we display the dynamics of a slow



(a)



(b)

Figure B.2. (a) Estimates of v_3 in time produced by 3DVAR using $\tau = 10^{-3}$, displayed against the true dynamics. (b) Estimates at each unit time obtained using 3DVAR with assimilation at $\tau = 10^0$ and $\tau = 10^{-3}$. Again the acronym 'dns' refers to direct numerical simulation. 3DVAR successfully synchronizes with the direct numerical simulation at the smaller value of τ but fails to do so as well when τ is larger. It is noteworthy that the synchronization takes place here in the context of model misspecification: the data is generated by the multiscale model, but 3DVAR is applied using the single-scale model.

variable and one associated fast variable of the Lorenz '96 multiscale model (B.1); at the parameter values chosen the system is chaotic.

We consider observations $\{y_n^\dagger\}_{n\in\mathbb{Z}^+}$ arising from the model

$$\begin{aligned}\left(v_{n+1}^\dagger, w_{n+1}^\dagger\right) &= \Psi_{\text{mult}}(v_n^\dagger, w_n^\dagger),\\ y_{n+1}^\dagger &= h\left(v_{n+1}^\dagger\right),\end{aligned}$$

where Ψ_{mult} is the solution operator to the multiscale model (B.1)–(B.3) over the observation time interval τ. We then take the data from the multiscale model

and assimilate it into the single-scale model (2.9), recalling the specific function m shown in Figure 2.1, and discussion in the preceding subsection concerning elimination of the fast variables in favour of a simple closure, using the averaging principle.

We now discuss data assimilation using this multiscale data. As in Example 2.3, we assume that the observation function is linear: $h(v) = Hv$ for matrix $H : \mathbb{R}^9 \to \mathbb{R}^6$ defined by (2.10). As before, we choose the gain $K : \mathbb{R}^6 \to \mathbb{R}^9$ to be defined by (2.12) and employ the 3DVAR algorithm (2.11) with Ψ the solution operator over time-interval τ for the single-scale model. Thus model misspecification is present, because data $Y^\dagger$ comes from the multiscale model.

As in Example 2.3 we display the results of 3DVAR on the unobserved component v_3. Figure B.2(a) shows the behaviour of the algorithm for $\tau = 10^{-3}$. Despite the fact that the data is generated from the multiscale model, whilst assimilation is conducted using the single-scale model, 3DVAR produces an accurate estimate of the true state with, after synchronization, the only discernible errors being slight under- and overshoots. Figure B.2(b) shows the effect of varying τ, the time between observations. As in Example 2.3, the assimilation is significantly worse when τ is larger. □

C. Mean-field maps

In Section C.1 we discuss the existence, form and properties of mean-field maps which carry out the program of approximate transport described in Section 2.5.4; these maps require access to simulated data and are hence referred to as stochastic second-order transport maps. In Section C.2 we discuss the existence, form and properties of mean-field maps which carry out the program of approximate transport described in Section 2.5.5; these maps do not require access to simulated noisy data and are hence referred to as deterministic second-order transport maps. The two approaches provide fundamental underpinnings of ensemble Kalman methods as we develop them in this paper, but were not adopted in its historical development. Section C.3 is devoted to the minimum variance approximation, a way of deriving the Kalman transport map (2.55), which is part of the historical development of the subject but does not play a central role in our presentation and analysis of the subject.

C.1. Mean-field maps: simulated data

We note that the maps of interest in this subsection take $Law(\widehat{v}_{n+1}, \widehat{y}_{n+1})$ into $Law(v_{n+1})$. Since the analysis is independent of any specific value of discrete-time n, we drop the subscript $n+1$ throughout the analysis; this streamlines the notation. Similar considerations apply in Sections C.2 and C.3.

Let $(\widehat{v}, \widehat{y}) \sim \pi$, where

$$\mathcal{G}\pi = \mathsf{N}\left(\begin{bmatrix}\widehat{m}\\ \widehat{o}\end{bmatrix}, \begin{bmatrix}\widehat{C} & \widehat{C}^{vy}\\ (\widehat{C}^{vy})^\top & \widehat{C}^{yy},\end{bmatrix}\right).$$

Define

$$m = \widehat{m} + \widehat{C}^{vy}(\widehat{C}^{yy})^{-1}(y^\dagger - \widehat{o}), \tag{C.1a}$$

$$C = \widehat{C} - \widehat{C}^{vy}(\widehat{C}^{yy})^{-1}(\widehat{C}^{vy})^\top. \tag{C.1b}$$

Note that these are the mean and covariance of the Gaussian random variable $\widehat{v}$ conditioned on $\widehat{y} = y^\dagger$ under the assumption that $(\widehat{v}, \widehat{y})$ is distributed according to the Gaussian measure $\mathcal{G}\pi$; see (2.41). The quantities in the above identities are as defined in Section 2.3.4, with the caveat that the subscripts $n+1$ have been dropped for clarity of exposition.

Our goal is to identify maps of the form

$$v = A\widehat{v} + B\widehat{y} + a, \tag{C.2}$$

so that if $(\widehat{v}, \widehat{y}) \sim \pi$, then v has mean m and covariance C given by (C.1). We make the following assumptions on the covariance under $\mathcal{G}\pi$ and on the matrices A, B and vector a.

Assumptions C.1. The covariance under $\mathcal{G}\pi$ is invertible. The matrices A, B and vector a may depend on $y^\dagger$ and measure π but not on the random variables $(\widehat{v}, \widehat{y})$; thus (C.2) takes the explicit form

$$v = A(\pi, y^\dagger)\widehat{v} + B(\pi, y^\dagger)\widehat{y} + a(\pi, y^\dagger). \qquad \square$$

Recall the discussion of pushforward of measures in the introduction to Section 2.5. Under Assumptions C.1, pushforward under the map (C.2), when chosen to match the desired first- and second-order statistics, defines a nonlinear map on the space of measures, and in particular on π itself. In what follows, all expectations are computed under π. Since the covariance under $\mathcal{G}\pi$ is strictly positive definite, so are the marginal covariances $\widehat{C}$ and $\widehat{C}^{yy}$ (see Stuart 2010, Lemma 6.21). Thus we may define the conditional mean and covariance by (C.1), as well as Kalman gain K, and conditional covariance $\widetilde{C}$, given by

$$K = \widehat{C}^{vy}(\widehat{C}^{yy})^{-1}, \tag{C.3}$$

$$\widetilde{C} = \widehat{C}^{yy} - (\widehat{C}^{vy})^\top(\widehat{C})^{-1}\widehat{C}^{vy}; \tag{C.4}$$

we note that the conditional covariances C and $\widetilde{C}$ are also strictly positive definite (see Stuart 2010, Lemma 6.21). From equations (C.1a) and (C.2), the following is immediate.

Lemma C.2. Let Assumptions C.1 hold and let $(\widehat{v}, \widehat{y}) \sim \pi$. If v given by (C.2) has mean given by (C.1a), then

$$a = (I - A)\mathbb{E}\widehat{v} + Ky^\dagger - (B + K)\mathbb{E}\widehat{y}. \qquad \diamond$$

As a consequence it follows that

$$v = \mathbb{E}\widehat{v} + A(\widehat{v} - \mathbb{E}\widehat{v}) + B(\widehat{y} - \mathbb{E}\widehat{y}) + K(y^\dagger - \mathbb{E}\widehat{y}), \tag{C.5}$$

and that

$$\mathbb{E}((v - \mathbb{E}v) \otimes (v - \mathbb{E}v)) = A\widehat{C}A^\top + B\widehat{C}^{yy}B^\top + A\widehat{C}^{vy}B^\top + B(\widehat{C}^{vy})^\top A^\top.$$

Thus, to match the covariance of the conditioned random variable, we obtain

$$C = A\widehat{C}A^\top + B\widehat{C}^{yy}B^\top + A\widehat{C}^{vy}B^\top + B(\widehat{C}^{vy})^\top A^\top. \tag{C.6}$$

Theorem C.3. Let Assumptions C.1 hold and let $(\widehat{v}, \widehat{y}) \sim \pi$. If a is given by Lemma C.2, then v defined by (C.2) has covariance (C.1b) if and only if real-valued matrices A and B are related by the identity

$$F\widehat{C}^{-1}F^\top = C - B\widetilde{C}B^\top, \tag{C.7}$$

where

$$F = A\widehat{C} + B(\widehat{C}^{vy})^\top. \tag{C.8}$$

◇

Proof. We complete the square on the right-hand side of (C.6) to obtain

$$(A\widehat{C} + B(\widehat{C}^{vy})^\top)\widehat{C}^{-1}(A\widehat{C} + B(\widehat{C}^{vy})^\top)^\top = C',$$

where

$$C' = C - B\widetilde{C}B^\top.$$

Rearranging gives the desired result. □

Define

$$\mathcal{B} = \{B \in \mathbb{R}^{d_v \times d_y} : C' \succ 0\},$$

noting that this set is non-empty and contains an open (and hence uncountable) set of B, since $C \succ 0$. For $B \in \mathcal{B}$ consider the eigenvalue problem

$$C'\varphi^{(i)} = (s^{(i)})^2\varphi^{(i)}, \quad \langle \varphi^{(i)}, \varphi^{(j)} \rangle = \delta_{ij}.$$

Note that this has d_v real solutions, up to sign changes in the eigenvectors and assuming the $s^{(i)}$ to be non-negative. We now seek to express F in terms of $B \in \mathcal{B}$. Writing the SVD given by $F\widehat{C}^{-1/2} = U\Xi V^\top$, where $U, V \in \mathbb{R}^{d_v \times d_v}$ are orthogonal matrices and $\Xi \in \mathbb{R}^{d_v \times d_v}$ is a diagonal matrix, we see from (C.7) that

$$U\Xi^2 U^\top = C',$$

so that U has columns given by the $\{\varphi^{(i)}\}_{i=1}^{d_v}$ and corresponding diagonal entries of Ξ, $\pm s^{(i)}$. We define

$$U = (\varphi^{(1)}, \ldots, \varphi^{(d_v)}), \quad \Xi = \mathrm{diag}(\pm s^{(1)}, \ldots, \pm s^{(d_v)}). \tag{C.9}$$

Corollary C.4. For every $B \in \mathcal{B}$, the choices of A such that the pair (A, B) satisfies the criterion of Theorem C.3 are defined as follows. For U, Ξ as given in (C.9), set

$$F = U\Xi V^\top \widehat{C}^{1/2},$$

where V is an arbitrary orthogonal matrix in $\mathbb{R}^{d_v \times d_v}$. Then

$$A = (F - B(\widehat{C}^{vy})^\top)\widehat{C}^{-1}. \qquad \diamond$$

Example C.5. Among the uncountably many possible solutions for pairs (A, B), we highlight two. The choice $B = 0$ is interesting because it does not require the data variable $\widehat{y}$ in the definition of v. The choice $A = I$ is interesting because it does not require action of an operator acting on $\widehat{v}$.

The first, with $B = 0$, allows the choice $F = C^{1/2}\widehat{C}^{1/2}$ and then $A = C^{1/2}\widehat{C}^{-1/2}$. Thus the map (C.5) becomes

$$v = \mathbb{E}\widehat{v} + C^{1/2}\widehat{C}^{-1/2}(\widehat{v} - \mathbb{E}\widehat{v}) + \widehat{C}^{vy}(\widehat{C}^{yy})^{-1}(y^\dagger - \mathbb{E}\widehat{y}) \tag{C.10}$$

$$= m + C^{1/2}\widehat{C}^{-1/2}(\widehat{v} - \mathbb{E}\widehat{v}). \tag{C.11}$$

The second comes from setting $B = -K$, which leads to the possible choice $F = C$ and $A = I$, under which the map (C.5) becomes

$$v = \widehat{v} + \widehat{C}^{vy}(\widehat{C}^{yy})^{-1}(y^\dagger - \widehat{y}).$$

We refer to this as the *Kalman transport* solution. □

Given the plethora of solutions for matrices (A, B), all of which effect the desired measure transport from π into the approximation for the conditional, it is natural to ask how to choose a specific pair (A, B). One possibility is to use optimal transport. To this end we define, for positive definite $W \in \mathbb{R}^{d_v \times d_v}$,

$$I_W(A, B) = \frac{1}{2}\mathbb{E}\langle (v - \widehat{v}), W(v - \widehat{v})\rangle.$$

Theorem C.6. Let v be given by (C.5). Consider the problem of finding minimizers of $I_W(A, B)$ over pairs (A, B) satisfying (C.7) and (C.8); we refer to the resulting map evaluated at such an (A, B) as an *optimal transport in the W-weighted Euclidean distance*. For any positive definite W, such minimizers satisfy $B = 0$. In particular, the Kalman transport solution is not an optimal transport solution. ◇

Proof. We formulate the optimization problem over the pair (F, B) since, because $\widehat{C}$ is invertible, there is a bijection between (A, B) and (F, B). Let : denote the Frobenius inner product. Then, under constraint (C.8),

$$I_W(A, B) = J_W(F) + \text{const.}, \tag{C.12a}$$

$$J_W(F) = -F : W, \tag{C.12b}$$

where const. denotes a matrix independent of A, B, F (with exact value changing from instance to instance). To see this, note that

$$v - \widehat{v} = -(\widehat{v} - \mathbb{E}\widehat{v}) + A(\widehat{v} - \mathbb{E}\widehat{v}) + B(\widehat{y} - \mathbb{E}\widehat{y}) + K(y^\dagger - \mathbb{E}\widehat{y}).$$

Now, using (C.6),

$$\mathbb{E}((v - \widehat{v}) \otimes (v - \widehat{v})) = \widehat{C} - \widehat{C}A^\top - A\widehat{C} - \widehat{C}^{vy}B^\top - B(\widehat{C}^{vy})^\top + C + \text{const.}$$

Noting that

$$I_W(A,B) = \frac{1}{2}\mathbb{E}((v-\widehat{v})\otimes(v-\widehat{v})) : W,$$

and that $D : W = D^\top : W$ for all D since W is symmetric, identity (C.12) follows from (C.8). It then also follows that minimization of $I_W(A,B)$ subject to the constraints given by (C.7) and (C.8) is equivalent to minimization of $J_W(F)$ subject to the constraint (C.7).

