BIT manuscript No.
(will be inserted by the editor)

A dynamical systems framework for intermittent data assimilation

Sebastian Reich

Received: 2010-07-05

Abstract We consider the problem of discrete time filtering (intermittent data assimilation) for differential equation models and discuss methods for its numerical approximation. The focus is on methods based on ensemble/particle techniques and on the ensemble Kalman filter technique in particular. We summarize as well as extend recent work on continuous ensemble Kalman filter formulations, which provide a concise dynamical systems formulation of the combined dynamics-assimilation problem. Possible extensions to fully nonlinear ensemble/particle based filters are also outlined using the framework of optimal transportation theory.

Keywords Data assimilation · ensemble Kalman filter · dynamical systems · nonlinear filters · optimal transportation

Mathematics Subject Classification (2000) 93E11 · 65L09 · 37M10 · 62F15 · 60G35

1 Introduction

We consider dynamical models given in form of ordinary differential equations (ODEs)

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

with state variable $\mathbf{x} \in \mathbb{R}^n$. Initial conditions at time t_0 are not precisely known and are treated as a random variable instead, i.e., we assume that

$$\mathbf{x}(t_0) \sim \pi_0,$$

where $\pi_0(\mathbf{x})$ denotes a given probability density function (PDF). The solution of (1) at time t with initial condition $\mathbf{x}_0$ at t_0 is denoted by $\mathbf{x}(t; t_0, \mathbf{x}_0)$.

To compensate for the resulting uncertainty in the solutions $\mathbf{x}(t; t_0, \mathbf{x}_o)$, $\mathbf{x}_0 \sim \pi_0$, of (1), $t > t_0$, we assume that we obtain measurements $\mathbf{y}(t_q) \in \mathbb{R}^k$ at discrete times $t_q \geq t_0$, $q = 1, 2, \ldots, M$, subject to measurement errors. The measurements are related to the state variable through a linear forward operator $\mathbf{H} \in \mathbb{R}^{k \times n}$ and the measurement errors are assumed to be Gaussian distributed with zero mean and covariance matrix $\mathbf{R} \in \mathbb{R}^{k \times k}$, i.e.

$$\mathbf{y}(t_q) - \mathbf{H}\mathbf{x}(t_q) \sim \mathrm{N}(\mathbf{0}, \mathbf{R}). \tag{2}$$

The evolution of the initial PDF π_0 under the ODE (1) up to the first measurement at t_1 is provided by the continuity equation

$$\frac{\partial \pi}{\partial t} = -\nabla_{\mathbf{x}} \cdot (\pi f), \tag{3}$$

which is also called Liouville's equation in the statistical mechanics literature [13]. Let us denote the solution of Liouville's equation at observation time $t_1 > t_0$ by $\pi_f(\mathbf{x}) = \pi(\mathbf{x}, t_1)$. In other words, solutions $\mathbf{x}(t_1; t_0, \mathbf{x}_0)$ with $\mathbf{x}_0 \sim \pi_0$ constitute a random variable with PDF π_f. The assimilation of the measurement $\mathbf{y}(t_1)$ leads now to a discontinuous change in the forecast PDF π_f to an analyzed PDF π_a. The precise relation between π_f and π_a will be summarized in the following section. Once the analyzed PDF π_a is available, Liouville's equation is solved to the next observation time t_2 with the analysed PDF π_a as the new initial condition at t_1. The sequence of discontinuous changes at assimilation times t_q and continuous propagation of the PDF in between observations under Liouville's equation (3) is then repeated for all $q \geq 2$.

Sebastian Reich
Universität Potsdam, Institut für Mathematik, Am Neuen Palais 10, D-14469 Potsdam, Germany
E-mail: sreich@math.uni-potsdam.de

In terms of practical implementations, one needs to replace the general class of PDFs by appropriate statistical models $\rho(\mathbf{x}|\mathbf{z})$, where ρ is a given function of phase space $\mathbf{x}$ and a set of parameters $\mathbf{z} \in \mathbb{R}^l$. The time evolution of a PDF π under a statistical model is then approximated by $\pi(\mathbf{x},t) = \rho(\mathbf{x}|\mathbf{z}(t))$ with the parameters $\mathbf{z}(t) \in \mathbb{R}^l$ evolving in time. Hence, once a statistical model has been chosen, the key challenge is to define appropriate evolution equations for the parameters $\mathbf{z}$. For linear differential equations, the statistical model can be chosen to be a Gaussian parametrized by its mean and covariance matrix. Update equations for the mean and the covariance matrix are provided by the celebrated Kalman filter (see, for example, [22,24]). For general differential equations, particle or sequential Monte Carlo filters have been proposed. These methods are based on empirical measures parametrized by particle locations and weights (see, for example, [4] and the following section). While particle filters can be shown to be asymptotically correct as the number of particles increases, they become computationally demanding for high dimensional problems. The extended Kalman filter (see, for example, [22]) represents an attempt to make the standard Kalman filter approach applicable to nonlinear problems under the assumption that the PDFs remain nearly Gaussian with small variance. A more recent addition to the family of filter algorithms is provided by the ensemble Kalman filter (EnKF), which combines empirical measures for approximating Liouville's equation (3) with a Kalman analysis update for the data assimilation step (see, for example, [11]). As for the extended Kalman filter, the EnKF relies on the assumption that the PDFs remain approximately Gaussian. However the assumption of small variance can be dropped since no linearization step is involved. The EnKF is now widely being used in atmosphere-ocean dynamics, but a rigorous analysis of approximation errors is not yet available.

Each of the above mentioned filter algorithms got its limitations when applied to complex physical models. A survey of these limitations and possible remedies are, for example, discussed in [20] in the context of geophysical fluid dynamics, which also provides a main motivation for the investigations presented in this paper. More specifically, we propose a novel formulation for the above described sequential data assimilation problem in form of a continuous time dynamical system in Section 3. We then demonstrate in Section 4 that popular EnKF techniques naturally fit into this framework. The continuous formulation of [5,6] for deterministic square root filters [26,11] is summarized first and is then extended to the EnKF with randomly perturbed observations [8,11]. As a novel result we obtain a stochastic differential equation formulation of the EnKF analysis step. We also discuss links to iterative regularization methods [18] and H_∞ filtering [24]. Possible extensions to non-Gaussian statistical models are outlined in Section 5. These extensions avoid the need for random re-sampling of particle locations as it is necessary for particle and sequential Monte Carlo methods [4] and, hence, should be more robust in applications where the particle numbers is smaller than or comparable to the dimension of the phase space $\mathbb{R}^n$.