To effect this latter minimization we introduce the Lagrange multiplier $L \in \mathbb{R}^{d_v \times d_v}$, which is symmetric because the constraint is symmetric, and define

$$\widetilde{J}_W(F,B,L) = -F : W + L : (F\widehat{C}^{-1}F^\top + B\widetilde{C}B^\top - C).$$

Differentiating with respect to F, B and L respectively gives

$$-W + 2LF\widehat{C}^{-1} = 0, \tag{C.13a}$$

$$2LB\widetilde{C} = 0, \tag{C.13b}$$

$$F\widehat{C}^{-1}F^\top + B\widetilde{C}B^\top = C. \tag{C.13c}$$

Since F, W and $\widehat{C}$ in (C.13a) are all invertible, the Lagrange multiplier L is necessarily invertible. Furthermore, since $\widetilde{C}$ is invertible because Σ is invertible, it follows from (C.13b) that $B = 0$ as required. □

Example C.7. Define symmetric matrix P, and from it symmetric A, by

$$P = (C^{1/2}\widehat{C}C^{1/2})^{-1/2}, \quad A = C^{1/2}PC^{1/2}.$$

We notice that if $B = 0$, as required for an optimal transport solution, then with this choice of the pair (A,B), equation (C.13c) has as a solution $F = A\widehat{C}$ and (C.13b) is satisfied. Furthermore, (C.13a) delivers the Lagrange multiplier L. Note that the solution is independent of the specific choice of positive definite W. □

C.2. Mean-field maps: no simulated data

Let $\widehat{v} \sim \widehat{\mu}$ and assume that $\widehat{y} = h(\widehat{v}) + \eta$, where $\eta \sim \mathsf{N}(0,\Gamma)$ is independent of $\widehat{v}$. Implicitly we have defined the joint distribution π of $(\widehat{v},\widehat{y})$. We note that then, expressed in terms of $\widehat{h} := h(\widehat{v})$ and quantities defined in (C.1),

$$\begin{aligned}\widehat{C}^{vh} &:= \mathbb{E}(\widehat{v}-\mathbb{E}\widehat{v})\otimes(\widehat{h}-\mathbb{E}\widehat{h}) = \widehat{C}^{vy},\\ \widehat{C}^{hh} &:= \mathbb{E}(\widehat{h}-\mathbb{E}\widehat{h})\otimes(\widehat{h}-\mathbb{E}\widehat{h}) = \widehat{C}^{yy} - \Gamma.\end{aligned}$$

Thus we may rewrite (C.1) as

$$m = \widehat{m} + \widehat{C}^{vh}(\widehat{C}^{hh}+\Gamma)^{-1}(y^\dagger - \mathbb{E}\widehat{h}), \tag{C.14a}$$

$$C = \widehat{C} - \widehat{C}^{vh}(\widehat{C}^{hh}+\Gamma)^{-1}(\widehat{C}^{vh})^\top. \tag{C.14b}$$

In so doing we have eliminated reference to $\widehat{y}$ and our goal becomes the identification of maps of the form

$$v = R\widehat{v} + S\widehat{h} + r, \tag{C.15}$$

so that if $\widehat{v} \sim \widehat{\mu}$ then v has mean m and covariance C given by (C.14). We make the following assumptions on the covariance under $\mathcal{G}\pi$ and on the matrices R, S and vector r.

Assumptions C.8. The covariance under $\mathcal{G}\pi$ is invertible. The matrices R, S and vector r may depend on $y^\dagger$ and measure $\widehat{\mu}$ but not on the random variable $(\widehat{v}, \widehat{h})$; thus (C.15) takes the explicit form

$$v = R(\pi, y^\dagger)\widehat{v} + S(\pi, y^\dagger)\widehat{h} + r(\pi, y^\dagger). \qquad \square$$

With these assumptions the pushforward under map (C.2), when constrained to match the desired first- and second-order statistics, defines a nonlinear map on the space of measures, and in particular on measure $\widehat{\mu}$. In what follows all expectations are computed under $\widehat{\mu}$. We note that matrix $\widehat{C}$ is invertible and that K in (C.3) may be rewritten as

$$K = \widehat{C}^{vh}(\widehat{C}^{hh} + \Gamma)^{-1}.$$

From equations (C.14a) and (C.15) the following is immediate.

Lemma C.9. Let Assumptions C.8 hold and let $\widehat{v} \sim \widehat{\mu}$. If v given by (C.15) has mean given by (C.14a), then

$$r = (I - R)\mathbb{E}\widehat{v} + Ky^\dagger - (S + K)\mathbb{E}\widehat{h}. \qquad \diamond$$

As a consequence it follows that

$$v = \mathbb{E}\widehat{v} + R(\widehat{v} - \mathbb{E}\widehat{v}) + S(\widehat{h} - \mathbb{E}\widehat{h}) + K(y^\dagger - \mathbb{E}\widehat{h}), \tag{C.16}$$

and that

$$\mathbb{E}((v - \mathbb{E}v) \otimes (v - \mathbb{E}v)) = R\widehat{C}R^\top + S\widehat{C}^{hh}S^\top + R\widehat{C}^{vh}S^\top + S(\widehat{C}^{vh})^\top R. \tag{C.17}$$

Thus, to match the covariance of the conditioned random variable, we obtain

$$C = R\widehat{C}R^\top + S\widehat{C}^{hh}S^\top + R\widehat{C}^{vh}S^\top + S(\widehat{C}^{vh})^\top R. \tag{C.18}$$

Define

$$\check{C} = \widehat{C}^{hh} - (\widehat{C}^{vh})^\top(\widehat{C})^{-1}\widehat{C}^{vh}.$$

Theorem C.10. Let Assumptions C.8 hold and let $\widehat{v} \sim \widehat{\mu}$. If s is given by Lemma C.9, then v defined by (C.15) has covariance (C.14b) if and only if real-valued matrices R and S are related by the identity

$$E\widehat{C}^{-1}E^\top = C - S\check{C}S^\top,$$

where

$$E = R\widehat{C} + S(\widehat{C}^{vh})^\top. \qquad \diamond$$

Proof. We complete the square on the right-hand side of (C.18) to obtain

$$(R\widehat{C} + S(\widehat{C}^{vh})^\top)\widehat{C}^{-1}(R\widehat{C} + S(\widehat{C}^{vh})^\top)^\top + S\check{C}S^\top = C.$$

Rearranging gives the desired result. □

Define

$$\mathcal{S} = \{S \in \mathbb{R}^{d_v \times d_y} : C - S\check{C}S^\top \succ 0\},$$

and for $S \in \mathcal{S}$ consider the eigenvalue problem

$$(C - S\check{C}S^\top)\psi^{(i)} = (o^{(i)})^2\psi^{(i)}, \quad \langle \psi^{(i)}, \psi^{(j)} \rangle = \delta_{ij},$$

noting that this has d_v real solutions, up to sign changes in the eigenvectors and assuming the $o^{(i)}$ to be non-negative. We now seek to express E in terms of $S \in \mathcal{S}$. Writing the SVD given by $E\widehat{C}^{-1/2} = W\Omega Z^\top$, where $W, Z \in \mathbb{R}^{d_v \times d_v}$ are orthogonal matrices and $\Omega \in \mathbb{R}^{d_v \times d_v}$ is a diagonal matrix, we see from (C.7) that

$$W\Omega^2 W^\top = C - S\check{C}S^\top,$$

so that W has columns given by the $\{\psi^{(i)}\}_{i=1}^{d_v}$ and corresponding diagonal entries of Ω, $\pm o^{(i)}$. We define

$$W = (\psi^{(1)}, \ldots, \psi^{(d_v)}), \quad \Omega = \mathrm{diag}(\pm o^{(1)}, \ldots, \pm o^{(d_v)}). \tag{C.19}$$

Corollary C.11. For every $S \in \mathcal{S}$, the choices of R such that the pair (R, S) satisfies the criterion of Theorem C.10 are defined as follows. For W, Ω as given in (C.19), set

$$E = W\Omega Z^\top \widehat{C}^{1/2},$$

where Z is an arbitrary orthogonal matrix in $\mathbb{R}^{d_v \times d_v}$. Then

$$R = (E - S(\widehat{C}^{vh})^\top)\widehat{C}^{-1}. \qquad \diamond$$

Example C.12. As in Example C.5, we highlight two examples, here corresponding to $S = 0$ and to $R = I$. The first, with $S = 0$, allows the choice $E = C^{1/2}\widehat{C}^{1/2}$ and hence $R = C^{1/2}\widehat{C}^{-1/2}$. We thus obtain (C.10) again.

The second comes from setting $R = I$. This leads from (C.17) to the following equation for S:

$$S\widehat{C}^{hh}S^\top + \widehat{C}^{vh}S^\top + S(\widehat{C}^{vh})^\top = -\widehat{C}^{vh}(\widehat{C}^{hh} + \Gamma)^{-1}(\widehat{C}^{vh})^\top.$$

We seek a solution for S in the form

$$S = -\widehat{C}^{vh}Y^{-1}$$

and determine Y. We obtain the equation

$$Y^{-1}\widehat{C}^{hh}Y^{-T} - Y^{-T} - Y^{-1} = -(\widehat{C}^{hh} + \Gamma)^{-1}.$$

Thus, pre-multiplying by Y and post-multiplying by $Y^\top$, we obtain

$$Y(\widehat{C}^{hh} + \Gamma)^{-1}Y^\top - Y - Y^\top + \widehat{C}^{hh} = 0,$$

which factorizes to give

$$(Y(\widehat{C}^{hh}+\Gamma)^{-1}-I)(\widehat{C}^{hh}+\Gamma)(Y(\widehat{C}^{hh}+\Gamma)^{-1}-I)^{\top}=\Gamma.$$

Thus, taking the symmetric square root, we have

$$(Y(\widehat{C}^{hh}+\Gamma)^{-1}-I)(\widehat{C}^{hh}+\Gamma)^{1/2}=\Gamma^{1/2}.$$

Hence

$$(Y(\widehat{C}^{hh}+\Gamma)^{-1}-I)=\Gamma^{1/2}(\widehat{C}^{hh}+\Gamma)^{-1/2}.$$

Rearranging gives

$$Y(\widehat{C}^{hh}+\Gamma)^{-1}=\Gamma^{1/2}(\widehat{C}^{hh}+\Gamma)^{-1/2}+(\widehat{C}^{hh}+\Gamma)^{1/2}(\widehat{C}^{hh}+\Gamma)^{-1/2}.$$

Thus

$$Y=(\Gamma^{1/2}+(\widehat{C}^{hh}+\Gamma)^{1/2})(\widehat{C}^{hh}+\Gamma)^{1/2}.$$

We obtain $S=-\widetilde{K}$, where

$$\widetilde{K}=\widehat{C}^{vh}((\widehat{C}^{hh}+\Gamma)+\Gamma^{1/2}(\widehat{C}^{hh}+\Gamma)^{1/2})^{-1}.$$

The map (C.16) becomes

$$\begin{aligned}v&=\widehat{v}-\widetilde{K}(\widehat{h}-\mathbb{E}\widehat{h})+\widehat{C}^{vh}(\widehat{C}^{hh}+\Gamma)^{-1}(y^{\dagger}-\mathbb{E}\widehat{h})\\&=\widehat{v}-\widetilde{K}(\widehat{h}-\mathbb{E}\widehat{h})+K(y^{\dagger}-\mathbb{E}\widehat{h})\\&=m+(\widehat{v}-\widehat{m})-\widetilde{K}(\widehat{h}-\mathbb{E}\widehat{h}).\end{aligned}$$

□

C.3. Minimum variance approximation

In the two preceding subsections we identified an uncountable set of transport maps that match the second-order statistics of true transport. Among all these, the *Kalman transport* (2.55) has a particular appeal because it is constructed around the *Kalman gain* familiar from filtering in the linear Gaussian setting. In this subsection we show how the principle of minimizing the variance within a class of *linear* estimators of the state, given observation, leads to this choice of transport map. We believe that the second-order transport approach, which we highlight in the main text, provides a more fundamental viewpoint on the mean-field models that underpin ensemble Kalman methods; however, the minimum variance perspective is widely adopted in the statistics and geophysics communities (see Section 2.7) and hence has an important place in the subject of ensemble Kalman methods.

Recall that the Kalman transport approach leads back to the map (2.15c), motivated by control-theoretic considerations, and identifies a specific choice of Kalman gain K_n, a choice which depends on the law of the predicted state and data. In the Gaussian case the Kalman transport map (2.55) exactly generates the desired transport, a fact that we discuss in Example 2.15. The general form of approximate stochastic second-order transport maps which we study using the second-order

transport approach in the main text is (2.50). Our goal here is to motivate a specific choice of $\widetilde{T}^S$ in (2.50c), in particular to derive (2.55c). In this subsection we achieve this by first defining, and identifying, the *best linear unbiased estimator*, BLUE for short. From this we will derive Kalman transport.

Lemma C.13. Assume that $\Gamma \succ 0$. Let all expectations be computed under $Law(\widehat{v}_{n+1}, \widehat{y}_{n+1})$, and consider m_{n+1}^{BL} in the form

$$m_{n+1}^{\mathrm{BL}} = a' + B\widehat{y}_{n+1}.$$

Define $\widehat{C}_{n+1}$, $\widehat{C}_{n+1}^{vy}$ and $\widehat{C}_{n+1}^{yy}$ by (2.32) and (2.33), and define

$$\mathsf{I}(a', B) := \mathbb{E}|\widehat{v}_{n+1} - m_{n+1}^{\mathrm{BL}}|^2.$$

Then m_{n+1}^{BL} is said to be the BLUE of $\widehat{v}_{n+1}$ given $\widehat{y}_{n+1}$ if vector a' and matrix B are independent of $(\widehat{v}_{n+1}, \widehat{y}_{n+1})$ but may depend on $Law(\widehat{v}_{n+1}, \widehat{y}_{n+1})$, and are chosen to minimize $\mathsf{I}(a', B)$. Then

$$m_{n+1}^{\mathrm{BL}} = \mathbb{E}\widehat{v}_{n+1} + \widehat{C}_{n+1}^{vy}\big(\widehat{C}_{n+1}^{yy}\big)^{-1}(\widehat{y}_{n+1} - \mathbb{E}\widehat{y}_{n+1}). \tag{C.20}$$

Furthermore, the estimator (C.20) is unbiased, that is,

$$\mathbb{E}\big(\widehat{v}_{n+1} - m_{n+1}^{\mathrm{BL}}\big) = 0,$$

and its covariance with respect to $Law(\widehat{v}_{n+1}, \widehat{y}_{n+1})$,

$$C_{n+1}^{\mathrm{BL}} := \mathbb{E}\Big(\big(\widehat{v}_{n+1} - m_{n+1}^{\mathrm{BL}}\big)\big(\widehat{v}_{n+1} - m_{n+1}^{\mathrm{BL}}\big)^{\top}\Big),$$

is given by

$$C_{n+1}^{\mathrm{BL}} = \widehat{C}_{n+1} - \widehat{C}_{n+1}^{vy}\big(\widehat{C}_{n+1}^{yy}\big)^{-1}\big(\widehat{C}_{n+1}^{vy}\big)^{\top}. \tag{C.21}$$

◇

Proof. Noting that we may, without loss of generality, reparametrize the proposed form of m_{n+1}^{BL} as

$$m_{n+1}^{\mathrm{BL}} = \mathbb{E}\widehat{v}_{n+1} + a + B(\widehat{y}_{n+1} - \mathbb{E}\widehat{y}_{n+1}),$$

we see that the desired objective to be minimized over vector–matrix pair (a, B) is

$$\begin{aligned}
\mathsf{J}(a, B) &:= \frac{1}{2}\mathbb{E}|\widehat{v}_{n+1} - \mathbb{E}\widehat{v}_{n+1} - a - B(\widehat{y}_{n+1} - \mathbb{E}\widehat{y}_{n+1})|^2 \\
&= \frac{1}{2}\mathbb{E}|\widehat{v}_{n+1} - \mathbb{E}\widehat{v}_{n+1}|^2 + \frac{1}{2}|a|^2 \\
&\quad + \frac{1}{2}\mathbb{E}|B(\widehat{y}_{n+1} - \mathbb{E}\widehat{y}_{n+1})|^2 - \mathbb{E}\langle\widehat{v}_{n+1} - \mathbb{E}\widehat{v}_{n+1}, B(\widehat{y}_{n+1} - \mathbb{E}\widehat{y}_{n+1})\rangle \\
&= \frac{1}{2}\mathbb{E}|\widehat{v}_{n+1} - \mathbb{E}\widehat{v}_{n+1}|^2 + \frac{1}{2}|a|^2 + \frac{1}{2}(BB^{\top}) : \widehat{C}_{n+1}^{yy} - B : \widehat{C}_{n+1}^{vy}.
\end{aligned}$$

Clearly the minimizer with respect to a is achieved by setting $a = 0$. Differentiating with respect to B, noting that $\widehat{C}_{n+1}^{yy}$ is symmetric, shows that $B = K_n$ given by (2.34).

That the resulting pair is indeed a minimizer, and not another critical point, follows from the fact that $\widehat{C}^{yy}_{n+1}$ is positive definite; this is implied by the assumption $\Gamma > 0$. Thus we obtain the estimator (C.20). A straightforward calculation shows that the estimator is unbiased and that its covariance is given by (C.21). □

We now connect the BLUE to an approximate (second-order) transport map. To this end, note that (C.20) may be viewed as mapping $\widehat{y}_{n+1}$ into $m^{\mathrm{BL}}_{n+1} = m^{\mathrm{BL}}(\widehat{y}_{n+1})$. With this notation we define $m^{\dagger}_{n+1}$ by

$$m^{\dagger}_{n+1} = m^{\mathrm{BL}}\big(y^{\dagger}_{n+1}\big) \tag{C.22a}$$

$$= \mathbb{E}\widehat{v}_{n+1} + \widehat{C}^{vy}_{n+1}\big(\widehat{C}^{yy}_{n+1}\big)^{-1}\big(y^{\dagger}_{n+1} - \mathbb{E}\widehat{y}_{n+1}\big). \tag{C.22b}$$

We may now make the following connection between BLUE and Kalman transport.

Theorem C.14. Let $(\widehat{v}_{n+1}, \widehat{y}_{n+1})$ be distributed according to measure π_{n+1}. Then the transport map (2.55c) has the properties

$$\mathbb{E}v_{n+1} = m^{\dagger}_{n+1}$$
$$\mathbb{E}\Big(\big(v_{n+1} - m^{\dagger}_{n+1}\big)\big(v_{n+1} - m^{\dagger}_{n+1}\big)^{\top}\Big) = C^{\mathrm{BL}}_{n+1},$$

where $m^{\dagger}_{n+1}$ and C^{BL}_{n+1} are given by (C.22) and (C.21) respectively. ◇

Proof. Note that the transport map (2.55c) can now be reformulated as

$$v_{n+1} = \widehat{v}_{n+1} + \big(m^{\dagger}_{n+1} - m^{\mathrm{BL}}_{n+1}\big).$$

Thus

$$v_{n+1} = m^{\dagger}_{n+1} + \widehat{v}_{n+1} - m^{\mathrm{BL}}_{n+1}.$$

The desired properties of the mean and covariance follow from Lemma C.13. □

Remark C.15. We now have two derivations of the Kalman gain, the approximate transport derivation from Section 2.5.4, and the minimum variance derivation given here. We include a third, dimensional, argument that motivates its form. Let `state` denote the physical units associated with the state variable and `data` those associated with the observation. Then the physical units of the gain matrix K should equal `state/data`. We note that $\widehat{C}^{yy}$ has units of `data` squared whilst the units of $\widehat{C}^{vy}$ are the product of `state` and `data`. If we then constrain the gain to be determined by covariance matrices involving the state and the data, it is natural to choose it to be formed as $\widehat{C}^{vy}(\widehat{C}^{yy})^{-1}$, where $\widehat{C}^{yy}$ is an estimate of covariance in the data, and $\widehat{C}^{vy}$ is an estimate of covariance between state and data. Making a choice of this type leads to the right units for the gain, since the innovation has units `data`, and relies only on use of first- and second-order statistics. This dimensional argument motivates the form (2.34), as derived in both Sections C.3 and 2.4. □

D. Stochastic calculus considerations

In Section D.1 we demonstrate how to change between Itô and Stratonovich integration in the Kushner–Stratonovich equation. Section D.2 contains the statement and proof of a lemma needed in the proof of Theorem 3.10. In Section D.3 we study the diamond form of stochastic integral discussed in Remark 3.11.

D.1. Derivation of the Kushner–Stratonovich equation

In Section 3, where we derived continuous-time limits from discrete models, we used the two small parameters δ and Δt. The first characterized the data frequency and the second a time-increment. In this appendix we make the choice $\delta = \Delta t$ and study the limit $\Delta t \to 0$ to derive continuum models. Studying this limit enables us to convert between different modes of stochastic integration in an explicit fashion; this may be helpful to some readers as it avoids the need to invoke abstract results on covariation. In particular we now restate Lemma 3.3 and then prove it from first principles.

Lemma D.1. Assume that $\Gamma \succ 0$ and that $z^\dagger$ is given by (3.9). The Itô and Stratonovich interpretations of the stochastic forcing term in (3.20) are related through

$$\mathrm{d}r = \langle \mathsf{h} - \mathbb{E}\mathsf{h}, \circ \mathrm{d}z^\dagger \rangle_\Gamma r - \frac{1}{2}\left\{ |\mathsf{h}|_\Gamma^2 - \mathbb{E}|\mathsf{h}|_\Gamma^2 \right\} r \, \mathrm{d}t \tag{D.1a}$$

$$= \langle \mathsf{h} - \mathbb{E}\mathsf{h}, \mathrm{d}z^\dagger - \mathbb{E}\mathsf{h} \, \mathrm{d}t \rangle_\Gamma r. \tag{D.1b}$$

◇

Proof. Consider (D.1a) and the term $r\langle \mathsf{h} - \mathbb{E}\mathsf{h}, \circ \mathrm{d}z^\dagger \rangle_\Gamma$ in particular. Define the corresponding Itô and Stratonovich integrals

$$I := \int_0^\top r \langle \mathsf{h} - \mathbb{E}\mathsf{h}, \mathrm{d}z^\dagger \rangle_\Gamma, \quad S := \int_0^\top r \langle \mathsf{h} - \mathbb{E}\mathsf{h}, \circ \mathrm{d}z^\dagger \rangle_\Gamma.$$

We aim to write S as the sum of I and an added correction. Consider the space

$$\mathcal{X} := \{ \varrho \in L^1(\mathbb{R}^{d_v}; \mathbb{R}^+) \colon \|\varrho\|_{L^1} = 1 \}$$

of probability density functions. We first choose an increasing sequence $(t_j)_{j=1,\dots,N}$ with $\Delta t := t_{j+1} - t_j$ such that $N\Delta t = T$. We next define

$$\Delta r_{t_j} := r_{t_{j+1}} - r_{t_j}, \quad \Delta \mathsf{h}_{t_j} := \mathsf{h}_{t_{j+1}} - \mathsf{h}_{t_j}, \quad \Delta z^\dagger_{t_j} := z^\dagger_{t_{j+1}} - z^\dagger_{t_j}.$$

Now recall the driving evolution equation (3.9b) for $z^\dagger$, namely

$$\mathrm{d}z^\dagger = \mathsf{h}(v^\dagger) \, \mathrm{d}t + \sqrt{\Gamma} \, \mathrm{d}B^\dagger. \tag{D.2}$$

From this, recalling the properties of Brownian motion $B^\dagger$, we deduce the discretization

$$\Delta z^\dagger_{t_j} = \mathsf{h}(v^\dagger_{t_j})\Delta t + \sqrt{\Gamma \Delta t}\xi_{t_j} + O(\Delta t^{3/2}), \tag{D.3}$$

where ξ_{t_j} are independent mean-zero normal random variables with variance I_{d_y} for each j. The fact that this discretization is accurate up to terms of $O(\Delta t^{3/2})$ follows from Itô–Taylor expansion (Kloeden and Platen 1991). Recalling the definition of the Stratonovich stochastic integral as a limit, we consider the following finite sum approximation of S:

$$S_{\Delta t} := \sum_{j=1}^{N} \frac{r_{t_{j+1}} + r_{t_j}}{2} \left\langle \mathsf{h} - \int \frac{r_{t_{j+1}} + r_{t_j}}{2} \mathsf{h}\, \mathrm{d}v, \Delta z_{t_j}^{\dagger} \right\rangle_{\Gamma}; \tag{D.4}$$

this converges in the $L^2_{\mathbb{P}}(\Omega; C([0,T]; \mathcal{X}))$ sense to S as $\Delta t \to 0$ (and thus as $N \to \infty$). Now expanding the sum in (D.4), we obtain

$$S_{\Delta t} = \sum_{j=1}^{N} \left(r_{t_j} + \frac{\Delta r_{t_j}}{2} \right) \left\langle \mathsf{h} - \int \left(r_{t_j} + \frac{\Delta r_{t_j}}{2} \right) \mathsf{h}\, \mathrm{d}v, \Delta z_{t_j}^{\dagger} \right\rangle_{\Gamma} \tag{D.5a}$$

$$= \sum_{j=1}^{N} r_{t_j} \left\langle \mathsf{h} - \int r_{t_j} \mathsf{h}\, \mathrm{d}v, \Delta z_{t_j}^{\dagger} \right\rangle_{\Gamma} + \frac{1}{2} \sum_{j=1}^{N} \Delta r_{t_j} \left\langle \mathsf{h} - \int r_{t_j} \mathsf{h}\, \mathrm{d}v, \Delta z_{t_j}^{\dagger} \right\rangle_{\Gamma} \tag{D.5b}$$

$$- \frac{1}{2} \sum_{j=1}^{N} r_{t_j} \left\langle \int \Delta r_{t_j} \mathsf{h}\, \mathrm{d}v, \Delta z_{t_j}^{\dagger} \right\rangle_{\Gamma} - \frac{1}{4} \sum_{j=1}^{N} \Delta r_{t_j} \left\langle \int \Delta r_{t_j} \mathsf{h}\, \mathrm{d}v, \Delta z_{t_j}^{\dagger} \right\rangle_{\Gamma}. \tag{D.5c}$$

Now, we note that by discretizing (D.1a) or (D.1b), we can write

$$\Delta r_{t_j} = r_{t_j} \left\langle \mathsf{h} - \int r_{t_j} \mathsf{h}\, \mathrm{d}v, \Delta z_{t_j}^{\dagger} \right\rangle_{\Gamma} + O(\Delta t).$$

By substituting this expression in the expanded sum in (D.5), we notice that the last term in (D.5c) is of order $O(N\Delta t^{3/2})$. We can thus write (D.5) as

$$S_{\Delta t} = I_{\Delta t} + J_{1,\Delta t} + J_{2,\Delta t} + O(N\Delta t^{3/2}), \tag{D.6}$$

where we have defined

$$I_{\Delta t} := \sum_{j=1}^{N} r_{t_j} \left\langle \mathsf{h} - \int r_{t_j} \mathsf{h}\, \mathrm{d}v, \Delta z_{t_j}^{\dagger} \right\rangle_{\Gamma},$$

$$J_{1,\Delta t} := \frac{1}{2} \sum_{j=1}^{N} r_{t_j} \left| \left\langle \mathsf{h} - \int r_{t_j} \mathsf{h}\, \mathrm{d}v, \Delta z_{t_j}^{\dagger} \right\rangle_{\Gamma} \right|^2,$$

$$J_{2,\Delta t} := -\frac{1}{2} \sum_{j=1}^{N} r_{t_j} \left\langle \int r_{t_j} \left\langle \mathsf{h} - \int r_{t_j} \mathsf{h}\, \mathrm{d}v, \Delta z_{t_j}^{\dagger} \right\rangle_{\Gamma} \mathsf{h}\, \mathrm{d}v, \Delta z_{t_j}^{\dagger} \right\rangle_{\Gamma}.$$

The first term in the sum (D.6), $I_{\Delta t}$, converges in the $L^2_{\mathbb{P}}$ sense to I as $\Delta t \to 0$. Now considering the term $J_{1,\Delta t}$ and using the discretization (D.3), we have that

$$\begin{aligned} J_{1,\Delta t} &= \frac{1}{2}\sum_{j=1}^{N} r_{t_j}\left|\left\langle \mathsf{h} - \int r_{t_j}\mathsf{h}\,\mathrm{d}v, \Delta z^{\dagger}_{t_j}\right\rangle_{\Gamma}\right|^2 \\ &= \frac{1}{2}\sum_{j=1}^{N} r_{t_j}\Delta t\left(\mathsf{h} - \int r_{t_j}\mathsf{h}\,\mathrm{d}v\right)^{\top}\Gamma^{-1/2}\xi_{t_j}\xi_{t_j}^{\top}\Gamma^{-1/2}\left(\mathsf{h} - \int r_{t_j}\mathsf{h}\,\mathrm{d}v\right) \\ &\quad + \sum_{j=1}^{N} O(\Delta t^{3/2}). \end{aligned}$$