2 Particle and ensemble Kalman filters

A first step to perform data assimilation for nonlinear ODEs (1) is to approximate solutions to the associated Liouville equation (3). Here we rely exclusively on particle methods [4] for which Liouville's equation is naturally approximated by the evolving empirical measure. More precisely, particle or ensemble filters rely on the simultaneous propagation of m independent solutions $\mathbf{x}_i(t)$, $i = 1,\ldots,m$, of (1) [11]. We associate the empirical measure

$$\pi_{\mathrm{em}}(\mathbf{x},t) = \sum_{i=1}^{m} \alpha_i \delta(\mathbf{x} - \mathbf{x}_i(t)) \tag{4}$$

with weights $\alpha_i > 0$ satisfying

$$\sum_{i=1}^{m} \alpha_i = 1.$$

Here $\delta(\cdot)$ denotes the Dirac delta function. Hence our statistical model is given by the empirical measure (4) and is parametrized by the particle weights $\{\alpha_i\}$ and the particle locations $\{\mathbf{x}_i\}$. In the absence of measurements, the empirical measure π_{em} with constant weights α_i is an exact (weak) solution to Liouville's equation (3) provided the $\mathbf{x}_i(t)$'s are solutions to the ODE (1).

At any observation point t_q, we need to merge the forecast or prior PDF $\pi_f(\mathbf{x})$ with a likelihood function [18]

$$\pi(\mathbf{y}|\mathbf{x}) \propto \exp\left(-\frac{1}{2}\left(\mathbf{H}\mathbf{x} - \mathbf{y}\right)^T \mathbf{R}^{-1}\left(\mathbf{H}\mathbf{x} - \mathbf{y}\right)\right),$$

which characterizes the conditional PDF induced by the measurements (2). Here $\propto$ stands for equality up to a constant scaling factor. For a given observation $\mathbf{y} = \mathbf{y}(t_q)$, Bayes' theorem [18] states that the posterior PDF $\pi_a(\mathbf{x})$ is provided by

$$\pi_a(\mathbf{x}) = \pi(\mathbf{x}|\mathbf{y}(t_q)) \quad \text{with} \quad \pi(\mathbf{x}|\mathbf{y}) \propto \pi(\mathbf{y}|\mathbf{x})\,\pi_f(\mathbf{x}). \tag{5}$$

Here the scaling factor hidden in the $\propto$ notation depends on $\mathbf{y}$, which however turns into a constant after setting $\mathbf{y} = \mathbf{y}(t_q)$. For a particle method, the prior PDF is given by (4) with $t = t_q$. A formal application of Bayes' formula (5) leads to a new set of weights α_i given by

$$\alpha_i^a = C\alpha_i^f \times \exp\left(-\frac{1}{2}\left(\mathbf{H}\mathbf{x}_i(t_q) - \mathbf{y}(t_q)\right)^T \mathbf{R}^{-1}\left(\mathbf{H}\mathbf{x}_i(t_q) - \mathbf{y}(t_q)\right)\right),$$

where C is a normalization constant to guarantee $\sum_{i=1}^m \alpha_i^a = 1$, while the particle locations $\mathbf{x}_i$ remain fixed.

To avoid a highly non-uniform distribution of the resulting particle weights after a number of data assimilation steps, a random re-sampling step is performed. As a result one obtains randomly re-sampled ensemble locations $\mathbf{x}_i(t_q)$, $i = 1, \ldots, m$, with equal weights $\alpha_i = 1/m$. See, for example, [4] for details. Convergence of particle filters to the analytic filtering solution can be shown in the limit $m \to \infty$ [4]. The stochastic re-sampling step is however problematic since it introduces a large amount of random noise for small ensemble sizes $m < n$, which are typical for many applications from geophysical fluid dynamics [20].

A notable exception to the above described procedure is provided by EnKFs, which lead to a dynamic change in ensemble locations $\mathbf{x}_i$ while keeping the weights fixed, i.e. $\alpha_i^a = \alpha_i^f$. This eliminates the need for stochastic re-sampling. The EnKF is based on the assumption that the particles $\mathbf{x}_i$ are drawn from a Gaussian PDF and relies on the standard Kalman filter variance minimizing methodology. The interpretation of the Kalman analysis step in terms of readjusted particle positions/ensemble members $\mathbf{x}_i$ is a key step in the derivation of the EnKF. In [5,6] it has been shown that the particle/ensemble readjustment step can be formulated as the solution of a differential equation in an artificial embedding parameter $s \in [0, 1]$. This reformulation can be interpreted as a continuous deformation of the associated PDF under the data assimilation step and we will provide a general dynamical systems framework for such a continuous deformation approach in the following section. The deformation approach formally leads to an additional source term for the ODE model (1). We mention that the same methodology appears in [9] in the context of time-continuous filtering problems. For such problems the evolving probability measure is already continuous in time and, contrary to the intermittent data assimilation problem considered in this paper, no artificial embedding process is required.

3 A novel continuous dynamical systems filter formulation

In this section, we derive a novel dynamical systems formulation of the data assimilation step (5). We first note that a single application of Bayes' formula (5) can be replaced by an N-fold recursive application with incremental likelihood

$$\pi^N(\mathbf{y}|\mathbf{x}) \propto \exp\left(-\frac{1}{2N}\left(\mathbf{H}\mathbf{x} - \mathbf{y}\right)^T \mathbf{R}^{-1}\left(\mathbf{H}\mathbf{x} - \mathbf{y}\right)\right), \tag{6}$$

i.e., we first write (5) as

$$\pi_a(\mathbf{x}) \propto \pi_f(\mathbf{x}) \prod_{j=1}^N \pi^N(\mathbf{y}(t_q)|\mathbf{x})$$

and then consider the implied iteration

$$\pi_{j+1}(\mathbf{x}) = \frac{\pi_j(\mathbf{x})\,\pi^N(\mathbf{y}(t_q)|\mathbf{x})}{\int \mathrm{d}\mathbf{x}\,\pi_j(\mathbf{x})\,\pi^N(\mathbf{y}(t_q)|\mathbf{x})}$$

with $\pi_0 = \pi_f$ and $\pi_a = \pi_N$. We may now expand the exponential function in (6) in the small parameter $\Delta s = 1/N$ to first obtain