Since $\sum_{j=1}^{N} O(\Delta t^{3/2}) = O(\Delta t^{1/2})$, taking the $\Delta t \to 0$ limit and using the independence of the ξ_{t_j} for each j, we see that

$$J_{1,\Delta t} \to \frac{1}{2}\int_0^{\top}\left[r(\mathsf{h} - \mathbb{E}\mathsf{h})^{\top}\Gamma^{-1/2}\mathbb{E}\xi\xi^{\top}\Gamma^{-1/2}(\mathsf{h} - \mathbb{E}\mathsf{h})\right]\mathrm{d}t = \int_0^{\top}\frac{r}{2}|\mathsf{h} - \mathbb{E}\mathsf{h}|^2_{\Gamma}\,\mathrm{d}t,$$

in the $L^2_{\mathbb{P}}$ sense. Similarly for $J_{2,\Delta t}$ we have that

$$\begin{aligned} J_{2,\Delta t} &= -\frac{1}{2}\sum_{j=1}^{N} r_{t_j}\left\langle \int r_{t_j}\left\langle \mathsf{h} - \int r_{t_j}\mathsf{h}\,\mathrm{d}v, \Delta z^{\dagger}_{t_j}\right\rangle_{\Gamma}\mathsf{h}\,\mathrm{d}v, \Delta z^{\dagger}_{t_j}\right\rangle_{\Gamma} \\ &= -\frac{1}{2}\sum_{j=1}^{N} r_{t_j}\int r_{t_j}\left\langle\left\langle \mathsf{h} - \int r_{t_j}\mathsf{h}\,\mathrm{d}v, \Delta z^{\dagger}_{t_j}\right\rangle_{\Gamma}\mathsf{h}, \Delta z^{\dagger}_{t_j}\right\rangle_{\Gamma}\mathrm{d}v \\ &= -\frac{1}{2}\sum_{j=1}^{N} r_{t_j}\Delta t\int r_{t_j}\mathsf{h}^{\top}\Gamma^{-1/2}\xi_{t_j}\xi_{t_j}^{\top}\Gamma^{-1/2}\left(\mathsf{h} - \int r_{t_j}\mathsf{h}\,\mathrm{d}v\right)\mathrm{d}v \\ &\quad + \sum_{j=1}^{N} O(\Delta t^{3/2}). \end{aligned}$$

Since $\sum_{j=1}^{N} O(\Delta t^{3/2}) = O(\Delta t^{1/2})$, taking the $\Delta t \to 0$ limit and using the independence of the ξ_{t_j} for each j, we see that, in the $L^2_{\mathbb{P}}$ sense,

$$\begin{aligned} J_{2,\Delta t} &\to -\frac{1}{2}\int_0^{\top} r\int r\mathsf{h}^{\top}\Gamma^{-1/2}\mathbb{E}(\xi\xi^{\top})\Gamma^{-1/2}(\mathsf{h} - \mathbb{E}\mathsf{h})\,\mathrm{d}v\,\mathrm{d}t \\ &= -\int_0^{\top}\frac{r}{2}\mathbb{E}|\mathsf{h} - \mathbb{E}\mathsf{h}|^2_{\Gamma}\,\mathrm{d}t. \end{aligned}$$

Finally, by taking the $\Delta t \to 0$ limit on both sides of (D.6), we conclude that

$$S = I + \int_0^{\top}\frac{r}{2}|\mathsf{h} - \mathbb{E}\mathsf{h}|^2_{\Gamma}\,\mathrm{d}t - \int_0^{\top}\frac{r}{2}\mathbb{E}|\mathsf{h} - \mathbb{E}\mathsf{h}|^2_{\Gamma}\,\mathrm{d}t.$$

Hence, by using the above Stratonovich-to-Îto correction in (D.1a), we obtain

$$\begin{aligned} \mathrm{d}r &= r\langle \mathsf{h} - \mathbb{E}\mathsf{h}, \circ \mathrm{d}z^\dagger\rangle_\Gamma - \frac{r}{2}\left\{|\mathsf{h}|_\Gamma^2 - \mathbb{E}|\mathsf{h}|_\Gamma^2\right\}\mathrm{d}t \\ &= r\langle \mathsf{h} - \mathbb{E}\mathsf{h}, \mathrm{d}z^\dagger\rangle_\Gamma + \frac{r}{2}|\mathsf{h} - \mathbb{E}\mathsf{h}|_\Gamma^2\,\mathrm{d}t - \frac{r}{2}\mathbb{E}|\mathsf{h} - \mathbb{E}\mathsf{h}|_\Gamma^2\,\mathrm{d}t - \frac{r}{2}\left\{|\mathsf{h}|_\Gamma^2 - \mathbb{E}|\mathsf{h}|_\Gamma^2\right\}\mathrm{d}t \\ &= r\langle \mathsf{h} - \mathbb{E}\mathsf{h}, \mathrm{d}z^\dagger - \mathbb{E}\mathsf{h}\,\mathrm{d}t\rangle_\Gamma. \end{aligned}$$

□

D.2. Lemma for proof of Theorem 3.10

We now establish the following lemma, the conclusions of which are used in the proof of Theorem 3.10.

Lemma D.2. The probability density for the solution v of (3.38), with respect to randomness induced by the law of $\widehat{z}$ and the initial condition $v(0)$, with $z^\dagger$ a fixed data sample path, is given by (3.39). ◇

Proof. Recall that in equation (3.38) the evolution of $z^\dagger$ is given by (D.2), with $B^\dagger$ and B independent draws from unit Brownian motion in $\mathbb{R}^{d_y}$. Thus

$$\mathrm{d}v = a(v;\rho)\,\mathrm{d}t + K(v;\rho)((\mathsf{h}(v^\dagger) - \mathsf{h}(v))\,\mathrm{d}t + \sqrt{\Gamma}\,\mathrm{d}B^\dagger - \sqrt{\Gamma}\,\mathrm{d}B).$$

We wish to find the probability density function $\rho(v,t)$ for v, with respect to randomness induced by the law of B and the initial condition $v(0)$, but with $v^\dagger, B^\dagger$ fixed signal and observational noise sample paths. Let $\phi\colon \mathbb{R}^{d_v} \to \mathbb{R}$ be smooth and note that the Itô formula shows that

$$\mathrm{d}\phi(v(t)) = (\mathcal{L}\phi)(v(t))\,\mathrm{d}t + \langle K(v(t);\rho(\cdot,t))\sqrt{\Gamma}(\mathrm{d}B^\dagger - \mathrm{d}B), \nabla\phi(v(t))\rangle,$$

where[29]

$$\mathcal{L}\psi(v) = \langle (a(v;\rho) + K(v;\rho)(\mathsf{h}(v^\dagger) - \mathsf{h}(v))), \nabla\psi\rangle + K(v;\rho)\Gamma K(v;\rho)^\top : \nabla\nabla\psi,$$

and the integrals are to be interpreted in the Itô sense. Note the factor 1 in the second-order term, arising because of the independent quadratic variation contributions from both B and $B^\dagger$. Taking expectation $\mathbb{E}$ with respect to B and initial condition $v(0)$, with $B^\dagger$ fixed, for a and K functions of v and ρ, yields, using again that $z^\dagger$ is given by (3.9),

$$\begin{aligned} \mathrm{d}\mathbb{E}\phi(v) &= \mathbb{E}\langle (a + K(\mathsf{h}(v^\dagger) - \mathsf{h})), \nabla\phi\rangle\mathrm{d}t + \mathbb{E}\langle K\sqrt{\Gamma}\,\mathrm{d}B^\dagger, \nabla\phi\rangle\rangle + \mathbb{E}K\Gamma K^\top : \nabla\nabla\phi\,\mathrm{d}t \\ &= \mathbb{E}\langle (a - K\mathsf{h}), \nabla\phi\rangle\,\mathrm{d}t + \mathbb{E}\langle K\,\mathrm{d}z^\dagger, \nabla\phi\rangle\rangle + \mathbb{E}K\Gamma K^\top : \nabla\nabla\phi\,\mathrm{d}t. \end{aligned}$$

Noting that the expectation operation corresponds to multiplication by $\rho(v,t)$ and integration over $v \in \mathbb{R}^{d_v}$, integrating by parts shows that ρ satisfies the desired equation (3.39), in a weak sense. □

[29] Here we let $\mathcal{L}$ denote the infinitesimal generator of the diffusion process.

D.3. Diamond integration

In this section we use notation analogous to that established in the proof of Lemma D.1. Consider equation (3.37), repeated here for convenience:

$$\mathrm{d}v = f(v)\,\mathrm{d}t + \sqrt{\Sigma}\,\mathrm{d}W + \nabla\cdot(K\Gamma K^\top)\,\mathrm{d}t - K\Gamma\nabla\cdot K^\top\,\mathrm{d}t + K(\mathrm{d}z^\dagger - \mathrm{d}\widehat{z}), \tag{D.7a}$$

$$\mathrm{d}\widehat{z} = \mathsf{h}(v)\,\mathrm{d}t + \sqrt{\Gamma}\,\mathrm{d}B. \tag{D.7b}$$

Our focus here is on the contribution $K(\mathrm{d}z^\dagger - \mathrm{d}\widehat{z})$, recalling that $z^\dagger$ is governed by (D.2). We first choose an increasing sequence $(t_j)_{j=0,\ldots,N-1}$ with $\Delta t := t_{j+1} - t_j$ such that $N\Delta t = T$. The Itô integral interpretation of this contribution, on time interval $(0,T)$, is as the $L^2_{\mathbb{P}}$ limit of

$$I_{\Delta t} := \sum_{j=0}^{N-1} K(v_{t_j};\rho_{t_j})\big(\Delta z^\dagger_{t_j} - \Delta\widehat{z}_{t_j}\big).$$

We define the Stratonovich integral interpretation as the $L^2_{\mathbb{P}}$ limit of

$$S_{\Delta t} := \sum_{j=0}^{N-1} K\left(\frac{v_{t_{j+1}} + v_{t_j}}{2}; \frac{\rho_{t_{j+1}} + \rho_{t_j}}{2}\right)\big(\Delta z^\dagger_{t_j} - \Delta\widehat{z}_{t_j}\big);$$

as is standard, we use $\circ$ to denote Stratonovich stochastic integration. Finally we define the diamond integral interpretation as the $L^2_{\mathbb{P}}$ limit of

$$R_{\Delta t} := \sum_{j=0}^{N-1} K\left(\frac{v_{t_{j+1}} + v_{t_j}}{2}; \rho_{t_j}\right)\big(\Delta z^\dagger_{t_j} - \Delta\widehat{z}_{t_j}\big).$$

Note that this is akin to a Stratonovich integral, but only with respect to variation of K with respect to v, not ρ. We use $\diamond$ to denote this unusual form of stochastic integration.

Throughout this subsection we move between these three forms of stochastic integration. To shorten the presentation we will sometimes use the = symbol when in fact we mean equality up to an additive constant which disappears in the $\Delta t \to 0$ limit. Note that the evolution equation (D.7) for v is driven by $z^\dagger$ and $\widehat{z}$, whereas the Kushner–Stratonovich equation (3.25) for ρ is driven only by $z^\dagger$. This difference will have implications for the calculations that follow, in which we compute interconversions between the different stochastic integrals. The first result highlights an interesting interpretation of the contribution

$$a = \nabla\cdot(K\Gamma K^\top) - K\Gamma\nabla\cdot K^\top$$

to the drift in (D.7), namely that it is simply the Itô-to-diamond correction, giving a compact re-interpretation of the mean-field model.

Lemma D.3. Let ρ solve the Kushner–Stratonovich equation (3.25) and let $z^\dagger$ be given by (3.9). The Itô interpretation of equation (3.37) and its interpretation with

respect to the $\diamond$ form of stochastic integration are related through the following equivalence. The system

$$\begin{aligned} \mathrm{d}v &= f(v)\,\mathrm{d}t + \sqrt{\Sigma}\,\mathrm{d}W + \nabla\cdot(K\Gamma K^\top)\,\mathrm{d}t - K\Gamma\nabla\cdot K^\top\,\mathrm{d}t + K(\mathrm{d}z^\dagger - \mathrm{d}\widehat{z}), \\ \mathrm{d}\widehat{z} &= \mathsf{h}(v)\,\mathrm{d}t + \sqrt{\Gamma}\,\mathrm{d}B, \end{aligned}$$

where $K = K(v,\rho)$, is equivalent to

$$\begin{aligned} \mathrm{d}v &= f(v)\,\mathrm{d}t + \sqrt{\Sigma}\,\mathrm{d}W + K\diamond(\mathrm{d}z^\dagger - \mathrm{d}\widehat{z}), \\ \mathrm{d}\widehat{z} &= \mathsf{h}(v)\,\mathrm{d}t + \sqrt{\Gamma}\,\mathrm{d}B. \end{aligned}$$ $\diamond$

Proof. We first recall the evolution equations for $z^\dagger$ and $\widehat{z}$:

$$\begin{aligned} \mathrm{d}z^\dagger &= \mathsf{h}(v^\dagger)\,\mathrm{d}t + \sqrt{\Gamma}\,\mathrm{d}B^\dagger, \\ \mathrm{d}\widehat{z} &= \mathsf{h}(v)\,\mathrm{d}t + \sqrt{\Gamma}\,\mathrm{d}B. \end{aligned}$$

Noting that v is driven by $(z^\dagger - \widehat{z})$ and that we are inter-converting between Itô and diamond integration, so that the noisy driving of the ρ equation does not play a role, we see that it suffices to consider the $L^2_{\mathbb{P}}$ limits of the two quantities

$$\begin{aligned} I_{\Delta t} &= \sum_{j=0}^{N-1} K(v_{t_j};\rho_{t_j})\sqrt{2\Gamma\Delta t}\xi_{t_j}, \\ R_{\Delta t} &= \sum_{j=0}^{N-1} K\left(\frac{v_{t_{j+1}}+v_{t_j}}{2};\rho_{t_j}\right)\sqrt{2\Gamma\Delta t}\xi_{t_j}, \end{aligned}$$

where the ξ_{t_j} are i.i.d. draws from a unit Gaussian. By adding and subtracting the Itô contribution, we obtain

$$\begin{aligned} R_{\Delta t} &= \sum_{j=0}^{N-1} K\left(\frac{v_{t_{j+1}}+v_{t_j}}{2};\rho_{t_j}\right)\sqrt{2\Gamma\Delta t}\xi_{t_j} \\ &= \sum_{j=0}^{N-1} K(v_{t_j};\rho_{t_j})\sqrt{2\Gamma\Delta t}\xi_{t_j} + \sum_{j=0}^{N-1}\left(K\left(\frac{v_{t_{j+1}}+v_{t_j}}{2};\rho_{t_j}\right) - K(v_{t_j};\rho_{t_j})\right)\sqrt{2\Gamma\Delta t}\xi_{t_j} \\ &= I_{\Delta t} + \sum_{j=0}^{N-1}\left(\frac{1}{2}D_v K(v_{t_j};\rho_{t_j})(v_{t_{j+1}}-v_{t_j}) + O(|v_{t_{j+1}}-v_{t_j}|^2)\right)\sqrt{2\Gamma\Delta t}\xi_{t_j}, \end{aligned}$$

where the last line follows from a first-order Taylor expansion. Using a discretization of the evolution for v, and neglecting the terms that do not contribute to the quadratic variation when computing the $L^2_{\mathbb{P}}$ limit, we substitute for $v_{t_{j+1}} - v_{t_j}$ to obtain

$$I_{\Delta t} + \sum_{j=0}^{N-1}\frac{1}{2}D_v K(v_{t_j};\rho_{t_j})(K(v_{t_j};\rho_{t_j})\sqrt{2\Gamma\Delta t}\xi_{t_j})\sqrt{2\Gamma\Delta t}\xi_{t_j}. \tag{D.8}$$

We now consider the correction term resulting from (D.8), which is determined by expectation of the summand with respect to the random increments ξ, scaled by Δt^{-1}. Dropping the v and ρ dependence of K and the t_j dependence for notational convenience, its kth component is given by[30]

$$\begin{aligned}\left[\mathbb{E}(D_v K)(K\Gamma^{1/2}\xi)\Gamma^{1/2}\xi\right]_k &= \left[\mathbb{E}(\partial_{v_i} K)\left[K\Gamma^{1/2}\xi\right]_i \Gamma^{1/2}\xi\right]_k \\ &= \left[\mathbb{E}(\partial_{v_i} K)(K_{il}(\Gamma^{1/2})_{lj}\xi_j)\Gamma^{1/2}\xi\right]_k \\ &= \mathbb{E}(K_{il}(\Gamma^{1/2})_{lj}\xi_j)(\partial_{v_i} K)_{kn}(\Gamma^{1/2})_{nm}\xi_m \\ &= (\partial_{v_i} K_{kn})\Gamma_{ln}K_{il}.\end{aligned}$$

But,

$$\begin{aligned}(\partial_{v_i} K_{kn})\Gamma_{ln}K_{il} &= \partial_{v_i}(K_{kn}\Gamma_{ln}K_{il}) - K_{kn}\Gamma_{ln}(\partial_{v_i} K_{il}) \\ &= [\nabla\cdot(K\Gamma K^\top)]_k - [K\Gamma\nabla\cdot(K^\top)]_k.\end{aligned}$$

Recalling that the only contributions in $R_{\Delta t}$ that do not vanish under the $\Delta t \to 0$ limit are the ones given by (D.8), taking the $\Delta t \to 0$ limit of $R_{\Delta t}$ yields

$$R = I + \int_0^\top (\nabla\cdot(K\Gamma K^\top) - K\Gamma\nabla\cdot(K^\top))\,\mathrm{d}t,$$

where the convergence is in the $L^2_{\mathbb{P}}(\Omega; C([0,T];\mathbb{R}^{d_v}))$ sense; this concludes the proof. □

The preceding lemma relates diamond integration to Itô integration; the following lemma relates it to Stratonovich integration.

Lemma D.4. Let ρ solve the Kushner–Stratonovich equation (3.25) and let $z^\dagger$ be given by (3.9). The interpretation with respect to the $\diamond$ form of stochastic integration of equation (3.37) and its Stratonovich interpretation are related through the following equivalence. The system

$$\begin{aligned}\mathrm{d}v &= f(v)\,\mathrm{d}t + \sqrt{\Sigma}\,\mathrm{d}W + K\diamond(\mathrm{d}z^\dagger - \mathrm{d}\widehat{z}),\\ \mathrm{d}\widehat{z} &= \mathsf{h}(v)\,\mathrm{d}t + \sqrt{\Gamma}\,\mathrm{d}B,\end{aligned}$$

where $K = K(v,\rho)$, is equivalent to

$$\begin{aligned}\mathrm{d}v &= f(v)\,\mathrm{d}t + b\,\mathrm{d}t + \sqrt{\Sigma}\,\mathrm{d}W + K\circ(\mathrm{d}z^\dagger - \mathrm{d}\widehat{z}),\\ \mathrm{d}\widehat{z} &= \mathsf{h}(v)\,\mathrm{d}t + \sqrt{\Gamma}\,\mathrm{d}B,\end{aligned}$$

where the term $b = b(v,\rho)$ satisfies

$$b(v,\rho) = \frac{\rho}{2} D_\rho K(\mathsf{h} - \mathbb{E}\mathsf{h}). \qquad \diamond$$

[30] Using Einstein summation convention, as described for example in Gonzalez and Stuart (2008), so that index repeated twice is summed over.

Proof. To make the inter-conversion between diamond and Stratonovich integration we need only consider the quadratic variation contribution induced by $z^\dagger$, since the equation for ρ is driven only by $z^\dagger$ and not by $\widehat{z}$. Recalling (D.2) and its discrete form (D.3), we see that it suffices to compare the $L^2_{\mathbb{P}}$ limits of

$$R_{\Delta t} = \sum_{j=0}^{N-1} K\left(\frac{v_{t_{j+1}} + v_{t_j}}{2}; \rho_{t_j}\right)\sqrt{\Gamma \Delta t}\xi^\dagger_{t_j},$$
$$S_{\Delta t} = \sum_{j=0}^{N-1} K\left(\frac{v_{t_{j+1}} + v_{t_j}}{2}; \frac{\rho_{t_{j+1}} + \rho_{t_j}}{2}\right)\sqrt{\Gamma \Delta t}\xi^\dagger_{t_j}.$$

Here $\xi^\dagger_{t_j}$ are i.i.d. draws from a unit Gaussian. By adding and subtracting the diamond contribution, we obtain

$$\begin{aligned} S_{\Delta t} &= \sum_{j=0}^{N-1} K\left(\frac{v_{t_{j+1}} + v_{t_j}}{2}; \frac{\rho_{t_{j+1}} + \rho_{t_j}}{2}\right)\sqrt{\Gamma \Delta t}\xi^\dagger_{t_j} \\ &= R_{\Delta t} + \sum_{j=0}^{N-1}\left(K\left(\frac{v_{t_{j+1}} + v_{t_j}}{2}; \frac{\rho_{t_{j+1}} + \rho_{t_j}}{2}\right) - K\left(\frac{v_{t_{j+1}} + v_{t_j}}{2}; \rho_{t_j}\right)\right)\sqrt{\Gamma \Delta t}\xi^\dagger_{t_j}. \end{aligned}$$

Thus

$$S_{\Delta t} - R_{\Delta t} = \sum_{j=0}^{N-1}\left(\frac{1}{2}D_\rho K\left(\frac{v_{t_{j+1}} + v_{t_j}}{2}; \rho_{t_j}\right)(\rho_{t_{j+1}} - \rho_{t_j}) + O(|\rho_{t_{j+1}} - \rho_{t_j}|^2)\right)\sqrt{\Gamma \Delta t}\xi^\dagger_{t_j},$$

where the last line follows from a first-order Taylor expansion in ρ. In the following we will use a discretization of the Kushner–Stratonovich equation (3.25) given in Theorem 3.4, and repeated here:

$$\mathrm{d}\rho = -\nabla \cdot (\rho f)\,\mathrm{d}t + \frac{1}{2}\nabla \cdot (\nabla \cdot (\rho \Sigma))\,\mathrm{d}t + \langle \mathsf{h} - \mathbb{E}\mathsf{h}, \mathrm{d}z^\dagger - \mathbb{E}\mathsf{h}\,\mathrm{d}t\rangle_\Gamma \rho. \tag{D.9}$$

Substituting an increment for ρ, in time, and summarizing the terms that do not contribute to the quadratic variation in the $L^2_{\mathbb{P}}$ limit as $O(\Delta t)$, we obtain

$$S_{\Delta t} - R_{\Delta t} = \sum_{j=0}^{N-1}\left(\frac{1}{2}D_\rho K\left(\frac{v_{t_{j+1}} + v_{t_j}}{2}; \rho_{t_j}\right)\left\langle \rho_{t_j}\Gamma^{-1}(\mathsf{h}_{t_j} - \mathbb{E}\mathsf{h}_{t_j}), \sqrt{\Gamma \Delta t}\xi^\dagger_{t_j}\right\rangle + O(\Delta t)\right)\sqrt{\Gamma \Delta t}\xi^\dagger_{t_j}.$$

We now perform a first-order Taylor expansion in v of $D_\rho K$. Disregarding $O(\Delta t^{3/2})$ terms which will vanish in the $\Delta t \to 0$ limit, we obtain

$$S_{\Delta t} = R_{\Delta t} + \sum_{j=0}^{N-1}\frac{1}{2}D_\rho K(v_{t_j}; \rho_{t_j})\left\langle \rho_{t_j}\Gamma^{-1}(\mathsf{h}_{t_j} - \mathbb{E}\mathsf{h}_{t_j}), \sqrt{\Gamma \Delta t}\xi^\dagger_{t_j}\right\rangle \sqrt{\Gamma \Delta t}\xi^\dagger_{t_j}. \tag{D.10}$$

We now consider the correction term resulting from (D.10), which is determined by expectation of the summand with respect to the random increments $\xi^\dagger$, scaled by Δt^{-1}. Indeed, dropping the v and ρ dependence of K, for each $j = 0, \dots, N-1$, its lth component is given by[31]

$$\begin{aligned}
&\mathbb{E}\left[\frac{1}{2} D_\rho K(v_{t_j}; \rho_{t_j}) \langle \rho_{t_j} \Gamma^{-1} (\mathsf{h}_{t_j} - \mathbb{E}\mathsf{h}_{t_j}), \sqrt{\Gamma} \xi^\dagger_{t_j} \rangle \sqrt{\Gamma} \xi^\dagger_{t_j} \right]_l \\
&\quad = \frac{1}{2} \rho_{t_j} \mathbb{E}\left(\left[\Gamma^{-1/2} (\mathsf{h}_{t_j} - \mathbb{E}\mathsf{h}_{t_j}) \right]_k \left[\xi^\dagger_{t_j} \right]_k [D_\rho K]_{lm} [\Gamma^{1/2}]_{mn} \left[\xi^\dagger_{t_j} \right]_n \right) \\
&\quad = \frac{1}{2} \rho_{t_j} [D_\rho K]_{lm} [\mathsf{h}_{t_j} - \mathbb{E}\mathsf{h}_{t_j}]_m .
\end{aligned}$$

Thus, taking the $\Delta t \to 0$ limit of $S_{\Delta t}$ yields

$$S = R + \int_0^\top \frac{1}{2} \rho D_\rho K(v; \rho)(\mathsf{h} - \mathbb{E}\mathsf{h}) \, \mathrm{d}t,$$

where the convergence is in the $L^2_{\mathbb{P}}(\Omega; C([0,T]; \mathbb{R}^{d_v}))$ sense; this concludes the proof. □

E. Flows in the Gaussian manifold

Recall that we obtained (5.26) as the continuous-time limit of its discrete-time formulation (4.32). Here we demonstrate that the same evolution equations can be derived from the gradient flow (5.20) through a sequence of approximations. To this end we first define, for any function g defined on state space $\mathbb{R}^{d_u}$,

$$C^{g\Phi} = \mathbb{E}(g(u)\Phi(u)) - \mathbb{E}(g(u))\mathbb{E}(\Phi(u)).$$

Now consider random variable u with probability density function ρ evolving according to (5.20). Then

$$\frac{\mathrm{d}}{\mathrm{d}t}(\mathbb{E}g(u)) = -C^{g\Phi}.$$

By making linear and quadratic choices for g, this identity leads to equations (5.74) for the mean m and the covariance matrix C under ρ, repeated here for convenience:

$$\begin{aligned}
\frac{\mathrm{d}m}{\mathrm{d}t} &= -\mathbb{E}(\Phi(u)(u-m)), \\
\frac{\mathrm{d}C}{\mathrm{d}t} &= -\mathbb{E}(\Phi(u)(u-m)(u-m)^\top) + C\mathbb{E}(\Phi(u)).
\end{aligned}$$

These equations do not define, in general, a closed evolution for the pair (m, C) because, in general, ρ is not Gaussian. In order to find closed evolution equations,

[31] Again using Einstein summation convention, but not with respect to index j, which simply denotes a fixed time.

we take the expectation on the right-hand side not with respect to ρ but with respect to the Gaussian $\mathsf{N}(m(t), C(t))$. The resulting closed evolution equation for the pair $(m(t), C(t))$ is different from the continuous-time Gaussian projected filter (5.26).

We now show that the two closed evolution equations (5.74) and (5.26) can be connected through a sequence of steps involving an integration by parts, which is exact under conditions regarding the tail behaviour of $\Phi(u)$, followed by a number of approximations.

The integration by parts[32] step in (5.74) results in

$$\frac{\mathrm{d}m}{\mathrm{d}t} = -C\,\mathbb{E}(\nabla\Phi(u)),$$
$$\frac{\mathrm{d}C}{\mathrm{d}t} = -C\,\mathbb{E}(D^2\Phi(u))\,C.$$

Now we use the explicit expression (5.2) for $\Phi(u)$ in terms of $G(u)$ to derive various approximations. First, using this expression, we utilize the Gauss–Newton approximation of the Hessian $D^2\Phi(u)$ to obtain

$$D^2\Phi(u) \approx DG(u)^\top \Gamma^{-1} DG(u).$$

We also replace $\nabla\Phi(u)$ with its explicit expression, resulting from (5.2), to obtain the modified evolution equations

$$\frac{\mathrm{d}m}{\mathrm{d}t} = -C\,\mathbb{E}(DG(u)^\top \Gamma^{-1}(G(u) - w^\dagger)),$$
$$\frac{\mathrm{d}C}{\mathrm{d}t} = -C\,\mathbb{E}(DG(u)^\top \Gamma^{-1} DG(u))\,C.$$

Secondly these equations are further modified by replacing expectations of product terms by products of expectations to yield

$$\frac{\mathrm{d}m}{\mathrm{d}t} = -C\,\mathbb{E}(DG(u))^\top \Gamma^{-1}\mathbb{E}(G(u) - w^\dagger),$$
$$\frac{\mathrm{d}C}{\mathrm{d}t} = -C\,\mathbb{E}(DG(u))^\top \Gamma^{-1}\mathbb{E}(DG(u))\,C.$$

The final step consists in eliminating the Jacobian $DG(u)$ using

$$C\,\mathbb{E}(DG(u))^\top = C^{uG},$$

which is obtained by yet another integration by parts under the Gaussian $\mathsf{N}(m(t), C(t))$. Thus we have recovered the Gaussian projected filter (5.26).

[32] Related to Stein's identity.

References

H. Abarbanel (2013), *Predicting the Future: Completing Models of Observed Complex Systems*, Springer.

A. Abdulle, E. Weinan, B. Engquist and E. Vanden-Eijnden (2012), The heterogeneous multiscale method, *Acta Numer.* **21**, 1–87.

W. Acevedo, J. de Wiljes and S. Reich (2017), Second-order accurate ensemble transform particle filters, *SIAM J. Sci. Comput.* **39**, A1834–A1850.

S. Agapiou, O. Papaspiliopoulos, D. Sanz-Alonso and A. Stuart (2017), Importance sampling: Intrinsic dimension and computational cost, *Statist. Sci.* **32**, 405–431.

O. Al-Ghattas and D. Sanz-Alonso (2023), Non-asymptotic analysis of ensemble Kalman updates: Effective dimension and localization, *Inform. Inference.* **13**, 1–66.

D. J. Albers, P.-A. Blancquart, M. E. Levine, E. E. Seylabi and A. Stuart (2019), Ensemble Kalman methods with constraints, *Inverse Problems* **35**, art. 095007.