$$\pi_{j+1}(\mathbf{x}) = \frac{\pi_j(\mathbf{x})\left\{1 - \frac{\Delta s}{2}\left(\mathbf{H}\mathbf{x} - \mathbf{y}(t_q)\right)^T \mathbf{R}^{-1}\left(\mathbf{H}\mathbf{x} - \mathbf{y}(t_q)\right)\right\}}{\int \mathrm{d}\mathbf{x}\,\pi_j(\mathbf{x})\left\{1 - \frac{\Delta s}{2}\left(\mathbf{H}\mathbf{x} - \mathbf{y}(t_q)\right)^T \mathbf{R}^{-1}\left(\mathbf{H}\mathbf{x} - \mathbf{y}(t_q)\right)\right\}} + \mathcal{O}(\Delta s^2)$$

and to then derive in the limit $N \to \infty$ the evolution equation

$$\frac{\partial \pi}{\partial s} = -\frac{1}{2}\left(\mathbf{H}\mathbf{x} - \mathbf{y}(t_q)\right)^T \mathbf{R}^{-1}\left(\mathbf{H}\mathbf{x} - \mathbf{y}(t_q)\right)\pi + \mu\pi \tag{7}$$

in the fictitious time $s \in [0, 1]$. The scalar Lagrange multiplier μ is equal to the expectation value of the negative log likelihood function

$$L(\mathbf{x}; \mathbf{y}(t_q)) = \frac{1}{2}\left(\mathbf{H}\mathbf{x} - \mathbf{y}(t_q)\right)^T \mathbf{R}^{-1}\left(\mathbf{H}\mathbf{x} - \mathbf{y}(t_q)\right) \tag{8}$$

with respect to π and ensures that $\int \mathrm{d}\mathbf{x}\,\partial\pi/\partial s = 0$. We also set $\pi(\mathbf{x}, 0) = \pi_f(\mathbf{x})$ and obtain $\pi_a(\mathbf{x}) = \pi(\mathbf{x}, 1)$. For further reference, we replace (7) by the equivalent but more compact formulation

$$\frac{\partial \pi}{\partial s} = -\pi\left(L - \mathbb{E}_\pi[L]\right). \tag{9}$$

Here $\mathbb{E}_\pi$ denotes expectation with respect to the PDF π. See [6] for a detailed discussion of a related derivation in the context of ensemble Kalman filters. It should be noted that the continuous embedding defined by (9) is not unique.

Eq. (9) defines the change (or transport) of the PDF π with respect to the fictitious time $s \in [0,1]$. Following an optimal transportation approach (see, for example, [27]), we can view this change alternatively as induced by a continuity (Liouville) equation

$$\frac{\partial \pi}{\partial s} = -\nabla_{\mathbf{x}} \cdot (\pi g) \tag{10}$$

for an appropriate vector field $g(\mathbf{x},s) \in \mathbb{R}^n$. At any time $s \in [0,1]$ the vector field $g(\cdot,s)$ is not uniquely determined by (9) and (10) unless we also require that it is the minimizer of the kinetic energy

$$\mathcal{T}(v) = \frac{1}{2}\int \mathrm{d}\pi\, v^T \mathbf{M} v$$

over all admissible vector fields $v \in L^2(\mathrm{d}\pi, \mathbb{R}^n)$, where $\mathbf{M} \in \mathbb{R}^{n\times n}$ is a positive definite mass matrix. Admissibility means that $g = v$ satisfies (10) for given π and $\partial\pi/\partial s$. Under these assumptions, minimization of the functional

$$\mathcal{L}[v,\phi] = \frac{1}{2}\int \mathrm{d}\pi\, v^T \mathbf{M} v + \int \mathrm{d}\mathbf{x}\, \phi \left\{ \frac{\partial \pi}{\partial s} + \nabla_{\mathbf{x}}\cdot(\pi v)\right\}$$

for given π and $\partial\pi/\partial s$ leads to the Euler-Lagrange equations

$$\pi \mathbf{M} g - \pi \nabla \psi = 0, \qquad \frac{\partial \pi}{\partial s} + \nabla_{\mathbf{x}}\cdot(\pi g) = 0$$

in the velocity field g and the potential ψ. Hence, provided that $\pi > 0$, the desired vector field is given by $g = \mathbf{M}^{-1}\nabla_{\mathbf{x}}\psi$, where the potential $\psi(\mathbf{x},s)$ is the solution of the elliptic partial differential equation (PDE)

$$\nabla_{\mathbf{x}} \cdot \left(\pi \mathbf{M}^{-1}\nabla_{\mathbf{x}}\psi\right) = -\frac{\partial \pi}{\partial s} = \pi\left(L - \mathbb{E}_\pi[L]\right). \tag{11}$$

We mention that the related formulation in [9] (i.e. eq. (19) in [9]) relies on an integral transform representation of the vector field g in (10), which could also be explored in the context of our intermittent data assimilation problem. More specifically, let $G(\mathbf{x},\mathbf{x}')$ denote the Greens function for Poisson's equation $\nabla_{\mathbf{x}}\cdot(\nabla_{\mathbf{x}} u) = f$ over $\mathbf{x} \in \mathbb{R}^n$, then

$$g(\mathbf{x},s) = \frac{1}{\pi(\mathbf{x},s)} \int_{\mathbb{R}^n} \nabla_{\mathbf{x}} G(\mathbf{x},\mathbf{x}') \left\{\pi(\mathbf{x}',s)\left(L(\mathbf{x}';\mathbf{y}(t_q)) - \mathbb{E}_\pi[L]\right)\right\} \mathrm{d}\mathbf{x}'.$$

On a more abstract level one could consider the Monge-Kantorovich problem for transporting the prior PDF π_f into the posterior PDF π_a under the quadratic cost function $c(\mathbf{x},\mathbf{y}) = \|\mathbf{x}-\mathbf{y}\|^2$ in phase space $\mathbb{R}^n$ [27]. General existence results of an optimal transportation map $\mathbf{y} = T(\mathbf{x})$; i.e.

$$\pi_f(\mathbf{x}) = \pi_a(T(\mathbf{x}))\,|\det \nabla_{\mathbf{x}} T(\mathbf{x})|,$$

and its displacement interpolation in form of a Liouville/continuity equation (10) with appropriate vector field g are stated, for example, in Theorems 2.12 and 5.51 of [27].