L. Ambrosio, N. Gigli and G. Savaré (2008), *Gradient Flows: In Metric Spaces and in the Space of Probability Measures*, Springer.

J. Amezcua, E. Kalnay, K. Ide and S. Reich (2014), Ensemble transform Kalman–Bucy filters, *Quart. J. R. Meteorol. Soc.* **140**, 995–1004.

B. D. O. Anderson and J. B. Moore (2012), *Optimal Filtering*, Courier Corporation.

J. L. Anderson (2001), An ensemble adjustment Kalman filter for data assimilation, *Mon. Weather Rev.* **129**, 2884–2903.

M. Asch, M. Bocquet and M. Nodet (2016), *Data Assimilation: Methods, Algorithms, and Applications*, SIAM.

P. Ashwin (2003), Synchronization from chaos, *Nature* **422**(6930), 384–385.

K. J. Åström and R. M. Murray (2021), *Feedback Systems: An Introduction for Scientists and Engineers*, Princeton University Press.

A. Azouani, E. Olson and E. S. Titi (2014), Continuous data assimilation using general interpolant observables, *J. Nonlinear Sci.* **24**, 277–304.

A. Bain and D. Crisan (2008), *Fundamentals of Stochastic Filtering*, Vol. 60 of Stochastic Modelling and Applied Probability, Springer.

K. Bergemann and S. Reich (2010*a*), A localization technique for ensemble Kalman filters, *Quart. J. R. Meteorol. Soc.* **136**, 701–707.

K. Bergemann and S. Reich (2010*b*), A mollified ensemble Kalman filter, *Quart. J. R. Meteorol. Soc.* **136**, 1636–1643.

K. Bergemann and S. Reich (2012), An ensemble Kalman–Bucy filter for continuous data assimilation, *Meteorol. Z.* **21**, 213–219.

A. Beskos, D. Crisan and A. Jasra (2014), On the stability of sequential Monte Carlo methods in high dimensions, *Ann. Appl. Probab.* **24**, 1396–1445.

A. Beskos, A. Jasra, E. A. Muzaffer and A. M. Stuart (2015), Sequential Monte Carlo methods for Bayesian elliptic inverse problems, *Statist. Comput.* **25**, 727–737.

P. Bickel, B. Li and T. Bengtsson (2008), Sharp failure rates for the bootstrap particle filter in high dimensions, in *Pushing the Limits of Contemporary Statistics: Contributions in Honor of Jayanta K. Ghosh*, Institute of Mathematical Statistics, pp. 318–329.

A. N. Bishop and P. Del Moral (2019), On the stability of matrix-valued Riccati diffusions, *Electron. J. Probab.* **24**, 1–40.

A. N. Bishop and P. Del Moral (2023), On the mathematical theory of ensemble (linear-Gaussian) Kalman–Bucy filtering, *Math. Control Signals Systems* **35**, 835–903.

A. N. Bishop, P. Del Moral and A. Niclas (2020), A perturbation analysis of stochastic matrix Riccati diffusions, *Ann. Inst. Henri Poincaré Probab. Statist.* **56**, 884–916.
A. N. Bishop, P. Del Moral and S. D. Pathiraja (2018), Perturbations and projections of Kalman–Bucy semigroups, *Stoch. Process. Appl.* **128**, 2857–2904.
A. N. Bishop, P. Del Moral, K. Kamatani and B. Remillard (2019), On one-dimensional Riccati diffusions, *Ann. Appl. Probab.* **29**, 1127–1187.
C. H. Bishop, B. J. Etherton and S. J. Majumdar (2001), Adaptive sampling with ensemble transform Kalman filter, Part I: Theoretical aspects, *Mon. Weather Rev.* **129**, 420–436.
C. M. Bishop (2011), *Pattern Recognition and Machine Learning*, second edition, Springer.
A. Biswas and M. Branicki (2024), A unified framework for the analysis of accuracy and stability of a class of approximate Gaussian filters for the Navier–Stokes equations, *Nonlinearity* **37**, art. 125013.
D. Blömker, K. Law, A. M. Stuart and K. C. Zygalakis (2013), Accuracy and stability of the continuous-time 3DVAR filter for the Navier–Stokes equation, *Nonlinearity* **26**, art. 2193.
D. Blömker, C. Schillings and P. Wacker (2018), A strongly convergent numerical scheme from ensemble Kalman inversion, *SIAM J. Numer. Anal.* **56**, 2537–2562.
D. Blömker, C. Schillings, P. Wacker and S. Weissmann (2022), Continuous time limit of the stochastic ensemble Kalman inversion: Strong convergence analysis, *SIAM J. Numer. Anal.* **60**, 3181–3215.
M. Bocquet and P. Sakov (2012), Combining inflation-free and iterative ensemble Kalman filters for strongly nonlinear systems, *Nonlinear Process. Geophys.* **19**, 383–399.
M. Bocquet and P. Sakov (2014), An iterative ensemble Kalman smoother, *Quart. J. R. Meteorol. Soc.* **140**, 1521–1535.
M. Bocquet, J. Brajard, A. Carrassi and L. Bertino (2020), Bayesian inference of chaotic dynamics by merging data assimilation, machine learning and expectation–maximization, *Found. Data Sci.* **2**, 55–80.
M. Bocquet, K. S. Gurumoorthy, A. Apte, A. Carrassi, C. Grudzien and C. K. Jones (2017), Degenerate Kalman filter error covariances and their convergence onto the unstable subspace, *SIAM/ASA J. Uncertain. Quantif.* **5**, 304–333.
G. Burgers, P. J. van Leeuwen and G. Evensen (1998), Analysis scheme in the ensemble Kalman filter, *Mon. Weather Rev.* **126**, 1719–1724.
D. Burov, D. Giannakis, K. Manohar and A. Stuart (2021), Kernel analog forecasting: Multiscale test problems, *Multiscale Model. Simul.* **19**, 1011–1040.
E. Calvello, P. Monmarché, A. M. Stuart and U. Vaes (2024), Accuracy of the ensemble Kalman filter in the near-linear setting. Available at arXiv:2409.09800.
J. A. Carrillo, Y.-P. Choi, C. Totzeck and O. Tse (2018), An analytical framework for consensus-based global optimization method, *Math. Models Methods Appl. Sci.* **28**, 1037–1066.
J. A. Carrillo, F. Hoffmann, A. M. Stuart and U. Vaes (2022), Consensus-based sampling, *Stud. Appl. Math.* **148**, 1069–1140.
J. A. Carrillo, F. Hoffmann, A. M. Stuart and U. Vaes (2024), The mean field ensemble Kalman filter: Near-Gaussian setting, *SIAM J. Numer. Anal.* **62**, 2549–2587.
N. K. Chada and X. T. Tong (2022), Convergence acceleration of ensemble Kalman inversion in nonlinear settings, *Math. Comp.* **91**, 1247–1280.
N. K. Chada, Y. Chen and D. Sanz-Alonso (2021), Iterative ensemble Kalman methods: A unified perspective with some new variants, *Found. Data Sci.* **3**, 331–369.

N. K. Chada, C. Schillings and S. Weissmann (2019), On the incorporation of box-constraints for ensemble Kalman inversion, *Found. Data Sci.* **1**, 433–456.
N. K. Chada, A. M. Stuart and X. T. Tong (2020), Tikhonov regularization within ensemble Kalman inversion, *SIAM J. Numer. Anal.* **58**, 1263–1294.
J. Chen, S. Jin and L. Lyu (2022*a*), A consensus-based global optimization method with adaptive momentum estimation, *Commun. Comput. Phys.* **31**, 1296–1316.
Y. Chen and D. Oliver (2012), Ensemble randomized maximum likelihood method as an iterative ensemble smoother, *Math. Geosci.* **44**, 1–26.
Y. Chen, D. Z. Huang, J. Huang, S. Reich and A. M. Stuart (2023), Sampling via gradient flows in the space of probability measures. Available at arXiv:2310.03597.
Y. Chen, D. Sanz-Alonso and R. Willett (2022*b*), Autodifferentiable Ensemble Kalman filters, *SIAM J. Math. Data Sci.* **4**, 801–833.
Y. Cheng and S. Reich (2015), Assimilating data into scientific models: An optimal coupling perspective, in *Nonlinear Data Assimilation*, Vol. 2 of Frontiers in Applied Dynamical Systems, Springer, pp. 75–118.
A. Chernov, H. Hoel, K. J. Law, F. Nobile and R. Tempone (2021), Multilevel ensemble Kalman filtering for spatio-temporal processes, *Numer. Math.* **147**, 71–125.
N. Chopin and O. Papaspiliopoulos (2020), *An Introduction to Sequential Monte Carlo*, Springer Nature.
N. Chustagulprom, S. Reich and M. Reinhardt (2016), A hybrid ensemble transform filter for nonlinear and spatially extended dynamical systems, *SIAM/ASA J. Uncertain. Quantif.* **4**, 592–608.
J. Clark and D. Crisan (2005), On a robust version of the integral representation formula of nonlinear filtering, *Probab. Theory Related Fields* **133**, 43–56.
E. Cleary, A. Garbuno-Inigo, S. Lan, T. Schneider and A. M. Stuart (2021), Calibrate, emulate, sample, *J. Comput. Phys.* **424**, art. 109716.
M. Coghi, T. Nilssen, N. Nüsken and S. Reich (2023), Rough McKean–Vlasov dynamics for robust ensemble Kalman filtering, *Ann. Appl. Probab.* **33**, 5693–5752.
A. Corenflos, J. Thornton, G. Deligiannidis and A. Doucet (2021), Differentiable particle filtering via entropy-regularized optimal transport, in *Proceedings of the 38th International Conference on Machine Learning* (M. Meila and T. Zhang, eds), PMLR, pp. 2100–2111.
C. Cotter and S. Reich (2013), Ensemble filter techniques for intermittent data assimilation, *Radon Ser. Comput. Appl. Math.* **13**, 91–134.
S. L. Cotter, G. O. Roberts, A. M. Stuart and D. White (2013), MCMC methods for functions: Modifying old algorithms to make them faster, *Statist. Sci.* **28**, 424–446.
D. Crisan and T. Lyons (1999), A particle approximation of the solution of the Kushner–Stratonovitch equation, *Probab. Theory Related Fields* **115**, 549–578.
D. Crisan and J. Xiong (2010), Approximate McKean–Vlasov representations for a class of SPDEs, *Stochastics* **82**, 53–68.
D. Crisan, P. Del Moral and T. Lyons (1999), Discrete filtering using branching and interacting particle systems, *Markov Process. Rel. Fields* **5**, 293–318.
M. Cuturi (2013), Sinkhorn distances: Lightspeed computation of optimal transport, in *Advances in Neural Information Processing Systems 27* (C. J. Burges *et al.*, eds), Curran Associates, pp. 2292–2300.
F. Daum, J. Huang and A. Noushin (2010), Exact particle flow for nonlinear filters, in *Signal Processing, Sensor Fusion, and Target Recognition XIX* (I. Kadar, ed.), International Society for Optics and Photonics (SPIE), pp. 92–110.

J. de Wiljes and X. T. Tong (2020), Analysis of a localised nonlinear ensemble Kalman Bucy filter with complete and accurate observations, *Nonlinearity* **33**, 4752–4782.

J. de Wiljes, S. Reich and W. Stannat (2018), Long-time stability and accuracy of the ensemble Kalman–Bucy filter for fully observed processes and small measurement noise, *SIAM J. Appl. Dyn. Syst.* **17**, 1152–1181.

P. Del Moral (1997), Nonlinear filtering: Interacting particle resolution, *C. R. Acad. Sci. Math.* **325**, 653–658.

P. Del Moral (2004), *Feynman–Kac Formulae: Genealogical and Interacting Particle Systems with Applications*, Springer.

P. Del Moral and A. Guionnet (2001), On the stability of interacting processes with applications to filtering and genetic algorithms, *Ann. Inst. Henri Poincaré Probab. Statist.* **37**, 155–194.

P. Del Moral and E. Horton (2023), A theoretical analysis of one-dimensional discrete generation ensemble Kalman particle filters, *Ann. Appl. Probab.* **33**, 1327–1372.

P. Del Moral and J. Tugaut (2018), On the stability and the uniform propagation of chaos properties of ensemble Kalman–Bucy filters, *Ann. Appl. Probab.* **28**, 790–850.

P. Del Moral, A. Doucet and A. Jasra (2006), Sequential Monte Carlo samplers, *J. R. Statist. Soc. Ser. B. Statist. Methodol.* **68**, 411–436.

Z. Ding and Q. Li (2021*a*), Ensemble Kalman inversion: Mean-field limit and convergence analysis, *Statist. Comput.* **31**, 1–21.

Z. Ding and Q. Li (2021*b*), Ensemble Kalman sampler: Mean-field limit and convergence analysis, *SIAM J. Math. Anal.* **53**, 1546–1578.

Z. Ding, Q. Li and J. Lu (2021), Ensemble Kalman inversion for nonlinear problems: Weights, consistency, and variance bounds, *Found. Data Sci.* **3**, 371–411.

A. Doucet, N. De Freitas and N. Gordon (2001), An introduction to sequential Monte Carlo methods, in *Sequential Monte Carlo Methods in Practice*, Springer, pp. 3–14.

O. R. A. Dunbar, A. B. Duncan, A. M. Stuart and M.-T. Wolfram (2022), Ensemble inference methods for models with noisy and expensive likelihoods, *SIAM J. Appl. Dyn. Syst.* **21**, 1539–1572.

O. R. A. Dunbar, A. Garbuno-Inigo, T. Schneider and A. M. Stuart (2021), Calibration and uncertainty quantification of convective parameters in an idealized GCM, *J. Adv. Model. Earth Systems* **13**, art. e2020MS002454.

M. L. Eaton (2007), *Multivariate Statistics: A Vector Space Approach*, Vol. 53 of Lecture Notes: Monograph Series, Institute of Mathematical Statistics.

A. Eberle (2019), Stochastic analysis. Lecture Notes, University of Bonn.

T. A. El Moselhy and Y. M. Marzouk (2012), Bayesian inference with optimal maps, *J. Comput. Phys.* **231**, 7815–7850.

A. A. Emerick and A. C. Reynolds (2013*a*), Ensemble smoother with multiple data assimilation, *Comput. Geosci.* **55**, 3–15.

A. Emerick and A. Reynolds (2013*b*), Investigation of the sampling performance of ensemble-based methods with a simple reservoir model, *Comput. Geosci.* **17**, 325–350.

O. G. Ernst, B. Sprungk and H.-J. Starkloff (2015), Analysis of the ensemble and polynomial chaos Kalman filters in Bayesian inverse problems, *SIAM/ASA J. Uncertain. Quantif.* **3**, 823–851.

L. C. Evans (2012), *An Introduction to Stochastic Differential Equations*, American Mathematical Society.

G. Evensen (1994), Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics, *J. Geophys. Res. Oceans* **99**, 10143–10162.
G. Evensen (2018), Analysis of iterative ensemble smoothers for solving inverse problems, *Comput. Geosci.* **22**, 885–908.
G. Evensen (2019), Accounting for model errors in iterative ensemble smoothers, *Comput. Geosci.* **23**, 761–775.
G. Evensen, F. C. Vossepoel and P. J. van Leeuwen (2022), *Data Assimilation Fundamentals: A Unified Formulation of the State and Parameter Estimation Problem*, Springer Nature.
I. Fatkullin and E. Vanden-Eijnden (2004), A computational strategy for multiscale systems with applications to Lorenz 96 model, *J. Comput. Phys.* **200**, 605–638.
E. J. Fertig, J. Harlim and B. R. Hunt (2007), A comparative study of 4D-VAR and a 4D ensemble Kalman filter: Perfect model simulations with Lorenz-96, *Tellus A* **59**, 96–100.
C. Foias and G. Prodi (1967), Sur le comportement global des solutions non-stationnaires des équations de Navier–Stokes en dimension 2, *Rend. Semin. Mat. Univ. Padova* **39**, 1–34.
C. Foias, C. F. Mondaini and E. S. Titi (2016), A discrete data assimilation scheme for the solutions of the two-dimensional Navier–Stokes equations and their statistics, *SIAM J. Appl. Dyn. Syst.* **15**, 2109–2142.
M. Fornasier, H. Huang, L. Pareschi and P. Sünnen (2020), Consensus-based optimization on hypersurfaces: Well-posedness and mean-field limit, *Math. Models Methods Appl. Sci.* **30**, 2725–2751.
M. Fornasier, T. Klock and K. Riedl (2024), Consensus-based optimization methods converge globally, *SIAM J. Optim.* **34**, 2973–3004.
M. Frei and H. R. Künsch (2013), Bridging the ensemble Kalman and particle filters, *Biometrika* **100**, 781–800.
A. Garbuno-Inigo, F. Hoffmann, W. Li and A. M. Stuart (2020*a*), Interacting Langevin diffusions: Gradient structure and ensemble Kalman sampler, *SIAM J. Appl. Dyn. Syst.* **19**, 412–441.
A. Garbuno-Inigo, N. Nüsken and S. Reich (2020*b*), Affine invariant interacting Langevin dynamics for Bayesian inference, *SIAM J. Appl. Dyn. Syst.* **19**, 1633–1658.
A. Gelman, W. R. Gilks and G. O. Roberts (1997), Weak convergence and optimal scaling of random walk metropolis algorithms, *Ann. Appl. Probab.* **7**, 110–120.
M. Gesho, E. Olson and E. S. Titi (2016), A computational study of a data assimilation algorithm for the two-dimensional Navier–Stokes equations, *Commun. Comput. Phys.* **19**, 1094–1110.
M. Ghil, S. Cohn, J. Tavantzis, K. Bube and E. Isaacson (1981), Applications of estimation theory to numerical weather prediction, in *Dynamic Meteorology: Data Assimilation Methods*, Springer, pp. 139–224.
M. Goldstein (2014), Bayes linear analysis. Wiley StatsRef: Statistics Reference Online, pp. 1–7.
M. Goldstein and J. Rougier (2006), Bayes linear calibrated prediction for complex systems, *J. Amer. Statist. Assoc.* **101**, 1132–1143.
M. Goldstein and D. Wooff (2007), *Bayes Linear Statistics: Theory and Methods*, Vol. 716 of Wiley Series in Probability and Statistics, Wiley.

O. Gonzalez and A. M. Stuart (2008), *A First Course in Continuum Mechanics*, Cambridge University Press.
C. González-Tokman and B. R. Hunt (2013), Ensemble data assimilation for hyperbolic systems, *Phys. D* **243**, 128–142.
I. Goodfellow, Y. Bengio and A. Courville (2016), *Deep Learning*, MIT Press.
J. Goodman and J. Weare (2010), Ensemble samplers with affine invariance, *Commun. Appl. Math. Comput. Sci.* **5**, 65–80.
G. A. Gottwald and A. J. Majda (2013), A mechanism for catastrophic filter divergence in data assimilation for sparse observation networks, *Nonlinear Process. Geophys.* **20**, 705–712.
G. A. Gottwald and S. Reich (2021), Supervised learning from noisy observations: Combining machine-learning techniques with data assimilation, *Phys. D* **423**, art. 132911.
I. Grooms (2021), Analog ensemble data assimilation and a method for constructing analogs with variational autoencoders, *Quart. J. R. Meteorol. Soc.* **147**, 139–149.
I. Grooms, Y. Lee and A. J. Majda (2014), Ensemble Kalman filters for dynamical systems with unresolved turbulence, *J. Comput. Phys.* **273**, 435–452.
I. Grooms, Y. Lee and A. J. Majda (2015), Ensemble filtering and low-resolution model error: Covariance inflation, stochastic parameterization, and model numerics, *Mon. Weather Rev.* **143**, 3912–3924.
Y. Gu and D. S. Oliver (2007), An iterative ensemble Kalman filter for multiphase fluid flow data assimilation, *SPE J.* **12**, 438–446.
M. E. Gurtin (1982), *An Introduction to Continuum Mechanics*, Academic Press.
K. S. Gurumoorthy, C. Grudzien, A. Apte, A. Carrassi and C. K. Jones (2017), Rank deficiency of Kalman error covariance matrices in linear time-varying system with deterministic evolution, *SIAM J. Control Optim.* **55**, 741–759.
P. A. Guth, C. Schillings and S. Weissmann (2022), Ensemble Kalman filter for neural network-based one-shot inversion, in *Optimization and Control for Partial Differential Equations* (R. Herzog *et al.*, eds), De Gruyter, pp. 393–418.
S.-Y. Ha, S. Jin and D. Kim (2021), Convergence and error estimates for time-discrete consensus-based optimization algorithms, *Numer. Math.* **147**, 255–282.
E. Haber, F. Lucka and L. Ruthotto (2018), Never look back: The EnKF method and its application to the training of neural networks without back propagation. Available at arXiv:1805.08034.
M. Hairer, A. M. Stuart and S. J. Vollmer (2014), Spectral gaps for a Metropolis–Hastings algorithm in infinite dimensions, *Ann. Appl. Probab.* **24**, 2455–2490.
J. Harlim and B. R. Hunt (2007*a*), Four-dimensional local ensemble transform Kalman filter: Numerical experiments with a global circulation model, *Tellus A* **59**, 731–748.
J. Harlim and B. R. Hunt (2007*b*), A non-Gaussian ensemble filter for assimilating infrequent noisy observations, *Tellus A* **59**, 225–237.
J. Harlim and A. J. Majda (2010), Filtering turbulent sparsely observed geophysical flows, *Mon. Weather Rev.* **138**, 1050–1083.
J. Harlim, A. Mahdi and A. J. Majda (2014), An ensemble Kalman filter for statistical estimation of physics constrained nonlinear regression models, *J. Comput. Phys.* **257**, 782–812.
K. Hayden, E. Olson and E. S. Titi (2011), Discrete data assimilation in the Lorenz and 2D Navier–Stokes equations, *Phys. D* **240**, 1416–1425.

D. J. Higham (2001), An algorithmic introduction to numerical simulation of stochastic differential equations, *SIAM Rev.* **43**, 525–546.
M. W. Hirsch, S. Smale and R. L. Devaney (2013), *Differential Equations, Dynamical Systems, and an Introduction to Chaos*, Academic Press.
Y. Ho and R. C. K. A. Lee (1964), A Bayesian approach to problems in stochastic estimation and control, *IEEE Trans. Automat. Control* **9**, 333–339.
H. Hoel, K. J. Law and R. Tempone (2016), Multilevel ensemble Kalman filtering, *SIAM J. Numer. Anal.* **54**, 1813–1839.
P. L. Houtekamer and H. L. Mitchell (2005), Ensemble Kalman filtering, *Quart. J. R. Meteorol. Soc.* **131**, 3269–3289.
P. L. Houtekamer and M. L. Mitchell (1998), Data assimilation using an ensemble Kalman filter techniques, *Mon. Weather Rev.* **126**, 796–811.
Y. Hu, G. Kallianpur and J. Xiong (2002), An approximation for the Zakai equation, *Appl. Math. Optim.* **45**, 23–44.
D. Z. Huang, J. Huang, S. Reich and A. M. Stuart (2022*a*), Efficient derivative-free Bayesian inference for large-scale inverse problems, *Inverse Problems* **38**, art. 125006.
D. Z. Huang, T. Schneider and A. M. Stuart (2022*b*), Iterated Kalman methodology for inverse problems, *J. Comput. Phys.* **463**, art. 111262.
B. R. Hunt, E. J. Kostelich and I. Szunyogh (2007), Efficient data assimilation for spatiotemporal chaos: A local ensemble transform Kalman filter, *Phys. D* **230**, 112–126.
M. A. Iglesias (2015), Iterative regularization for ensemble data assimilation in reservoir models, *Comput. Geosci.* **19**, 177–212.
M. A. Iglesias (2016), A regularizing iterative ensemble Kalman method for PDE-constrained inverse problems, *Inverse Problems* **32**, art. 025002.
M. A. Iglesias, K. J. Law and A. M. Stuart (2013), Ensemble Kalman methods for inverse problems, *Inverse Problems* **29**, art. 045001.
M. Iglesias and Y. Yang (2021), Adaptive regularisation for ensemble Kalman inversion, *Inverse Problems* **37**, art. 025008.
A. H. Jazwinski (2007), *Stochastic Processes and Filtering Theory*, Courier Corporation.
S. Julier, J. Uhlmann and H. F. Durrant-Whyte (2000), A new method for the nonlinear transformation of means and covariances in filters and estimators, *IEEE Trans. Automat. Control* **45**, 477–482.
J. Kaipio and E. Somersalo (2006), *Statistical and Computational Inverse Problems*, Vol. 160, Springer.
R. E. Kalman (1960), A new approach to linear filtering and prediction problems, *J. Basic Eng.* **82**, 35–45.
R. E. Kalman and R. S. Bucy (1961), New results in linear filtering and prediction theory, *J. Basic Eng.* **83**, 95–108.
N. Kantas, A. Beskos and A. Jasra (2014), Sequential Monte Carlo methods for high-dimensional inverse problems: A case study for the Navier–Stokes equations, *SIAM/ASA J. Uncertain. Quantif.* **2**, 464–489.
D. Kelly, A. J. Majda and X. T. Tong (2015), Concrete ensemble Kalman filters with rigorous catastrophic filter divergence, *Proc. Nat. Acad. Sci. USA* **112**, 10589–10594.
D. T. B. Kelly, K. J. H. Law and A. M. Stuart (2014), Well-posedness and accuracy of the ensemble Kalman filter in discrete and continuous time, *Nonlinearity* **27**, art. 2579.
J. W. Kim and P. G. Mehta (2024*a*), Duality for nonlinear filtering I: Observability, *IEEE Trans. Automat. Control* **69**, 699–711.

J. W. Kim and P. G. Mehta (2024*b*), Duality for nonlinear filtering II: Optimal control, *IEEE Trans. Automat. Control* **69**, 712–725.
P. K. Kitanidis (1995), Quasi-linear geostatistical theory for inversion, *Water Resour. Res.* **31**, 2411–2419.
P. Kloeden and E. Platen (1991), *Numerical Methods for Stochastic Differential Equations*, Springer.
N. B. Kovachki and A. M. Stuart (2019), Ensemble Kalman inversion: A derivative-free technique for machine learning tasks, *Inverse Problems* **35**, art. 095005.
H. Kushner and G. G. Yin (2003), *Stochastic Approximation and Recursive Algorithms and Applications*, Vol. 35 of Applications of Mathematics: Stochastic Modelling and Applied Probability, Springer.
L. Kuznetsov, K. Ide and C. K. R. T. Jones (2003), A method for assimilation of Lagrangian data, *Mon. Weather Rev.* **131**, 2247–2260.
E. Kwiatkowski and J. Mandel (2015), Convergence of the square root ensemble Kalman filter in the large ensemble limit, *SIAM/ASA J. Uncertain. Quantif.* **3**, 1–17.
T. Lange and W. Stannat (2021*a*), On the continuous time limit of ensemble square root filters, *Commun. Math. Sci.* **19**, 1855–1880.
T. Lange and W. Stannat (2021*b*), On the continuous time limit of the ensemble Kalman filter, *Math. Comp.* **90**, 233–265.
A. Larios and Y. Pei (2024), Nonlinear continuous data assimilation, *Evol. Equ. Control Theory* **13**, 329–348.
J. Latz (2016), Bayes linear methods for inverse problems. Master's thesis, University of Warwick.
K. J. H. Law and A. M. Stuart (2012), Evaluating data assimilation algorithms, *Mon. Weather Rev.* **140**, 3757–3782.
K. J. H. Law, D. Sanz-Alonso, A. Shukla and A. M. Stuart (2016*a*), Filter accuracy for the Lorenz 96 model: Fixed versus adaptive observation operators, *Phys. D* **325**, 1–13.
K. J. H. Law, A. Shukla and A. M. Stuart (2014), Analysis of the 3DVAR filter for the partially observed Lorenz'63 model, *Discrete Contin. Dyn. Syst.* **34**, 1061–1078.
K. J. H. Law, H. Tembine and R. Tempone (2016*b*), Deterministic mean-field ensemble Kalman filtering, *SIAM J. Sci. Comput.* **38**, A1251–A1279.
K. J. Law, A. Stuart and K. Zygalakis (2015), *Data Assimilation: A Mathematical Introduction*, Vol. 214 of Texts in Applied Mathematics, Springer.
F. Le Gland, V. Monbet and V.-D. Tran (2011), Large sample asymptotics for the ensemble Kalman filter, in *The Oxford Handbook of Nonlinear Filtering*, Oxford University Press, pp. 598–631.
D. J. Lea, M. R. Allen and T. W. Haine (2000), Sensitivity analysis of the climate of a chaotic system, *Tellus A* **52**, 523–532.
Y. Lee, A. J. Majda and D. Qi (2017), Preventing catastrophic filter divergence using adaptive additive inflation for baroclinic turbulence, *Mon. Weather Rev.* **145**, 669–682.
J. Lei and P. Bickel (2011), A moment matching ensemble filter for nonlinear non-Gaussian data assimilation, *Mon. Weather Rev.* **139**, 3964–3973.
B. Leimkuhler, C. Matthews and J. Weare (2018), Ensemble preconditioning for Markov chain Monte Carlo simulation, *Statist. Comput.* **28**, 277–290.
G. Li and A. C. Reynolds (2009), An iterative ensemble Kalman filter for data assimilation, *SPE J.* **14**, 496–505.