In light of the above discussions, we may now replace (9) by the continuity or Liouville equation

$$\frac{\partial \pi}{\partial s} = -\nabla_{\mathbf{x}} \cdot \left(\pi \mathbf{M}^{-1}\nabla_{\mathbf{x}}\psi\right) \tag{12}$$

with an underlying ODE formulation

$$\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}s} = \mathbf{M}^{-1}\nabla_{\mathbf{x}}\psi(\mathbf{x},s) \tag{13}$$

in the fictitious time $s \in [0,1]$. As for the ODE (1) and its associated Liouville equation (3), we may approximate (13) and its associated Liouville equation (12) by an empirical measure of type (4). Furthermore, one and the same empirical measure approximation can now be used for both the ensemble propagation step under the model dynamics (1) and the data assimilation step (12) using constant and equal weights $\alpha_i = 1/m$. The particle filter approximation is closed by finding an appropriate numerical solution to the elliptic PDE (11). This is the crucial step which will lead to different nonlinear particle filter algorithms. In the following section, we will solve (11) under the simplifying assumption that the PDF π in (11) is taken to be a Gaussian PDF with mean equal to the ensemble mean and covariance equal to the ensemble covariance matrix. Other obvious choices, such as π in (11) being represented by a Gaussian mixture model, will be explored in forthcoming publications.

4 Continuous ensemble Kalman filter formulations

In this section, we again take the particle approximation (4) as a starting point, assume that all ensemble members have equal weights, i.e. $\alpha_i = 1/m$, and discuss continuous EnKF formulations both in a deterministic and stochastic setting. We start by discussing (6) and Bayes' theorem in case the prior distributions π_f is Gaussian parametrized by its mean $\overline{\mathbf{x}}_f$ and its covariance matrix $\mathbf{P}_f$. It should be noted that the product of two Gaussian PDFs is again a Gaussian PDF which allows for a more direct derivation of equation (13). To do so we introduce a generalized square root $\widehat{\mathbf{Y}}_f$ of $\mathbf{P}_f$ such that $\mathbf{P}_f = \widehat{\mathbf{Y}}_f \widehat{\mathbf{Y}}_f^T$. Then the continuous deformation of the Gaussian prior into its posterior can be characterized by the differential equation

$$\frac{\mathrm{d}\overline{\mathbf{x}}}{\mathrm{d}s} = -\mathbf{P}\mathbf{H}^T\mathbf{R}^{-1}(\mathbf{H}\overline{\mathbf{x}} - \mathbf{y}) \tag{14}$$

for the mean and the differential equation

$$\frac{\mathrm{d}\widehat{\mathbf{Y}}}{\mathrm{d}s} = -\frac{1}{2}\mathbf{P}\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}\widehat{\mathbf{Y}} \tag{15}$$

for the generalized square root of $\mathbf{P}$, respectively. The initial conditions are $\overline{\mathbf{x}}(0) = \overline{\mathbf{x}}_f$ and $\widehat{\mathbf{Y}}(0) = \widehat{\mathbf{Y}}_f$. The posterior values of the mean and the square root are provided by the solutions of (14) and (15), respectively, at $s = 1$. We also have $\mathbf{P}_a = \widehat{\mathbf{Y}}(1)\widehat{\mathbf{Y}}(1)^T$. See [6] for a derivation of (14) and (15).

It should be noted that (14) and (15) formally arise from a continuous time Kalman filter formulation with trivial model dynamics $\dot{\mathbf{x}} = \mathbf{0}$ [24]. Furthermore, (15) implies the well-known Riccati equation

$$\frac{\mathrm{d}\mathbf{P}}{\mathrm{d}s} = -\mathbf{P}\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}\mathbf{P} \tag{16}$$

for the covariance matrix $\mathbf{P}$.

Continuous EnKF formulations in the sense of Section 3 rely on a reformulation of (14) and (15) in terms of dynamical equations in the ensemble members $\mathbf{x}_i$, $i = 1, \ldots, m$. We discuss two specific formulations in the following two subsections.

4.1 Ensemble square root filter

In this section we focus on ensemble square root filter implementations of an EnKF [26,11]. For notational convenience, the ensemble members $\{\mathbf{x}_i(s)\}_{i=1}^m$ are collected in a matrix $\mathbf{X}(s) \in \mathbb{R}^{n\times m}$. In terms of $\mathbf{X}(s)$, the ensemble mean is given by

$$\overline{\mathbf{x}}(s) = \frac{1}{m}\mathbf{X}(s)\mathbf{e} \in \mathbb{R}^n$$

and we introduce the ensemble deviation matrix

$$\mathbf{Y}(s) = \mathbf{X}(s) - \overline{\mathbf{x}}(s)\mathbf{e}^T \in \mathbb{R}^{n\times m}, \tag{17}$$

where $\mathbf{e} = (1, \ldots, 1)^T \in \mathbb{R}^m$. The implied covariance matrix is

$$\mathbf{P}(s) = \frac{1}{m-1}\sum_{i=1}^{m}(\mathbf{x}_i(s) - \overline{\mathbf{x}}(s))(\mathbf{x}_i(s) - \overline{\mathbf{x}}(s))^T = \frac{1}{m-1}\mathbf{Y}(s)\mathbf{Y}^T(s). \tag{18}$$

Note that $\widehat{\mathbf{Y}} = \mathbf{Y}/\sqrt{m-1}$ defines a generalized square root of $\mathbf{P}$.

As shown in [5,6], an equivalent formulation of (14) and (15) in terms of the ensemble members $\mathbf{x}_i(s)$ is provided by

$$\frac{\mathrm{d}\mathbf{x}_i}{\mathrm{d}s} = -\mathbf{P}\nabla_{\mathbf{x}_i}\mathcal{V}_q(\mathbf{X}) = -\frac{1}{2}\mathbf{P}\mathbf{H}^T\mathbf{R}^{-1}(\mathbf{H}\mathbf{x}_i + \mathbf{H}\overline{\mathbf{x}} - 2\mathbf{y}(t_q)), \tag{19}$$

$i = 1, \ldots, m$, where $\mathcal{V}_q$ is the potential

$$\mathcal{V}_q(\mathbf{X}) = \frac{m}{2}\left\{L(\overline{\mathbf{x}}; \mathbf{y}(t_q)) + \frac{1}{m}\sum_{i=1}^{m}L(\mathbf{x}_i; \mathbf{y}(t_q))\right\} \tag{20}$$

with the negative log likelihood function $L(\mathbf{x}; \mathbf{y}(t_q))$ given by (8) for a measurement at t_q. The solutions of (19) over a unit time interval provide a particular analysis step for an ensemble square root filter.