S. Liu, S. Reich and X. T. Tong (2025), Dropout ensemble Kalman inversion for high dimensional inverse problems, *SIAM J. Numer. Anal.* **63**, 685–715.
Z. Liu, A. Stuart and Y. Wang (2022), Second order ensemble Langevin method for sampling and inverse problems. Available at arXiv:2208.04506.
D. M. Livings, S. L. Dance and N. K. Nichols (2008), Unbiased ensemble square root filters, *Phys. D* **237**, 1021–1028.
E. N. Lorenz (1996), Predictability: A problem partly solved, in *Predictability of Weather and Climate*, Cambridge University Press.
D. G. Luenberger (1964), Observing the state of a linear system, *IEEE Trans. Military Electronics* **8**, 74–80.
D. G. Luenberger (1971), An introduction to observers, *IEEE Trans. Automat. Control* **16**, 596–602.
A. J. Majda and X. T. Tong (2018), Performance of ensemble Kalman filters in large dimensions, *Commun. Pure Appl. Math.* **71**, 892–937.
J. Mandel, L. Cobb and J. D. Beezley (2011), On the convergence of the ensemble Kalman filter, *Appl. Math.* **56**, 533–541.
R. I. McLachlan and G. R. W. Quispel (2002), Splitting methods, *Acta Numer.* **11**, 341–434.
C. F. Mondaini and E. S. Titi (2018), Uniform-in-time error estimates for the postprocessing Galerkin method applied to a data assimilation algorithm, *SIAM J. Numer. Anal.* **56**, 78–110.
A. J. F. Moodey, A. S. Lawless, R. W. E. Potthast and P. J. van Leeuwen (2013), Nonlinear error dynamics for cycled data assimilation methods, *Inverse Problems* **29**, art. 025002.
L. Nerger (2022), Data assimilation for nonlinear systems with a hybrid nonlinear Kalman ensemble transform filter, *Quart. J. R. Meteorol. Soc.* **148**, 620–640.
D. J. Nott, L. Marshall and T. M. Ngoc (2012), The ensemble Kalman filter is an ABC algorithm, *Statist. Comput.* **22**, 1273–1276.
N. Nüsken and S. Reich (2019), Note on interacting Langevin diffusions: Gradient structure and ensemble Kalman sampler by Garbuno-Inigo, Hoffmann, Li and Stuart. Available at arXiv:1908.10890.
B. Øksendal (2013), *Stochastic Differential Equations: An Introduction with Applications*, Springer.
D. S. Oliver, L. B. Cunha and A. C. Reynolds (1997), Markov chain Monte Carlo methods for conditioning a permeability field to pressure data, *Math. Geol.* **29**, 61–91.
D. S. Oliver, A. C. Reynolds and N. Liu (2008), *Inverse Theory for Petroleum Reservoir Characterization and History Matching*, Cambridge University Press.
E. Olson and E. S. Titi (2003), Determining modes for continuous data assimilation in 2D turbulence, *J. Statist. Phys.* **113**, 799–840.
F. Parzer and O. Scherzer (2022), On convergence rates of adaptive ensemble Kalman inversion for linear ill-posed problems, *Numer. Math.* **152**, 371–409.
S. Pathiraja, S. Reich and W. Stannat (2021), McKean–Vlasov SDEs in nonlinear filtering, *SIAM J. Control Optim.* **59**, 4188–4212.
G. A. Pavliotis and A. M. Stuart (2008), *Multiscale Methods: Averaging and Homogenization*, Springer.
G. A. Pavliotis, A. M. Stuart and U. Vaes (2022), Derivative-free Bayesian inversion using multiscale dynamics, *SIAM J. Appl. Dyn. Syst.* **21**, 284–326.
L. M. Pecora and T. L. Carroll (1990), Synchronization in chaotic systems, *Phys. Rev. Lett.* **64**, art. 821.

G. Peyré and M. Cuturi (2019), Computational optimal transport: With applications to data science, *Found. Trends Mach. Learn.* **11**, 355–607.
J. Pidstrigach and S. Reich (2023), Affine-invariant ensemble transform methods for logistic regression, *Found. Comput. Math.* **23**, 675–708.
R. Pinnau, C. Totzeck, O. Tse and S. Martin (2017), A consensus-based model for global optimization and its mean-field limit, *Math. Models Methods Appl. Sci.* **27**, 183–204.
M. Pulido, P. Tandeo, M. Bocquet, A. Carrassi and M. Lucini (2018), Stochastic parameterization identification using ensemble Kalman filtering combined with maximum likelihood methods, *Tellus A* **70**, art. 1442099.
E. Qian and C. Beattie (2024), The fundamental subspaces of ensemble Kalman inversion. Available at arXiv:2409.08862.
P. Rebeschini and R. Van Handel (2015), Can local particle filters beat the curse of dimensionality?, *Ann. Appl. Probab.* **25**, 2809–2866.
S. Reich (2011), A dynamical systems framework for intermittent data assimilation, *BIT Numer. Math.* **51**, 235–249.
S. Reich (2013), A nonparametric ensemble transform method for Bayesian inference, *SIAM J. Sci. Comput.* **35**, A2013–A2024.
S. Reich (2019), Data assimilation: The Schrödinger perspective, *Acta Numer.* **28**, 635–711.
S. Reich and C. Cotter (2015), *Probabilistic Forecasting and Bayesian Data Assimilation*, Cambridge University Press.
S. Reich and S. Weissmann (2021), Fokker–Planck particle systems for Bayesian inference: Computational approaches, *SIAM/ASA J. Uncertain. Quantif.* **9**, 446–482.
G. O. Roberts and J. S. Rosenthal (1998), Optimal scaling of discrete approximations to Langevin diffusions, *J. R. Statist. Soc. Ser. B. Statist. Methodol.* **60**, 255–268.
G. O. Roberts and J. S. Rosenthal (2001), Optimal scaling for various Metropolis–Hastings algorithms, *Statist. Sci.* **16**, 351–367.
G. Robinson, I. Grooms and W. Kleiber (2018), Improving particle filter performance by smoothing observations, *Mon. Weather Rev.* **146**, 2433–2446.
P. Sakov and P. R. Oke (2008), A deterministic formulation of the ensemble Kalman filter: An alternative to ensemble square root filters, *Tellus A* **60**, 361–371.
P. Sakov, D. S. Oliver and L. Bertino (2012), An iterative EnKF for strongly nonlinear systems, *Mon. Weather Rev.* **140**, 1988–2004.
H. Salman, L. Kuznetsov, C. K. R. T. Jones and K. Ide (2006), A method for assimilating Lagrangian data into a shallow-water-equation ocean model, *Mon. Weather Rev.* **134**, 1081–1101.
C. Sampson, A. Carrassi, A. Aydoğdu and C. K. R. T. Jones (2021), Ensemble Kalman filter for nonconservative moving mesh solvers with a joint physics and mesh location update, *Quart. J. R. Meteorol. Soc.* **147**, 1539–1561.
D. Sanz-Alonso and A. M. Stuart (2015), Long-time asymptotics of the filtering distribution for partially observed chaotic dynamical systems, *SIAM/ASA J. Uncertain. Quantif.* **3**, 1200–1220.
D. Sanz-Alonso, A. Stuart and A. Taeb (2023), *Inverse Problems and Data Assimilation*, Vol. 107 of LMS Student Texts, Cambridge University Press.
S. Särkkä and L. Svensson (2023), *Bayesian Filtering and Smoothing*, second edition, Cambridge University Press.
C. Schillings and A. M. Stuart (2017), Analysis of the ensemble Kalman filter for inverse problems, *SIAM J. Numer. Anal.* **55**, 1264–1290.

C. Schillings and A. M. Stuart (2018), Convergence analysis of ensemble Kalman inversion: The linear, noisy case, *Appl. Anal.* **97**, 107–123.
T. Schneider, S. Lan, A. M. Stuart and J. Teixeira (2017), Earth system modeling 2.0: A blueprint for models that learn from observations and targeted high-resolution simulations, *Geophys. Res. Lett.* **44**, 12–396.
S. A. Sisson, Y. Fan and M. Beaumont (2018), *Handbook of Approximate Bayesian Computation*, CRC Press.
C. Snyder (2014), Introduction to the Kalman filter, in *Advanced Data Assimilation for Geosciences: Lecture Notes of the Les Houches School of Physics, June 2012*, Oxford University Press.
C. Snyder, T. Bengtsson, P. Bickel and J. Anderson (2008), Obstacles to high-dimensional particle filtering, *Mon. Weather Rev.* **136**, 4629–4640.
E. D. Sontag (2013), *Mathematical Control Theory: Deterministic Finite Dimensional Systems*, Vol. 6 of Texts in Applied Mathematics, Springer.
A. Spantini, R. Baptista and Y. Marzouk (2022), Coupling techniques for nonlinear ensemble filtering, *SIAM Rev.* **64**, 921–953.
A. S. Stordal, H. A. Karlsen, G. Nævdal, H. J. Skaug and B. Vallés (2011), Bridging the ensemble Kalman filter and particle filters: The adaptive Gaussian mixture filter, *Comput. Geosci.* **15**, 293–305.
G. Strang (1968), On the construction and comparison of difference schemes, *SIAM J. Numer. Anal.* **5**, 506–517.
A. M. Stuart (2010), Inverse problems: A Bayesian perspective, *Acta Numer.* **19**, 451–559.
A.-S. Sznitman (1991), Topics in propagation of chaos, *Éc. Été Probab. St.-Flour XIX–1989* **1464**, 165–251.
A. Taghvaei and B. Hosseini (2022), An optimal transport formulation of Bayes' law for nonlinear filtering algorithms, in *2022 IEEE 61st Conference on Decision and Control (CDC)*, pp. 6608–6613.
A. Taghvaei and P. G. Mehta (2020), An optimal transport formulation of the ensemble Kalman filter, *IEEE Trans. Automat. Control* **66**, 3052–3067.
A. Taghvaei and P. G. Mehta (2023), A survey of feedback particle filter and related controlled interacting particle systems (CIPS), *Ann. Rev. Control* **55**, 356–378.
A. Taghvaei, J. de Wiljes, P. G. Mehta and S. Reich (2017), Kalman filter and its modern extensions for the continuous-time nonlinear filtering problem, *ASME. J. Dyn. Syst. Meas. Control* **140**, art. 030904.
A. Taghvaei, P. G. Mehta and S. P. Meyn (2020), Diffusion map-based algorithm for gain function approximation in the feedback particle filter, *SIAM/ASA J. Uncertain. Quantif.* **8**, 1090–1117.
R. Temam (2012), *Infinite-Dimensional Dynamical Systems in Mechanics and Physics*, Vol. 68 of Applied Mathematical Sciences, Springer.
M. K. Tippett, J. L. Anderson, C. H. Bishop, T. M. Hamill and J. S. Whitaker (2003), Ensemble square root filters, *Mon. Weather Rev.* **131**, 1485–1490.
J. Tödter and B. Ahrens (2015), A second-order exact ensemble square root filter for nonlinear data assimilation, *Mon. Weather Rev.* **143**, 1347–1367.
X. T. Tong and M. Morzfeld (2023), Localized ensemble Kalman inversion, *Inverse Problems* **39**, art. 064002.
X. T. Tong, A. J. Majda and D. Kelly (2016*a*), Nonlinear stability and ergodicity of ensemble based Kalman filters, *Nonlinearity* **29**, art. 657.

X. T. Tong, A. J. Majda and D. Kelly (2016*b*), Nonlinear stability of the ensemble Kalman filter with adaptive covariance inflation, *Commun. Math. Sci.* **14**, 1283–1313.

K. I. Tsianos, S. Lawlor and M. G. Rabbat (2012), Consensus-based distributed optimization: Practical issues and applications in large-scale machine learning, in *2012 50th Annual Allerton Conference on Communication, Control, and Computing*, IEEE, pp. 1543–1550.

P. J. van Leeuwen (2020), A consistent interpretation of the stochastic version of the ensemble Kalman filter, *Quart. J. R. Meteorol. Soc.* **146**, 2815–2825.

P. J. van Leeuwen and G. Evensen (1996), Data assimilation and inverse methods in terms of a probabilistic formulation, *Mon. Weather Rev.* **124**, 2898–2913.

P. J. van Leeuwen, H. R. Künsch, L. Nerger, R. Potthast and S. Reich (2019), Particle filters for high-dimensional geoscience applications: A review, *Q. J. Royal Meteorol. Soc.* **145**, 2335–2365.

E. Vanden-Eijnden (2003), Fast communications: Numerical techniques for multi-scale dynamical systems with stochastic effects, *Commun. Math. Sci.* **1**, 385–391.

S. Vetra-Carvalho, P. J. van Leeuwen, L. Nerger, A. Barth, M. U. Altaf, P. Brasseur, P. Kirchgessner and J.-M. Beckers (2018), State-of-the-art stochastic data assimilation methods for high-dimensional non-Gaussian problems, *Tellus A* **70**, 1–43.

C. Villani (2008), *Optimal Transport: Old and New*, Vol. 338 of Grundlehren der mathematischen Wissenschaften, Springer.

C. Villani (2021), *Topics in Optimal Transportation*, Vol. 58 of Graduate Studies in Mathematics, American Mathematical Society.

X. Wang, C. H. Bishop and S. J. Julier (2004), Which is better, an ensemble of positive-negative pairs or a centered spherical simplex ensemble?, *Mon. Weather Rev.* **132**, 1590–1605.

S. Weissmann, N. K. Chada, C. Schillings and X. T. Tong (2022), Adpative Tikhonov strategies for stochastic ensemble Kalman inversion, *Inverse Problems* **38**, art. 045009.

G. Welch and G. Bishop (1995), An introduction to the Kalman filter. Report TR 95-041, Department of Computer Science, University of North Carolina at Chapel Hill, NC.

J. S. Whitaker and T. M. Hamill (2002), Ensemble data assimilation without perturbed observations, *Mon. Weather Rev.* **130**, 1913–1924.

L. M. Yang and I. Grooms (2021), Machine learning techniques to construct patched analog ensembles for data assimilation, *J. Comput. Phys.* **443**, art. 110532.

T. Yang, H. A. P. Blom and P. G. Mehta (2014), The continuous-discrete time feedback particle filter, in *American Control Conference*, IEEE, pp. 648–653.

T. Yang, P. G. Mehta and S. P. Meyn (2013), Feedback particle filter, *IEEE Trans. Automat. Control* **58**, 2465–2480.

J. Zech and Y. Marzouk (2022), Sparse approximation of triangular transports, Part I: The finite-dimensional case, *Constr. Approx.* **55**, 919–986.

C. Zhang (2013), A particle system for global optimization, in *52nd IEEE Conference on Decision and Control*, pp. 1714–1719.

C. Zhang, A. Taghvaei and P. G. Mehta (2017), A controlled particle filter for global optimization. Available at arXiv:1701.02413.

M. Zupanski (2005), Maximum likelihood ensemble filter: Theoretical aspects, *Mon. Weather Rev.* **133**, 1710–1726.