Upon comparison with (13) we find that (19) fits into the dynamical systems framework developed in Section 3 with mass matrix $\mathbf{M}(s) = \mathbf{P}^{-1}(s)$ and $\nabla_{\mathbf{x}}\psi(\mathbf{x}_i, s) = -\nabla_{\mathbf{x}_i}\mathcal{V}_q(\mathbf{X}(s))$.

In practice, (19) needs to be discretized by an appropriate time-stepping method such as the forward Euler method. It should be noted that the forward Euler method leads to an iterative procedure similar to a pre-conditioned Landweber-Fridman iteration for ill-posed problems of type

$$\mathbf{H}\mathbf{x} = \mathbf{y}. \tag{21}$$

See, for example, [18]. In our case, the pre-conditioner is provided by the forecast ensemble covariance matrix $\mathbf{P}_f$ and the adjoint of $\mathbf{H}$ is based on the measurment error covariance matrix $\mathbf{R}$. More precisely, the analyzed mean $\overline{\mathbf{x}}_a$ is provided as solution to the linear system

$$\mathbf{P}_f^{-1}(\overline{\mathbf{x}}_a - \overline{\mathbf{x}}_f) + \mathbf{H}^T\mathbf{R}^{-1}(\mathbf{H}\overline{\mathbf{x}}_a - \mathbf{y}) = \mathbf{0}, \tag{22}$$

which provides a regularized solution to (21). One can formally replace $\mathbf{P}_f^{-1}$ by $\alpha\mathbf{P}_f^{-1}$ in (22) and treat α as a regularization parameter in the sense of Tikhonov [18]. While it is then common practice to stop the Landweber-Fridman iteration by a discrepancy principle for unknown parameter value α [18], we integrate (19) over the finite time-interval $s \in [0,1]$ under the assumption that $\alpha = 1$ is optimal. It would be of interest to explore the relation between these two "stopping" criteria and to also investigate the duality between solving (22) by simple optimization algorithms and iterative approaches based on underlying differential equations in articficial time (see, for example, [3] and [2]). It should, however, be kept in mind that a Kalman filter requires not only the computation of the most likely state but in addition the update of a covariance matrix. Furthermore, the general data assimilation problem requires the update of a complete PDF.

As proposed in [6], the continuous formulation (19) allows for a concise formulation of a sequences of observations at time instances t_q, $q = 1, \ldots, M$, and intermittent propagation of the ensemble under the dynamics (1). Specifically, we obtain the differential equation

$$\dot{\mathbf{x}}_i = f(\mathbf{x}_i, t) - \sum_{q=1}^{M} \delta(t - t_q)\, \mathbf{P}\nabla_{\mathbf{x}_i}\mathcal{V}_q(\mathbf{X}) \tag{23}$$

in each ensemble member, where $\delta(\cdot)$ denotes again the standard Dirac delta function. The mathematical interpretation of (23) relies on replacing the Dirac delta function by a family of compactly supported smooth functions δ_ϵ which approach the Dirac delta function δ in the limit $\epsilon \to 0$. See [6] and the following subsection for more details.

It should be noted that (23) is equivalent to a standard Kalman filter (and hence is optimal) if the model is linear and if the number of ensemble members m is larger than the dimension of phase space n. Since neither of the two assumptions are satisfied for most geophysical applications, EnKF formulations need to be modified to make them robust with respect to sampling errors and/or weakly non-Gaussian PDFs. Two popular techniques are localization [16,15] and ensemble inflation [1]. However, we first discuss the more recently introduced mollification approach [6].

4.1.1 Mollification: a seamless data assimilation approach

The Dirac delta functions in (23) lead to discontinuous changes in the ensemble members at assimilation times t_q, which can lead to artificial readjustment processes under the subsequent model dynamics. Hence it makes sense to "mollify" the discontinuous analysis adjustments and we obtain

$$\dot{\mathbf{x}}_i = f(\mathbf{x}_i, t) - \sum_{q=1}^{M} \delta_\epsilon(t - t_q)\, \mathbf{P}\nabla_{\mathbf{x}_i}\mathcal{V}_q(\mathbf{X}), \tag{24}$$

where

$$\delta_\epsilon(s) = \frac{1}{\epsilon}\psi(s/\epsilon),$$

$\psi(s)$ is the standard hat function

$$\psi(s) = \begin{cases} 1 - |s| & \text{for } |s| \le 1, \\ 0 & \text{else}, \end{cases} \tag{25}$$

and $\epsilon > 0$ is an appropriate parameter. The hat function (25) could, of course, be replaced by another B-spline. We note that the term mollification was introduced to denote families of compactly supported smooth functions δ_ϵ which approach the Dirac delta function δ in the limit $\epsilon \to 0$. Mollification via convolution turns non-standard functions (distributions) into smooth functions [12]. Here we relax the smoothness assumption and allow for any non-negative, compactly supported family of functions that can be used to approximate the Dirac delta function.

Formulation (24) can be solved numerically by any standard ODE solver and leads to a seamless data assimilation approach. Of course, the ODE formulation becomes increasingly stiff as $\epsilon \to 0$. Formulation (24) has been shown in [6] to avoid the generation of unbalanced waves in multi-scale wave-advection equations, which is a problem with standard EnKF implementations. See, e.g. [17,19]. Related "mollification" approaches are provided by [7,21].

4.1.2 Covariance localization

Due to the fact that the number of ensemble members m is often much smaller than the dimension n of phase space, the empirical covariance matrix (18) is highly rank deficient and spurious correlations arise due to under-sampling. Schur-product covariance localization [16,15] has become a popular and powerful technique to deal with these issues. We find that Schur-product-based localizations can easily be applied to (24) to obtain, e.g.

$$\dot{\mathbf{x}}_i = f(\mathbf{x}_i, t) - \sum_{q=1}^{M} \delta_\epsilon(t - t_q)\, \widetilde{\mathbf{P}} \nabla_{\mathbf{x}_i} \mathcal{V}_q(\mathbf{X}), \tag{26}$$

with

$$\widetilde{\mathbf{P}} = \mathbf{C}_{\mathrm{loc}} \circ \mathbf{P},$$

and $\mathbf{C}_{\mathrm{loc}} \in \mathbb{R}^{n\times n}$ an appropriate localization matrix [16,15,14]. See [5] for implementation details.

Calculations can be simplified by freezing the time-dependent covariance matrix $\mathbf{P}(t)$ for the duration of a single assimilation step, i.e.

$$\dot{\mathbf{x}}_i = f(\mathbf{x}_i, t) - \sum_{q=1}^{M} \delta_\epsilon(t - t_q)\, \widetilde{\mathbf{P}}(t_q) \nabla_{\mathbf{x}_i} \mathcal{V}_q(\mathbf{X}).$$

4.1.3 Ensemble inflation and H_∞ filtering

Ensemble inflation [1] is another popular technique to correct for poor statistics from small ensemble sizes, i.e $m \ll n$. Ensemble inflation is performed either before or after a data analysis step and consists in replacing $\mathbf{P}$ by $\delta\mathbf{P}$ with factor $\delta > 1$. A corresponding adjustment of the ensemble members $\{\mathbf{x}_i\}$ is required to reflect the change in the ensemble covariance matrix.

We now point to a link between ensemble inflation and H_∞ filtering [24]. We first introduce a (negative) cost function

$$D(\mathbf{x}) = -\frac{\theta}{2} (\mathbf{L}\mathbf{x})^T \mathbf{S}^{-1} \mathbf{L}\mathbf{x}$$

for given matrices $\mathbf{L}$ and $\mathbf{S}$ and parameter $\theta > 0$. The matrix $\mathbf{S}$ is assumed to be symmetric positive definite. We also introduce the (negative) potential

$$\mathcal{W}(\mathbf{X}) = \frac{m}{2} \left\{ \frac{1}{m} \sum_{i=1}^{m} D(\mathbf{x}_i) - D(\overline{\mathbf{x}}) \right\}$$

and the modified data assimilation equations

$$\begin{aligned} \dot{\mathbf{x}}_i &= f(\mathbf{x}_i, t) - \sum_{q=1}^{M} \delta_\epsilon(t - t_q) \mathbf{P} \left[\nabla_{\mathbf{x}_i} \mathcal{V}_q(\mathbf{X}) + \nabla_{\mathbf{x}_i} \mathcal{W}(\mathbf{X}) \right] \\ &= f(\mathbf{x}_i, t) - \sum_{q=1}^{M} \delta_\epsilon(t - t_q) \mathbf{P} \left[\nabla_{\mathbf{x}_i} \mathcal{V}_q(\mathbf{X}) - \frac{\theta}{2} \mathbf{L}^T \mathbf{S}^{-1} \mathbf{L} \left(\mathbf{x}_i - \overline{\mathbf{x}} \right) \right]. \end{aligned}$$

The additional term does not directly affect the evolution of the mean. Instead an additional contribution to the ensemble deviations is introduced that exactly mirrors the term found in continuous H_∞ filter formulations [24].

4.2 Ensemble Kalman filter with perturbed observations

In this section, we derive a novel stochastic differential equation (SDE) formulation of the EnKF with randomly perturbed observations [8,11]. The perturbed ensemble Kalman filter step is given by

$$\mathbf{x}_i^a = \mathbf{x}_i^f - \mathbf{K} \left(\mathbf{H}\mathbf{x}_i^f + \mathbf{d}_i \right),$$

where

$$\mathbf{K} = \mathbf{P}_f \mathbf{H}^T \left(\mathbf{H}\mathbf{P}_f \mathbf{H}^T + \mathbf{R} \right)^{-1}$$

is the Kalman gain matrix and

$$\mathbf{d}_i = \mathbf{y}(t_q) + \mathbf{r}_i$$

are randomly perturbed observations, i.e. $\mathbf{r}_i \sim \mathrm{N}(\mathbf{0}, \mathbf{R})$, $i = 1, \ldots, m$. Here $\mathbf{x}_i^f$ denotes the forecast ensemble value at time t_q and $\mathbf{x}_i^a$ the analyzed values for $i = 1, \ldots, m$.

The perturbed ensemble Kalman filter step motivates us to introduce the stochastic differential equation (SDE)

$$\mathrm{d}\mathbf{x}_i = -\mathbf{P}\mathbf{H}^T\mathbf{R}^{-1}\left([\mathbf{H}\mathbf{x}_i - \mathbf{y}(t_q)]\,\mathrm{d}s + \mathbf{R}^{1/2}\mathrm{d}\mathbf{W}_i \right) \tag{27}$$

in the ensemble members $\mathbf{x}_i(s)$, $i = 1, \ldots, m$, where $\mathbf{W}_i(s) \in \mathbb{R}^k$ denotes standard k-dimensional Brownian motion [23]. It should be noted that the covariance matrix $\mathbf{P}$ depends on the ensemble members $\mathbf{x}_i(s)$ and the noise is therefore of multiplicative nature. We use the Itô interpretation of (27) and the SDE is solved over the interval $s \in [0, 1]$ with initial condition $\mathbf{x}_i(0) = \mathbf{x}_i^f$ [23]. The analysed ensemble value is provided by $\mathbf{x}_i^a = \mathbf{x}_i(1)$.

We now verify that (27) is indeed the correct continuous stochastic EnKF formulation. A standard time discretization, which is compatible with the Itô interpretation of (27), is given by the forward Euler method

$$\mathbf{x}_i^{n+1} = \mathbf{x}_i^n - \mathbf{P}^n\mathbf{H}^T\mathbf{R}^{-1}\left(\left[\mathbf{H}\mathbf{x}_i^n - \mathbf{y}(t_q)\right] \Delta s + \mathbf{R}^{1/2}\mathbf{z}_i^n \Delta s^{1/2} \right), \tag{28}$$

where $\mathbf{z}_i^n \in \mathrm{N}(\mathbf{0}, \mathbf{I}_k)$ and $\Delta s > 0$ is the step-size. We now take the limit $m \to \infty$ and obtain the update

$$\overline{\mathbf{x}}^{n+1} = \overline{\mathbf{x}}^n - \mathbf{P}^n\mathbf{H}^T\mathbf{R}^{-1}\left(\mathbf{H}\overline{\mathbf{x}}^n - \mathbf{y}(t_q) \right) \Delta s$$

for the ensemble mean, which becomes

$$\frac{\mathrm{d}}{\mathrm{d}s}\overline{\mathbf{x}} = -\mathbf{P}\mathbf{H}^T\mathbf{R}^{-1}\left(\mathbf{H}\overline{\mathbf{x}} - \mathbf{y}(t_q) \right) \tag{29}$$

in the limit $\Delta s \to 0$. We next study the update of the covariance matrix $\mathbf{P}$. In terms of the ensemble deviations (17), Euler's method (28) leads to

$$\mathbf{Y}^{n+1} = \mathbf{Y}^n - \mathbf{P}^n\mathbf{H}^T\mathbf{R}^{-1}\left(\mathbf{H}\mathbf{Y}^n \Delta s + \mathbf{R}^{1/2}\mathbf{Z}^n \Delta s^{1/2} \right)$$

with random matrices

$$\mathbf{Z}^n = \left[\mathbf{z}_1^n | \mathbf{z}_2^n | \cdots | \mathbf{z}_m^n \right] - \overline{\mathbf{z}}^n \mathbf{e}^T \in \mathbb{R}^{n \times m}.$$

We use the definition

$$\mathbf{P}^{n+1} = \frac{1}{m-1}\mathbf{Y}^{n+1}(\mathbf{Y}^{n+1})^T$$

to obtain

$$\mathbf{P}^{n+1} = \mathbf{P}^n - 2\Delta s \mathbf{P}^n\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}\mathbf{P}^n + \frac{\Delta s}{m-1}\mathbf{P}^n\mathbf{H}^T\mathbf{R}^{-1/2}\mathbf{Z}^n(\mathbf{Z}^n)^T\mathbf{R}^{-1/2}\mathbf{H}\mathbf{P}^n,$$

where we have already dropped all terms linear in $\mathbf{Z}^n$ because of statistical independence to $\mathbf{Y}^n$, i.e., $\mathbf{Z}^n(\mathbf{Y}^n)^T/(m-1) \to \mathbf{0}$ as $m \to \infty$. Since also $\mathbf{Z}^n(\mathbf{Z}^n)^T/(m-1) \to \mathbf{I}_k$ as $m \to \infty$, we derive

$$\mathbf{P}^{n+1} = \mathbf{P}^n - \Delta s \mathbf{P}^n\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}\mathbf{P}^n$$

in the limit of $m \to \infty$ ensemble members. Hence, under the additional limit $\Delta s \to 0$, we recover the desired Riccati equation (16) for the ensemble covariance matrix $\mathbf{P}$.

In summary, we may conclude that the proposed SDE formulation (27) is consistent with the deterministic formulations (29) and (16) in the limit $m \to \infty$. Eqs. (29) and (16), on the other hand, are equivalent to the standard Kalman filter step.

Unlike (19), formulation (27) does not directly fit into the dynamical systems framework of Section 3. However, if we replace the continuity equation (12) by a Fokker-Planck equation [13]

$$\frac{\partial \pi}{\partial s} = -\nabla_{\mathbf{x}} \cdot \left(\pi \mathbf{M}^{-1}\nabla_{\mathbf{x}}\psi - \frac{1}{2}\mathbf{D}\nabla_{\mathbf{x}}\pi \right)$$

with an appropriate diffusion tensor $\mathbf{D} \in \mathbb{R}^{n \times n}$ and also replace the elliptic PDE (11) by

$$\nabla_{\mathbf{x}} \cdot \left(\pi \mathbf{M}^{-1}\nabla_{\mathbf{x}}\psi \right) = \pi \left(L - \mathbb{E}_\pi[L] \right) + \frac{1}{2}\nabla_{\mathbf{x}} \cdot \left(\mathbf{D}\nabla_{\mathbf{x}}\pi \right),$$

then we can treat stochastic particle filters within the dynamical systems framework of Section 3 with the ODE (13) now being replaced by the associated SDE [13]. In case of (27), the diffusion tensor is given by $\mathbf{D} = \mathbf{P}\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}\mathbf{P}$.

We finally mention two approaches to increase the robustness of ensemble Kalman filters with randomly perturbed observations. First, a complete ensemble Kalman filter formulation in terms of a mollification approach is given by

$$\mathrm{d}\mathbf{x}_i = f(\mathbf{x}_i,t)\mathrm{d}t - \sum_{q=1}^{M} \delta_\epsilon(t-t_q)\,\mathbf{P}\mathbf{H}^T\mathbf{R}^{-1}\left([\mathbf{H}\mathbf{x}_i - \mathbf{y}(t_q)]\,\mathrm{d}t + \mathbf{R}^{1/2}\mathrm{d}\mathbf{W}_i\right). \tag{30}$$

Secondly, multiple ensemble formulations have been introduced by [16,17] to make the EnKF with perturbed observations more robust with respect to random sampling errors. The key idea is to use different covariance matrices for each ensemble member. For example, the covariance matrix $\mathbf{P}_{ii}$ for the ith particle is given by

$$\mathbf{P}_{ii} = \frac{1}{m-2}\sum_{l\neq i}(\mathbf{x}_l - \overline{\mathbf{x}}_i)(\mathbf{x}_l - \overline{\mathbf{x}}_i)^T, \tag{31}$$

where

$$\overline{\mathbf{x}}_i = \frac{1}{m-1}\sum_{l\neq i}\mathbf{x}_l.$$

In other words, $\mathbf{P}_{ii}$ is the covariance matrix for the reduced ensemble $\mathbf{X}_i = (\mathbf{x}_1,\ldots,\mathbf{x}_{i-1},\mathbf{x}_{i+1},\ldots,\mathbf{x}_m)$, which is obtained from the full ensemble by eliminating its ith ensemble member. The multiple ensemble continuous Kalman filter with perturbed observations is now given by

$$\mathrm{d}\mathbf{x}_i = f(\mathbf{x}_i,t)\mathrm{d}t - \sum_{q=1}^{M} \delta_\epsilon(t-t_q)\,\mathbf{P}_{ii}\mathbf{H}^T\mathbf{R}^{-1}\left([\mathbf{H}\mathbf{x}_i - \mathbf{y}(t_q)]\,\mathrm{d}t + \mathbf{R}^{1/2}\mathrm{d}\mathbf{W}_i\right)$$

with $\mathbf{P}_{ii}$ given by (31).

5 Statistical models and generalized ensemble filters

We finally outline a general framework for data assimilation which generalizes the continuous EnKF approach discussed previously. We start from an empirical measure (4) with equal weights $\alpha_i = 1/m$ and particle locations $\mathbf{x}_i(t)$ being propagated under the mollified filter equations

$$\dot{\mathbf{x}}_i = f(\mathbf{x}_i,t) + \sum_{q=1}^{M}\delta_\epsilon(t-t_q)\,\mathbf{M}^{-1}\nabla_{\mathbf{x}}\psi_q(\mathbf{x};\mathbf{X})_{|\mathbf{x}=\mathbf{x}_i},$$

where the potentials ψ_q are obtained as numerical solutions to the associated elliptic PDE (13) over the interval $t \in [t_q-\epsilon, t_q+\epsilon]$. We use the notation $\psi_q(\mathbf{x};\mathbf{X})$ to indicate that the statistical model for the PDF π in (13) depends on the ensemble $\mathbf{X}$ and on the measurement at time $t = t_q$. There are many possible choices for numerical approximations. For example, one could approximate π by a Gaussian mixture model [1]. The identification of the appropriate mixture components from the ensemble $\mathbf{X}$ could be achieved by the expectation maximization (EM) algorithm [10,25]. The vector field $g = \mathbf{M}^{-1}\nabla_{\mathbf{x}}\psi_q$ could be approximated by an appropriate linear combination of continuous ensemble Kalman filter updates for the individual Gaussian mixture components. Other possible choices could, of course, be explored.

6 Summary

We have summarized previous work on continuous EnKF formulations by [5,6] and have extended it to the EnKF formulation with perturbed observations, which leads to stochastic differential equations in the ensemble members. We have also outlined a general dynamical systems framework for continuous ensemble filtering, which can serve as a base for deriving nonlinear ensemble/particle filters based on the numerical approximation of an elliptic PDE. The proposed approach avoids the random re-sampling step necessary for traditional particle filters [4].

Acknowledgment. We would like to thank Yann Brenier and Andrew Stuart for inspiring discussions on the subject of this paper.

References

1. Anderson, J., Anderson, S.: A Monte Carlo implementation of the nonlinear filtering problem to produce ensemble assimilations and forecasts. Mon. Wea. Rev. **127**, 2741–2758 (1999)
2. Ascher, U., van den Doel, K., Huang, H., Svaiter, B.: Gradient descent and fast artificial time integration. M2AN **43**, 689–708 (2009)
3. Ascher, U., Huang, H., Doel, K.v.d.: Artificial time integration. BIT **47**, 3–25 (2007)
4. Bain, A., Crisan, D.: Fundamentals of stochastic filtering, *Stochastic modelling and applied probability*, vol. 60. Springer-Verlag, New-York (2009)
5. Bergemann, K., Reich, S.: A localization technique for ensemble Kalman filters. Q. J. R. Meteorological Soc. **136**, 701–707 (2010)
6. Bergemann, K., Reich, S.: A mollified ensemble Kalman filter. Q. J. R. Meteorological Soc. **136**, 1636–1643 (2010)
7. Bloom, S., Takacs, L., Silva, A.D., Ledvina, D.: Data assimilation using incremental analysis updates. Q. J. R. Meteorological Soc. **124**, 1256–1271 (1996)
8. Burgers, G., van Leeuwen, P., Evensen, G.: On the analysis scheme in the ensemble Kalman filter. Mon. Wea. Rev. **126**, 1719–1724 (1998)
9. Crisan, D., Xiong, J.: Approximate McKean-Vlasov representation for a class of spdes. Stochastics **82**, 53–68 (2010)
10. Dempster, A., Laird, N., Rubin, D.: Maximum likelihood from incomplete data via the EM algorithm. J. Royal Statistical Soc. **39B**, 1–38 (1977)
11. Evensen, G.: Data assimilation. The ensemble Kalman filter. Springer-Verlag, New York (2006)
12. Friedrichs, K.: The identity of weak and strong extensions of differential operators. Trans. Am. Math. Soc. **55**, 132–151 (1944)
13. Gardiner, C.: Handbook on stochastic methods, 3rd edn. Springer-Verlag (2004)
14. Gaspari, G., Cohn, S.: Construction of correlation functions in two and three dimensions. Q. J. Royal Meteorological Soc. **125**, 723–757 (1999)
15. Hamill, T., Whitaker, J., Snyder, C.: Distance-dependent filtering of background covariance estimates in an ensemble Kalman filter. Mon. Wea. Rev. **129**, 2776–2790 (2001)
16. Houtekamer, P., Mitchell, H.: A sequential ensemble Kalman filter for atmospheric data assimilation. Mon. Wea. Rev. **129**, 123–136 (2001)
17. Houtekamer, P., Mitchell, H.: Ensemble Kalman filtering. Q. J. Royal Meteorological Soc. **131**, 3269–3289 (2005)
18. Kaipio, J., Somersalo, E.: Statistical and computational inverse problems. Springer-Verlag, New York (2005)
19. Kepert, J.: Covariance localisation and balance in an ensemble Kalman Filter. Q. J. Royal Meteorological Soc. **135**, 1157–1176 (2009)
20. Leeuwen, P.V.: Particle filtering in geophysical systems. Monthly Weather Review **137**, 4089–4114 (2009)
21. Lei, L., Stauffer, D.: A hybrid ensemble Kalman filter approach to data assimilation in a two-dimensional shallow-water model. In: 23rd Conference on Weather Analysis and Forecasting/19th Conference on Numerical Weather Prediction, AMS Conference proceedings, p. 9A.4. American Meteorological Society, Omaha, NE (2009)
22. Lewis, J., Lakshmivarahan, S., Dhall, S.: Dynamic data assimilation: A least squares approach. Cambridge University Press, Cambridge (2006)
23. Øksendal, B.: Stochastic Differential Equations, 5th edn. Springer-Verlag, Berlin-Heidelberg (2000)
24. Simon, D.: Optimal state estimation. John Wiley & Sons, Inc., New York (2006)
25. Smith, K.: Cluster ensemble Kalman filter. Tellus **59A**, 749–757 (2007)
26. Tippett, M., Anderson, J., Bishop, G., Hamill, T., Whitaker, J.: Ensemble square root filters. Mon. Wea. Rev. **131**, 1485–1490 (2003)
27. Villani, C.: Topics in Optimal Transportation. American Mathematical Society, Providence, Rhode Island, NY (2003)